Reality-based interaction affecting mental workload in virtual reality mental arithmetic training

Patrick Jost[a]*, Sue Cobb[b] and Isabella Hämmerle[c]

[a]Department of Computer Science, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology; [b]Human Factors Research Group, Department of Mechanical, Materials and Manufacturing Engineering, Faculty of Engineering, University of Nottingham, Nottingham, England; [c]User-Centred Technologies Research Centre, University of Applied Sciences Vorarlberg, Dornbirn, Austria

* Department of Computer Science,
Faculty of Information Technology and Electrical Engineering
IT-bygget, 116, Gløshaugen, Sem Sælands vei 9
E-Mail: patrick.jost@ntnu.no

Reality-based interaction affecting mental workload in virtual reality mental arithmetic training

The concept of digital game-based learning (DGBL) evolves rapidly together with technological enhancements of virtual reality (VR) and smart phones. However, the mental workload (MWL) that VR-training applications demand and motivational qualities originating from user experience (UX) should be identified in order to create effective and enjoyable training/learning challenges that fit with individual users' capabilities.

This study examined the effects of reality-based interaction (RBI) and VR on measures of student motivation and MWL, in a mental arithmetic game for secondary school pupils. In a randomized controlled trial with sixty school children a mental arithmetic game was tested with three different interaction and two different presentation methods – VR RBI, VR head-mounted-display tapping and tablet flick-gesture.

Results found a significant effect of RBI on MWL but no differences in enjoyment of training were found between VR-experience and tablet training-experience. In fact, adding the gaming-context to the mental arithmetic task created an enjoyable, motivating experience regardless of presentation or interaction-style.

Keywords: mental workload; virtual reality; digital game-based learning; reality-based interaction, education; user experience

1 Introduction and purpose

Table 1. Key terminology and abbreviations used in this paper

Term	Abbreviation	Definition
Digital game-based learning	DGBL	Concept of using game-based approaches with desktop, mobile computing or virtual reality devices to impart knowledge or skills

Virtual Reality	VR	Computer generated simulation of lifelike environment that can be interacted with
User Experience	UX	Perceptions and responses resulting from the use and or anticipated use of a product, system or service
Reality-based interaction	RBI	Concept of interacting in VR in the same way as in reality (e.g. throwing gesture to throw a ball).
Mental Workload	MWL	Cognitive demand that is additively generated through intrinsic, germane and extraneous cognitive load during performance of a task (Sweller, Ayres, and Kalyuga 2011)
Working Memory	WM	Limited cognitive capacity of humans during task performance (Baddeley 2002; Kahneman 1973; Wickens 2008)
Virtual Reality Mental Arithmetic Game	VRMAG	Game in VR to train mental arithmetic
Tablet Mental Arithmetic Game	TMAG	Game on a tablet computer to train mental arithmetic
Head-Mounted-Display	HMD	Hardware for presenting VR with a device worn on the head. Sometimes these devices also include buttons or swipe areas to interact with content
Virtual Reality Game for training mental arithmetic with reality-based interaction (i.e. throwing ball gesture)	Variant A or VRMAG-RBI	Experiment condition where the learning game was persented on a Virtual Reality HMD and interaction was implemented as throwing gesture
Virtual Reality Game for training mental arithmetic with tap-touch interaction on HMD	Variant B or VRMAG-HMD	Experiment condition where the learning game was persented on a Virtual Reality HMD and interaction was implemented by tapping at the integrated touchpad
Tablet Computer Game for training mental arithmetic with touch-swipe interaction	Variant C or TMAG	Experiment condition where the learning game was persented on a tablet computer and interaction was implemented by touch-flick gesture on the tablet screen

1.1 Virtual reality training and mental workload

Virtual Reality (VR) can be defined as a concept of total immersion of an individual in a computed – synthetic – environment. Key features are real-time response of the computer-simulation to user movement and interaction (Burdea and Coiffet 2003). Fully-immersive systems use body-tracking sensors to ensure natural response of the virtual world to user movement (Rothbaum et al. 2001). The Oxford Dictionary denotes VR as: "A computer-generated simulation of a lifelike environment that can be interacted with in a seemingly real or physical way by a person, esp. by means of responsive hardware such as a visor with screen or gloves with sensors…" (OED online, 2018). Therefore, VR is not only a visual experience but can encompass all human senses. It must therefore be regarded holistically as perceptions via any sensory channel may impact the user expereince of presence and immersion. The potential and possibilities of VR for training and education are enormous and well investigated in recent studies focusing on knowledge/skill-acquisition or knowledge/skill-training and learning outcome. VR-training has been investigated for example in surgery (Rahm et al. 2016; Alaraj et al. 2015; Jensen et al. 2015; Piedra et al. 2016), military (Bhagat, Liou, and Chang 2016),and business/social training with adults (Kiss et al. 2015; Froese, Iizuka, and Ikegami 2014) as well as in education contexts with students or children (Ijaz, Bogdanovych, and Trescak 2017). However, most of these studies utilized VR in a non-immersive form, relied on specific stationary simulators/hardware for training or did not consider immersive VR-interaction or cognitive demands (e.g. Alhalabi 2016; Jimeno-Morenilla et al. 2016).

In immersive VR training contexts the concepts of reality-based interaction (RBI) or natural user interfaces (NUI) are implicitly involved in three-dimensional virtual experiences (Jacob et al. 2008; Wigdor and Wixon 2010) as touch and gestures are

essential human behaviours in any reality, virtual or other. Although it has long been known that educational applications demand a clear understanding of the cognitive factors that may impair or enhance learning-outcome and learning-experience (Wickens 1992), the cognitive demands of RBI in VR education or training have not been widely explored in present research. While numerous studies have demonstrated benefits of VR training on learning/skill performance (Cannon et al. 2014; Cates, Lönn, and Gallagher 2016; Saleh et al. 2013; Zhang et al. 2017; Jia, Bhatti, and Nahavandi 2014) only a few studies have specifically examined the cognitive demands of VR training (Lee and Lee 2017; Chao et al. 2017). The work of Sweller et al. (2011a; 1998) on cognitive load in instructional design identified three cognitive load types (intrinsic, germane and extraneous cognitive load) which additively generate the aggregated cognitive demand, i.e. the mental workload (MWL). MWL demands, not involved in processing the learning content and knowledge acquisition, for example interacting with learning-media, are extraneous cognitive loads and should be avoided or reduced for an effective learning experience (Paas, Renkl, and Sweller 2003). It is therefore imperative to investigate the scientific gap in VR-interaction regarding educational application and the effects on MWL.

1.2 Working memory and mental arithmetic

Specifying MWL as a measure implies that mental efforts are measureable on a model of mental resources. For this study we follow the suggestions of Kahneman (1973) that human cognitive resources are limited and task performance requires capacity (Norman and Bobrow 1979) in accordance with the established working memory (WM) model by Baddeley and Hitch (1974) in its latest conception (Baddeley 2002). In this respect, MWL cannot be regarded as a single dimension since not all tasks compete for one resource-pool (Wickens 2008). In fact, data originating from divided attention tasks

(Kantowitz 2000; Kantowitz and Knight 1974; Kantowitz and Simsek 2001; Wickens 1976) provide evidence of differentiated resource-pools for auditory and visuospatial subsystems in WM. Consequently, the current understanding of WM involves two separate subsystems that both are suggested to have certain temporary storage-capacity – the phonological loop and the visuospatial sketchpad – which will be referred to as auditory and visuospatial WM in this paper. Besides those independent subsystems, WM is believed to have a central executive (CE) component with no storage-capability but functioning as attention control where attentional capacity is focused, switched or divided and a component termed episodic buffer to address questions of long-term memory (LTM) retrieval into WM and interaction between subsystems (Baddeley 2002). Further, it is believed to also play a role to facilitate chunking using LTM access (Baddeley 2002; Miller 1956). Naturally, with regards to measuring MWL in such a distinct framework of WM it has to be assured that tasks compete on the same resources if specific impacts are to be assessed. The key elements to consider for MWL-assessment and resource-distribution in respect of this study of MWL in a VR mental arithmetic experience are VR-perception/interaction on one hand and mental arithmetic processing on the other.

As studies on spatial sound perception in VR have shown positive effects on MWL with attribution of load reduction concerning visuospatial WM overload (Flanagan et al. 1998; Nelson et al. 1998), it can be inferred that both, auditory and visuospatial WM are substantially involved in MWL with VR-perception.

Aspects of WM and mental arithmetic tasks have been investigated thoroughly over the last few decades. Undoubtedly, WM is involved in simple arithmetic processes (Adams and Hitch 1997; Ashcraft 1995; Ashcraft and Krause 2007). Studies have confirmed the involvement of all WM subsystems/components in mental arithmetic, CE and episodic buffer (Lemaire 1996; Logie and Baddeley 1987), auditory WM (Fürst and

Hitch 2000; Hecht et al. 2001; Logie, Gilhooly, and Wynn 1994; Otsuka and Osaka 2015; Seitz and Schumann-Hengsteler 2002) and visuospatial WM (Clearman, Klinger, and Szűcs 2017; Geary et al. 2000; Kyttälä and Lehto 2008; Li and Geary 2013) while pointing out varying involvement from childhood to adulthood (Kyttälä, Aunio, and Hautamäkki 2010; Metcalfe et al. 2013; Szűcs et al. 2014).

Additionally, Best and colleagues (2011) argue that different types of mathematic tasks are related differently to WM-components. For instance, calculating is suggested to be linked more strongly to fact retrieval and would therefore require less attentional control through the CE (Best, Miller, and Naglieri 2011). However, fact retrieval is associated with demand on the episodic buffer which would suggest a shift of demand in WM rather than a reduction (Baddeley 2003). It is, in this regard, also of importance if the mental arithmetic task is one of production or verification (i.e. the result is to calculate, or it is presented with questioning true or false), as a proposed answer facilitates fact retrieval (Lemaire 1996). It can be concluded from the present evidence that auditory WM, visuospatial WM, CE and episodic buffer are all to varying degrees implicated in mental arithmetic, a conclusion also drawn by DeStefano and LeFevre (2004).

1.3 Motivation and digital game-based learning

MWL not only represents a determinant for learning outcome regarding efficiency but may also be linked to the motivational capacity of VR training. The connection between MWL and intrinsic motivation has been demonstrated through empirically tested theories of Csíkszentmihályi (2014), who found that a state of intrinsic motivation resulting in ideal concentration and holistic immersion (flow-experience) is likely to be reached when a task meets the appropriate level of challenge between overload and boredom. Thus, it requires a specifically designed MWL-level that represents a person or group specific

challenge-skill balance to induce intrinsic motivation.

