International Journal of Psychological Studies; Vol. 10, No. 4; 2018
ISSN 1918-7211 E-ISSN 1918-722X
Published by Canadian Center of Science and Education

Pilot Study on the Relationship of Test Anxiety to Utilizing Self-Testing in Self-Regulated Learning

Sophia C. Weissgerber[1] & Marc-André Reinhard[1]

[1] Institute of Psychology, University of Kassel, Holländische Str. 36-38, 34127 Kassel, Germany

Correspondence: Sophia C. Weissgerber, Institut für Psychologie, Universität Kassel, Holländische Str. 36-38, 34127 Kassel, Germany. E-mail: sophiaweissgerber@gmail.com

Received: October 27, 2018 Accepted: November 25, 2018 Online Published: November 29, 2018
doi:10.5539/ijps.v10n4p95 URL: https://doi.org/10.5539/ijps.v10n4p95

Abstract

Whether or not test-anxious students leverage the power of testing as potent learning tool is unclear. In a pilot study we investigated the relation of test anxiety to the utilization of testing activities and academic performance in self-regulated learning. We hypothesized that increased cognitive test anxiety would relate to less self-reported use of self-testing in favor of repetition strategies, and thus relate in turn to lower self-reported exam grades. To examine this idea, we created a scale contrasting self-testing and repetition strategies, which showed sufficient preliminary reliability and validity. The findings support our notion with respect to the cognitive interference component of test anxiety: More interference was associated with less self-testing, and the link of interference with exam grades was fully mediated by the reported degree of self-testing. Although our findings are preliminary and limited in generalizability due to small sample size and lack of factor analysis of the created scale, the results hint at one potential reason why test-anxious students may underperform. Consequently, educators might motivate their test-anxious students to rely more on effective study approaches.

Keywords: test anxiety, self-testing, testing activities, testing effect, retrieval practice, exam grades

1. Introduction

Recitation, testing-effect, practice testing, retrieval practice, and test-enhanced learning are all labels pertaining to the same remarkable memory phenomenon: actively recalling rather than passively repeating information aids in long-term retention (e.g., McDaniel, Roediger, & McDermott, 2007). Consequently, retrieval beats restudying, and implications are clear: learners should apply re-testing instead of re-reading when information is to be remembered long-term. From a practical point of view, the phenomenon's effectivity (meta-analytic average weighted g = .61; Adesope, Trevisan, & Sundararajan, 2017; g = .50; Rowland, 2014) and broad applicability to educational settings makes it a promising candidate to boost learning throughout all educational sectors (Roediger & Karpicke, 2006).

The testing-effect is empirically-well supported and robust across varying study conditions (e.g. with and without feedback; Adesope, Trevisan, & Sundararajan, 2017). Although some researchers argue that testing is not beneficial for complex materials (Van Gog & Sweller, 2015), this concern is not shared by others (Karpicke & Aue, 2015). Even transfer and deductive inferences can be fostered under certain circumstances (Eglington & Kang, 2016; Smith & Karpicke, 2014). Moreover, testing holds indirect advantages, such as monitoring what has and has not been learned to guide future studying (Dunlosky, Kubat-Silman, & Hertzog, 2003), and it facilitates the acquisition of future material (Pastötter & Bäuml, 2014). Consequently, multiple researchers emphasize testing's value for learning, and they advocate a move away from a means of formal assessment to a potent learning tool (e.g. McDaniel, Roediger, & McDermott, 2007; Roediger, Agarwal, Kang, & Marsh, 2010).

The advantage of testing however can be attenuated when testing is accompanied with performance pressure during retrieval (Hinze & Rapp, 2014), and the mnemonic effectivity seems to be reduced for learners with test-anxiety ((Tse & Pu, 2012). *Test anxiety* as state and/or trait phenomenon occurs before and during test-taking (for a review see Zeidner, 2014) and encompasses negative *physio-emotional* and cognitive reactions like *worry* (concerns about one's test performance), *lack of confidence* (low self-expectations), and *interference* (intrusive and disturbing thoughts that disrupt task-specific information processing; cf. Hodapp, Rohrmann & Ringeisen, 2011). To which degree learners with test anxiety may choose to utilize the testing effect in their self-directed studying has not yet been directly investigated.

Research on the testing-effect has only recently begun to stimulate an interest in self-regulated aspects of learning. Within this new research avenue, the link of engaging in testing activities to academic achievement is underexplored. The integration of test anxiety's influence on the link between self-testing (Note 1) and achievement is pending. Thus, it is not only unclear to which degree students with higher levels of test anxiety use self-testing, but also how this translates to their academic achievement. Until now, this relationship was *indirectly* covered by broader research on learning strategies, without explicitly focusing on self-testing. The goal of the present study is to twofold. First, we aim to directly investigate whether test anxiety relates to the self-reported use of self-testing activities. Second, we examine whether and how test-anxiety may be related to exam performance via self-testing activities.

1.1 Effectivity and Use of Testing in Self-Regulated Contexts

The benefits of testing for personal learning may be crucially tied to the learner itself, because when organizing the 'how-to' of studying, the learner may insufficiently apply self-testing. Students learning activities often involve extensive highlighting and underlining, note taking, summary writing, and frequent re-reading (e.g. Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Hartwig & Dunlosky, 2012). The activities also encompass flashcards, quizzes, answering test questions, solving exercise problems, and completing practice exams (e.g. Karpicke, Butler, & Roediger, 2009). Both sets of activities can be seen to reflect a decision to either repeat or retrieve information, and of course, repetition and testing are often alternated. Moreover, retrieval or testing is often used to check one's understanding (Karpicke, Butler, & Roediger, 2009). Which of these study approaches do students report to utilize more, and how does it affect memory and performance?

Self-testing in flashcard-style is favored over re-reading for simple materials (Wissman, Rawson & Pyc, 2012) and a half-half split was effective for memory (Kornell & Son, 2009). Students often start self-testing later in the learning phase and stop prematurely, which negatively affected long-term memory, while repetitions were unrelated (Karpicke, 2009). For more complex content, re-rereading is the most frequent strategy (Karpicke, Butler, & Roediger, 2009). Self-testing (quizzes/practice testing) is also applied (Kornell & Bjork, 2007), although the amount varies up until a half-half split (Bartoszewski & Gurung, 2015; Susser & McCabe, 2013). Nevertheless, self-testing is never the prime method. While greater levels of self-testing with practice problems and quizzes were positively related to higher academic achievement (Gurung, 2005; Hartwig & Dunlosky, 2012; Morehead, Rhodes, & DeLozier, 2016), re-reading was not (Bartoszewski & Gurung, 2015). Note, the association of self-testing and academic achievement was weak; nevertheless, a reduced frequency of self-testing in favor of re-reading should decrease its merit for learning and achievement. The question then arises: How might test anxiety affect decisions for self-testing or repetition activities and achievement?

