

Cognition and Neurosciences

Does mathematics anxiety moderate the effect of problem difficulty on cognitive effort?

TERJE ULV THRONSDEN,¹  MARCUS LINDSKOG,²  MARKKU NIEMIVIRTA³  and RIIKKA MONONEN^{1,4} ¹Department of Special Needs Education, University of Oslo, Oslo, Norway²Department of Psychology and Department of Education, Uppsala University, Uppsala, Sweden³School of Applied Educational Science and Teacher Education, University of Eastern Finland and Department of Education, University of Helsinki, Helsinki, Finland⁴Teachers, Teaching and Educational Communities, University of Oulu, Oulu, FinlandThronsdén, T. U., Lindskog, M., Niemivirta, M. & Mononen, R. (2022). Does mathematics anxiety moderate the effect of problem difficulty on cognitive effort? *Scandinavian Journal of Psychology*, 63, 601–608.

A negative relationship between mathematics anxiety (MA) and mathematics performance is well documented. One suggested explanation for this relationship is that MA interferes with the cognitive processes needed when solving mathematics problems. A demand for using more cognitive effort (e.g., when performing harder mathematics problems), can be traced as an increase in pupil dilation during the performance. However, we lack knowledge of how MA affects this relationship between the problem difficulty and cognitive effort. This study investigated, for the first time, if MA moderates the effect of arithmetic (i.e., multiplication) problem difficulty on cognitive effort. Thirty-four university students from Norway completed multiplication tasks, including three difficulty levels of problems, while their cognitive effort was also measured by means of pupil dilation using an eye tracker. Further, the participants reported their MA using a questionnaire, and arithmetic competence, general intelligence, and working memory were measured with paper-pencil tasks. A linear mixed model analysis showed that the difficulty level of the multiplication problems affected the cognitive effort so that the pupil dilated more with harder multiplication problems. However, we did not find a moderating effect of MA on cognitive effort, when controlling for arithmetic competence, general intelligence, and working memory. This suggests that MA does not contribute to cognitive effort when solving multiplication problems.

Key words: Cognitive effort, mathematics anxiety, arithmetic, working memory, pupillometry, eye tracking.

Terje Ulv Thronsdén, Department of Special Needs Education, P.O. Box 1140, Blindern, 0318 Oslo, Norway. E-mail: t.u.thronsdén@isp.uio.no

INTRODUCTION

Proficiency in mathematics is important to succeed in modern society and individuals who do not possess adequate skills in mathematics are at greater risk of unemployment, as well as physical and mental illness (Duncan, Dowsett, Claessens *et al.*, 2007; Parsons & Bynner, 2006). A negative association between mathematics performance and mathematics anxiety (MA) has consistently been found across studies and in meta-analyses looking at students all the way from primary school to college levels (Barroso, Ganley, McGraw, Geer, Hart & Daucourt, 2021; Hart & Ganley, 2019; Hembree, 1990; Ma, 1999; Maldonado Moscoso, Anobile, Primi & Arrighi, 2020; Namkung, Peng & Lin, 2019; Ramirez, Shaw & Maloney, 2018). MA has been defined as feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems, both in everyday life and in academic situations (Richardson & Suinn, 1972). Further, previous research has shown that high school, college and university students with high MA enjoy mathematics less, avoid taking mathematics courses, and get lower grades in mathematics compared to their peers with low MA (Andrews & Brown, 2015; Ashcraft & Kirk, 2001; Ashcraft & Moore, 2009; Jameson & Fusco, 2014; Núñez-Peña, Suárez-Pellicioni & Bono, 2013).

One suggested explanation for the negative relationship between MA and mathematics performance is that the working memory (WM) capacity needed for mathematical problem solving is occupied by worrying thoughts (Ashcraft & Kirk, 2001;

Ashcraft & Krause, 2007) leading to lower performance. WM can be defined as a cognitive control system with a limited capacity in both storage and ability to process (Baddeley, 1992). Originally, WM was assumed to consist of a central executive system and two slave systems; the phonological loop and the visuospatial sketchpad (Baddeley, 1992; Baddeley & Hitch, 1974), but more recently, a new component of WM has been proposed, the episodic buffer (Baddeley, 2000). In previous studies, the influence of MA on WM has been inferred by showing that individuals high in MA perform worse on mathematical problems that are purportedly high in WM demands (Ashcraft & Kirk, 2001; Ashcraft & Krause, 2007) and hence that MA increases the total cognitive effort used during problem-solving. However, no previous study has investigated whether MA directly increases the total cognitive effort when measured online during mathematical problem solving. One reason for the lack of such studies is that measuring cognitive effort online during mathematical problem solving is not always straightforward. From a methodological viewpoint, some studies have shown that pupil dilation can be used to index WM load, or more generally as a proxy for cognitive effort, as the pupil dilates more with increasing cognitive processing (Heitz, Schrock, Payne & Engle, 2008; Johnson, Miller Singley, Peckham, Johnson & Bunge, 2014; Kahneman & Beatty, 1966; Siegle, Steinhauer, Stenger, Konecky & Carter, 2003; Unsworth & Robison, 2018).

By investigating whether MA moderates the effect of mathematics task difficulty on cognitive effort, the current study

aimed to test if MA interferes with cognitive processing while performing mathematics tasks. We reasoned that if MA indeed generates worrying thoughts that reduce WM capacity, an increase in cognitive effort as a result of task difficulty should be steeper for individuals high in MA than for those low in MA. In terms of pupil dilation, we thus expected a larger effect of mathematics task difficulty for individuals with high compared to low MA.

