Effects of Narrative in the Context of Digital Game-based Learning for Young Children

Vliv příběhu v digitálních výukových hrách pro mladší děti

Bc. Tomáš Sýkora

New Media Studies
Institute of Information Studies and Librarianship
Faculty of Arts, **Charles University**

This diploma thesis is submitted for the degree of
Master of Arts

Prague, December 2018 doc. Mgr. Cyril Brom, Ph.D.

Acknowledgements

I am very grateful for all the help and support I received from my fiancée, my parents, other family members, friends, and colleagues from the research team. I would especially like to express my gratitude to my supervisor doc. Mgr. Cyril Brom, Ph.D. I have learnt a lot thanks to his knowledge, experience and critical view. The experiment described in my thesis, which involved a narrative educational game, took place at several schools - I would hereby like to acknowledge the help of all the teachers, parents and participating children we received in our search for answers to the question of how to design better games for learning. This research was partially supported by funding provided by the Charles University Grant Foundation (GA UK, nr. 684218) and by project PRIMUS/HUM/03 at Charles University. Finally, I would like to mention TechSophia s.r.o., the company which provided not only versions of their games *Matemág* and *Spojovačky* and devices for families which needed to borrow one, but also a bundle of games as a reward for the participants. The following chapters are, as demonstrated by this acknowledgement, the fruit of my cooperation with many helpful people and institutions. I extend my sincere thanks to them all.

Declaration
Prohlášení

I declare that this diploma thesis was composed solely by myself, that I have cited all sources used in the writing of this thesis, and that the work has not been used in any other university programme or to apply for the same or a different degree.

Prohlašuji, že jsem diplomovou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

Bc. Tomáš Sýkora
December 2018

Abstract

Research in the digital game-based learning domain has so far shown mixed results as to the use of narrative in educational games. The aim of this thesis is to help to answer the question of whether and to which extent it is reasonable to employ the narrative feature in educational games for young children. In addition to a literature review, the thesis presents an experimental study comparing two versions of a maths game that are the same except for for the richness of the game's story (a value-added study). The participating children (N = 67) from school year 2 and 3 (mean = 8.67 years, SD = 0.4 years) were given the opportunity to play a game for two weeks on touch devices: a game version with a simple narrative frame, or a game version with a rich story narrated through an interactive voiced comic, or a "placebo" game (control group). No significant effect of the rich narrative on the children's engagement, as reported by parents, was found (d = 0.45, p = .245). Furthermore, the two narrative condition groups did not differ in terms of in-game progress (Cliff's δ = 0.01), and the difference was neither significant for the number of solved game tasks (d = 0.08, p = .857), nor the learning gains (d = -0.25, p = .691) measured using a near-transfer maths skill test (pre-post design). Both narrative groups had significantly greater learning gains than the control group (narrative frame: d = 1.00, $p < 0.01$; rich story: d = 0.74, $p < 0.05$). The conclusion, based on the findings and the literature review, is that providing a basic context for game mechanics of an educational game in the form of a simple story (as opposed to a rich story) may engage young learners just enough to reap the benefits without increasing the risk of distraction from learning. The thesis describes the conducted experiment in detail and discusses the results in the context of other related research.

Keywords: game-based learning, digital games, serious games, narrative, mathematics, engagement, children