1.2 Test Anxiety, Study Strategies, and Academic Performance

If a learning situation is seen as threatening by test anxious-students, they may adopt rote learning (Fransson, 1977) and less deep learning (Helmke & Tuyet, 1999). Deep learning strategies foster a thorough understanding, while surface learning strategies aid rote learning (e. g. Biggs, 2003). Deep-learning can be linked to better grades (Zhang, 2000; but see also Campbell & Cabrera, 2014). Moreover, direct detrimental effects of test anxiety on performance have been shown by a plethora of studies (Hembree, 1988; Seipp, 1991; overall r = -.20.). Thereby, the meta-analytic association between performance and worry (r = -.29) is much stronger than between performance and emotionality (r = -.15; Seipp, 1991). The cognitive aspects of test-anxiety play an important role in the explanation of the negative anxiety-performance link: either short-term by working memory impairments in the test situation or by mechanism unfolding long-term like avoidance of evaluative situations and study deficits (cf. Carey, Devine, & Szücs, 2016), as described in the following.

1.3 Test Anxiety's Influence on the Use and Effectivity of Self-testing and Achievement in Self-Regulated Learning

Test anxiety is characterized by a propensity for higher degrees of worry, self-doubt, and interfering thoughts. If a highly test-anxious learner is engaging in self-testing, older interference accounts postulate cognitive impairments by intrusive worries evoked during the self-test situation, which hinder retrieval (Sarason, 1980) and attentional allocation (Wine, 1971). Newer accounts hold that trait test anxiety interferes with attentional control (Eysenck, Derakshan, Santos, & Calvo, 2007) and efficient cognitive processing in working memory, especially with inhibition (Berggren & Derakshan, 2013). Since self-testing entails retrieval practice and retrieval itself is impaired by unwarranted thoughts and worries due to working memory taxation (Hayes, Hirsch, & Mathews, 2008), retrieval practice itself becomes less effective, if not compensated by higher working-memory capacity (Tse & Pu, 2012; Owen, Stevenson, Hadwin, & Norgate, 2012). Simultaneously, interfering thoughts reduce the cognitive capacity available to work on the test problems at hand, acting like extraneous cognitive load leaving less working-memory capacity to the processing of the concurrent learning task (Richard, French, Keogh, & Carter, 2000; Sweller, Ayres, & Kalyuga, 2011). Moreover, test anxiety can negatively impact performance by greater distractibility to task−irrelevant material (Keogh & French, 200; Unsworth, Spillers, Brewer, & McMillan, 2011). Deficient attentional and cognitive processing as well as increased distractibility therefore reduce the efficiency in performing the ongoing self-testing activity. Cognitive processing inefficacy may go along with greater retrieval problems and difficulties to correctly answer test questions, which may stimulate premature termination of self-testing.

Experiences difficulties or incorrect answers can feed negatively into self-perceptions (Schunk & Ga, 1982), triggering avoidance motivations with respect to the failure-inducing activity (Heimpel, Elliot, & Wood, 2006; McGregor & Elliot, 2002). The result would be a lower propensity to engage in further self-testing and to switch to other learning strategies, such as repetition (Wittmaier, 1972). For example, some studies have demonstrated that highly test anxious learner study extensively but often use highly repetitive activities that are less effective (Cassady, 2004; Culler & Holohan, 1980).

The adoption of less effective study habits intersects with deficit accounts (Birenbaum, & Nasser, 1994; Naveh-Benjamin, McKeachie & Lin, 1987), which emphasize impairments related to domain-specific abilities and study habits in the learning phase (for a discussion of deficits during test preparation phase see also Cassady, 2004). The initial acquisition of the learning content is lessened due to cognitive interference when uncompensated by higher abilities, impairing conceptual encoding and representations of the learning materials (Mueller & Courtois, 1980; Naveh-Benjamin, 1991). The consequence is a deficient mastery of the content (Birenbaum & Pinku, 1997), which would be revealed in future self-tests. As students often use self-testing to diagnose their learning instead of improving it (Kornell & Son, 2009), the subsequent utilization of self-testing would heighten one's meta-cognitive awareness of poor learning. For test-anxious students, increased meta-cognitive awareness could heightening the saliency of failure, which may be experienced as an aversive cognition. The meta-cognitive awareness of a potentially upcoming failure intensifies the threat-potential to one's self-efficacy and thus lowers motivation to self-test to protect one's self-concept (Hancock, 2010; Roick & Ringeisen, 2017). In this respect, learners with lower competence expectations for successful performance tend to adopt avoidance goals perceiving evaluative situations as threatening events to be avoided (Cassady, 2004). "The perception of a testing event as threatening impacts student performance in the preparation phase generally through the engagement of ineffective preparatory strategies driven by feelings of helplessness, attempts to avoid or reduce the occurrence of failure […]. Naturally, ineffective strategies during the preparation phase will lead to eventual failure in test performance." (Cassady, 2004, p. 572).

As the previous literature on self-testing and memory suggests, premature termination and insufficient use or processing efficiency as well as avoidance of self-testing lowers its effect on memory and performance. This is especially detrimental, as highly test-anxious students would need, as compensation, higher frequencies of self-testing rather than less. A deficient mastery of the learning contents may perpetuate lower domain-specific abilities (Cassady & Johnson, 2002), feeding into further avoidance of test events.

In summary, the previous argumentation stresses deficits related to the cognitive component of test-anxiety (interference, worries, low confidence). On the one hand, it suggests that test anxiety may be associated with cognitive deficits during self-testing. On the other hand, it suggests that text anxiety is associated with further deficits in the study approach (for example, an avoidance of self-testing, premature termination of self-testing and the use of repetition strategies). Therefore, the effectivity of self-testing and the degree of usage as a study strategy should be related to the cognitive component of test anxiety. Less usage and an underutilization or adoption of alternative repetition approaches should be negatively related to exam performance. We hypothesize the following: 1) students with higher levels of cognitive test anxiety should be less prone to use self-testing as personal study strategy; 2) students with higher levels of cognitive test anxiety should have worse exams grade due to decreased self-testing (mediation).