THE RELATIONSHIP BETWEEN MATHEMATICS ANXIETY AND MATHEMATICS PERFORMANCE

With an estimated prevalence of MA being between 2% and 33%, depending on the cut-off-scores used and how it is measured (Ashcraft & Moore, 2009; Barroso *et al.*, 2021; Chinn, 2009; Hart & Ganley, 2019; Johnston-Wilder, Brindley & Dent, 2014; Luttenberger, Wimmer & Paechter, 2018), a considerable number of people experience MA and its possible negative consequences. Multiple meta-analyses have documented MA's negative relationship ($r = -0.21$ to -0.40) with mathematics performance in individuals from early grades to adulthood (Barroso *et al.*, 2021; Hembree, 1990; Ma, 1999; Namkung *et al.*, 2019; Zhang, Zhao & Kong, 2019). The directionality between MA and mathematics performance has been examined in numerous studies across all educational levels (Ashcraft & Kirk, 2001; Carey, Hill, Devine & Szücs, 2016; Zhang *et al.*, 2019). The Debilitating Anxiety Model (also known as the Cognitive Interference Theory, Wine, 1980) proposes that MA leads to low mathematics performance. This model suggests that MA impacts three stages of performance; pre-processing, processing, and retrieval of mathematics knowledge (Carey *et al.*, 2016). Further, the model proposes that MA obstructs mathematical problem solving by reducing the processing and storage capacity of WM due to worrying (Ashcraft & Moore, 2009; Eysenck & Calvo, 1992). This assumption has been supported by a line of empirical studies investigating the relationship between MA and performance (e.g., Ashcraft & Kirk, 2001; Justicia-Galiano, Martín-Puga, Linares & Pelegrina, 2017; Ramirez *et al.*, 2018; Skagerlund, Östergren, Västfjäll & Träff, 2019; Vukovic, Kieffer, Bailey & Harari, 2013). Ashcraft and Krause (2007), for example, reported that individuals with high MA performed significantly lower in arithmetic problems of high WM demand than their peers with low or moderate MA. Thus, it seems that MA compromises the functioning of WM, which leads to poorer mathematics performance in individuals with high MA.

The significance of WM is substantiated in an abundant amount of research with results that supports the role of WM in mathematical cognition (Ashcraft & Krause, 2007; Logie, Gilhooly & Wynn, 1994; Miller & Bichsel, 2004; Skagerlund *et al.*, 2019). It is also evident that there is a link between the complexity of the mathematics problem and the processing demands on WM, where more complex problems impose a higher load on WM (Ashcraft & Krause, 2007). Supporting this, research where WM capacity has been deliberately strained by irrelevant disturbance (e.g., given a secondary task), mathematics performance has been negatively affected (e.g., Logie *et al.*, 1994). However, even though previous studies show individuals with high MA to have inferior performance on arithmetic problems with high WM demands, there is little direct

evidence that these individuals are in fact exerting higher cognitive effort during problem solving.

PUPIL DILATION AS A MEASURE OF COGNITIVE EFFORT

Pupillometry is to study the changes in pupil size as a function of cognitive processing (Sirois & Brisson, 2014). Although the pupil's principal function is to regulate how much light gets into the eye, it is well established that pupils also dilate independent of changes in light conditions (e.g., Balkenius, Fawcett, Falck-Ytter, Gredebäck & Johansson, 2019; Mathôt, 2018). For instance, numerous studies have shown that pupil diameter is a reliable indicator of mental effort, cognitive intensity, or cognitive load (Ahern & Beatty, 1979; Beatty, 1982; Einhäuser, 2017; Hess & Polt, 1964; Kahneman & Beatty, 1966; Kiefer, Giannopoulos, Duchowski & Raubal, 2016). Despite the use of different terms, the overall findings are coherent, and show the pupil to dilate as a reaction to cognitive activation (Mathôt, 2018). Furthermore, there is a direct link between task difficulty and pupil dilation; the harder the task, the more the pupil dilates (Ahern & Beatty, 1979; Boersma, Wilton, Barham & Muir, 1970; Hess & Polt, 1964).

This effect has seemingly been moderated by various factors. In their seminal study, Ahern and Beatty (1979) found that higher arithmetic skills and higher general intelligence moderated the effect of task difficulty on pupil dilation in a sample of undergraduates solving mathematical problems. More specifically, participants with higher arithmetic skills and higher general intelligence had smaller pupil dilation compared to participants with lower arithmetic skills and intelligence when solving identical mathematical problems. A suggested explanation for this is that individuals with higher intelligence are processing information more efficiently and require less effort, while the group with lower intelligence has to work harder to solve the same problems (Ahern & Beatty, 1979). Regarding WM, it has been found that individual differences in WM capacity also moderate the connection between problem difficulty and pupil dilation, where participants with higher WM capacity show relatively lower pupil dilation than those with low WM capacity (Heitz *et al.*, 2008). This means that individuals with higher WM capacity use relatively less cognitive effort compared to individuals with lower WM capacity. To our knowledge, no prior studies have investigated MA as a moderator for the effect of task difficulty on cognitive effort, when also controlling for arithmetic competence, WM, and general intelligence.

PRESENT STUDY

Prior research has shown that MA negatively affects students' mathematics performance by reducing cognitive resources needed for manipulation of mathematical problems (Ashcraft & Kirk, 2001; Justicia-Galiano *et al.*, 2017; Ramirez *et al.*, 2018). Consequently, this may have unfavourable effects on different areas in student's life (Duncan *et al.*, 2007), unless appropriate intervention is provided. From a methodological viewpoint, much of the research has focused on using only behavioral measures when investigating this relationship (Douglas & LeFevre, 2017; Maldonado Moscoso *et al.*, 2020). Here, we will also apply pupillometry, which can give us more precise information about

the cognitive processing during mathematical problem solving. Combined with self-reported MA, this study will expand the current state of research by examining if MA moderates the effect of arithmetic problem difficulty on cognitive effort.