Abstrakt

Výsledky výzkumů v oblasti digitálních vzdělávacích her zatím nevypovídají jednoznačně o užitečnosti použití narativu ve výukových hrách. Cílem této diplomové práce je zodpovědět otázku, zda-li a do jaké míry může být opodstatněné zapojení příběhu ve vzdělávacích hrách pro mladší děti. Mimo přehledu odborné literatury práce prezentuje experimentální studii porovnávající dvě verze matematické hry lišící se bohatostí herního příběhu. Děti z druhých a třetích ročníků ZŠ ($N = 67$, průměrný věk = 8.67 let, $SD = 0.4$ roku) měly možnost po dva týdny hrát videohru na dotekových zařízeních. Jedna skupina dostala verzi s jednoduchým příběhovým zarámováním, druhá skupina obdržela verzi obsahující bohatý namluvený příběh ve formě interaktivního komiksu. Třetí, kontrolní skupina obdržela „placebo“ hru. Při srovnání skupin dle rodičovských dotazníků nebyl zjištěn signifikantní vliv příběhu na zapojení dětí do hry ($d = 0.45$, $p = .245$). Dále se skupiny s jednoduchým/bohatým příběhem nelišily co do postupu ve hře (Cliff's $\delta = 0.01$). Též nebyla zjištěna signifikantní diference ani v počtu vyřešených úloh ($d = 0.08$, $p = .857$), ani v rozdílu výsledků matematického pretestu a posttestu ($d = -0.25$, $p = .691$) měřícího přenos učené látky. Obě skupiny hrající hru s jednoduchým/bohatým příběhem se lišily signifikantně od kontrolní skupiny dle měřeného rozdílu v matematickém pretestu a posttestu (jednoduchý příběh: $d = 1.00$, $p < 0.01$; bohatý příběh: $d = 0.74$, $p < 0.05$). Na základě výsledků výzkumu a přehledu odborné literatury je možné shrnout, že i jednoduchý příběh (v porovnání s bohatším) může zaujmout mladší děti dostatečně pro účely vzdělávacích her, aniž by zvyšoval riziko odvedení pozornosti od učení. Práce detailně popisuje provedený experiment, rozebírá jeho výsledky a dává je do kontextu s ostatními souvisejícími výzkumy.

Klíčová slova: vzdělávací hry, videohry, příběh, matematika, děti, Matemág

Table of contents

List of appendices

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

CATLM Cognitive-Affective Theory of Learning with Media

DGBL Digital Game-based Learning

RCTs Randomised Control Trials

URL Uniform Resource Locator

Introduction

The focus of this thesis is on digital educational games. In particular, it examines the role of narrative as one of the possible design features of games for learning.

Digital game-based learning (DGBL for short) research has shown mixed results in terms of narrative; many are not in favour of the feature. Learners might be more engaged thanks to a story (Cordova and Lepper, 1996; Marsh et al., 2011; McQuiggan et al., 2008; Rowe et al., 2011) and enjoy the experience more (Koenig, 2008; Marsh et al., 2011). Nevertheless, even though some may associate story with positive influence on learning (Cordova and Lepper, 1996; Jimenez, 2014; Rowe et al., 2011), multiple recent studies have shown that the feature itself does not seem to yield significantly different learning outcomes (Echeverría et al., 2012; Jimenez, 2014; Johnson-Glenberg and Megowan-Romanowicz, 2017; Marsh et al., 2011). Meta-analytical results indicate that games with no story or a thin story depth have significantly larger effects on learning than games with a story evolving over the course of the game (Clark et al., 2016), and studies have also found other instructional methods (such as a slide presentation) to perform better than certain narrative serious games (Adams et al., 2012; McQuiggan et al., 2008). Some blank spots still remain within the domain of DGBL research if we take into account, for example, the variety of target audiences and differences in social context.

Children younger than 12 years of age have so far participated only in a study authored by Cordova and Lepper (1996) and a study conducted by Jimenez (2014). Both studies suggest that story might play a positive role in educational games. Experimental designs of the studies focused on narrative games mentioned in the previous paragraph usually involved only one session shorter than two hours, with the exception of Jimenez (2014), whose experiment involved several sessions in a school setting. A study that tests the effects of narrative in the context of home environment does not appear to be available. Thus, the goal of the research introduced here is to help to identify the extent to which designers should employ narrative in educational leisure-time games for young children with playtime exceeding two hours.

After laying down the theoretical foundations, the thesis presents a value-added[1] research study comparing two newly created versions of a commercial game for maths skills training. The conducted experiment examined whether a richer narrative could better support learning in a home environment or if a simple narrative frame could yield the same results. Different player preferences concerning narrative were also explored.

One game version employed in the study uses a simple narrative frame of a journey of two sibling heroes through a magical land, where the players solve puzzle tasks standing in their way. In addition to the journey theme, the other game version contains a rich story told through voiced interactive comics at the start and the end of the game. Children participating in the study were 7–9 years old and they played the game at home for two weeks according to their will. The text describes the experiment in detail. The reported results are discussed and they are aligned with the current body of research.