To examine our hypotheses, we created a scale that covers the various ways (e.g., collaborative testing) that students might apply testing when learning; there was no ready-to-use instrument since the exploration of testing in self-regulated learning is a new research line. Thus, like the few existing surveys on the use of testing as personal learning tool, we had to create our own items. We constructed the items in line with the scheme of laboratory research, contrasting the testing condition with the repetition condition (e.g., re-reading or summaries). Although we were inspired by existing items in previous surveys (see Introduction), we had to compile or own, because not all formulations were specific enough, for example, 'make and use flashcards,' or 'do practice problems' (Susser & McCabe, 2013) does not specify the use as a test tool. 'Use' could also be understood as inscribing flashcards with short summaries to re-read. ‘Do’ does not exclude the option of preparing with worked-out practice problems that are already solved (like case studies).

2. Method

2.1 Participants and Procedure

Participants (N = 81; 27 males; average age of 23.62, SD = 4.97; 60% were in the 2nd semester/freshman) signed up for a study advertised to be about learning strategies, personality, and exam grades via the internal study platform (SONA) at the University of Kassel (Germany). After registration, participants received an internal online link to the survey platform, where participants completed the survey. Almost all participants were psychology students (n = 73) and were compensated with partial course credit.

At the beginning, participants were asked to list all exams they completed in the previous semester, which was intended to help students to remember their study behaviors; they were also asked to record the achieved exam grades at the end of the study. We then inquired about learning strategies via our constructed self-testing scale (ST-scale, see below), followed by other items (e.g., generation, spaced/interleaved learning), which we included for distraction (and for reasons not relevant to the present study: replication of a previous published relationship of NFC and goals orientations with learning activities by the first author). Thereafter, items of the German Learning Strategy Inventory LIST (Schiefele & Wild, 1994) followed, which served to validate our newly constructed scale. We embedded the measurement of students’ test anxiety (PAF; Hodapp, Rohrmann & Ringeisen, 2011) within other questionnaires (e.g. need for cognition, goal orientations). Finally, we listed all exams participants had indicated at the beginning of the study. We emphasized that it is important for the study to completely and truthfully enter received exam grades, or to indicate if the grade was not yet known to them. The survey concluded with demographic questions.

2.2 Measures

Self-testing scale (STS). We constructed various items to assess the degree to which students used self-testing or relied on re-reading/non-testing activities (see Table 1 in the Appendix for the items; Cronbach’s α = .76). For example, “I used quiz and exercise questions at the end of a chapter to… A) re-read the respective text passage B) simulate an exam.” All items were measured on 7-point scales (-3 = only A, -2 = mostly A, -1 = rather A, 0 = A equally as B, 1 = rather B, 2 = mostly B, 3 = only B) with an additional answer option of ‘neither A nor B’ (coded as missing). The items were constructed to contrast two learning strategies: A(= repetition/non-testing) and B(= self-testing) mirroring experimental research contrasting rereading or other activities (e.g. note-taking: Rummer, Schweppe, Gerst, & Wagner, 2017) with testing (cf. Kornell, Rabelo, & Klein, 2012). Due to some of the items being reversely formulated, they were recoded in a way that -3 to -1 represented the degree of re-reading/non-testing, while 1 to 3 represented the degree of self-testing. Zero served as the midpoint and represented the use of both strategies to a similar extent. The items were presented in randomized order in one block.

German Learning Strategy Inventory (LIST). The inventory (Wild & Schiefele, 1994) distinguishes three groups of strategies – cognitive, metacognitive, and resource management strategies, which are comprised of several subscales: organization (α = .83), elaboration (α = .81), repetition (α = .63), planning (α = .77), monitoring (α = .66), regulation (α = .78), effort (α = .70), time management (α = .86), and joint learning (α = .87). All items were measured on 5-point scales (1 = very seldom, 5 = very often). Psychometric properties were described by Wild and Schiefele (1994).

Test anxiety (PAF). We used the German version of the test anxiety inventory (PAF; Hodapp, Rohrmann & Ringeisen, 2011), which encompasses four facets: Lack of confidence (α = .89), worry (α = .82), emotionality (α = .89), interference (α = .83). All items were answered on 4-point scales (1 = almost never to 4 = almost always). The psychometric properties are described by Hodapp, Rohrmann, and Ringeisen (2011).

Exam grades. Exam grades ranged from 1.0 (A^+), 1.3(A), 1,7 (A-) and so forth to 4.0 (D^-); 5.0 (F) as a failing grade was recoded to 4.3 to preserve the mathematical distance between grades.

For each person, we computed the mean grade of all reported exam grades to represent the averaged exam performance in this semester. Note, most of the samples had completed three (n = 39) or four (n =11) exams, while 13 students had not completed any exams in the previous semester and therefore did not report any exam grades. The most frequently reported exams were developmental psychology (n = 38), general psychology (n = 42), and quantitative methods (n = 43).

3. Results

Prior to testing our hypotheses, we examined the validity of our newly created scale on the use of self-testing versus repetition/non-testing. We correlated (two-tailed) the sub-scales of the established learning strategy inventory (LIST; Wild & Schiefele, 1994) with our created instrument. It would speak to the scale's validity to observe a) a negative relationship with the LIST-subscale of repetition and b) a positive relationship to the sub-scales of monitoring and regulation, because students often use testing for monitoring reasons (Kornell & Son, 2009), which can help to regulate their learning (Fernandez & Jamet, 2017) as well as c) no relationship to organization, planning, or study time, which are aspects that tap more into management and scheduling. The results support the convergent and discriminant validity of our scale (see Table 1). We found the negative correlation with repetition (r = .37, p < .001) and the anticipated positive correlations with monitoring (r = .48, p < .001) and regulation (r = .23, p < .05). Our repetition vs. self-testing scale was not related to organization, planning, or study time, all ps = ns. Like the monitoring (r = -.46, p < .001). The regulation (r = -.43, p < .001) sub-scale (related to testing aspects) of the established inventory and our ST-scale (r = -.42, p < .001) were both associated with the averaged exam performance as outcome variable; the coefficients are comparable in size and direction, supporting the criterion validity.

Table 1. Bivariate Correlations, Means, and Standard Deviations of the Self-Testing Scale (STS) and the Learning Strategy Inventory (LIST)

	Variable	1	2	3	4	5	6	7	8	9	10	11	*M*	*SD*
1.	STS (*N* = 81)	--											0.46	0.75
2.	Organization (*N* = 81)	.10	--										3.15	0.82
3.	Elaboration (*N* = 76)	.32**	.16	--									3.45	0.63
4.	Repetition (*N* = 76)	-.35**	.02	-.25*	--								3.36	0.60
5.	Planning (*N* = 81)	.06	.26*	.07	.18	--							3.37	0.67
6.	Monitoring (*N* = 76)	.47***	.35**	.26*	-.12	.06	--						3.26	0.61
7.	Regulation (*N* = 81)	.24*	.40***	.25*	.14	.15	.56***	--					3.83	0.55
8.	Effort (*N* = 76)	.22†	.27*	.20†	.14	.14	.42***	.59***	--				3.58	0.61
9.	Time (*N* = 76)	.00	.14	.01	.09	.59***	.07	.12	.19†	--			2.67	1.00
10.	Peer learning (*N* = 76)	.22†	.23*	.09	-.26*	.26*	.36***	.14	.12	.28*	--		2.89	0.76
11.	Grades (*N* = 64)	-.42***	-.07	-.17	.21	-.34**	-.46***	-.43***	-.40**	-.47***	-.43***	--	2.30	0.85

Note. Bivariate correlation coefficients are based on Spearman's Rho. † p = .10, * p < .05, ** p < .01, *** p < .001 (two-tailed). Grades: higher numbers denote worse averaged exam grades.