We addressed these issues with the following two research questions: RQ1) *Does problem difficulty influence cognitive effort?* and RQ2) *If there is a relationship between problem difficulty and cognitive effort, is this relationship moderated by MA when controlling for arithmetic competence, working memory, and general intelligence?* As this, to the best of our knowledge, is the first time the effect of MA on cognitive effort has been investigated using pupillometry, we chose a similar sample (i.e., sample size and university students), as Ahern and Beatty (1979), because they successfully found both an effect of problem difficulty and a moderating effect by a covariate on cognitive effort. We expected that participants' cognitive effort would increase as a function of task difficulty, thus replicating the findings of Ahern and Beatty (1979). Furthermore, in line with predictions derived from the debilitating anxiety model, we expected MA to moderate the effect of task difficulty on cognitive effort during mathematical problem solving. We predicted an increase in cognitive effort as a result of task difficulty to be steeper for individuals higher in MA than for those lower in MA. Arithmetic competence, WM and general intelligence were used as control variables, as these factors in previous research have shown to moderate the effect (i.e., problem difficulty on pupil size) we are investigating (e.g., Ahern & Beatty, 1979; Hess & Polt, 1964; Johnson *et al.*, 2014; Kahneman & Beatty, 1966).

METHODS

Participants

We used a convenience sample with participants recruited from a large university in eastern Norway, using posters to advertise for our study. Students interested in participating made contact by email. This resulted in a sample of 37 students (27 females; M age = 26.62 y.), with most participants studying educational sciences. Participants represented all study levels (i.e., BA, MA, and PhD students). Participation was voluntary, and all participants were compensated with a gift certificate for a movie ticket. All participants that initially made contact chose to take part in the study and finished all measures, but data from three participants were excluded due to calibration errors with the eye-tracker. Hence, our final sample consisted of 34 participants (26 females). The Norwegian Centre for Research Data gave ethical approval for the study, and the participants gave an active consent.

Measures

Cognitive effort. Cognitive effort was measured while the participants were performing mental arithmetic problems (i.e., multiplication) on a laptop. The task consisted of a total of 36 multiplication problems of three different difficulty levels taken from the study by Ahern and Beatty (1979); 12 *easy* problems (i.e., multiplicand from 6 to 9 with multiplier from 12 to 14), 12 *medium-difficulty* problems (i.e., multiplicand from 6 to 9 with multiplier from 16 to 19) and 12 *hard* problems (i.e., multiplicand from 11 to 14 with multiplier from 16 to 19).

To measure pupil dilation, we used a Tobii X2-60 compact eye tracker with a 60 Hz sampling rate (Tobii, Stockholm, Sweden). The eye tracker was mounted on a HP Elitebook G3 laptop with a 15.6 inch 16:9, 1920 × 1,080-pixel screen. The software OpenSesame (Mathôt, Schreij & Theeuwes, 2012) was used to design and run the experiment. The

instructions appeared in the font "mono," black on a white background in size 24. The digits were in the same font and color, but in size 48. Both text and digits were centred on the screen. During the multiplication tasks, participants were presented with only one digit (i.e., the multiplier or the multiplicand) at the time on the screen. First, a fixation dot appeared for 2000 ms, then the multiplier was visible for 500 ms, followed by a blank screen for 1,500 ms, then the multiplicand was visible 500 ms, followed by a fixation dot for 5,000 ms. The participants were instructed to keep their eyes on the last fixation dot until it disappeared when performing the mental arithmetic. After the fixation dot disappeared, participants were prompted by instructions on the screen to type in their answer of the multiplication problem.

Gaze data was exported as raw data to MatLab (Version 9.5) for further analysis. Following Nyström, Falck-Ytter and Gredebäck (2016), we pre-processed the eye-tracking data in four steps before extracting our dependent measure of pupil size. In a first step, individual samples with a pupil size outside of the range 0–5 mm were removed. Next, we rejected trials with less than 60% data. In a third step, gaps shorter than five samples were linearly interpolated. Finally, in a fourth step we applied a moving average filter with a window of five samples on the time series. To extract a dependent measure, we defined for each trial two time-windows relative to the start of the trial. The baseline period was set as the first 2,000 ms of the trial when a fixation dot was presented. The analysis period was set as the 5,000 ms interval after the multiplicand had been presented. The dependent measure for each trial was calculated as the mean change in pupil size from the baseline to the analysis period.

Mathematics anxiety. Mathematics anxiety was measured with six items from the Achievement Emotions Questionnaire (AEQ; Pekrun, Goetz, Frenzel & Perry, 2011), translated into Norwegian. All questions relate to how anxious the students feel in situations where mathematics performance is required. To fit the questions to university students who no longer take part in mathematics lectures, we made a small alteration to the questions and changed the wording from present to past tense. For example, a change was made from "When taking a math test, I *am* tense and nervous" to "When taking a math test, I *was* tense and nervous." The participants gave their response on a five-point Likert-type scale ranging from 1 (Not at all) to 5 (Very much). Given the evidence of cross-cultural usability and validity of different versions of the AEQ (Frenzel, Thrash, Pekrun & Goetz, 2007; Lichtenfeld, Pekrun, Stupnisky, Reiss & Murayama, 2012; Sánchez Rosas, 2015), we considered the scales to be valid in the Norwegian context as well. The McDonald's omega (ω)

Table 1. *Descriptive statistics and reliability of raw scores for all measures*

Variable	min–max	M	SD	Skew	Kurt	Reliability (ω)
Cognitive effort	–1.12–0.30	0.05	0.08			
Math anxiety	1–5	2.50	1.09	0.55	–0.70	0.90
Addition (CMAT)	6–16	12.32	2.57	–0.23	–0.45	0.75
Subtraction (CMAT)	4–14	9.65	2.53	–0.34	–0.04	0.74
Multiplication (CMAT)	0–14	5.82	2.84	0.45	1.16	0.82
Division (CMAT)	0–13	5.09	3.11	0.8	0.14	0.82
Forward digit recall	6–14	9.47	2.06	0.68	–0.13	0.81
Backwards digit recall	2–10	6.41	1.92	0.13	0.15	0.76
General intelligence	2–18	8.91	3.77	0.48	–0.05	0.75

Note: CMAT = the Comprehensive Mathematical Abilities Test.

reliability coefficient for AEQ was 0.90 (see Table 1), hence, within what is considered acceptable.