A possible conflict of interest should also be mentioned. I am a co-author of the commercial game *Matemág* (TechSophia s.r.o., 2017a) - the altered versions of which were used in the research - as well as a co-owner of the company which holds the rights to the game; this made the research possible and explains my deep curiosity about the topic. I would like to state that I kept my integrity as a researcher and that the results should not be influenced in any way.

[1] The value-added research approach within the DGBL domain "seeks to identify features that enhance learning by comparing a base version of a game to a version with an added feature."(Pilegard and Mayer, 2016)

Chapter 1

Theoretical Background

1.1 Educational Games

The idea of using the charm of video games for educational purposes has been around since their inception. Already the first commercial gaming console for home use with a bundle of games - the *Magnavox Odyssey* (Magnavox, 1972) - was advertised as coming with “educational experiences” (Magnavox, 1973). And the *Carmen Sandiego* franchise, which started back in 1985 with a game designed to teach about geography (Broderbund, 1985), has continued to this day in newer titles such as *Carmen Sandiego Returns* (The Learning Company, 2015). The online app stores of today offer hundreds of games labelled as ’educational’ (Apple Inc., 2018; Google LLC, 2018; Microsoft Corporation , 2018). The vision of games that strive to fulfill educational purposes apart from mere entertainment, the so-called ‘serious games’ (Abt, 1987), is indeed still alive.

Researchers have not been idle all this time either. They have been testing bold claims about the benefits of educational video games. Drawing on literature reviews (Hainey et al., 2016; Mayer, 2014; Wouters et al., 2009), meta-analyses (Clark et al., 2016; Sitzmann, 2011; Vogel et al., 2006; Wouters et al., 2013; Wouters and van Oostendorp, 2013) and other related academic works (Mayer, 2011, 2016; Moreno and Mayer, 2007), both the theory and the key findings are summarised in the paragraphs to follow. The findings mostly confirm that it is possible to support learning through video games, but as Mayer puts it: “[A]n educational revolution based on gaming is not indicated” (Mayer, 2016, p. 1).

How are games thought to influence learning? At the beginning of their meta-analysis, Wouters et al. (2013) name two ways: by changing the cognitive processes involved in learning, or by affecting the learner’s motivation as well. If learning is successful, there should be some type of learning outcome. The possible learning outcomes can be divided using the taxonomy proposed by Wouters et al. (2009), into these categories:

- Cognitive learning outcomes - subdivided into knowledge and cognitive skills
- Motor skills
- Affective learning outcomes - subdivided into attitude and motivation
- Communicative learning outcomes

Academic literature may provide an answer as to whether there is any evidence that serious games can have impact on any of these types of learning outcomes. Wouters et al. (2009) themselves present their own review. Although they point out that it is not possible to make definitive conclusions due to the low number of studies (n = 29), they do, in the end, state that educational games seem to be effective in terms of cognitive learning outcomes, and they call serious games ‘promising’ as to the training of motor skills and attitudinal change. Little substantiation was found in terms of the effect on motivation and communicative learning outcomes (Wouters et al., 2009). More recent literature reviews (Hainey et al., 2016; Mayer, 2014) not only help us to extend our view, they also demonstrate that there is still need for more research in the field.

Mayer (2014) examines the cognitive consequences of playing video games in his review. His conclusions about learning outcomes, which may be classified as cognitive learning outcomes according to the taxonomy mentioned above, suggest that there is no substantial evidence (for Mayer, that is an effect size greater than 0.4 based on six or more comparisons) to support the claim that playing computer games can, in general, improve one’s mind. According to Mayer (2014), only two types of games were found to perform consistently and sufficiently well in training of specific cognitive skills : first-person shooter games (with a large effect size across 18 comparisons) and puzzle games such as *Tetris* (with a large effect size across six comparisons). While *Tetris* improved spatial cognitive skills in a very limited way - only mental rotation of Tetris-like shapes - first-person shooters had an effect on a variety of perceptual attention skills, such as useful field of view or multiple object tracking (Mayer, 2014).