Given a satisfactory reliability (Note 2) score of α = .76 and support for internal validity, we scrutinized the relationship between test anxiety, learning strategies, and exam performances. Table 2 shows the correlations (two-tailed) of the test-anxiety inventory and all test anxiety sub-scales with our self-testing scale (ST-scale) and its association with the averaged exam grades achieved. In line with Hypothesis 1, test anxiety indeed related to the degree to which students reported to utilize self-testing as a learning strategy: Overall increased levels of test-anxiety covary with a decreased use of self-testing in favor of repetition/non-testing strategies, r = -.23, p < .05. When considering the sub-dimensions of test anxiety, as expected, the interference component was moderately and negatively related to the ST-scale, r = -.40, p < .001, explaining 20% of the variance in the degree of reported self-testing. The more interference is experienced, the less self-testing is chosen as a learning strategy. The subscales worry, $r = -.15$, $p = .18$, and lack of confidence, $r = -.18$, $p = .12$, were not related to the ST-scale. On a descriptive level, however, they also show a similar trend. The subscale of emotionality was not associated with the ST-scale, $r = -.04$, $p = .72$.

Table 2. Bivariate Correlations, Means, and Standard Deviations of Test Anxiety Inventory, Test Anxiety Sub-scales, Self-Testing Scale (STS) and Averaged Exam Grades

	Variable	1	2		3	4	5	6	7	*M*	*SD*
1.	Test anxiety inventory [a] (*N* = 80)	--								2.34	0.36
2.	Test anxiety inventory [b] (*N* = 80)	.94***	--							2.22	0.53
3.	Lack of confidence (*N* = 80)	.43***	.63***	--						2.43	0.70
4.	Worry (*N* = 80)	.71***	69***	.28*	.--					2.70	0.64
5.	Emotionality (*N* = 80)	.77***	.83***	.56***	.34**	--				2.08	0.76
6.	Interference (*N* = 80)	.64***	.75***	.63***	.29**	.49***	--			1.88	0.63
7.	STS (*N* = 80)	-.23*	-.24*	-.18	-.15	-.04	-.40***	--		0.46	0.75
8.	Grades (*N* = 64)	-.02	.02	.13	-.05	-.14	.26*	-.42***	--	2.30	0.85

Note. Bivariate correlation coefficients are based on Spearman's Rho. * $p < .05$, ** $p < .01$, *** $p < .001$ (two-tailed). Grades: higher numbers denote worse averaged exam grades. [a] Test anxiety inventory (Hodapp, Rohrmann & Ringeisen, 2011) with all 4 sub-scales. [b] Test anxiety inventory without the sub-scale lack of confidence, because some models on the link of anxiety and performance see low competence as a cause of test anxiety and low performance (Carey, Hill, Devine, & Szücs, 2016).

When considering the relationship of test anxiety and learning approaches to the performance criteria, we found that the undifferentiated test anxiety inventory was unrelated to the averaged achieved exam grades in the previous semester, $r = -.02$, $p = .87$. Notably, when considering all subscales, only cognitive interference was associated with exam performances, $r = .26$, $p < .05$. As postulated, the more interference is experienced during exam preparation, the worse the averaged exam performances. However, unlike expectations, the cognitive components of worry and lack of confidence showed no associations. Thus, only the interference component will be considered for Hypothesis 2.

Hypothesis 2 assumes a mediational model by which the cognitive aspects of test anxiety, the ST-scale, and the exam performances are all intertwined. In accordance with this notion, a regression analyses with anxiety-driven interference should predict the degree to which self-testing is reported as a learning strategy during the exam preparation. In turn, the used learning approach should predict the average achieved exam performances (see Figure 1). Conducting regression analyses and a Zobel-test (Note 3), we examined whether a full or partial mediation model was supported.

First, we tested all direct effects: The direct effect of interference on the ST-scale (mediator) was significant, $B = -.43$, $t(78) = -3.57$, $p < .001$, 95% CI [-.68, -.19], $sr^2 = .14$; Model 1, $F(1, 78) = 12.77$, $p < .001$, $R^2 = .14$. The direct effect of the ST-scale (mediator) on averaged exam performances was also significant, $B = -.46$, $t(62) = -3.37$, $p = .001$, 95% CI [-.74, -.19], $sr^2 = .15$; Model 2, $F(1, 62) = 11.37$, $p = .001$, $R^2 = .16$. Similarly, the direct effect of interference on averaged exam grade was significant, too, $B = .35$, $t(62) = 2.16$, $p < .05$, 95% CI [.03, .68], $sr^2 = .07$; Model 3, $F(1, 62) = 4.68$, $p < .05$, $R^2 = .07$.

Second, we tested whether a full or partial mediation model was supported by looking at the direct effect of the interference sub-scale when controlling for the mediator ST-scale. We found that the direct effect of the interference sub-scale on averaged exam performance is rendered non-significant, $B = .19$, $t(61) = 1.14$, $p = .26$, 95% CI [-.14, .52], sr^2=.02, Model 4, $F(2, 61) = 6.37$, $p < .01$, R^2=.17, when controlling for the mediator ST-scale in the regression model. In contrast, the effect of the mediator ST-scale remained significant, $B = -.40$, $t(61) = 2.75$, $p < .01$, 95% CI [-.70, -.11], sr^2=.10, Model 4, $F(2, 61) = 6.37$, $p < .01$, R^2=.17.

The pattern of results suggested full mediation, supported by the Sobel-test statistic (Soper, 2018), $Z_{Sobel} = 2.24$, $p = .026$ (two-tailed). Higher reported levels of anxiety-induced interference are associated with lower degrees of reported self-testing, which is associated with worse academic achievement (represented by averaged exam grades in this semester). Consequently, Hypothesis 2 was partly supported with respect to the cognitive interference component (but not worry or lack of confidence).