Related studies (Bhagat, Liou, and Chang 2016; Smith and Ericson 2009) indeed found applying VR in a training or learning context results in added motivational benefits (i.e. an enjoyable experience) but did not assess the origin of motivational effects or any cognitive demands. However, which specific factors contribute to motivational benefits is debatable as interaction style within different VR training studies is variable and scenarios are frequently entangled with gaming elements or include game-based interaction-sequences. Whereas some scholars argue motivational effects transpire from immersion and feelings of presence in virtual environments (Psotka 2013), serious gaming literature shows that digital game-based learning (DGBL) (Prensky 2003) represents in itself a concept which is proven to be capable of inducing intrinsic motivation (Dörner et al. 2016). Empirical studies have established a fair evidence-base that DGBL improves student engagement and motivation (Hamari et al. 2016; Papastergiou 2009; Hung, Huang, and Hwang 2014; Erhel and Jamet 2013) and also content learning (Lester et al. 2014; Perrotta et al. 2013). In addition, there are some studies which have investigated 2D, 3D mobile and non-mobile video games in combination with mathematics training in early education with mixed outcomes. Whereas newer empirical studies on DGBL with mathematics context focused on investigating motivation as well as learning performance and could sometimes acknowledge positive effects (Hsiao et al. 2016; Hung, Huang, and Hwang 2014; Sutopo and Pamungkas 2017) and sometimes not (Ke 2013), it is striking that an early study of Elliot et al. (2002) underlined how interaction-difficulties (i.e. extraneous cognitive load) impaired the transfer of motivational effects with VR mathematics training. Recent systematic reviews of the empirical database by Boyle et al. (2016) and Hainey et al. (2016) still point out the need for further empirical investigation of differences between 2D and 3D-immersive

DGBL. Indeed, as mentioned, most of the existing research within DGBL/VR training does not provide insight into the origination of discovered effects regarding VR and/or gaming concepts.

With respect to the origins of motivational effects it is worth considering Spitzer's theory of learning (1996; 1995) that suggests any activity can be made motivating by adding a motivational context, i.e. the addition of motivators including action (interactivity), fun, variety, choice and challenge. Besides being linked to the pre-conditions for the aforementioned flow-experience, these motivators also suggest that interaction-modality affects motivation through physical interaction-challenges or eliciting emotional reactions, which essentially adds further intricacy to the question about origin of motivation in VR educational training. Furthermore, Hassenzahl (2010) points out, emotion, cognition, motivation and action are inseparably entwined in a user experience (UX). Consequently, the emotional experience resulting from RBI with virtual environments may create a different level of intrinsic motivation than interacting in a non RBI-way with buttons on a gamepad.

1.4 Empirical research approach

Recent related studies have focused on comparing traditional training-methods with VR (Chao et al. 2017; Ijaz, Bogdanovych, and Trescak 2017) or usability in VR with non-RBI interaction (Xu and Ke 2016; Xiong et al. 2016; Shin, Biocca, and Choo 2013). Our view is that, to create a holistic, motivating and also effective VR learning experience, the effects of RBI in VR specifically with respect to MWL and UX must be identified. Our study, therefore, in addressing the MWL of RBI and RBI effects on UX, respectively, explored the potential connection between physical and psychological challenge and demand as illustrated in Figure 1.

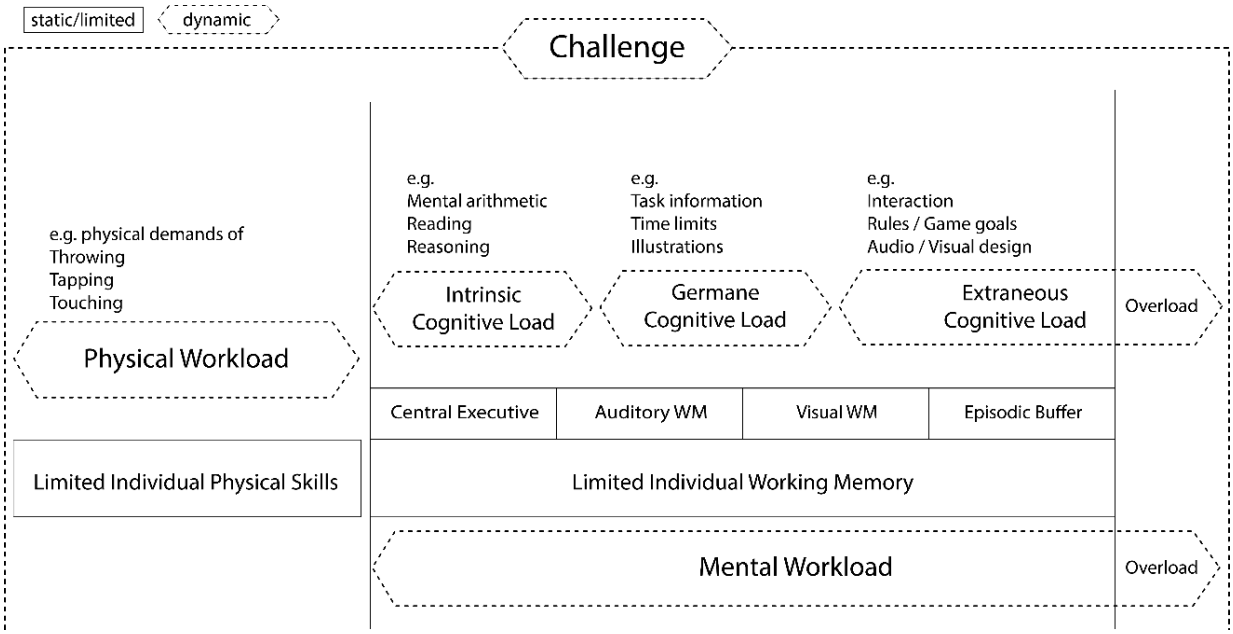

Figure 1. Challenge–demand balance in VR game-based learning experiences as elaborated from Kahneman (1973), Baddeley (2002) and Sweller et al. (2011b)

The purpose of the empirical study consequently was twofold:

(1) assess the MWL of RBI in VR game-based training by comparing measures of MWL (perceived difficulty of the mental arithmetic task and mental arithmetic performance) when performing a mental arithmetic game (MAG) under one of three presentation/interaction-configurations (between-subjects study-design):

(a) virtual reality display with reality-based interaction, natural interaction (VRMAG-RBI)

(b) virtual reality display with interaction controlled by tapping the VR headset (VRMAG-HMD)

(c) non-VR conventional tablet display with touchscreen interaction (TMAG)

(2) identify differences in learner-motivation through joy-of-use originating from the presentation/interaction-differences.

A game-based mobile-app was developed for training mental arithmetic skills for . secondary education level at key stage 3. The mental arithmetic game was produced in both mobile VR and a tablet-computer format. The VR mental arithmetic training game (VRMAG) offers two different interaction-modalities; one representing a natural ball-throwing gesture (VRMAG-RBI) and the other representing conventional head mounted display (HMD)-interaction by tapping at the designated touch-area on the HMD to 'throw' a ball (VRMAG-HMD). The tablet-variant (TMAG) was visually identical except for the fact that it is non-VR and presented on a 2D-screen with touchscreen input such as swiping/flipping to 'throw' a ball.

2 Methodology and materials

2.1 Evaluation methods and hypothesis

There are four basic empirical strategies to assess MWL (Wickens et al. 2015, Fourth edition:530–548): primary task (PT) performance, secondary task (ST) performance, psychophysiological measurement or subjective rating scales (SRS). As this study involved intensive movement by executing throwing gestures repeatedly, validity impairment of psychophysiological measures (e.g. heart rate variability, electrodermal activity, eye-blink rate) would be unavoidable and so these measures not applied.

ST-measurement requires a clear distinction on which WM subsystem (i.e. auditory or visual) the demands are competing. As the task in this study was not channel-exclusive, this would provide insufficient insight on total added MWL originating from the three different presentation/interaction-models (Sharples and Megaw 2015, 532).

In consequence, this study followed the recommendations of Sharples and Megaw (2015, 543) by combining subjective MWL rating with objective MWL–performance-measurement. As argued by DeStefano and LeFevre (2004) primary–task-measurement is

advisable due to involvement of all WM components in cognitive arithmetic. Therefore, primary task metrics including total points (correct solutions), wrong solutions, balls thrown, balls missed, falsely thrown balls and others were logged within the game. In order to measure MWL effects in performance, the PT has to be of high cognitive demand as otherwise reserve capacity (Wickens et al. 2013, 350) and coping strategies (Hancock and Warm 2003) will prevent performance decrease. Thus, for the mental arithmetic tasks in this study teachers were asked to provide appropriate levels of calculation demands suitable for the education level of the children in the study (see Table 2).

With respect to measurement of ease-of-application, we applied a simple post-task question taken from - the subjective mental effort question (SMEQ) (Salvendy 2012; Sauro and Dumas 2009) focusing on the question of mental arithmetic calculation difficulty was applied for this research.

To assess perceived challenge and motivation 14 bipolar adjective pairs of the hedonic qualities of the AttrakDiff2 questionnaire (Hassenzahl et al., 2003, Hassenzahl 2010) were rated on a 7-point semantic differential scale. Perceived physical effort and perceived fun, as well as willingness to train with the game, were additionally evaluated on a 7-point rating scale. All questions were presented to participants on a printed questionnaire sheet post-task.

To investigate MWL and UX this study established three null hypotheses:

> H_{0a}: 'There are no significant differences between the three interaction/presentation variants in the mental arithmetic performance.'
>
> H_{0b}: 'There are no significant differences between the three interaction/presentation variants in perceived difficulty of the mental arithmetic task.'

H_{0c}: 'There are no significant differences between the three interaction/presentation variants in perceived user experience.'

2.2 Game design and interaction variants

2.2.1 Concept of mental arithmetic game

The underlying concept of the mental arithmetic game was a mathematical verification-task with a true answer result and two false confusion results. For example, Figure 2 displays a mathematical calculation task (147+50=) and three given answers, only one of which is correct. The player is required to select the correct answer by throwing a ball in the gate with the correct answer. Table 1 shows the resulting types and level of cognitive math-tasks included in the game. The two false results were computed by randomly adding or subtracting a number between 1 and 5 to or from the correct result.

Table 2. Determined mental arithmetic tasks as established with secondary school math teacher

Operation	Range	Restrictions	Example
Addition	Summands between: 1 – 200	Max. summand = 200	196 + 42
Subtraction	Minuend / subtrahend between: 1 – 200	Max. minuend and subtrahend = 200	57 - 111
Multiplication	Factors between: 1 – 200	Max. multiplicand = 20; Product divisible by 10	115 x 18
Division	Dividend between: 1 – 200 Divisor between: 1 – 100	Max. divisor = 100; Without a remainder	176 : 88
Exponentiation	Base between: 1 – 15	Max. exponent = 2	9^2
Square Root	Radicand between: 1 – 200	Only integer results	$\sqrt{121}$

The game-environment displays an outdoor setting at a lake surrounded by mountains where three goals are presented with one correct result and two obfuscated results for a mental arithmetic task shown above the gates (Figure 2). Acoustic atmosphere and effects for the ball-throwing-interaction were implemented to form an enjoyable experience and provide auditory feedback for the throwing-interaction.