Covariates. *Arithmetic competence* was measured with the Comprehensive Mathematical Abilities Test (CMAT; Hresko, Schlieve, Herron, Sawain & Sherbenou, 2003). CMAT is a paper-and-pencil test consisting of four timed subtests measuring arithmetic skills (i.e., 16 items for addition; 14 items for subtraction; 17 items for multiplication; 16 items for division), with 10 min given for each subtest. The participants were told to solve as many problems as possible within the time limit. Some modifications were done to the layout of the division problems, as the symbol for long division “√,” is not used in Norway. Consequently, for example, $29\sqrt{2,929}$ was modified to 2,929:29. For each correct answer, participants were given one point. To the best of our knowledge, no standardized Norwegian arithmetic test for this age group was available at the time of our study, hence we chose to use CMAT, which is reported as a reliable and valid measure. Apart from the small changes in the presentation of the division problems, and the fact that the instructions were translated from English to Norwegian, the CMAT was kept as close to its original format as possible. The McDonald’s Omega reliability scores (see Table 1) for the addition and subtractions measures were all acceptable, but somewhat lower (i.e. 0.75 and 0.74) than the reliability of the multiplication and division (i.e. 0.82). This is likely due to few easy (and thus extremely skewed) items in the beginning of the addition and subtraction tests, which, in turn, attenuates the item-total correlations for those items, and, consequently, the reliability estimate. However, as we wanted to keep the measure as intended, we included all items in the sum score.

Working memory was measured with The Digit Span subtest of the WAIS-III (Wechsler, 1997). In this subtest a maximum of 16-digit sequences are orally presented forward (two trials per item, 2 to 9 digits) and a maximum of 14-digit sequences backwards (two trials per item, 2 to 8 digits). After each number sequence was presented, with an approximately 1 s pause between each digit, the participant was to repeat the number sequence. Participants were first given the forward sequences with two consecutive mistakes as the stopping criteria, followed by backwards sequences with the same stopping criteria. One point was given for each correct answer, and zero for incorrect ones. The McDonald’s omega (ω) reliability coefficients for these measures were 0.81 and 0.76 (see Table 1), respectively, and thus within the range of acceptable.

General intelligence was measured with 18 items from Raven’s Standard Progressive Matrices (Raven, 1998). We followed Lindskog, Winman and Poom (2017) and Stanovich and West (1998), and removed the six most difficult and the 12 easiest items and added a 15-min time limit. The participants were instructed to solve as many of the 18 problems as possible within the time limit. For each correct answer one point was given. This measure had an acceptable McDonald’s omega (ω) reliability coefficient of 0.75 (see Table 1).

Procedure

The participants were tested in two sessions. In the first group session, participants were administered measures of MA, arithmetic competence, and general intelligence, in that order. The group session lasted for approximately one hour. In the following two weeks, each of the participants had an individual session (mean 3.4 days [$SD = 3.0$] between the group and individual session), in which they completed the measures of WM, followed by the multiplication task. For the multiplication task, the participant was seated approximately 60 cm in front of a laptop screen. A five-point calibration was conducted before the multiplication task instructions appeared on the screen. The participants were told to keep their eyes on the screen and press “space” when ready. The individual session was conducted for each participant in the same room under the same light conditions.

Method of analysis

All the pupil dilation data from the multiplication task were prepared for the analysis with TimeStudio (Nyström *et al.*, 2016) in MATLAB (Version

9.5). SPSS (IBM Corp, Armonk, NY, USA) and Jamovi (2020) were used for the statistical analyses. Due to lack of pupil measures caused by off-screen fixations, and/or eye-tracker malfunction, three participants were excluded. Separate principal component analyses on the four arithmetic subtests and the two WM subtests, respectively, were used to extract principal components as proxies for arithmetic competence and WM.

We used linear mixed model (LMM) analysis to investigate the research questions. In the LMM analysis, cognitive effort was the dependent variable, and event type (i.e., easy, medium, and hard problems) was a fixed factor. Further, MA, arithmetic competence, general intelligence, and WM were used as covariates, and participant ID as a cluster variable. In addition to the mentioned factor and covariates, the interaction between event type and MA was added as a fixed effect.

RESULTS

Table 1 summarizes descriptive statistics and reliability of raw scores for all measures, except cognitive effort. Table 2 displays the correlations between all observed variables where the variable *cognitive effort* is the mean of all pupil dilation measurements for each participant. Although the association between MA and *cognitive effort* was not significant ($p = 0.180$), MA correlated strongly with *arithmetic competence* ($r = -0.74$, $p < 0.001$), and moderately with WM ($r = -0.50$, $p = 0.003$) and *general intelligence* ($r = -0.58$, $p < 0.001$), thus indicating a connection between higher MA and lower arithmetic competence, WM and general intelligence.

Our first research question was concerning whether multiplication problem difficulty influences cognitive effort measured by pupil dilation; *does problem difficulty influence cognitive effort?* The LMM analysis indicated that difficulty level of the multiplication problems affected cognitive effort, $F(2, 1138.160) = 7.089$, $p < 0.001$ (we used the Satterthwaite method for degrees of freedom). Furthermore, post hoc comparisons using Bonferroni correction indicated that the cognitive effort for hard problems ($M = 0.09$, $SD = 0.13$) was significantly different from the cognitive effort for medium problems ($M = 0.05$, $SD = 0.12$), $t(1138.150) = -2.576$, $p = 0.030$ (Table 3) and from the cognitive effort on the easy problems ($M = 0.05$, $SD = 0.12$), $t(1138.18) = -3.66$, $p < 0.001$ (Table 3). The difference between easy and medium multiplication problems was not significant ($p = 0.833$).