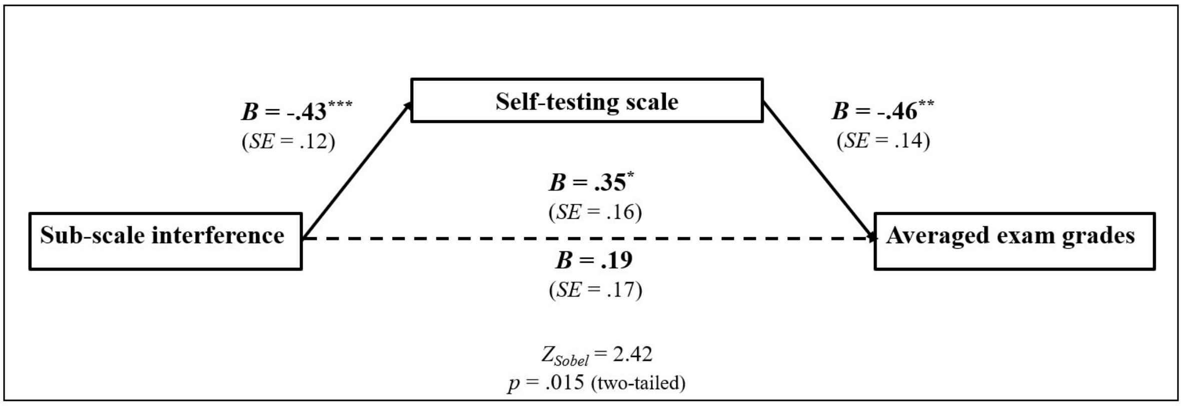

Figure 1. Full mediation of test anxiety induced interference on averaged exam grades by self-testing. Self-testing scale: higher number represent more self-testing. Averaged exam grade left in its original form: higher numbers denote to worse exam grades. With respect to averaged exam grades, $n = 62$.

4. Discussion

Is test anxiety related to the degree of utilizing self-testing strategies (e.g., practice problems, mock exams, collaborative testing, retrieval practice) and achieved exam grades in self-regulated learning? We argued that cognitive aspects of test anxiety go along with multiple deficits while self-testing (cf. interference accounts; e.g. Eysenck, Derakshan, Santos, & Calvo, 2007) and throughout the learning phase (cf. deficit models, e.g. Birenbaum, & Nasser, 1994), in line with processual accounts (cf. Cassady, 2004). In short, cognitive interference while self-testing may reduce processing efficiency, for example by imposing an unfavorable cognitive load on working memory or deficient attentional regulation. Retrieval difficulties and errors as self-concept threatening meta-cognitive feedback of may lead to premature termination and an avoidance of self-testing. Moreover, worries and low competence beliefs in evaluative situations also trigger avoidance reactions of the anxiety inducing activity (self-testing). Consequently, students are less prone to use self-testing and may switch to less anxiety-evoking strategies like repetition, which may not prove as effective.

Based on the rationale above, we predicted that students with higher levels of cognitive test anxiety should report less utilization of self-testing as a personal learning strategy (H1) and as a result, report worse exams grades (H2). Supporting our notion, we found that test-anxious students with higher levels of cognitive interference reported less usage of self-testing and lower averaged exam grades. This relationship was fully mediated, meaning the reduced frequency of self-testing relative to less efficient activities like repetition accounted for the lower exam grades by higher anxiety-related interference. The reduced frequency of self-testing in favor of other activities may be one mechanism underlying the negative test anxiety-performance link.

We did not find a significant relationship between self-testing and the cognitive component lack of confidence nor between self-testing and the cognitive component of worry. This may be due to the small sample size and/or that these components of test anxiety may play a lesser role with less sizable influences. The small sample size is also a caveat with respect to the newly constructed scale, which was not sufficient to conduct factor analyses to explore its structure. However, the scale preliminarily shows sufficient reliability and good discriminant and convergent validity. The obtained relationship of our scale with averaged exam grades can also be interpreted as an indication of predictive validity. A deficient instrument would not meaningfully relate to the expected relationships. Although the scale has room for improvement, it demonstrates construct validity. The creation of this instrument was necessary, because no instrument exists (to our knowledge) that specifically focused on a variety of self-testing activities (e.g., quizzes/flashcards/test problems/mock exams/practice tests/collective testing) relative to repetition/non-testing activities students may use when learning.

An aspect worth mentioning concerns the replication of the association between self-testing and academic achievement in real-world educational settings. So far, few studies have assessed how self-testing with quizzes/practice problems and flashcards related to GPA (see introduction: Hartwig & Dunlosky, 2012; Morehead, Rhodes, & DeLozier, 2016). However, GPA reflects cumulated achievements for various academic performances, like essays, for which self-testing may not be useful. It is possible, therefore, that the effect of self-testing on GPA is underestimated. Actual exam grades may be more accurate, which to date, have barely been examined. In line with this idea, we found a stronger association (r^2 = .18) when comparing our obtained association to the previous studies with GPA (Hartwig & Dunlosky, 2012: $r^2 \approx .08$; Morehead et al. 2016: $r^2 \approx .03$) and exam performance in introductory courses (Gurung, 2005; Bartoszewski & Gurung, 2015; both $r^2 \approx .08$). A comparison of both correlations indicates a small difference of Cohen's q = .17 (Cohen, 1988). We sampled a wider range of exam performances and included multiple aspects (e.g., collaborative testing; Wissman & Rawson, 2018), with more specific questions about the use of self-testing. This may have potentially contributed to the improved association. However, it also could be simply methodologic (e.g., we have a small and homogenous sample of a selective group).

The (meta-analytic) weighted mean effect size g of testing compared to re-studying is .51, stemming from regulated contexts like laboratory or classroom (Adesope et al., 2017). A transformation of our obtained r to d yields ≈ .93 (Cohen, 1988). Interestingly, the effect of self-testing may be higher in self-regulated learning contexts than in regulated contexts. To put it another way, for personal learning, self-testing could be even more important, but this is preliminary.

Another flaw of the present study concerns the correlational nature and the reliance on self-report: Students' retrospective judgments about learning activities may not accurately represent actual learning but remembered learning. Maybe test anxiety influences the remembered degree of self-testing and not the actual use. It is therefore necessary to observe actual choices of test-anxious students and effects on performance criteria, especially with respect to establishing the causal influence of the mediator. Nevertheless, the study is an important, albeit limited, attempt to explore whether test anxiety negatively influences the use of self-testing activities and academic achievements in self-regulated learning.