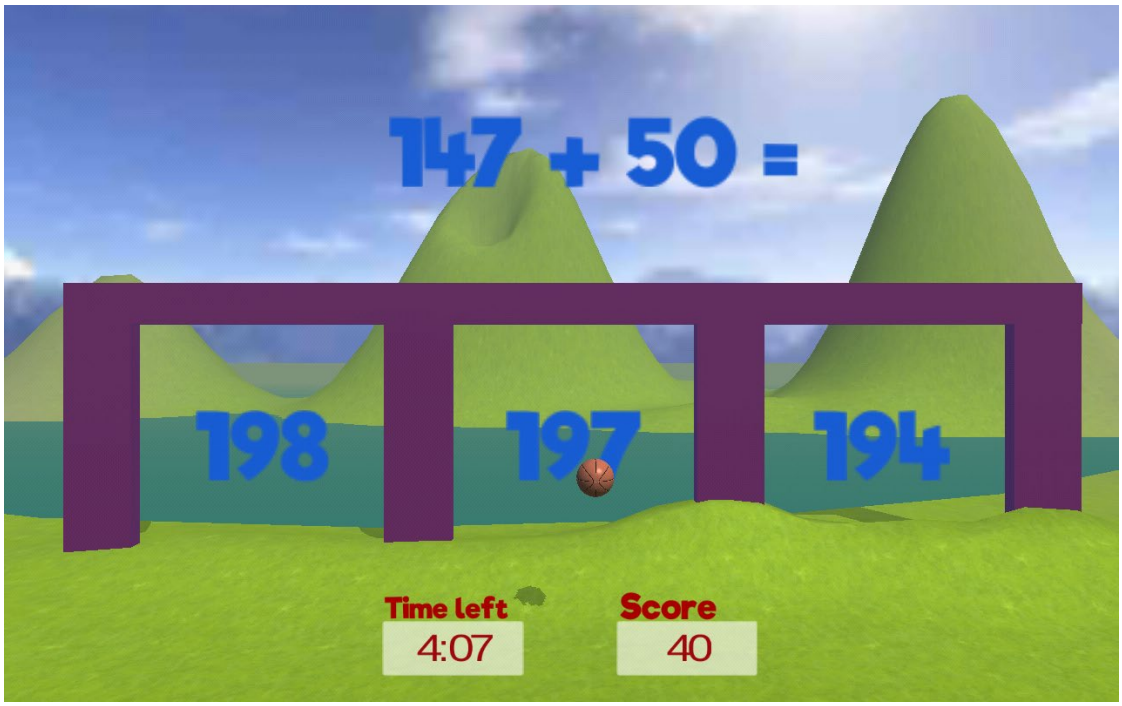

Figure 2. Illustration of the mental arithmetic game as presented in VRMAG conditions. The player is required to throwing a ball through the gate in which the correct result is shown.

The primary goal for the game was to throw a ball through the gate with the correct solution to each arithmetic task presented. A timer displayed on the screen counted down from 5 minutes to zero and the player was instructed to achieve a high score by correctly answering as many arithmetic tasks as possible in the time limit. For every ball thrown into the correct gate, the player was awarded 10 points; whereas for each ball thrown into an incorrect gate 10 points were deducted and missing all gates had no impact on the score. It was therefore possible to end with a negative score. Throwing the ball into an incorrect gate could result from selection of an incorrect answer or a poor throw of the ball hitting an

incorrect answer by accident. Furthermore, it was possible to miss the gate or hit the border requiring the player to throw another ball.

The mental arithmetic game was created in the Unity 3D game development environment (Version 5.3.2f1) and programmed in C# (Mono .NET 2.0). The same game was used in all three experimental conditions with differences only in 2D-screen/3D-VR-display presentation. Whereas the visual design on a smart phone or tablet is presented from a static (camera) viewpoint in 2D, the smart phone can be inserted in a portable HMD mount such as the Samsung GEAR VR, which subsequently enables an immersive 3D-VR-experience of the same visual presentation allowing free rotation of the viewpoint through head-tracking. In order to assess influence of interaction method on MWL the interaction mechanisms described below were used.

2.2.2 Experimental condition A: VR reality-based interaction (VRMAG-RBI)

This throwinggesture should match the natural interaction in having a controllable release moment so that it matches the real experience to release the ball with the appropriate velocity.

A gesture ring by Nod, Inc., Mountain View, USA was used to develop the RBI. The version used for this study was an unobtrusive finger-ring with a touch-sensitive area. It connects to the smartphone via Bluetooth LE and is capable of tracking acceleration and rotations of the finger/hand through integrated inertial sensors with an accuracy of millimetre resolution. However, the ring was designed for adults and was too big for the children in the study.

Figure 3. Nod gesture ring with thumb-facing touch-area / 3D-printed flexible adapter

Therefore, we designed flexible ring-size adapters to enable comfortable fitting of the ring on a child's index finger (Figure 3), resulting in EU-ring-size range of 51 to 59 mm.

Figure 4. Girl throwing virtual ball in *variant A* - VRMAG RBI

Use of the gesture ring enabled players to throw the ball in the VRMAG RBI variant with the same action as they would throw a ball in reality. Instead of holding a ball and releasing it at the desired moment, they hold the touch-area with the thumb and release it at the same moment they would release the ball in reality. Figure 4 shows a participant performing the throwing sequence. The speed of the throwing gesture was tracked using

the accelerometer sensor-data, thus, the velocity of the ball was determined from the throwing motion intensity. The physics engine of Unity 3D calculated the flying curve of the virtual ball in VR according to its mass and the tracked acceleration. Gaze direction of the player was assessed through head-tracking of the GEAR VR and was taken as the desired throwing direction.

2.2.3 Experimental condition B: VR HMD conventional tap interaction (VRMAG-HMD)

Game variant B utilised a more conventional VR-interaction paradigm using the integrated touchpad on the GEAR VR HMD. Players could initiate a ball-throw by tapping on the 4 x 5 cm touch area on the right side of the HMD. A tangible elevation of approximately half a millimetre in diameter in the centre of the touch area facilitated positioning of the finger on the device. The tap interaction (Figure 5) initiated a ball-throw in the direction of aim as determined by the player's head orientation / eye gaze, identical to variant A.

Figure 5. Girl "throwing" virtual ball in *variant B* - VRMAG HMD-tap interaction

However, the velocity of the throw could not be influenced by the players in this interaction variation. Instead, the virtual ball was tossed at the time of tap recognition with a constant force (i.e. initial velocity of 40 m/sec), which was assessed and determined in user testing during development. This velocity was also reached by the children with variant A when throwing with the expected swing of the arm.

2.2.4 Experimental condition C: Tablet computer swipe/flick gesture (TMAG)

Game variant C was presented on a tablet computer. The Samsung Galaxy Note 8 features an LCD-TFT display with 1280 x 800 px resolution resulting in 189 ppi. While the game on the tablet was presented from a 2D fixed point-of-view, all other visual and auditory properties were identical to the above described VR variants. To "throw" a ball, players had to perform a flick-gesture on the touchscreen of the tablet from the centre of the viewpoint in the direction of the goal intended (Figure 6). Thereby, the speed of the thrown ball was derived from the speed of the executed flick-gesture. More specifically, the time-span between start and end of the flick-gesture was tracked and correlated with the speed of the thrown ball. Just as with the other variants, balls could miss the gates or hit the posts and bounce back according to physics calculation from the game engine.

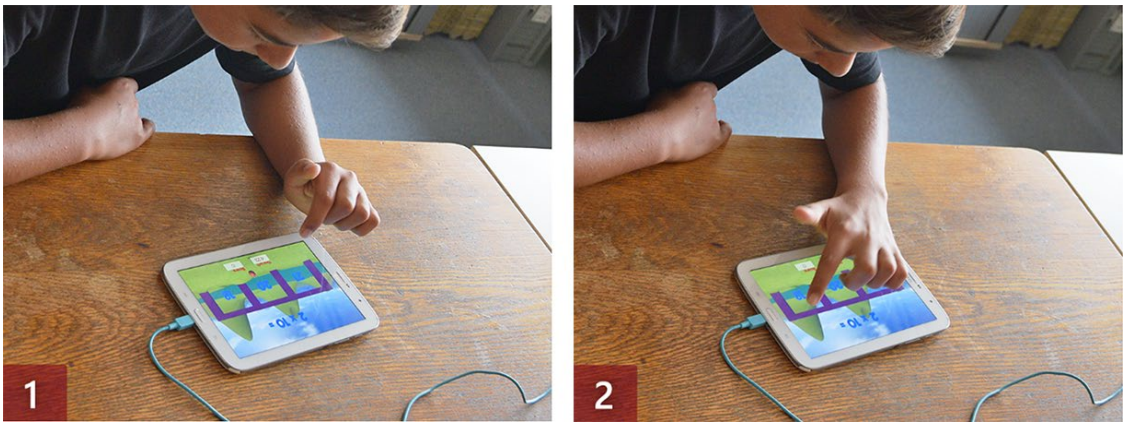

Figure 6. Boy "throwing" virtual ball in *variant C* - TMAG touch-flick interaction

2.3 Study design and procedure

2.3.1 Participants

Participants were secondary school students (n = 60; 32 girls, 28 boys) from two different schools (henceforth referred to as school m and school g) aged between 12 and 14 years (mean age: 12.95 years) enrolled in various different classes at key stage 3 level. The two

schools were differentiated by general socio-economic background with pupils in school g tending to come from families with a more academic background and pupils in school m originating from working/middle class families. Participating pupils were healthy and had no physical or psychological disabilities. All participating students gave their written consent as did their legal representatives and their form teacher and school management. Participants did not receive financial compensation but were given the mental arithmetic game for home usage if desired. Ethical approval for the study was granted by the Faculty of Engineering Ethical Committee, University of Nottingham.

2.3.2 Procedure and experiment location

The study was designed as a randomised controlled trial (RCT) and took place at each of the two schools to allow the children to participate within an accustomed educational setting.

The students were randomly assigned to the experimental conditions which resulted in the demographic distribution for the study as shown in Table 3.

Table 3. Participant allocation to each experimental condition (game variants A, B, C)

		Study Condition – Game variant			Total *n*
		A	B	C	
Gender	w	10	10	12	32
	m	11	10	7	28
Total		21	20	19	60

For assessing MWL the game performance measures were logged as PTM with implemented logging algorithms. The adapted AttrakDiff2 and SMEQ scales were assessed using a paper-based post-task. Students were also asked to rate their experience

with the game with respect to fun, physical demand and their willingness to train with the game on a 7-point rating scale (1, not at all; 7, very much) (Albert and Tullis 2013, 128) and had the option to add their own comments at the end of the questionnaire.

Participating students were called in the randomly assigned order from their ongoing classes to the test-rooms. Each room was supervised by a research assistant who instructed the pupil with the specifics of the interaction/presentation-variant and the assessment-procedure before the training. Ring-size and HDM were fitted according to the needs of each participant before starting the game.

3 Results

3.1 Mental Workload – Primary Task Measurements

For analysis of PTM three participants (two in group A and one in group B) were excluded from the ball-throwing measures as the logging metrics could not be retrieved from their training session. Furthermore, two pupils (one of school m and one of school g) omitted valuating one attribute-pair in the semantic differential which was also considered in the analysis.