Our second research question was; *if there is a relationship between problem difficulty and cognitive effort, is this relationship moderated by MA when controlling for arithmetic competence, working memory, and general intelligence?* To answer this

Table 2. Correlation matrix of variables included in the LMM

Variable	1	2	3	4	5
1. Cognitive effort	—				
2. Math anxiety	-0.24	—			
3. Arithmetic competence ¹	0.31	-0.75***	—		
4. Working memory ¹	0.19	-0.50**	0.65***	—	
5. General intelligence	0.30	-0.58***	0.60***	0.48**	—

** $p < 0.01$,

*** $p < 0.001$.

¹Principal component.

Table 3. Post hoc comparisons of cognitive effort during solving of arithmetic problems in the three difficulty levels

Comparison		Difference	SE	t	df	Pbonferroni
Event type	Event type					
Medium	Hard	-0.027	0.011	-2.576	1138.150	0.030
Easy	Hard	-0.039	0.011	-3.667	1138.178	< 0.001
Easy	Medium	-0.011	0.011	-1.086	1138.149	0.833

research question, we took a stepwise approach, with an initial model comprising *event type* (i.e., easy, medium, and hard multiplication problems, *event type 1* being easy-medium and *event type 2* being hard-easy), *MA*, and the interaction between *MA* and *event type* (*event type*MA*), as fixed factors, and *cognitive effort* as the dependent variable. The LMM analysis of this model showed the event type to still predict *cognitive effort*, $F(2, 57.045) = 4.567$, $p < 0.05$ (we used the Satterthwaite method for degrees of freedom). Neither the main effect of *MA* ($p = 0.300$), nor the interaction between *MA* and *event type* ($p = 0.909$) predicted *cognitive effort*. In the next model, we added *arithmetic competence* as a covariate, thus this model included the following variables; *event type*, *MA*, *arithmetic competence*, and *event type*MA*. This model showed that the effect of *event type* still affected *cognitive effort*, $F(2, 1136.18) = 7.077$, $p < 0.001$. None of the other variables had a significant effect on *cognitive effort*. In the third and final model, we added *WM* and *general intelligence* as covariates. *Event type* (and more specifically the difference between hard and easy problems, and between hard and medium problems) was still the only variable that affected *cognitive effort* $F(2, 1136.163) = 7.078$, $p < 0.001$. The results from the final model are reported in Table 4.

DISCUSSION

The aim of this study was to investigate if *MA* moderates the effect of mathematical problem difficulty on cognitive effort, and by that gaining knowledge on the mechanisms behind the relationship between *MA* and mathematics performance. Our first research question was to investigate if problem difficulty influences cognitive effort. We predicted that cognitive effort would be different as a function of problem difficulty, so that individuals are using more cognitive effort with more difficult problems. Replicating previous research (e.g., Ahern & Beatty, 1979; Hess & Polt, 1964), our results clearly show that cognitive effort measured by pupil dilation was affected by the level of arithmetic problem difficulty, so that the pupil dilated more with harder problems. This difference in cognitive effort was seen both between hard and easy problems, as well as between hard and medium problems.

Contrary to our expectations based on previous findings (e.g., Ahern & Beatty, 1979), we did not observe a significant difference in cognitive effort between the medium and the easy problems. There could be several reasons for this. One possible explanation is that the difference in difficulty level between the easy and the medium problems simply was too small, and hence it did not demand significantly more effort to be allocated to solve the medium problems compared to the easy ones. On one hand, it

can also be argued that these problems can be solved by following a set of learnt algorithms, and even that they are somewhat overlearned by students (Lee *et al.*, 2015). Thus, it is quite possible that the medium problems, even though they were considered demanding as compared to easy tasks, did not present a sufficient challenge. On the other hand, the failure to capture a difference between easy and medium problems could also be due to the easy problems being more difficult than expected. That could lead to the same pattern of results. In any case, this could entail the need to develop other tasks in the future that better distinguish the different levels of difficulty. Although Ahern and Beatty (1979) did find a difference in pupil dilation also between the easy and the medium problems, and the sample in this study had similarities with their study, it is possible to argue that there are also some qualitative differences, which may have affected the result. Our study took place more than 40 years later, and in a Nordic country rather than in the USA. This will most probably mean that there are differences in the participants' mathematical training; hence, there can be substantial dissimilarities between the mathematics courses both over time and between the places.

Our second research question was *if there is a relationship between problem difficulty and cognitive effort, is this relationship moderated by MA when controlling for arithmetic competence, working memory, and general intelligence?* Although we provided empirical evidence that *MA* has a negative relationship with mathematics performance, we did not find the effect of problem difficulty on cognitive effort to be different as a function of *MA*, and as such our findings are inconsistent with our expectation based on previous research (Ashcraft & Kirk, 2001; Ashcraft & Moore, 2009; Skagerlund *et al.*, 2019). According to the Debilitating Anxiety Model, *MA* should affect math ability directly as a result of worrying thoughts that drain the *WM* resources needed for problem solving. We found no support for this account in our results. On the one hand, one, and possibly the most elementary explanation for this, could be that *MA* does not lead to increased cognitive effort when solving arithmetic problems. If so, the mechanisms behind the *MA*-mathematics performance relationship are seemingly not due to *MA* taking up *WM* resources. On the other hand, to our knowledge, this is the first time this relationship has been investigated with online measures of pupil dilations as a proxy for cognitive effort. The discrepancy between current and previous findings may be due to differences in measurements, and the explanation for our lack of support of the Debilitation Anxiety Model may not be straightforward.