This notion is interesting for theoretical and practical reasons, for example, theoretically with respect to meta-cognitions. While students often tend towards overconfidence, this bias seems to be reduced in test-anxious students leaning towards under confidence (Miesner & Maki, 2007). This may have interesting consequences for calibration accuracy and ongoing meta-cognitive processes and subsequent self-regulation decisions (e.g., change study strategy, spent less time, etc.). Think-aloud protocols could help to illuminate dynamics between meta-cognitions and study decisions (cf. Fernandez & Jamet, 2017). These self-regulation processes may be one cause for the well-established negative test anxiety-performance link, which speaks to the practical reasons: advising students with test anxiety about how to improve their test preparation and their test-taking skills.

Intervention could entail encouragement to utilize self-testing for those with deficient study habits or desensitization/habituation to testing situations for a reduction of cognitive interference (Mowbray, 2012; Naveh-Benjamin, 1991). Promising training programs, coaching, or therapies in groups or individual settings exist and use biofeedback, relaxation, or skill-building interventions (for an overview see Von der Embse, Barterian & Segool, 2013), as well as mindfulness techniques (Carsley, Heath, & Fajnerova, 2015).

Could the decreased motivation to use self-testing on their own by test-anxious students be counteracted by making testing an integral part of a course? In the classroom, Agarwal and collequgues (2016) showed that practice testing can be especially advantageous for students with lower working memory. While Brewer and Unsworth (2012) did not obtain a moderation by working memory capacity, they found that weaker students with lower memory abilities and lower general-fluid intelligence profited more from testing (but see also Pan, Pashler, Potter, & Rickard, 2015). Moreover, Messineo, Gentile, and Allegra (2015) showed that the negative link of test anxiety on performance in class is stronger for re-reading materials; thus, re-testing acted as a buffer for more test anxious students. Testing must take place, however, under low-threat conditions, such as low performance pressure (Khanna & Cortese, 2016).

It may be worthwhile to offer voluntary, ungraded self-testing options, for example, by e-learning as an integral part of a course, and to encourage its usage by those with a propensity to avoid self-testing. In this respect, the benefits of self-testing should be emphasized for test-anxious students: Repeated low-stakes testing can help to habituate to the testing situation and reduce test anxiety (e.g. Nyroos, Schéle & Wiklund-Hörnqvist, 2016). In conclusion, students and instructors would both be well-advised to increase the frequency of testing activities.

References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87,* 659–701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K., Finley, J. R., Rose, N. S. & Roediger, H. L. (2016). 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

Bartoszewski, B. L., & Gurung, R. A. R. (2015). Comparing the relationship of learning techniques and exam score. *Scholarship of Teaching and Learning in Psychology*, *1*(3), 219–228. https://doi.org/10.1037/stl0000036

Berggren, N., & Derakshan, N. (2013). Attentional control deficits in trait anxiety: why you see them and why you don't. *Bio Psychology, 92(3),* 440–446. https://doi.org/10.1016/j.biopsycho.2012.03.007

Biggs, J. (2003). *Teaching for quality learning at University* (2nd ed.). London: The Society for Research into Higher Education & Open University Press.

Birenbaum, M., & Nasser, F. (1994). On the relationship between test anxiety and performance. *Measurement and Evaluation in Counseling and Development, 27,* 293–301. https://doi.org/10.1016/j.sbspro.2011.04.372

Birenbaum, M. & Pinku, P. (1997). Effects of test anxiety, information organization, and testing situation on performance on two test formats. *Contemporary Educational Psychology, 22,* 23–38. https://doi.org/10.1080/096582196388960

Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language, 66,* 407–415. https://doi.org/10.1016/j.jml.2011.12.009

Campbell, C. M., & Cabrera, A. F. (2014). Making the mark: Are grades and deep learning related? *Research in Higher Education*, *55*(5), 494–507. https://doi.org/10.1007/s11162-013-9321-8

Carey, E., Hill, F., Devine, A., & Szücs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Frontiers in Psychology*, *6*, 1–6. https://doi.org/10.3389/fpsyg.2015.01987

Carsley, D., N, Heath, L. & Fajnerova, S. (2015). Effectiveness of a classroom mindfulness coloring activity for test anxiety in children. *Journal of Applied School Psychology, 31,* 239–255. https://doi.org/10.1080/15377903.2015.1056925

Cassady, J. C. (2004). The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure. *Applied Cognitive Psychology, 18*(3), 311–325. https://doi.org/10.1002/acp.968.

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, *27*(2), 270–295. https://doi.org/10.1006/ceps.2001.1094.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Dunlosky, J., Kubat-Silman, A. K., & Hertzog, C. (2003). Training monitoring skills improves older adults' self-paced associative learning. *Psychology and Aging*, *18*(2), 340–345. https://doi.org/10.1037/0882-7974.18

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*(1), 4–58. https://doi.org/10.1177/1529100612453266

Eglington, L. G., & Kang, S. H. K. (2016). Retrieval practice benefits deductive inference. *Educational Psychology Review, 30(1),* 215–228. https://doi.org/10.1007/s10648-016-9386-y

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*, 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Fernandez, J. & Jamet, E. (2017). Extending the testing effect to self-regulated learning. *Metacognition Learning, 12,* 131–156. https://doi.org/10.1007/s11409-016-9163-9

Fransson, A. (1977). On qualitative differences in learning: IV-effects of intrinsic motivation and extrinsic test anxiety on process and outcome. *British Journal of Educational Psychology*, *47*(3), 244–257. https://doi.org/10.1111/j.2044-8279.1977.tb02353.x

Gurung, R. A. R. (2005). How do students really study (and does it matter)? *Teaching of Psychology, 32,* 367–372.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis.* (2nd Ed.). New York: The Guilford Press.

Hayes, S., Hirsch, C., & Mathews, A. (2008). Restriction of working memory capacity during worry. *Journal of Abnormal Psychology, 117,* 712–717. https://doi.org/10.1037/a0012908.

Hancock, D. R. (2010). Effects of test anxiety and evaluative threat on students' achievement and motivation. *The Journal of Educational Research*, *94*(5), 284–290. https://doi.org/10.1080/00220670109598764

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review*, *19*(1), 126–134. https://doi.org/10.3758/s13423-011-0181-y

Heimpel, S. A., Elliot, A. J., & Wood, J. V. (2006). Basic personality dispositions, self-esteem, and personal goals: An approach-avoidance analysis. *Journal of Personality*, *74*(5), 1293–1320. https://doi.org/10.1111/j.1467-6494.2006.00410.x

Helmke, A., & Tuyet, V. T. A. (1999). Do Asian and Western students learn in different way? An empirical study on motivation, study time, and learning strategies of German and Vietnamese university students. *Asia Pacific Journal of Education*, *19*(2), 30–44. https://doi.org/10.1080/0218879990190204

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, *58*(1), 47–77. https://doi.org/10.3102/00346543058001047