Table 4. Primary task performance measures: mean, standard error and variance-analysis results

	Variant							
	C / Tablet flick		B / VR HMD-tap		A / VR RBI			
	Mean	SE	Mean	SE	Mean	SE		
Total points	**300.00**	28,55	**223.00**	24.46	**170.48**	24.60	F (2,57) = 6.31	p = .003
Balls thrown	45.00	2.95	42.00	2.57	42.26	3.78	F (2,54) = 0.28	p = .757
Balls correct %	**77.91**	3.36	**74.84**	3.70	**59.57**	3.86	F (2,54) = 7.27	p = .002
Balls false %	**10.94**	1.82	**18.22**	2.69	**16.28**	2.53	F (2,54) = 2.51	p = .090
Balls missed %	**11.16**	3.02	**6.94**	2.11	**24.61**	2.76	F (2,54) = 12.05	p < .001

Variance-analysis of PTM revealed significant differences between the interaction/presentation-variants in key performance metrics (Table 4) with large effect sizes. Total points score differed significantly among the three interaction/presentation-variants, F (2, 57) = 6.31, p = .003, η_p^2 = .181. Similarly, the quota (i.e. percentage) of correctly thrown balls, F (2, 54) = 7.27, p = .002, η_p^2 = .212 and quota of balls missing all target-goals F (2, 54) = 12.05, p < .001, η_p^2 = .309 displayed considerable differences with the latter showing the highest significance and effect size.

The specific significant differences between the three variants concerning PTM were subsequently exposed by Games-Howell post-hoc testing. Mean performance of the learners with VR RBI variant A was significantly (p = .004) lower compared to the tablet-variant C with a mean decrease of approx. 130 points (*95%-CI* [-221.58, -37.46]) which is illustrated in Figure 7.

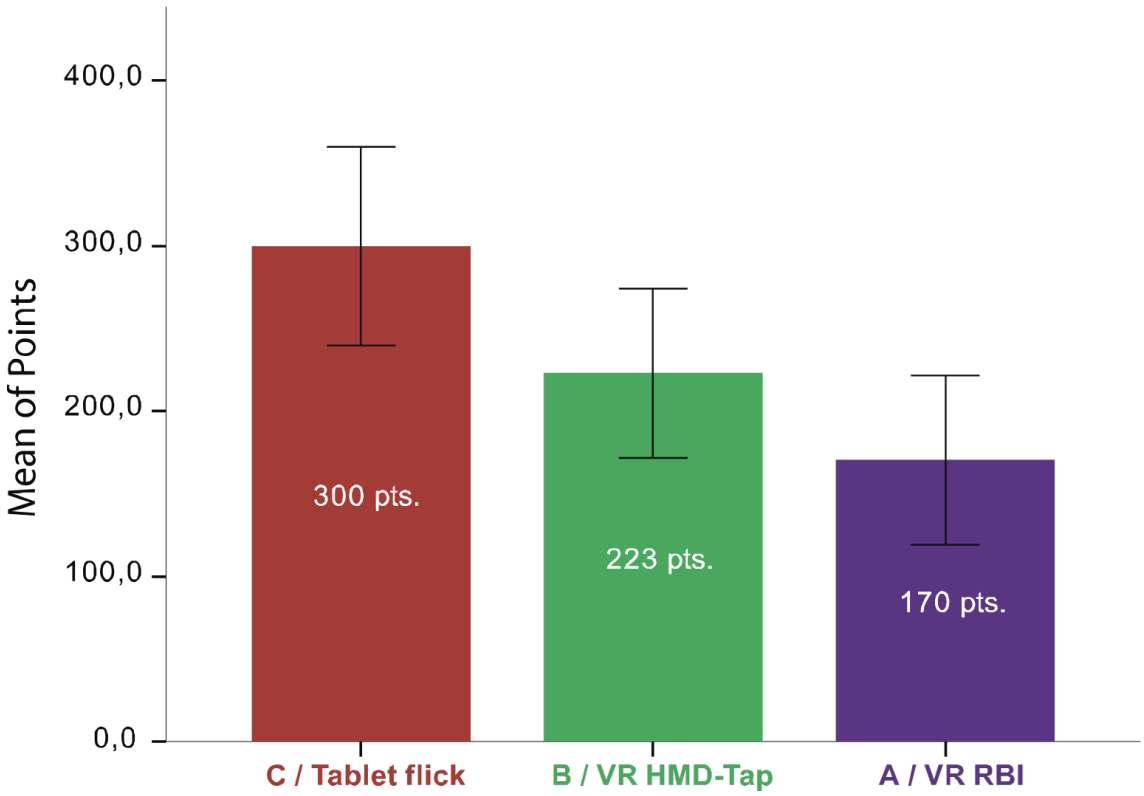

Figure 7: Primary task performance measures: mean total points in the mental arithmetic game with *95%-CI* indication

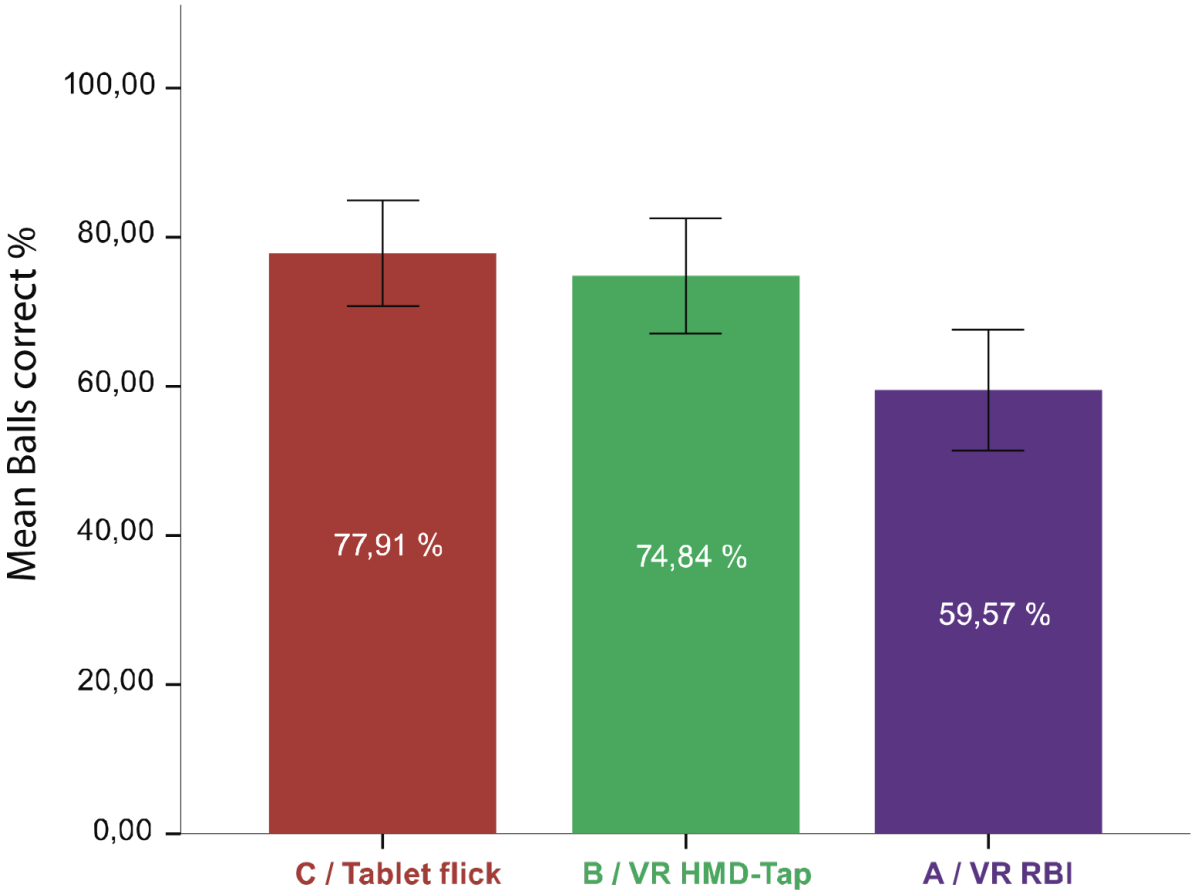

Figure 8: Primary task performance measures: mean percentage of balls hit at the correct solution with *95%-CI* indication

Considering the quota of correctly thrown balls, Figure 8 displays the significant (p = .003) mean decrease of 18.34% (*95%-CI* [-30.85, -5.83]) between participants from group A to C a measure that also differed between A to B significantly (p = .019) with 15.27% (*95%-CI* [-28.3, -2.21]) less correctly identified results in the VR RBI-group. Students performing in variant B were, on average, almost as good (-3.07%, *95%-CI* [-15.28, 9.15]) in hitting the correct solution as the tablet control-group C with no significant difference (p = .814).

Comparatively, another PTM-measure – the thrown balls that missed all targets – was identified as statistically differing through Games Howell analysis (Figure 9).

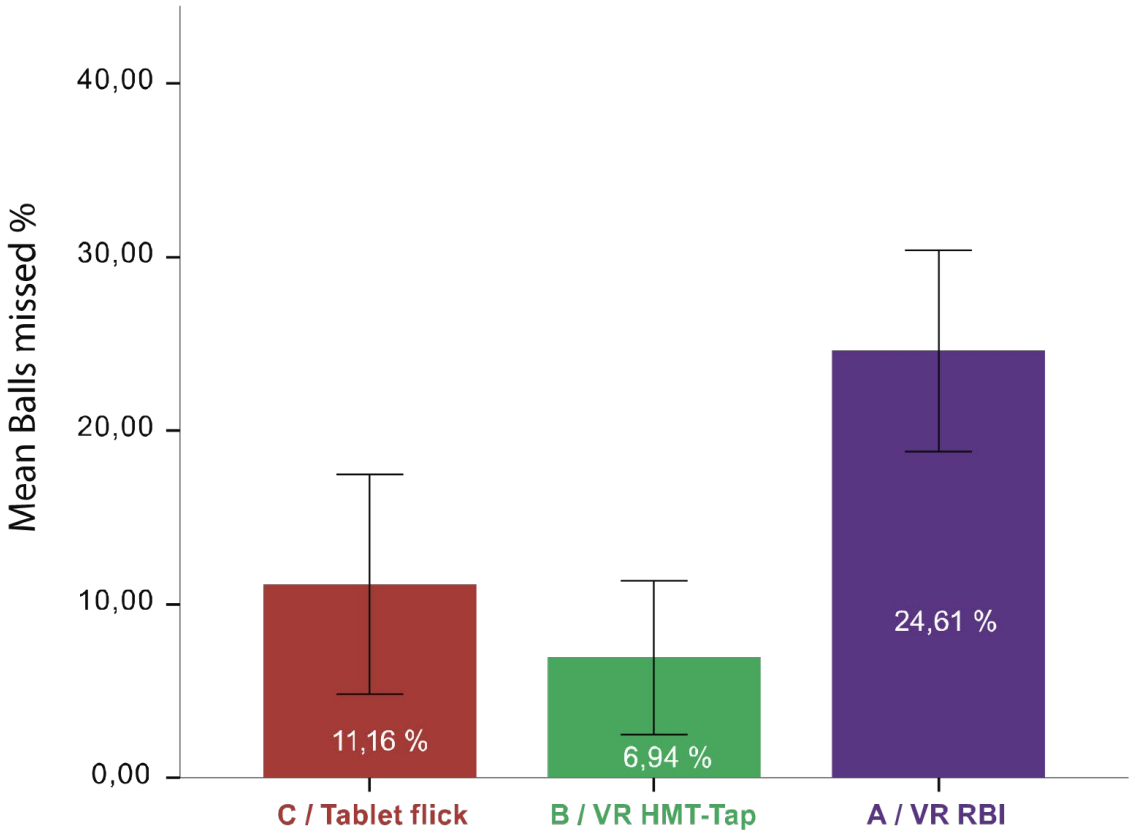

Figure 9: Primary task performance measures: mean percentage of balls missing all targets with 95%-CI indication

Students in RBI-group A did, on average, miss all targets significantly (p = .006) more often than the pupils in group C (13.4%, *95%-CI* [3.45, 23.46]) and group B (17.6%, p < .001, *95%-CI* [-9.15, 26.19]). Learners in VR-group B performed best on average regarding accuracy but this was statistically (p = .494) better than tablet control-group C (-4.3%, *95%-CI* [-13.27,4.83]).