However, there are several other plausible explanations for the lack of moderation effect by *MA* that should be further investigated in future studies. One being that the individuals with

Table 4. Linear mixed model fixed effects parameter estimates with all variables and control variables

Names	Effect	Estimate	SE	df	t	p
(Intercept)	(Intercept)	0.053	0.016	28,995	3.282	0.003
Event type 1	Medium - easy	0.011	0.011	1136.147	1.084	0.279
Event type 2	Hard - easy	0.039	0.011	1136.180	3.664	< 0.001
Math anxiety	Math Anxiety	0.009	0.023	28,932	0.413	0.683
Arithmetic competence	Arithmetic competence	0.050	0.029	28,902	1.757	0.089
Working memory	Working memory	-0.009	0.022	28,909	-0.390	0.699
General intelligence	General intelligence	-0.001	0.006	28,955	-0.176	0.861
Event Type 1*Math Anxiety	Medium-Easy*Math Anxiety	0.004	0.010	1136.101	0.441	0.660
Event Type 2*Math Anxiety	Hard-Easy*Math Anxiety	0.004	0.010	1136.077	0.432	0.666

low MA are enjoying mathematics more, compared to the individuals with high MA (Ashcraft & Kirk, 2001; Núñez-Peña *et al.*, 2013), and thus, they are devoting a higher intensity of attention to the task. As pupil dilation has shown to increase with intensity of attention (Miller *et al.*, 2019), this may have contributed to washing out the possible effect MA has on cognitive effort. Moreover, there was a significant and strong negative association between MA and arithmetic competence. It is a plausible assumption that individuals with higher arithmetic skills also enjoy mathematics more (Pekrun, Lichtenfeld, Marsh, Murayama & Goetz, 2017), and therefore, are more likely to focus more on the task, compared to individuals with lower arithmetic competence. Individuals with higher arithmetic skills would then show a larger increase in cognitive effort compared to the individuals with lower skills. There was also a significant negative correlation between MA and general intelligence that may cause a washing out effect, as people that are more intelligent are more likely to engage in and enjoy problem solving compared to individuals with lower intelligence (Cacioppo & Petty, 1982; Fleischhauer, Enge, Brocke, Ullrich, Strobel & Strobel, 2010; Furnham & Thorne, 2013; Hill, Foster, Elliott, Shelton, McCain & Gouvier, 2013). Another aspect is that older students are better at emotion regulation enabling them to inhibit the negative feelings caused by their MA (Gross, 1998; Gross & Levenson, 1997), and this would also make it harder to discover a possible effect of MA in this age group.

Although not being able to contribute to provide evidence for MA hindering math performance by occupying WM resources, the present study could, when performing correlation analyses, confirm the well-documented negative relationship between MA and arithmetic performance (Barroso *et al.*, 2021; Hembree, 1990; Ma, 1999; Namkung *et al.*, 2019; Zhang *et al.*, 2019). As it is well established that individuals who do not possess adequate skills in mathematics run a greater risk for not succeeding in life (Duncan *et al.*, 2007), efforts to support these students should be made. Related to MA, there are some promising intervention studies, which suggest that focused breathing exercises can be used as a method to reduce MA's negative impact on mathematics performance (Brunyć *et al.*, 2013).

LIMITATIONS

There are some limitations in our study that should be considered when interpreting the results and in guidance for future studies.

First, we chose a sample similar to the sample of Ahern and Beatty (1979) because they had successfully found both an effect of problem difficulty and a moderating effect by a covariate on cognitive effort. As our study is exploratory, it was our assessment that using a similar sample should be sufficient to detect an effect. However, this also meant that our sample size was relatively small and included only students who had previously taken an academic track in upper secondary school involving studying mathematics. Thus, in future studies, to have a greater chance at discovering the variability in MA and mathematics performance, a more heterogeneous group of students with varying backgrounds in their mathematics training should be considered. This could for instance, be a group of unselected school-aged children where a greater variation in MA is likely to be found. Furthermore, it would be of interest to investigate the stability of MA and its possible effect on cognitive effort in a longitudinal design where school-aged children are measured in these characteristics on two or more time points. Another issue concerns the measures used in our study. In self-reporting of MA, the participants evaluated their MA retrospectively. The quality of such retrospective evaluations has been problematized for decades (e.g., Bartlett, 1932; Ross & Newby-Clark, 1998). The possible unreliability of this measure may have biased the results as what we have assumed are high MA individuals may in fact not be. However, as the questions are all related to negative feelings towards mathematics, it is more likely that participants who experience high MA as adults report of high MA in the past, rather than trivialize their past MA (Ross & Newby-Clark, 1998). Future research should therefore include larger samples and more concurrent measures of MA. Another aspect that needs to be considered is the sensitivity of the measurement. The eye tracker used in this study records at a rate of 60 measurements per second, which is rather low in comparison to some other eye trackers. It could be argued that this low rate has failed to capture the changes in pupil size during the measurement. However, this is not very likely, as pupil dilation and contraction are much slower processes than other types of eye movement, for example saccades (Holmqvist *et al.*, 2011). The method of analysis, a linear mixed model (LMM) also has some limitations. An LMM is commonly used to analyse multilevel data, but one of the shortcomings with this model is that it does not account for measurement error. In future studies, with a bigger sample size, this could be better accounted for by, for instance, taking a structural equation modelling approach to the analyses.

CONCLUSION

Despite the mentioned potential limitations, this study has several notable strengths. The present study was among the first to investigate if MA moderates the effect of problem difficulty on cognitive effort. Our results substantiate previous studies that attribute pupil dilation to problem difficulty level; the harder that problem, the larger the dilation of pupil and thus more cognitive effort is used. However, in contrast to what one would expect from the explanation of poor mathematics performance as a result of MA occupying WM, we did not find a moderating effect of MA on this relationship. The most obvious reason may be that the relationship between MA and low performance is not due to MA taking up WM resources. However, as, to the best of our knowledge, being the first study to explore this phenomenon using an eye tracker, more research is needed to confirm this finding with a larger, more heterogeneous sample.