Hinze, S. R., & Rapp, D. N. (2014). Retrieval (sometimes) enhances learning: Performance pressure reduces the benefits of retrieval practice. *Applied Cognitive Psychology, 28*(4), 597–606. https://doi.org/10.1002/acp.3032

Hodapp, V., Rohrmann, S. and Ringeisen, T. (2011). *Prüfungsangstfragebogen PAF-S. Hogrefe Verlag, Gottingen.*

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General*, *138*(4), 469–486. https://doi.org/10.1037/a0017341

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., & 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

Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practise retrieval when they study on their own? *Memory*, *17*(4), 471–479. https://doi.org/10.1080/09658210802647009

Khanna, M. M., & Cortese, M. J. (2016). The benefits of quizzing in content-focused versus skills-focused courses. *Scholarship of Teaching and Learning in Psychology, 2(1),* 87–97. https://doi.org/10.1002/acp.968

Keogh, E., Bond, F. W., French, C. C., Richards, A. & Davis, R. E. (2004). Test anxiety, susceptibility to distraction and examination performance. *Anxiety, Stress & Coping, 17(3)*, 241–252. https://doi.org/10.1080/10615300410001703472

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. https://doi.org/10.12691/education-5-8-5

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory*, *17*(5), 493–501. https://doi.org/10.1080/09658210902832915

McDaniel, M. A., Roediger, H. L., & McDermott, K. B. (2007). Generalizing test-enhanced learning from the laboratory to the classroom. *Psychonomic Bulletin & Review*, *14*(2), 200–206. https://doi.org/10.3758/BF03194052

McGregor, H. A., & Elliot, A. J. (2002). Achievement goals as predictors of achievement-relevant processes prior to task engagement. *Journal of Educational Psychology, 94(2),* 381–395.

https://doi.org/10.1037/h0043445

Messineo, L., Gentile, M., & Allegra, M. (2015). Test-enhanced learning: analysis of an experience with undergraduate nursing students. *BMC Medical Education, 15,* 182. https://doi.org/10.1186/s12909-015-0464

Miesner, M. T. & Maki, R. H. (2007). The role of test anxiety in absolute and relative metacomprehension accuracy. *European Journal of Cognitive Psychology,* 650–670. https://doi.org/10.1080/09541440701326196

Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, *24*(2), 257–271. https://doi.org/10.1080/09658211.2014.1001992

Mowbray, T. (2012). Working memory, test anxiety and effective interventions: a review. *The Australian Educational and Developmental Psychologist*, *29*(02), 141–156. https://doi.org/10.1017/edp.2012.16

Mueller, J. H., & Courtois, M. R. (1980). Test anxiety and breadth of encoding experiences in free recall. *Journal of Research in Personality*, *14*(4), 458–466. https://doi.org/10.1016/0092-6566(80)90004-5

Naveh-Benjamin, M. (1991). A comparison of training programs intended for different types of test-anxious students: Further support for an information-processing model. *Journal of Educational Psychology*, *83*(1), 134–139. https://doi.org/10.1037/0022-0663.83.1.134

Naveh-Benjamin, M., McKeachie, W. J., & Lin, Y.-g. (1987). Two types of test-anxious students: Support for an information processing model. *Journal of Educational Psychology*, *79*(2), 131–136. https://doi.org/10.1037/0022-0663.79.2.131

Nyroos, M., Schele, I., & Wiklund-Hornqvist, C. (2016). Implementing Test Enhanced Learning: Swedish Teacher Students' Perception of Quizzing. *International Journal of Higher Education, 5(4),* 1–12. https://doi.org/10.5430/ijhe.v5n4p1

Owens, M., Stevenson, J., Hadwin, J. A., & Norgate, R. (2014). When does anxiety help or hinder cognitive test performance? The role of working memory capacity. *British Journal of Psychology, 105*(1), 92-101. https://doi.org/10.1111/bjop.12009

Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology*, *5*, 1–5. https://doi.org/10.3389/fpsyg.2014.00286.

Pan, S. C., Pashler, H., Potter, Z. E., Rickard, T. C. (2015). Testing enhances learning across a range of episodic memory abilities. *Journal of Memory and Language, 83,* 53–61. https://doi.org/10.1016/j.jml.2015.04.001

Richards, A., French, C. C., Keogh, E. & Carter, C. (2000). Test-Anxiety, inferential reasoning and working memory load. *Anxiety, Stress & Coping, 13(1),* 87-109. https://doi.org/10.1080/10615800008248335

Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied*, *17*(4), 382–395. https://doi.org/10.1037/a0026252

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

Roick, J., & Ringeisen, T. (2017). Self-efficacy, test anxiety, and academic success: A longitudinal validation. *International Journal of Educational Research*, *83*, 84–93. https://doi.org/10.1016/j.ijer.2016.12.006

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

Sarason, I. G. (1980). *Test anxiety*. Hillsdale, N.J.: L. Erlbaum Associates.

Schunk, H. D.,& Gaa, J. P. (1981). Goal-setting influence on learning and self-evaluation. *The Journal of Classroom Interaction,16*(2),38-44. https://doi.org/10.1080/10573560390143094

Seipp, B. (1991). Anxiety and academic performance: A meta-analysis of findings. *Anxiety Research*, *4*(1), 27–41. https://doi.org/10.1080/08917779108248762

Smith, M. A., & Karpicke, J. D. (2014). Retrieval practice with short-answer, multiple-choice, and hybrid tests. *Memory, 22,* 784–802. https://doi.org/10.1080/09658211.2013.831454

Soper, D.S. (2018). Sobel Test Calculator for the Significance of Mediation [Software]. Available from http://www.danielsoper.com/statcalc

Susser, J. A., & McCabe, J. (2013). From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science, 41(2),* 345–363. https://doi.org/10.1007/s11251-012-9231-8.

Sweller, J., Ayres, P. L., & Kalyuga, S. (2011). *Cognitive load theory.* New York: Springer.