3.2 Mental Workload – Perceived Mental Workload

While PTM as one part of the MWL-assessment has exposed several significant findings between the three variants, variance analysis of the perceived difficulty of the game also revealed significant differences, F (2, 57) = 5.46, p = .007 with large effect size, η_p^2 = .161 in support of PTM-results (Table 5).

Table 5. Perception of arithmetic difficulty, fun and physical demand: mean, standard error and variance-analysis results

	Variant							
	C / Tablet flick		B / VR HMD-tap		A / VR RBI			
	Mean	SE	Mean	SE	Mean	SE		
Perceived difficulty of mental calculation (SMEQ)	**15.05**	2.66	**26.80**	3.14	**31.62**	4.57	F (2,57) = 5.46	p = .007
Training with game was fun	6.37	0.19	6.45	0.15	6.19	0.28	F (2,57) = 0.38	p = .687
I can imagine using the game to train mental arithmetic	6.42	0.25	6.20	0.28	6.57	0.21	F (2,57) = 0.59	p = .558
Training was physically demanding	1.42	0.27	1.60	0.22	2.05	0.30	F (2,57) = 1.46	p = .241

Note. SMEQ difficulty of calculation task (0, not at all;85, very hard to do);

Rating scale fun to train, physically demanding and desire to train with game: (1, not at all; 7, very much)

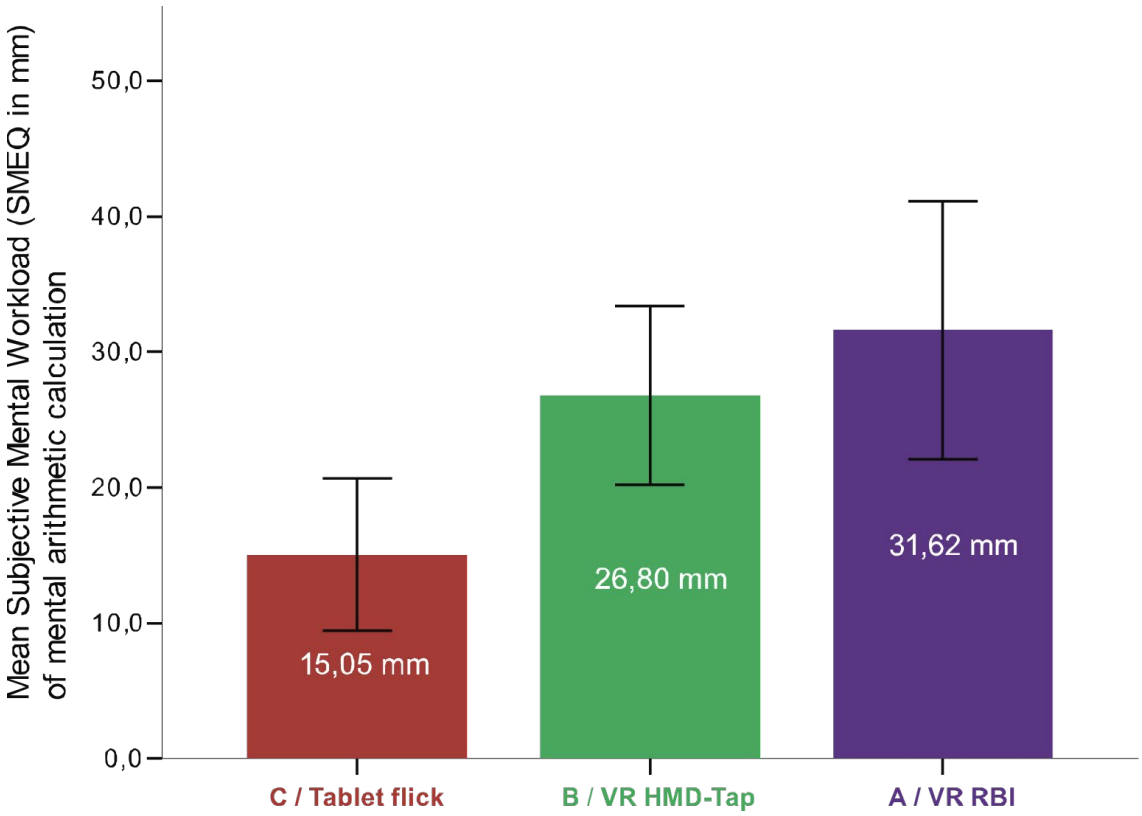

Figure 10: Mean perceived difficulty of mental arithmetic calculation (SMEQ – [0, not at all hard to do; 85 very hard to do]) with *95%-CI* indication

Figure 10 illustrates the findings from post-hoc tests and show that learners in RBI-group A perceived the arithmetic calculations as significantly (p = .010) more difficult than students in tablet control-group C with a mean increase of 16.57 mm on the SMEQ-scale (*95%-CI* [3.57, 29.56]). This represents a mean increase in the perceived MWL-scale from "not very hard to do" to "fairly hard to do". Furthermore, students in VR-group B controlling the game with HMD-tap also significantly (p = .019) felt the arithmetic tasks were more demanding than group C with an average increase of 11.75 mm (p = .019, *95%-CI* [1.681, 21.81]) which represents MWL-intensification from "not very hard to do" to "a bit hard to do" on the SMEQ-scale.

Supplementary statistical analysis between gender and schools on assessed MWL-data with Welch's t-test revealed no significant differences of PTM or subjective MWL between boys and girls. However, school m did not perform as well as school g, with a 68.14 points (*95%-CI* [-130.80, -5.48]) lower average score than school g, t (57.75) = -2.18, p = .034, d = -0.57, r = 0.27 and an average of 5.71 (p = .036, *95%-CI* [-11.03, -0.39]) fewer correctly verified solutions over all variants A, B and C, t (51.09) = -2.16, p = .036, d = -0.60, r = 0.29.

3.3 Motivation – Joy-of-Use

With regard to motivational attributes, the semantic differential (Figure 11) displayed relative consistent mean values over the variants. However, one-way ANOVA of the data – perception of attractiveness and stimulation – confirmed statistical differences for the attributes challenging/undemanding, F (2, 57) = 6.52, p = .003, η_p^2 = .186 and attractive/ugly, F (2, 57) = 4.09, p = .022, η_p^2 = .125.

Post-hoc testing revealed that VR RBI variant A was perceived as significantly (p = .005) more challenging than control variant C (-1.78, *95%-CI* [-3.1, -0.50]) but not

significantly (p = .826) more challenging than VR-control-group B (-3, *95%-CI* [-1.52, 0.93]).

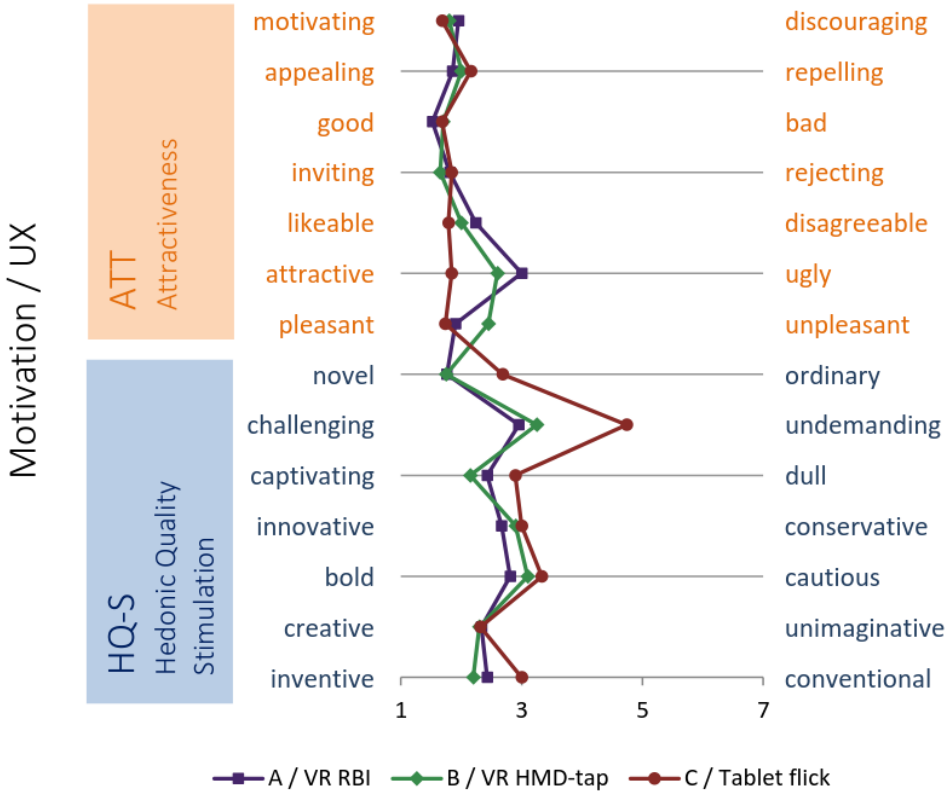

Figure 11: Semantic differential – motivational qualities: mean perception A, B, C

Additionally, learners had on average perceived the tablet-interaction/presentation version C significantly (p = .020) more attractive than VR RBI variant A (-1,16, *95%-CI* [-2.15, -0.16]) but not significantly (p = .142) more attractive than VR-control-group B (-0.76, *95%-CI* [-1.71, -0.20]).

Further analysis between gender and schools with respect to the gathered data on UX-dimensions did not reveal relevant disparities between girls and boys but showed again significant differences between the two participating schools. Participants from school m reported on average a higher perception of fun when training with one of the

game variants than school g. As rated on a scale from one to seven, children of school m perceived on average 8 percent (0.56, *95%-CI* [0.08, 1.03]) more fun in training with the game, t (43.52) = 2.30, p = .023, d = 0.71, r = 0.34 and indicated an approx. 10 percent (0.69, *95%-CI* [0.16, 1.23]) higher mean aspiration to use the game for mental arithmetic training by themselves, t (53.34) = 2.60, p = .012, d = 0,71, r = 0.34.

In general, statistical analysis suggests to reject all three null hypotheses ($H_{0a, b, c}$) as testing revealed significant differences among the three interaction/presentation-variants concerning mental arithmetic performance, perceived difficulty of the mental arithmetic task and perceived user experience.