Hainey et al. (2016) provide a review of literature focusing on empirical evidence for game-based learning in primary education. The articles labelled as ‘high quality’ by the reviewers mostly support game-based learning. A narrower summary of randomised control trials (RCT), which are according to Hainey and collegues “the best way of demonstrating the effectiveness of novel educational approaches” (Hainey et al., 2016, p. 211), show that seven RCT studies in the *Knowledge acquisition/content understanding* category report positive effects of DGBL and two RCT studies report negative ones. Finally, four of the RCT studies fall in the category of *Affective and motivational outcomes*, and all report positive effects (Hainey et al., 2016).

These literature reviews show that there are some studies which report positive results as to the effects of educational games on learning, but they do not provide enough evidence to draw a solid conclusion just yet. More information on the topic appears in the next section, which summarises findings from meta-analytical studies by Vogel et al. (2006), Sitzmann (2011), Wouters et al. (2013), Mayer (2014), and Clark et al. (2016).

1.2 Media Comparison Research

A question may be raised as to what performs better - serious games or conventional instruction methods (such as lectures, readings, etc.)? Research comparing games and other instructional methods would be a research genre classified by Mayer (2011) as 'media comparison'. Several meta-analyses (Clark et al., 2016; Mayer, 2014; Sitzmann, 2011; Vogel et al., 2006; Wouters et al., 2013) offer an overview of the matter.

The meta-analysis by Vogel et al. (2006) reports significantly higher cognitive gain outcomes ($z = 6.051$, $p < .0001$ ($N = 8549$)) and significantly better attitudes toward learning ($z = 13.74$, $p < .0001$ ($N = 2378$)) when games or interactive simulations were used versus traditional teaching methods[1]. However, in the meta-analysis dedicated solely to computer-based simulation games for adults by Sitzmann (2011), the author claims that Vogel et al. (2006) might have overestimated the cognitive gains from simulation games due to publication bias. In her paper, Sitzmann reports that with simulation games, self-efficacy was 20 % higher, declarative knowledge was 11 % higher, procedural knowledge was 14 % higher, and retention was 9 % higher relative to the comparison group, but she also provides strong evidence of publication bias in simulation games research (Sitzmann, 2011). Interestingly, Sitzmann discovered that when the comparison group received actively engaging treatment, they outperformed the simulation game group (Sitzmann, 2011), which illustrates that one cannot simply recommend games in every situation. Using a combination of games and of other instructional methods seems to be a promising approach, as both meta-analyses - by Sitzmann (2011) and by Wouters et al. (2013) - show that it appears to be more effective than using games alone.

Wouters et al. (2013) worked with multiple hypotheses in their meta-analysis, and they tried to identify the moderators for learning and motivation as well. They included 39 studies published between 1990 and 2012 comparing game-based education with conventional

[1] The z-score (or standard score, normal score, z-value) specifies the number of standard deviations a value is from the population mean (McKillup, 2011).

instruction methods. Cohen's d was used as the indicator for effect size[2]. The results suggest that serious games are more effective in terms of learning (d = 0.29, $p < .01$) and retention (d = 0.36, $p < .01$). The games were found to be only marginally[3] more motivating (d = 0.26, $.076 > p > .05$) than the other teaching methods. The authors offer a hypothesis for the reported lack of motivation: the games were often played in a context where the players did not get to choose the game nor the time, therefore their inner motivation might have been limited; secondly, the design choices could have been made which prioritised education over entertainment; finally, the tools for measuring motivation itself could have been imprecise (Wouters et al., 2013). Although the main findings speak in favour of games, the results seem to be influenced by methodological rigour - serious games were not, in fact, found to be more effective than conventional instruction methods in randomised studies (Wouters et al., 2013).