This work was supported by the Norwegian Research Council [Grant number: 283396], for RM.

The data that support the findings of this study are available from the corresponding author upon reasonable request

REFERENCES

- Ahem, S. & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. *Science (New York, N.Y.)*, 205, 1289–1292.
- Andrews, A. & Brown, J. (2015). The effects of math anxiety. *Education*, 135(3), 362–370.
- Ashcraft, M.H. & Kirk, E.P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, 130, 224.
- Ashcraft, M.H. & Krause, J.A. (2007). Working memory, math performance, and math anxiety. *Psychonomic Bulletin & Review*, 14, 243–248.
- Ashcraft, M.H. & Moore, A.M. (2009). Mathematics anxiety and the affective drop in performance. *Journal of Psychoeducational Assessment*, 27, 197–205.
- Baddeley, A. (1992). Working memory. *Science*, 255, 556–559.
- Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417–423.
- Baddeley, A.D. & Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 8, pp. 47–89). New York: Academic Press.
- Balkenius, C., Fawcett, C., Falck-Ytter, T., Gredebäck, G., & Johansson, B. (2019). Pupillary Correlates of Emotion and Cognition: A Computational Model. *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, 903–907. <https://doi.org/10.1109/NER.2019.8717091>
- Barroso, C., Ganley, C.M., McGraw, A.L., Geer, E.A., Hart, S.A. & Daucourt, M.C. (2021). A meta-analysis of the relation between math anxiety and math achievement. *Psychological Bulletin*, 147, 134.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276.
- Boersma, F., Wilton, K., Barham, R. & Muir, W. (1970). Effects of arithmetic problem difficulty on pupillary dilation in normals and educable retardates. *Journal of Experimental Child Psychology*, 9, 142–155.
- Brunyé, T.T., Mahoney, C.R., Giles, G.E., Rapp, D.N., Taylor, H.A. & Kanarek, R.B. (2013). Learning to relax: Evaluating four brief interventions for overcoming the negative emotions accompanying math anxiety. *Learning and Individual Differences*, 27, 1–7.
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116.
- Carey, E., Hill, F., Devine, A. & Szűcs, D. (2016). The chicken or the egg? The direction of the relationship between mathematics anxiety and mathematics performance. *Frontiers in Psychology*, 6, 81–86. <https://doi.org/10.3389/fpsyg.2015.01987>.
- Chinn, S. (2009). Mathematics anxiety in secondary students in England. *Dyslexia*, 15(1), 61–68.
- Douglas, H. P., & LeFevre, J.-A. (2017). Exploring the influence of basic cognitive skills on the relation between math performance and math anxiety. *Journal of Numerical Cognition*, 3(3), 642–666.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428–1446.
- Einhäuser, W. (2017). The pupil as marker of cognitive processes. In Q. Zhao (Ed.), *Computational and cognitive neuroscience of vision* (pp. 141–169). Singapore: Springer.
- Eysenck, M.W. & Calvo, M.G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion*, 6, 409–434.
- Fleischhauer, M., Enge, S., Brocke, B., Ullrich, J., Strobel, A. & Strobel, A. (2010). Same or different? Clarifying the relationship of need for cognition to personality and intelligence. *Personality and Social Psychology Bulletin*, 36, 82–96.
- Frenzel, A.C., Thrash, T.M., Pekrun, R. & Goetz, T. (2007). Achievement emotions in Germany and China: A cross-cultural validation of the Academic Emotions Questionnaire—Mathematics. *Journal of Cross-Cultural Psychology*, 38, 302–309.
- Furnham, A. & Thorne, J.D. (2013). Need for Cognition. *Journal of Individual Differences*, 34(4), 230–240.
- Gross, J.J. (1998). The emerging field of emotion regulation: An integrative review. *Emotion Regulation*, 2(3), 271–299.
- Gross, J.J. & Levenson, R.W. (1997). Hiding feelings: The acute effects of inhibiting negative and positive emotion. *Journal of Abnormal Psychology*, 106, 95–103.
- Hart, S.A. & Ganley, C.M. (2019). The nature of math anxiety in adults: prevalence and correlates. *Journal of Numerical Cognition*, 5, 122–139.
- Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, 45(1), 119–129.
- Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 21, 33–46.
- Hess, E.H. & Polt, J.M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, 143, 1190–1192.
- Hill, B.D., Foster, J.D., Elliott, E.M., Shelton, J.T., McCain, J., Gouvier, W. et al. (2013). Need for cognition is related to higher general intelligence, fluid intelligence, and crystallized intelligence, but not working memory. *Journal of Research in Personality*, 47, 22–25.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.
- Hresko, W., Schlieve, P.L., Herron, S.R., Sawain, C. & Sherbenou, R. (2003). *Comprehensive math abilities test*. Austin, TX: Pro-Ed.
- Jameson, M. M., & Fusco, B. R. (2014). Math anxiety, math self-concept, and math self-efficacy in adult learners compared to traditional undergraduate students. *Adult Education Quarterly*, 64(4), 306–322.
- Johnson, E.L., Miller Singley, A.T., Peckham, A.D., Johnson, S.L. & Bunge, S.A. (2014). Task-evoked pupillometry provides a window into the development of short-term memory capacity. *Frontiers in Psychology*, 5, 450–457. <https://doi.org/10.3389/fpsyg.2014.00218>.
- Johnston-Wilder, S., Brindley, J. & Dent, P. (2014). *A survey of mathematics anxiety and mathematical resilience among existing apprentices*. London: The Gatsby Foundation.
- Justicia-Galiano, M.J., Martín-Puga, M.E., Linares, R. & Pelegrina, S. (2017). Math anxiety and math performance in children: The mediating roles of working memory and math self-concept. *British Journal of Educational Psychology*, 87, 573–589.
- Kahneman, D. & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583–1585.