4 Discussion and interpretation of the results

Observed behaviour, commentary of pupils as well as self-reported perception of excitement indicate that the participants of this randomised controlled trial generally enjoyed the training with all three presented variants to a very high degree. In fact, the question on perceived fun by training with the game was, on average, rated above 6 on a scale of 1 (not enjoyable) to 7 (very enjoyable) for all interaction/presentation-variants. The same findings were obtained for reported willingness to train with the mental arithmetic game (Table 4) with only marginal differences between schools. Most notably in this regard, neither immersive VR nor RBI improved the fun experience significantly as compared to the 2D tablet. On the contrary, while most of the perceived UX-qualities did not differ between variants, the results of data analysis revealed that pupils evaluated the tablet-variant as more attractive on average (approx. 17%) than VR-variant A. It is not clear whether the children rated attractiveness solely on visual design/presentation or also included other factors in this judgement; however, it could be inferred that when totally immersed in a virtual experience with no other visual sensations, the visual design has to be more appealing than when presented on a 2D display. Obviously, the intense

magnification of the display in the VR HMD reveals a pixilation effect which potentially adds to the reported difference. However, as the difference from control-group C to VR-variant B was not significant this remains to be investigated further.

On the other hand, both VR-variants presented a significantly higher challenge (approx. 24% higher for A than C) to the pupils as rated on UX-dimension stimulation. This would seem to suggest that immersion in VR does indeed have an isolated effect on stimulation/motivation (Psotka 2013) in representing a challenge. An immersive VR-content – even a solitary, passively observed 360° picture – would, according to this assumption, present a challenge to perceiving individuals. However, it is probable that the challenging effect in the case of this study originates rather from interaction than visual perception, immersion or presence. In fact, both VR-variants included the challenge to aim with the head/gaze while visually verifying the correct math solution from three possibilities. This competition on the visual WM channel could cause a higher demand or perceived challenge for the pupils. Participants in control-group C, on the other hand, could verify one result and flick the ball at another without regulating their gaze. This MWL-explanation of challenge is more concordant to the findings in this trial as no further differences in joy-of-use of the three interaction/presentation-variants could be detected. Indeed, it can be concluded that neither immersive VR nor RBI provided a substantial negative or positive effect on UX as compared to a tablet mental arithmetic game. Adding the gaming-context to the mental arithmetic task, however, created an experience that students reported as having fun training with and would like to train with regardless of interaction-modality or presentation-form specifics. Thus, the results on joy-of-use clearly support the supermotivation theory of Spitzer (1996) in this respect.

As far as MWL is concerned, the trial results on this key question are very clear and identified congruent findings observed in the students' performance and their

subjective perception of the arithmetic task-difficulty. The learners' performance declined significantly between tablet-group C and VR RBI A variant with a more than 43% lower average in total points; an approx. 18% lower mean in thrown balls hitting the correct arithmetic solution and around 13% higher average in tossed balls missing all targets in the VR RBI version. At the same time, the subjectively perceived difficulty of solving the mental arithmetic task was assessed on average as being more than twice as high by students in the VR RBI-group rising from "not very hard to do" to "fairly hard to do" compared to control-group C.

It is therefore inferred, that RBI does indeed significantly increase MWL while training or learning in VR compared to a non-VR tablet-variant mental arithmetic training. This finding contrasts with the view of Wickens (1992) that cognitive effort could be reduced by a "natural" interface. In fact, the findings suggest RBI presents a more extraneous cognitive load to the learners and should be avoided when aiming for effective learning outcomes (Paas, Renkl, and Sweller 2003; Sweller, Ayres, and Kalyuga 2011b) within a mental training task. Nonetheless, as studies on cognitive impact of learning-media interaction have shown (Holst, Churchill, and Gilmore 1997), it is required to determine exactly what cognitive demands of interaction could be a beneficial part of intrinsic cognitive load. For instance, motor skills training such as VR surgery simulation requires RBI for interaction as it is an essential training goal to perfect hand movements and RBI cognitive processes are thereby part of the intrinsic cognitive load.

Notably, as with the reported level of fun experienced in all variants, the pupils did not perceive any difference in physical demand when interacting with the three variations. Although the throwing gesture could be considered to be much more demanding than flicking a finger, the students did not say so. Additionally, we observation that the RBI produced very engaged, active learner participation with holistic

body movement whereas variant C, and sometimes also B. were performed in physically unfavourable sitting postures.

However, we can assume that part of the extraneous overhead in MWL is attributed to immersive VR-presentation regardless of interaction-mode as some results displayed significant differences in this regard. Calculation difficulty in variant B with HMD-tap interaction was subjectively perceived as more difficult compared to control-group C, and significant different, but to a lesser extent, was observed RBI-variant A and control-group C, but there was no difference in perceived difficulty between variants A and B. However, PTM-parameters showed that students in variant B performed best at hitting targets and almost as good at hitting correct arithmetic solutions as tablet-group C. Thus, the perceived cognitive overhead in variant B is likely to originate from other factors than interaction. It is possible that these reported higher demands are linked with the findings on higher challenge/lower attractiveness in UX and result from differences such as the visually more pixelated presentation. This needs further investigation to verify.

4.1 Study limitations

This study revealed and described significant differences in MWL and motivational effects regarding a mental arithmetic game with three different interaction/presentation-modes designed and developed for the research process. However, a few limiting aspects on these findings have to be considered. The assessment of MWL in this research is partly based on subjective perception as a post-task measure. Although supplemented and verified by PTM it might still be advisable in further studies to utilize methods that can actually address real-time changes in MWL during training such as psychophysiological measurements (e.g. eye-blink parameters). In this respect it is also worth considering methods that can clearly determine which subcomponent of WM is involved in the

MWL overhead between VR and non-VR as well as RBI and conventional interaction.

The findings of this research also do not distinguish the composition of MWL as the primary task – mental arithmetic – involves every WM-subcomponent to some degree. This should be altered for future VR RBI investigation to gather more detailed insight in the exact composition of MWL. Furthermore, although the findings in UX showed substantial perceived enjoyment within all groups and displayed congruence to the MWL-measures, more detailed examination of UX in VR RBI compared to other interaction-methods is required to determine specific influences.

The subjective post-task questioning in this study relied on common understanding of the presented bipolar adjective pairs. However, we consider that this understanding could be varying to some extent in a young group of participants. In its nature as a randomized controlled trial the study did not address for long-term effects in learning or possible habituation effects considering MWL of RBI. Furthermore, the study did not address extraneous MWL originating from game context as all tested variants incorporated the same game mechanics / DGBL concept, but instead assessed MWL focused on the added cognitive load, generated from RBI and VR-presentation.

5 Conclusions and implications for future research

This study explored the MWL and motivational effects of RBI and VR within an educational setting of mental arithmetic DGBL. The results clearly indicate a higher MWL for students with RBI-interaction when training with a mental arithmetic game in VR than training with a tablet game and flick interaction. The findings further suggest that immersive VR-presentation itself represents an extraneous cognitive demand, albeit to a smaller extent than RBI. These effects are to be regarded in an instructional design if learning is desired to be effective, particularly if intrinsic cognitive load is already high due to challenging learning content. Importantly, those differences in MWL should

further be considered when aiming for an appropriate challenge-skill balance to allow for flow-experience and intrinsic motivation. Moreover, the presented results suggest that neither VR-presentation nor RBI contribute significantly to the UX in mental arithmetic DGBL. In fact, the research outcome on UX leads to the conclusion that, in accordance with Spitzer (1996), adding an adequate gaming context to the task of mental arithmetic training creates a fun and motivating experience regardless of VR or non-VR presentation and RBI or conventional interaction. Nonetheless, the results of our study also reflected the basic model of Virtual Learning Environments of Dalgarno and Lee (2010) as both, representational fidelity and learner interactions proved as significant differntiators in the learning experience. This is an important finding concerning the soft issues of VR in education and paedagogy as finding ideal levels in these characteristics can maximise learning (Fowler 2015).

Our investigation of interaction within a learning context displayed clearly that DGBL approaches can result in positive training experiences on multiple platforms. We could demonstrate that a mental arithmetic application with a high user experience and motivation quality is easily adaptable for tablet computing and to mobile VR HMD but MWL is to be regarded considering cognitive demands of interaction. As key outcome for VR research our investigation showed significant impact of RBI on MWL in a VR training scenario and verified that a DGBL context can create fun learning experiences with little impact of presentation or interaction style. However, with respect to cognitive computing, big data and other emerging ICT research domains our study was also able to display an easy and usable approach for digitalisation in education settings that displayed excellent technology acceptance with a young target group. Creating a VRMAG based on our outlined approach that is widely used or distributed over social media can thereby for example provide the basis for deep learning scenarios (Lytras, Raghavan, and Damiani

2017; Lytras et al. 2018) or social media analysis (Lytras et al. 2015) and in consequence lead to improved effectiveness through personalisation of mental arithmetic training.

Future research should investigate the effects of RBI on MWL in a VR training context to distinguish MWL composition according to the individual WM-components and allow for creating advanced challenge-skill balanced instruction designs. Furthermore, familiarisation effects regarding MWL of RBI should be targeted in prospective investigations as well as long-term learning-outcomes with VR DGBL. Equally, possible advantages of VR-presentation compared to other DGBL presentation forms should be explored in more detail with respect to motivational benefits as this study did not indicate such effects. Ultimately, educators should feel strongly advised to incorporate DGBL in their tutoring schedule by the results of the presented study regardless of presentation form. However, identifying the appropriate challenge-skill balance in this endeavour remains imperative.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work has been partially funded by the Austrian Research Promotion Agency FFG through the COMET programme (5th call) research project LiTech sponsored by the Austrian Ministry for Transport, Innovation and Technology (BMVIT) and the Austrian Ministry of Science, Research and Economy (BMWFW).

ORCiD

Patrick Jost https://orcid.org/0000-0002-4972-9586

Sue Cobb https://orcid.org/0000-0002-4600-6235

References

Adams, John W., and Graham J. Hitch. 1997. 'Working Memory and Children's Mental Addition'. *Journal of Experimental Child Psychology* 67 (1): 21–38.

Alaraj, Ali, Cristian J. Luciano, Daniel P. Bailey, Abdussalam Elsenousi, Ben Z. Roitberg, Antonio Bernardo, P. Pat Banerjee, and Fady T. Charbel. 2015. 'Virtual Reality Cerebral Aneurysm Clipping Simulation With Real-Time Haptic Feedback': *Neurosurgery*, January, 1. doi:10.1227/NEU.0000000000000583.

Albert, Bill, and Tom Tullis. 2013. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. 2. Auflage. Amsterdam, Boston: Morgan Kaufmann.

Alhalabi, Wadee S. 2016. 'Virtual Reality Systems Enhance Students' Achievements in Engineering Education'. *Behaviour & Information Technology* 35 (11): 919–925. doi:10.1080/0144929X.2016.1212931.

Ashcraft, Mark H. 1995. 'Cognitive Psychology and Simple Arithmetic: A Review and Summary of New Directions'. *Mathematical Cognition* 1 (1): 3–34.

Ashcraft, Mark H., and Jeremy A. Krause. 2007. 'Working Memory, Math Performance, and Math Anxiety'. *Psychonomic Bulletin & Review* 14 (2): 243–248.