In a later meta-analysis, Clark et al. (2016) report a notably larger average effect size in the quasi-experimental studies than in the randomised controlled trials. However, after performing a correlation analysis, they offer an explanation: "the trend of smaller observed effects among the randomized controlled trials may be due to variations in the types of games used" (Clark et al., 2016, p. 107). No publication/small-study bias was found in this more recent meta-analysis. The overall results summarising findings from 57 media comparison studies from 2000 to 2012, which focused on educational games for students between 6 and 25 years of age, indicate that "digital games improved students' learning outcomes by approximately 0.3 standard deviations relative to typical instruction" (Clark et al., 2016, p. 97). According to the authors, the results underline the affordances of games for learning.

Let us close the part of this chapter dedicated to the media comparison approach with a few notes from review by Mayer (2014). Mayer identifies two promising academic domains where games were shown to outperform conventional media: science (d = 0.69 based on 16 comparisons) and second-language learning (d = 0.96 based on five comparisons). He also names two other domains as 'not-yet-promising': language arts (d = 0.32 based on three comparisons) and social studies (d = 0.62 based on three comparisons). An unpromising one is, according to him, the domain of mathematics (d = 0.03 based on five comparisons) (Mayer, 2014).

These literature reviews and meta-analyses give reason to choose games over other instructional methods in some cases. Mayer (2016) recommends using games for targeted

[2]The descriptors for magnitudes of Cohen's d help to interpret the values: with d = 0.01 standing for a very small effect size, d = 0.2 for small, d = 0.5 for medium, d = 0.8 for large, up until d = 1.2 and d = 2.0 for a very large and a huge effect size respectively (Cohen, 2013; Sawilowsky, 2009).

[3]Although the p value is not stated precisely in Wouters et al. (2013), we may find it in a habilitation thesis by Brom (2017) as the author was in communication with Wouters and colleagues' research team.

learning of well-specified objectives in those situations for which there is some evidence grounded in DGBL research.

1.3 Value-added Research

Which features of a serious game could support learning? Which could hinder the attainment of educational goals? The ‘value-added’ research approach (Mayer, 2011) can be applied to find the answers to these questions. The idea is simple - researchers examine two versions of a game in the same fashion. The versions are similar but for the feature in question. The results then indicate whether the feature had a positive, a negative or a neutral effect on learning (or another desired purpose). For example Mayer (2014) identified, based on his literature review, several game features that substantially and consistently improved student performance on tests assessing learning outcomes: using conversational style (personalisation), presenting words in spoken form (modality), adding prompts to explain (self-explanation), adding explanations or advice (coaching), and adding pregame descriptions of key components (pretraining). These features overlap with those reported by Wouters and van Oostendorp (2013) in their meta-analytical review of the role of instructional support in GBL; additionally, the authors mention feedback ($d = .49$, $p < .001$), modelling ($d = .46$, $p < .001$), and collaboration ($d = .14$, $p < .05$). Apart from the inside features of a serious game, Clark et al. (2016) report that scaffolding provided outside of the game by a teacher leads to a significantly larger effect on learning outcomes than conditions where the learner receives only a simple success/failure/score feedback.

How does narrative fare as a game feature? Wouters and van Oostendorp (2013) identified, through meta-analysis, a positive, but not significant effect size. Clark et al. (2016) show that a simple or no story could prove to be more beneficial than a deeper storyline. Mayer (2016) concludes that no strong recommendations can be made at this stage and that more research is needed. This particular topic is discussed further in the following sections.

1.4 Cognitive-affective Theory of Learning with Media

The *Cognitive-affective theory of learning with media* (or CATLM for short) by Moreno and Mayer (2007) provides great support in trying to understand the way educational games impact cognitive and affective-motivational processes. It can be applied to interactive multimodal learning environments, including serious games. ‘Multimodal’ means that the learning environment uses two (or more) different modes to represent content knowledge (Moreno and Mayer, 2007), such as text or spoken word (verbal), pictures, animations (non-verbal), or even some less traditional modes which employ senses other than vision and hearing.