Tse, C.-S., & Pu, X. (2012). The effectiveness of test-enhanced learning depends on trait test anxiety and working-memory capacity. *Journal of Experimental Psychology: Applied*, 18(3), 253–264. https://doi.org/10.1037/a0029190

Unsworth, N., Spillers, G. J., Brewer, G. A., & McMillan, B. (2011). Attention control and the antisaccade task: A response time distribution analysis. *Acta Psychologica, 137,* 90–100. https://doi.org/10.1016/j.actpsy.2011.03.004

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

von der Embse, N., Barterian, J., & Segool, N. (2013). Test anxiety interventions for children and adolescents: A systematic review of treatment studies from 2000–2010. *Psychology in the Schools, 50,* 57–71. https://doi.org/10.1002/pits.21660

Wild, K.-P. & Schiefele, U. (1994). Lernstrategien im Studium: Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie, 15,* 185–200. https://doi.org/10.1026/0049-8637.37.1.17

Wine, J. (1971). Test anxiety and direction of attention. *Psychological Bulletin*, *76*(2), 92–104. https://doi.org/10.1037/10037-000

Wissman, K. T., & Rawson, K. A. (2018). Collaborative testing for key-term definitions under representative conditions: Efficiency costs and no learning benefits. *Memory & Cognition, 46(1), 148–157* https://doi.org/10.3758/s13421-017-0752-x

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2012). How and when do students use flashcards? *Memory*, *20*(6), 568–579. https://doi.org/10.1080/09658211.2012.687052

Wittmaier, B. C. (1972). Test anxiety and study habits. *The Journal of Educational Research*, *65*(8), 352–354. https://doi.org/10.2466/pr0.1967.20.3.975

Zeidner, M. (2014). *Test anxiety: The state of the art.* Heidelberg: Springer.

Zhang, L. F. (2000). University students' learning approaches in three cultures: An investigation of Biggs's 3P model. *The Journal of Psychology*, *134*(1), 37–55. https://doi.org/10.1080/00223980009600847

Notes

Note 1. Self-testing in this paper refers to students' use of activities that check one's understanding and mastery of the learning content and/or practicing retrieval of the content from memory.

Note 2. Since only a limited number of participants signed-up and participated, we regret not being able to perform factor analyses to test the structure of the scale. We think however that this pilot study entails interesting findings worth not to vanish in a file-drawer, especially given the scarce literature on test anxiety and (self-)testing.

Note 3. For readers preferring the mediation test statistics with the SPSS PROCESS procedure version 3.0 (Hayes, 2018), see Appendix Table 2 to Table 6.

Appendix

Table 1. Items of the Self-testing Scale (STS)

Self-testing items	Mean (SD)	Frequency			
		more repetition	equal	more testing	neither
1. To remember learning content of classes/slides/texts… A I repeatedly re-read (or listened/or watched) the learning contents …B I repeatedly tested myself regarding the learning contents	-.04 (1.22)	28	30	23	0
2. To remember the subject matter, I…A…(re-) copied the important contents multiple times and created summaries. B…rehearsed the important contents from memory over and over again	0.21 (1.49)	23	32	26	0
3. For my exam preparation, I met with fellow students or learning groups, to…A…check the contents and the understanding of the subject matter. B…share the reprocessing of the material to several persons **(R)**	1.90 (1.20)	2	2	46	21
4. For my exam preparation, I met with fellow students or learning groups, to…A…exchange learning materials and to complete contents. B…reciprocally test ourselves regarding the learning contents	0.21 (1.70)	20	20	23	18
5. For my exam preparation, I met with fellow students or learning groups, to…A…collate the contents and its understanding. B…split die material to several persons for a joint summary **(R)**	2.08 (.89)	0	3	57	21
6. To memorize the subject matter, I…A…repeated and re-read summaries, slides and relevant passages as much as possible. B…created my own exercise and exam questions and answered them	-0.99 (1.45)	43	19	9	0
7. I used flashcards, to…A…re-read the content multiple times. B…retrieve the contents from memory repeatedly	0.94 (1.37)	6	17	36	27
8. I used test- and exercise questions at the end of a chapter, to…A…re-read the concerning text passage as repetition. B…simulate an examination	0.31 (1.44)	23	15	32	11
9. I used test- and exercise questions at the end of a chapter, that did not have any solution or right answers, to…A…create and write down the correct answers directly based on the text. B…test my understanding without consulting the text while answering the questions	0.66 (1.32)	12	17	33	19
10. I prepared for the exam by…A…repeating slides, important texts, or exercises as much as possible. B…checking my understanding of the slides, texts or exercises with several tests	-0.82 (1.32)	49	18	12	2
11. I learned by…A…solving test questions, exercises or previous exams. B…memorizing correct answers of test question, exercises or previous	0.78 (1.50)	9	24	43	5

exams questions **(R)**					
12. Quizzes and mock exams …A I used during studying for actual learning of the materials …B I used after studying to check my mastery over the materials	0.62 (1.47)	17	22	42	0
13. In my exam preparation, I…A…worked through many similar and preferably known exercises to memorize the solutions. B…solved many different and preferably unknown exercises to check my understanding	0.53 (1.40)	18	20	35	8
14. Questions, exercises or tests about the subject matter…A…I read B…I solved myself	0.74 (1.48)	14	22	45	0
15. I used the model answers of test questions …A…to compare my own answers with the model answers. B…learn the model answers by heart (R)	0.89 (1.34)	8	17	50	6

Note. Items were translated into English by the authors and originally in German language. The (R) indicates items that were reversely phrased. Higher numbers reflected self-testing, while lower numbers reflected repetition/non-testing, the 0 of the scale meant both strategies were used equally often.

Mediation analyses by PROCESS procedure

Reported values below based on the PROCESS procedure by Hayes (2018).

Table 2. Regression of the Interference Sub-Scale Predicting Scores on the Self-testing Scale (STS)

Dependent variable STS

Variable	*B*	*SE B*	*t*	*p*	*95 % CI*
Constant	1.22	.27	4.51	.000	.67, 1.76
Interference	-.40	.14	-3.01	.004	-.67, -.14

Note. $F(1, 62) = 9.07$, $p < .01$, $R^2 = .13$

Table 3. Regression of the Interference Sub-Scale and STS predicting Averaged Exam Grades

Dependent variable Averaged Exam Grades

Variable	*B*	*SE B*	*t*	*p*	*95 % CI*
Constant	2.11	.36	5.84	.000	1.39, 2.83
Interference	.19	.17	1.14	.257	-.14, .52
ST-scale	-.4	.15	-2.75	.008	-.70, -11

Note. $F(2, 61) = 6.37$, $p < .01$, $R^2 = .17$; averaged exam grades is left in its original form with higher values meaning worse grades.

Table 4. Total Effect of Interference on Averaged Exam Grades

B	*SE B*	*t*	*p*	*95 % CI*
.35	.16	2.16	.034	.03, .68

Table 5. Direct Effect of Interference on Averaged Exam Grades

B	*SE B*	*t*	*p*	*95 % CI*
.19	.17	1.14	.257	-.14, .52

Note. Averaged exam grades are left in its original form with higher values meaning worse grades.

Table 6. Indirect Effect of Interference on Averaged Exam Grades via STS

B	*Boot SE*	*Boot 95 % CI*
.16	.17	.04, .32

Note. Bootstrapping based on 10000 samples. Averaged exam grades are left in its original form with higher values meaning worse grades.