Baddeley, Alan D. 2002. 'Is Working Memory Still Working?' *European Psychologist* 7 (2): 85–97. doi:10.1027//1016-9040.7.2.85.

Baddeley, Alan D. 2003. 'Working Memory: Looking Back and Looking Forward'. *Nature Reviews Neuroscience* 4 (10): 829–839. doi:10.1038/nrn1201.

Baddeley, Alan D., and Graham Hitch. 1974. 'Working Memory'. *Psychology of Learning and Motivation* 8: 47–89.

Best, John R., Patricia H. Miller, and Jack A. Naglieri. 2011. 'Relations between Executive Function and Academic Achievement from Ages 5 to 17 in a Large, Representative National Sample'. *Learning and Individual Differences* 21 (4): 327–336.

Bhagat, Kaushal Kumar, Wei-Kai Liou, and Chun-Yen Chang. 2016. 'A Cost-Effective Interactive 3D Virtual Reality System Applied to Military Live Firing Training'. *Virtual Reality* 20 (2): 127–140. doi:10.1007/s10055-016-0284-x.

Boyle, Elizabeth A., Thomas Hainey, Thomas M. Connolly, Grant Gray, Jeffrey Earp, Michela Ott, Theodore Lim, Manuel Ninaus, Claudia Ribeiro, and João Pereira. 2016. 'An Update to the Systematic Literature Review of Empirical Evidence of the Impacts and Outcomes of Computer Games and Serious Games'. *Computers & Education* 94: 178–192.

Burdea, Grigore C., and Philippe Coiffet. 2003. *Virtual Reality Technology*. John Wiley & Sons.

Cannon, W. Dilworth, William E. Garrett, Robert E. Hunter, Howard J. Sweeney, Donald G. Eckhoff, Gregg T. Nicandri, Mark R. Hutchinson, et al. 2014. 'Improving Residency Training in Arthroscopic Knee Surgery with Use of a Virtual-Reality Simulator'. *J Bone Joint Surg Am* 96 (21): 1798–1806.

Cates, Christopher U., Lars Lönn, and Anthony G. Gallagher. 2016. 'Prospective, Randomised and Blinded Comparison of Proficiency-Based Progression Full-Physics Virtual Reality Simulator Training versus Invasive Vascular Experience for Learning Carotid Artery Angiography by Very Experienced Operators'. *BMJ Simulation and Technology Enhanced Learning* 2 (1): 1–5.

Chao, Chin-Jung, Sheng-Yu Wu, Yi-Jan Yau, Weng-Yan Feng, and Feng-Yi Tseng. 2017. 'Effects of Three-Dimensional Virtual Reality and Traditional Training Methods on Mental Workload and Training Performance'. *Human Factors and Ergonomics in Manufacturing & Service Industries* 27 (4): 187–196. doi:10.1002/hfm.20702.

Clearman, Jack, Vojtěch Klinger, and Dénes Szűcs. 2017. 'Visuospatial and Verbal Memory in Mental Arithmetic'. *The Quarterly Journal of Experimental Psychology* 70 (9): 1837–1855.

Csíkszentmihályi, Mihály. 2014. *Flow and the Foundations of Positive Psychology*. Dordrecht: Springer Netherlands.

Dalgarno, Barney, and Mark JW Lee. 2010. 'What Are the Learning Affordances of 3-D Virtual Environments?' *British Journal of Educational Technology* 41 (1): 10–32.

DeStefano, Diana, and Jo-Anne LeFevre. 2004. 'The Role of Working Memory in Mental Arithmetic'. European Journal of Cognitive Psychology 16 (3): 353–386.

Dörner, Ralf, Stefan Göbel, Wolfgang Effelsberg, and Josef Wiemeyer, eds. 2016. *Serious Games*. Cham: Springer International Publishing.

Erhel, S., and E. Jamet. 2013. 'Digital Game-Based Learning: Impact of Instructions and Feedback on Motivation and Learning Effectiveness'. *Computers & Education* 67 (September): 156–167. doi:10.1016/j.compedu.2013.02.019.

Flanagan, Patrick, Ken I. McAnally, Russell L. Martin, James W. Meehan, and Simon R. Oldfield. 1998. 'Aurally and Visually Guided Visual Search in a Virtual Environment'. *Human Factors* 40 (3): 461–468.

Fowler, Chris. 2015. 'Virtual Reality and Learning: Where Is the Pedagogy?: Learning Activities in 3-D Virtual Worlds'. *British Journal of Educational Technology* 46 (2): 412–422. doi:10.1111/bjet.12135.

Froese, Tom, Hiroyuki Iizuka, and Takashi Ikegami. 2014. 'Embodied Social Interaction Constitutes Social Cognition in Pairs of Humans: A Minimalist Virtual Reality Experiment'. *Scientific Reports* 4 (January). doi:10.1038/srep03672.

Fürst, Ansgar J., and Graham J. Hitch. 2000. 'Separate Roles for Executive and Phonological Components of Working Memory in Mental Arithmetic'. *Memory & Cognition* 28 (5): 774–782.

Geary, David C., Scott J. Saults, Fan Liu, and Mary K. Hoard. 2000. 'Sex Differences in Spatial Cognition, Computational Fluency, and Arithmetical Reasoning'. *Journal of Experimental Child Psychology* 77 (4): 337–353.

Hainey, Thomas, Thomas M. Connolly, Elizabeth A. Boyle, Amanda Wilson, and Aisya Razak. 2016. 'A Systematic Literature Review of Games-Based Learning Empirical Evidence in Primary Education'. *Computers & Education* 102: 202–223.

Hamari, Juho, David J. Shernoff, Elizabeth Rowe, Brianno Coller, Jodi Asbell-Clarke, and Teon Edwards. 2016. 'Challenging Games Help Students Learn: An Empirical Study on Engagement, Flow and Immersion in Game-Based Learning'. *Computers in Human Behavior* 54: 170–179.

Hancock, Peter A., and Joel S. Warm. 2003. 'A Dynamic Model of Stress and Sustained Attention'. *Journal of Human Performance in Extreme Environments* 7 (1): 4.

Hassenzahl, Marc. 2010. 'Experience Design: Technology for All the Right Reasons'. *Synthesis Lectures on Human-Centered Informatics* 3 (1): 1–95.

Hassenzahl, Marc, Michael Burmester, and Franz Koller. 2003. 'AttrakDiff: A Questionnaire to Measure Perceived Hedonic and Pragmatic Quality'. In *Mensch & Computer*, 187–196.

Hassenzahl, Marc, and Noam Tractinsky. 2006. 'User Experience - a Research Agenda'. *Behaviour & Information Technology* 25 (2): 91–97. doi:10.1080/01449290500330331.

Hecht, Steven A., Joseph K. Torgesen, Richard K. Wagner, and Carol A. Rashotte. 2001. 'The Relations between Phonological Processing Abilities and Emerging Individual Differences in Mathematical Computation Skills: A Longitudinal Study from Second to Fifth Grades'. *Journal of Experimental Child Psychology* 79 (2): 192–227.

Holst, Shirley J., Elizabeth F. Churchill, and David J. Gilmore. 1997. 'Balloons, Boats and Ponies: Interface Manipulation Style and Learning in a Constraint-Based Planning Task'. In *Human-Computer Interaction INTERACT '97*, edited by Steve Howard, Judy Hammond, and Gitte Lindgaard, 180–187. Boston, MA: Springer US. doi:10.1007/978-0-387-35175-9_33.

Hsiao, Indy Y. T., Stephen J. H. Yang, Tzi-Li Chang, Yu-Heng Wei, and Yu-Ju Lan. 2016. 'Creating a 3D Game-Based Learning System in a Virtual World for Low-Achieving Students in Mathematics'. In *2016 IEEE 16th International Conference on Advanced Learning Technologies (ICALT)*, 518–519. IEEE. doi:10.1109/ICALT.2016.37.

Hung, Chun-Ming, Iwen Huang, and Gwo-Jen Hwang. 2014. 'Effects of Digital Game-Based Learning on Students' Self-Efficacy, Motivation, Anxiety, and Achievements in Learning Mathematics'. *Journal of Computers in Education* 1 (2–3): 151–166.

Ijaz, Kiran, Anton Bogdanovych, and Tomas Trescak. 2017. 'Virtual Worlds vs Books and Videos in History Education'. *Interactive Learning Environments* 25 (7): 904–929. doi:10.1080/10494820.2016.1225099.

Jacob, Robert JK, Audrey Girouard, Leanne M. Hirshfield, Michael S. Horn, Orit Shaer, Erin Treacy Solovey, and Jamie Zigelbaum. 2008. 'Reality-Based Interaction: A Framework for Post-WIMP Interfaces'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 201–210. ACM.

Jensen, Katrine, Flemming Bjerrum, Henrik Jessen Hansen, René Horsleben Petersen, Jesper Holst Pedersen, and Lars Konge. 2015. 'A New Possibility in Thoracoscopic Virtual Reality Simulation Training: Development and Testing of a Novel Virtual Reality Simulator for Video-Assisted Thoracoscopic Surgery Lobectomy'. *Interactive CardioVascular and Thoracic Surgery* 21 (4): 420–426. doi:10.1093/icvts/ivv183.

Jia, Dawei, Asim Bhatti, and Saeid Nahavandi. 2014. 'The Impact of Self-Efficacy and Perceived System Efficacy on Effectiveness of Virtual Training Systems'. *Behaviour & Information Technology* 33 (1): 16–35. doi:10.1080/0144929X.2012.681067.

Jimeno-Morenilla, Antonio, José Luis Sánchez-Romero, Higinio Mora-Mora, and Rafael Coll-Miralles. 2016. 'Using Virtual Reality for Industrial Design Learning: A Methodological Proposal'. *Behaviour & Information Technology* 35 (11): 897–906. doi:10.1080/0144929X.2016.1215525.

Kahneman, Daniel. 1973. *Attention and Effort*. Prentice-Hall Inc.

Kantowitz, Barry H. 2000. 'Attention and Mental Workload'. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 44:3–456. SAGE Publications.

Kantowitz, Barry H., and James L. Knight. 1974. 'Testing Tapping Time-Sharing.' *Journal of Experimental Psychology* 103 (2): 331.

Kantowitz, Barry H., and Ozgur Simsek. 2001. 'Secondary-Task Measures of Driver Workload'. In *Stress, Workload and Fatigue*, edited by Pertar A. Hancock and Paula A. Desomond. Human Factors in Transportation. Mahwah, N.J: Lawrence Erlbaum Associates.

Ke, Fengfeng. 2013. 'Computer-Game-Based Tutoring of Mathematics'. *Computers & Education* 60 (1): 448–457. doi:10.1016/j.compedu.2012.08.012.

Kiss, Laura, Balázs Péter Hámornik, Máté Köles, Péter Baranyi, Péter Galambos, and György Persa. 2015. 'Training of Business Skills in Virtual Reality'. In *Cognitive Infocommunications (CogInfoCom), 2015 6th IEEE International Conference On*, 215–216.

Kyttälä, Minna, Pirjo Aunio, and Jarkko Hautamäkki. 2010. 'Working Memory Resources in Young Children with Mathematical Difficulties'. *Scandinavian Journal of Psychology* 51 (1): 1–15.