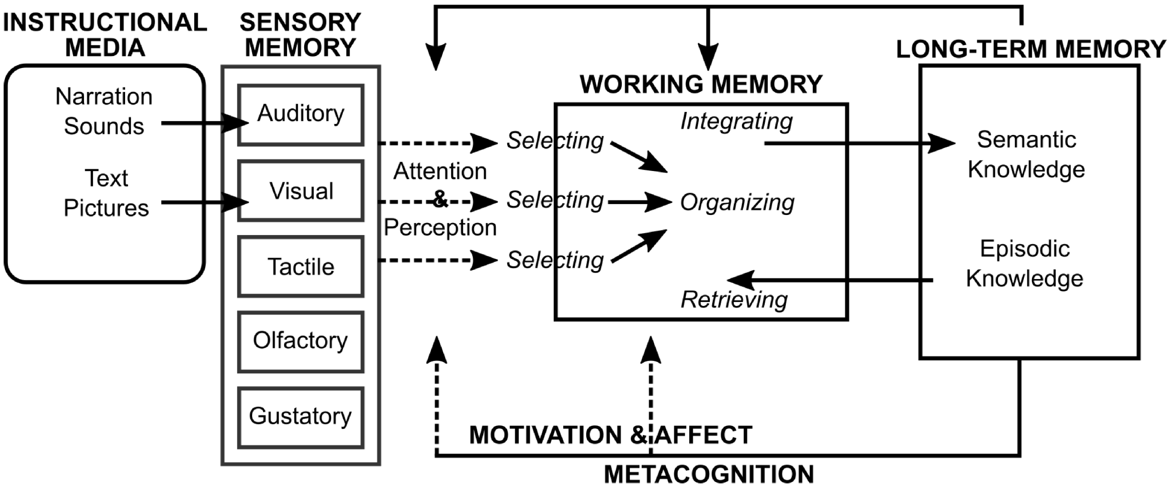


Fig. 1.1 A cognitive-affective model of learning with media as introduced by Moreno and Mayer (2007)

The CATLM acknowledges separate channels for processing of different information, be it sounds or spoken word, images or written texts, or even touches and smells. It stresses that only a limited amount of information can be processed in working memory at any one time as our cognitive resources are not infinite. From the CATLM point of view, proper learning can occur only if the learner is actively engaged in selecting, organizing, and integrating new information with existing knowledge (stored in long-term memory). The theory also further identifies several important factors that influence learning - the learner's prior knowledge and abilities are one key component, motivation mediates learning by increasing or decreasing cognitive engagement, and lastly, there are metacognitive skills, thanks to which the learner might be able to regulate cognitive processing and affect (Moreno and Mayer, 2007).

Good instructional material should limit processing of unnecessary information (so called 'extraneous processing') and try to support both 'essential processing' needed for selecting new information and 'generative processing', which happens when the brain is trying to make sense of the new information (Moreno and Mayer, 2007). Designers and educators should be aware of pupils' limited cognitive capacities when creating a serious game. The games should convey an appropriate amount of relevant information at a time and avoid content useless for learning. Otherwise the cognitive load may be too great and the learning harmed (Adams et al., 2012; McQuiggan et al., 2008; Schrader and Bastiaens, 2012).

The processes highlighted by the CATLM should be respected by designers as well. Games need to properly direct the learner's attention to the most important things (supporting the selection process), the information needs to be structured coherently (supporting the organisation of it into mental models), and the learning environment should offer some guidance on how to connect new knowledge to previous experience (supporting the integration with prior knowledge). Importantly, well-designed games may keep learners engaged in the educational activity through the leverage of motivation. However, game designers ought to be aware that interesting design elements could become distracting as well (Adams et al., 2012; Harp and Mayer, 1998), therefore, the designers must keep in mind to try to reduce extraneous processing, which is unfruitful for learning.

1.5 Stories & Human Communication

The words 'narrative' and 'story' are used very often throughout this thesis - let us start with their formal definition. Multiple authors have proposed a definition of 'narrative' (Abbott, 2002; Genette et al., 1982; Prince, 1982; Richardson, 2000; Ryan, 2007), but this thesis employs the one provided by Abbott (2002) as it offers an explanation for the relationship between 'narrative' and 'story' as well: "Narrative is the representation of events, consisting of story and narrative discourse, story is an event or sequence of events (the action), and narrative discourse is those events as represented" (Abbott, 2002, p. 16). From this perspective, as soon as one tries to convey a story through a representation (oral, textual, pictorial, or procedural, such as a game), it can be called a narrative. As this thesis is concerned with digital game-based learning, the stories discussed are those narrated via video games; this is precisely why an exact distinction between 'narrative' and 'story' is not crucial for the purposes of this text. It is more important to distinguish the qualitative aspects of different narratives. There is the (rich) narrative, where the story can work just as well on its own, and then there is the 'narrative frame', which is a term that, in this thesis, refers to cases where a simple story is used only to enwrap some other content (such as game mechanics).