- Kiefer, P., Giannopoulos, I., Duchowski, A. & Raubal, M. (2016). Measuring cognitive load for map tasks through pupil diameter. In J.A. Miller, D. O'Sullivan & N. Wiegand (Eds.), *Geographic information science* (pp. 323–337). Cham: Springer International Publishing.
- Lee, G., Ojha, A., Kang, J.-S. & Lee, M. (2015). Modulation of resource allocation by intelligent individuals in linguistic, mathematical and visuo-spatial tasks. *International Journal of Psychophysiology*, *97*, 14–22.
- Lichtenfeld, S., Pekrun, R., Stupnisky, R.H., Reiss, K. & Murayama, K. (2012). Measuring students' emotions in the early years: The Achievement Emotions Questionnaire-Elementary School (AEQ-ES). *Learning and Individual Differences*, *22*, 190–201.
- Lindskog, M., Winman, A. & Poom, L. (2017). Individual differences in nonverbal number skills predict math anxiety. *Cognition*, *159*, 156–162.
- Logie, R.H., Gilhooly, K.J. & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, *22*, 395–410.
- Luttenberger, S., Wimmer, S. & Paechter, M. (2018). Spotlight on math anxiety. *Psychology Research and Behavior Management*, *11*, 311–322.
- Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, *30*, 520–540.
- Maldonado Moscoso, P. A., Anobile, G., Primi, C., & Arrighi, R. (2020). Math anxiety mediates the link between number sense and math achievements in high math anxiety young adults. *Frontiers in Psychology*, *11*. <https://www.frontiersin.org/article/10.3389/fpsyg.2020.01095>
- Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, *1*, 16.
- Mathôt, S., Schreij, D. & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*, 314–324.
- Miller, A.L., Gross, M.P. & Unsworth, N. (2019). Individual differences in working memory capacity and long-term memory: The influence of intensity of attention to items at encoding as measured by pupil dilation. *Journal of Memory and Language*, *104*, 25–42.
- Miller, H. & Bichsel, J. (2004). Anxiety, working memory, gender, and math performance. *Personality and Individual Differences*, *37*, 591–606.
- Namkung, J.M., Peng, P. & Lin, X. (2019). The relation between mathematics anxiety and mathematics performance among school-aged students: A meta-analysis. *Review of Educational Research*, *89*, 459–496.
- Núñez-Peña, M.I., Suárez-Pellicioni, M. & Bono, R. (2013). Effects of math anxiety on student success in higher education. *International Journal of Educational Research*, *58*, 36–43.
- Nyström, P., Falck-Ytter, T. & Gredebäck, G. (2016). The TimeStudio Project: An open source scientific workflow system for the behavioral and brain sciences. *Behavior Research Methods*, *48*, 542–552.
- Parsons, S., & Bynner, J. (2006). Does numeracy matter more? Retrieved XXXX from <https://discovery.ucl.ac.uk/id/eprint/1566245/1/parsons2006does.pdf>
- Pekrun, R., Goetz, T., Frenzel, A. C., & Perry, R. P. (2011). *Academic Emotions Questionnaire (AEQ). User's manual* (2nd ed.) [User's manual]. Munich: Department of Psychology, University of Munich.
- Pekrun, R., Lichtenfeld, S., Marsh, H.W., Murayama, K. & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development*, *88*, 1653–1670.
- Ramirez, G., Shaw, S.T. & Maloney, E.A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*, 145–164.
- Raven, J. C. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford: Oxford psychologists Press.
- Richardson, F. C., & Suinn, R. M. (1972). The mathematics anxiety rating scale: Psychometric data. *Journal of Counseling Psychology*, *19*(6), 551.
- Ross, M. & Newby-Clark, I.R. (1998). Construing the past and future. *Social Cognition*, *16*(1), 133–150.
- Sánchez Rosas, J. (2015). The Achievement Emotions Questionnaire-Argentine (AEQ-AR): Internal and external validity, reliability, gender differences and norm-referenced interpretation of test scores. *Revista Evaluar*, *15*, 41–74.
- Siegle, G.J., Steinhauer, S.R., Stenger, V.A., Konecky, R. & Carter, C.S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *NeuroImage*, *20*, 114–124.
- Sirois, S. & Brisson, J. (2014). Pupillometry. *WIREs. Cognitive Science*, *5*, 679–692.
- Skagerlund, K., Östergren, R., Västfjäll, D. & Träff, U. (2019). How does mathematics anxiety impair mathematical abilities? Investigating the link between math anxiety, working memory, and number processing. *PLoS One*, *14*, e0211283.
- Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*(2), 161.
- Unsworth, N., & Robison, M. K. (2018). Tracking working memory maintenance with pupillometry. *Attention, Perception, & Psychophysics*, *80*(2), 461–484.
- Vukovic, R. K., Kieffer, M. J., Bailey, S. P., & Harari, R. R. (2013). Mathematics anxiety in young children: Concurrent and longitudinal associations with mathematical performance. *Contemporary Educational Psychology*, *38*(1), 1–10.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-III (WAIS-III, UK)*. London: Harcourt Assessment.
- Wine, J. D. (1980). Cognitive-attentional theory of test anxiety. *Test anxiety: Theory, Research, and Applications*, 349–385.
- Zhang, J., Zhao, N. & Kong, Q.P. (2019). The relationship between math anxiety and math performance: A meta-analytic investigation. *Frontiers in Psychology*, *10*(1613). <https://doi.org/10.3389/fpsyg.2019.01613>.

Received 4 November 2021, Revised 13 April 2022, accepted 31 May 2022