Kyttälä, Minna, and Juhani E. Lehto. 2008. 'Some Factors Underlying Mathematical Performance: The Role of Visuospatial Working Memory and Non-Verbal Intelligence'. *European Journal of Psychology of Education* 23 (1): 77–94.

Lee, Gyusung I., and Mija R. Lee. 2017. 'Can a Virtual Reality Surgical Simulation Training Provide a Self-Driven and Mentor-Free Skills Learning? Investigation of the Practical Influence of the Performance Metrics from the Virtual Reality Robotic Surgery Simulator on the Skill Learning and Associated Cognitive Workloads'. *Surgical Endoscopy*, June. doi:10.1007/s00464-017-5634-6.

Lemaire, Patrick. 1996. 'The Role of Working Memory Resources in Simple Cognitive Arithmetic'. *European Journal of Cognitive Psychology* 8 (1): 73–104.

Lester, James C., Hiller A. Spires, John L. Nietfeld, James Minogue, Bradford W. Mott, and Eleni V. Lobene. 2014. 'Designing Game-Based Learning Environments for Elementary Science Education: A Narrative-Centered Learning Perspective'. *Information Sciences* 264: 4–18.

Li, Y., and D. C. Geary. 2013. 'Developmental Gains in Visuospatial Memory Predict Gains in Mathematics Achievement'. *PLoS ONE* 8 (7): e70160.

Logie, Robert H., and Alan D. Baddeley. 1987. 'Cognitive Processes in Counting'. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13 (2): 310–326.

Logie, Robert H., Kenneth J. Gilhooly, and Valerie Wynn. 1994. 'Counting on Working Memory in Arithmetic Problem Solving'. *Memory & Cognition* 22 (4): 395–410.

Lytras, Miltiadis D., Naif Radi Aljohani, Amir Hussain, Jiebo Luo, and Jacky Xi Zhang. 2018. 'Cognitive Computing Track Chairs' Welcome & Organization'. In *Companion of the The Web Conference 2018 on The Web Conference 2018*, 247–250. International World Wide Web Conferences Steering Committee.

Lytras, Miltiadis D., Hassan I. Mathkour, Hassan Abdalla, Wadee Al-Halabi, Cornelio Yanez-Marquez, and Sean Wolfgand Matsui Siqueira. 2015. *An Emerging–Social and Emerging Computing Enabled Philosophical Paradigm for Collaborative Learning Systems: Toward High Effective next Generation Learning Systems for the Knowledge Society*. Elsevier.

Lytras, Miltiadis D., Vijay Raghavan, and Ernesto Damiani. 2017. 'Big Data and Data Analytics Research: From Metaphors to Value Space for Collective Wisdom in Human Decision Making and Smart Machines'. *International Journal on Semantic Web and Information Systems (IJSWIS)* 13 (1): 1–10.

Metcalfe, Arron WS, Sarit Ashkenazi, Miriam Rosenberg-Lee, and Vinod Menon. 2013. 'Fractionating the Neural Correlates of Individual Working Memory Components Underlying Arithmetic Problem Solving Skills in Children'.

Miller, George A. 1956. 'The Magical Number Seven, plus or Minus Two: Some Limits on Our Capacity for Processing Information.' *Psychological Review* 63 (2): 81.

Nelson, W. Todd, Lawrence J. Hettinger, James A. Cunningham, Bart J. Brickman, Michael W. Haas, and Richard L. McKinley. 1998. 'Effects of Localized Auditory Information on Visual Target Detection Performance Using a Helmet-Mounted Display'. *Human Factors* 40 (3): 452–460.

Norman, Donald A., and Daniel G. Bobrow. 1979. 'Descriptions: An Intermediate Stage in Memory Retrieval'. *Cognitive Psychology* 11 (1): 107–123. doi:10.1016/0010-0285(79)90006-9.

Otsuka, Yuki, and Naoyuki Osaka. 2015. 'High-Performers Use the Phonological Loop Less to Process Mental Arithmetic during Working Memory Tasks'. *The Quarterly Journal of Experimental Psychology* 68 (5): 878–886.

Paas, Fred, Alexander Renkl, and John Sweller. 2003. 'Cognitive Load Theory and Instructional Design: Recent Developments'. *Educational Psychologist* 38 (1): 1–4.

Papastergiou, Marina. 2009. 'Digital Game-Based Learning in High School Computer Science Education: Impact on Educational Effectiveness and Student Motivation'. *Computers & Education* 52 (1): 1–12.

Perrotta, Carlo, Gill Featherstone, Helen Aston, and Emily Houghton. 2013. 'Game-Based Learning: Latest Evidence and Future Directions'. *NFER Research Programme: Innovation in Education. Slough: NFER*.

Piedra, Jose A., Juan J. Ojeda-Castelo, Florencio Quero-Valenzuela, and Inmaculada Piedra-Fernandez. 2016. 'Virtual Environment for the Training of the Hands in Minimally Invasive Thoracic Surgery'. In *Games and Virtual Worlds for Serious Applications (VS-Games), 2016 8th International Conference On*, 1–4. IEEE.

Prensky, Marc. 2003. 'Digital Game-Based Learning'. *Comput. Entertain.* 1 (1): 21–21. doi:10.1145/950566.950596.

Psotka, Joseph. 2013. 'Educational Games and Virtual Reality as Disruptive Technologies.' *Educational Technology & Society* 16 (2): 69–80.

Rahm, Stefan, Karl Wieser, Ilhui Wicki, Livia Holenstein, Sandro F. Fucentese, and Christian Gerber. 2016. 'Performance of Medical Students on a Virtual Reality Simulator for Knee Arthroscopy: An Analysis of Learning Curves and Predictors of Performance'. *BMC Surgery* 16 (1). doi:10.1186/s12893-016-0129-2.

Rothbaum, Barbara O., Larry F. Hodges, David Ready, Ken Graap, and Renato D. Alarcon. 2001. 'Virtual Reality Exposure Therapy for Vietnam Veterans with Posttraumatic Stress Disorder.' *The Journal of Clinical Psychiatry*.

Saleh, George M, Julia Lamparter, Paul M Sullivan, Fiona O'Sullivan, Badrul Hussain, Ioannis Athanasiadis, Andre S Litwin, and Stewart N Gillan. 2013. 'The International Forum of Ophthalmic Simulation: Developing a Virtual Reality Training Curriculum for Ophthalmology'. *British Journal of Ophthalmology* 97 (6): 789–792. doi:10.1136/bjophthalmol-2012-302764.

Salvendy, Gavriel. 2012. *Handbook of Human Factors and Ergonomics*. John Wiley & Sons.

Sauro, Jeff, and Joseph S. Dumas. 2009. 'Comparison of Three One-Question, Post-Task Usability Questionnaires'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1599–1608. ACM.

Seitz, Katja, and Ruth Schumann-Hengsteler. 2002. 'Phonological Loop and Central Executive Processes in Mental Addition and Multiplication'. *Psychological Test and Assessment Modeling* 44 (2): 275.

Sharples, Sarah, and Ted Megaw. 2015. 'Definition and Measurement of Human Workload'. In *Evaluation of Human Work*, edited by John R. Wilson and Sarah Sharples, 516–544. CRC Press.

Shin, Dong-Hee, Frank Biocca, and Hyunseung Choo. 2013. 'Exploring the User Experience of Three-Dimensional Virtual Learning Environments'. *Behaviour & Information Technology* 32 (2): 203–214. doi:10.1080/0144929X.2011.606334.

Smith, Shana, and Emily Ericson. 2009. 'Using Immersive Game-Based Virtual Reality to Teach Fire-Safety Skills to Children'. *Virtual Reality* 13 (2): 87–99. doi:10.1007/s10055-009-0113-6.

Spitzer, Dean R. 1995. *SuperMotivation: A Blueprint for Energizing Your Organization from Top to Bottom*. New York: AMACOM.

Spitzer, Dean R. 1996. 'Motivation: The Neglected Factor in Instructional Design.' *Educational Technology* 36 (3): 45–49.

Sutopo, H., and W. Pamungkas. 2017. 'Developing Mathematics Mobile Game to Enhance Learning for Children'. In *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, 1:191–197. doi:10.1109/CSE-EUC.2017.41.

Sweller, John, Paul Ayres, and Slava Kalyuga. 2011a. 'Measuring Cognitive Load'. In *Cognitive Load Theory*, by John Sweller, Paul Ayres, and Slava Kalyuga, 71–85. New York, NY: Springer New York.

Sweller, John, Paul Ayres, and Slava Kalyuga. 2011b. 'Intrinsic and Extraneous Cognitive Load'. In *Cognitive Load Theory*, by John Sweller, Paul Ayres, and Slava Kalyuga, 57–69. New York, NY: Springer New York.

Sweller, John, Jeroen JG Van Merrienboer, and Fred GWC Paas. 1998. 'Cognitive Architecture and Instructional Design'. *Educational Psychology Review* 10 (3): 251–296.

Szűcs, Dénes, Amy Devine, Fruzsina Soltesz, Alison Nobes, and Florence Gabriel. 2014. 'Cognitive Components of a Mathematical Processing Network in 9-Year-Old Children'. *Developmental Science* 17 (4): 506–524. doi:10.1111/desc.12144.

Wickens, Christopher D. 1976. 'The Effects of Divided Attention on Information Processing in Manual Tracking.' *Journal of Experimental Psychology: Human Perception and Performance* 2 (1): 1.

Wickens, Christopher D. 1992. 'Virtual Reality and Education'. In *Systems, Man and Cybernetics, 1992., IEEE International Conference On*, 842–847. IEEE.

Wickens, Christopher D. 2008. 'Multiple Resources and Mental Workload'. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50 (3): 449–455.

Wickens, Christopher D., Justin G. Hollands, Simon Banbury, and Raja Parasuraman. 2013. *Engineering Psychology & Human Performance*. Psychology Press.

Wickens, Christopher D., R. Parasuraman, Simon Banbury, and Justin G. Hollands. 2015. *Engineering Psychology and Human Performance*. Vol. Fourth edition. London: Psychology Press.

Wigdor, Daniel, and Dennis Wixon. 2010. *Brave NUI World : Designing Natural User Interfaces for Touch and Gesture*. Burlington, MA: Morgan Kaufmann.

Xiong, Wei, Qing-Hui Wang, Zhong-Dong Huang, and Zhi-Jia Xu. 2016. 'A Framework for Interactive Assembly Task Simulation in Virtual Environment'. *The International Journal of Advanced Manufacturing Technology* 85 (5–8): 955–969. doi:10.1007/s00170-015-7976-3.

Xu, Xinhao, and Fengfeng Ke. 2016. 'Designing a Virtual-Reality-Based, Gamelike Math Learning Environment'. *American Journal of Distance Education* 30 (1): 27–38.

Zhang, Xi, Shan Jiang, Patricia Ordóñez de Pablos, Miltiadis D. Lytras, and Yongqiang Sun. 2017. 'How Virtual Reality Affects Perceived Learning Effectiveness: A Task–Technology Fit Perspective'. *Behaviour & Information Technology* 36 (5): 548–556. doi:10.1080/0144929X.2016.1268647.