Humankind has used stories for millenia for entertainment and to spread information (Rubin and others, 1995). The capacity of a story to communicate and to share an experience[4]

[4]Thanks to the use of modern methods such as fMRI in brain research in recent decades, the way stories work on the neural level might be understood better relatively soon. Examples of how they work arise, for instance, from the discovery that metaphors may invoke somatosensory brain response (Lacey et al., 2012) or that the speaker's and the listener's brains may exhibit joint response patterns while one is telling an unrehearsed real-life story to the other - the process is called 'neural coupling' (Stephens et al., 2010). The meta-analysis of fMRI studies by Mar (2011) suggests that there seems to be a substantial overlap in the brain networks used to understand stories and the part of the brain responsible for understanding that others have beliefs, desires, intentions, and perspectives that are different from one's own. These findings are interesting and important, but further commentary in this thesis is limited due to the lack of expertise in neuroscience.

has been used in multiple fields, including health education (Berkley-Patton et al., 2009; Duveskog et al., 2009), moral education (Tappan and Brown, 1989), or the teaching of history (Bage, 2012; Šisler, 2016). Let us discuss which other particular features of narratives relevant to education have been researched academically.

It is not very surprising that literacy and language skills in general can be supported through stories (Feitelson et al., 1993; Gillam et al., 2012; Hudson and Test, 2011; Morrow et al., 1990). Ryokai et al. (2003) have demonstrated that even a virtual storyteller is able to help young children with language development.

Stories may also be useful for memorising and recalling information. Graesser and Ottati (1995) report results of Graesser's experiments from the 80's, that compared reading of narrative texts and of expository texts. The texts containing some sort of a story were recalled approximately twice as well as the purely informative expository texts. The author comments further: "There was a robust .92 correlation between narrativity and the amount of information recalled from a text. In contrast, topic familiarity and interestingness had very small effects on reading time and recall for these texts" (Graesser and Ottati, 1995, p. 124). In addition to better recall, the narrative texts were read approximately twice as fast as the expository texts (ibid). These results are aligned with an earlier experiment by Bower and Clark (1969) that showed a very effective way of memorising unrelated words by chaining them into a story.

Creators of educational materials may benefit from the fact that stories can help create a meaningful context for information. For example, Mares and Pan (2013) shows, based on data from multiple countries, that the popular programme for preschoolers Sesame Street, which uses stories and narrative framing, has significant positive effects aggregated across learning outcomes in cognitive skills, knowledge, and social reasoning. Casey et al. (2008) have helped to improve geometry skills in kindergarten settings through intervention with a storytelling-context, and they have shown de-contextualised intervention to be less successful.

The capacity of a narrative to enwrap various content seems also to be often used in the domain of video games (in this case, not only serious games). Fullerton (2014) notes that without a story, or at least some story elements, a lot of a game's appeal is lost. Which sounds more interesting: a dwarf fighting an evil wizard or sets of data transforming according to an algorithm? (Fullerton, 2014, p. 93-94) There have been various discussions among game designers regarding the extent to which a story should be included (only as a framing device or as a whole storyline) as there are certain trade-offs between narrativity and interactivity (Juul, 1999). Nevertheless, a story is regarded as one of the possible game features today and it mainly depends on the game designer's choice if and how it is going to be employed. The following section examines whether a similar conclusion can be made concerning stories in the DGBL domain.

1.6 Narrative & DGBL

The consideration of narrative elements within the narrower focus of educational games may benefit from previous relevant research (Adams et al., 2012; Barab et al., 2007; Cordova and Lepper, 1996; Echeverría et al., 2012; Jimenez, 2014; Johnson-Glenberg and Megowan-Romanowicz, 2017; Koenig, 2008; Marsh et al., 2011; McQuiggan et al., 2008; Rowe et al., 2011; Ryokai et al., 2003; Wouters et al., 2011). Although the list is relatively long, there are still some blank spaces and unanswered questions left. Using a narrative to frame game mechanics seems to enhance engagement (Cordova and Lepper, 1996; Marsh et al., 2011; McQuiggan et al., 2008), but multiple research studies suggest that a story makes no significant difference for learning outcomes (Echeverría et al., 2012; Jimenez, 2014; Johnson-Glenberg and Megowan-Romanowicz, 2017; Marsh et al., 2011); some results indicate that a game with a rich narrative is less effective than a game with a simpler narrative (Clark et al., 2016) or than other instructional methods (Adams et al., 2012; McQuiggan et al., 2008). The paragraphs below summarise the research and begin to outline the answer as to whether, when, how, and why a story should might be used in DGBL.

Firstly, a paper by Cordova and Lepper (1996) about intrinsic motivation and learning is partially relevant to the questions above. Many other studies refer to this particular study as its results appear to be promising for the use of a story. Cordova and Lepper compare several versions of an educational game about arithmetics for primary school children, using conditions different to the control group. The conditions were: contextualisation, personalisation, and provision of choices. 'Contextualisation' meant using a simple fantasy story to frame the abstract mathematical operations (a space-themed story about saving the Earth and a nautical treasure-hunt story). The fantasy game versions were liked significantly more than the non-fantasy control version, embedding the activities in a fantasy context yielded significantly higher levels of learning, and children even felt more competent. The results were better when the story was personalised and not generic (Cordova and Lepper, 1996). A paper by Pilegard and Mayer (2016) also mentions the positive effect of using a story for contextualisation established by other researchers in studies of non-educational games: "Early ethnographic studies and analyses of video game playing noted that players appear to become engaged in game playing through the storyline of games, even when the stories are quite simple such as in the case of PacMan" (Pilegard and Mayer, 2016, p. 1).

There have been hopes for integration of narrative into serious games, and some theoretical works have advocated for the idea. For example, one of the 25 recommendations on how to design learning environments by Halpern et al. (2007) was to employ a story. Lee et al. (2006) argue that narratives are important in interactive media like computer games. Barab et al. (2007) define narrative as one of the three cornerstones of their framework for supporting socio-scientific inquiry, and, using the framework, they have built an educational game. Rowe et al. (2011) claim, based on their experiment where middle-schoolers (primary school) were taught about microbiology using the game *Crystal Islands*, that "[n]arrative-centered learning environments offer a promising vehicle for delivering experiences that are both effective and engaging" (Rowe et al., 2011, p. 16).

However, multiple media comparison and value-added experimental studies focused on narrative educational games are not very much in favour of the use of the narrative for educational purposes. Koenig (2008) used an action adventure game *Cache 17*, in which a player learns about electromechanics while searching for a lost valuable painting. The outcomes were better in the narrative group, but not significantly so, although participants reported enjoying the game more. The same game was used in another experiment conducted by Adams et al. (2012). They found that students learnt better about the electromagnetic devices from a *PowerPoint* presentation than from the narrative game. In the same article, Adams and colleagues also present another experiment which involved the game *Crystal Islands* (mentioned in the previous paragraph), that lead to a similar conclusion: learning through a narrative game is less efficient than learning through a presentation. This experiment confirmed earlier results published by McQuiggan et al. (2008), who used the adventure game *Crystal Islands* as well. Wouters et al. (2011) succeeded in supporting curiosity by adding back story and foreshadowing to the game they used for their research, but learning was not influenced. A space-themed game for introducing teenagers to physical concepts of electrostatics redesigned by Echeverría et al. (2012) was tested in both the fantasy and non-fantasy version. The results show no statistically significant difference neither in learning nor in engagement. Jimenez (2014) reports a strong correlation between fun and learning gains for a story version, with characters, of a digital card game used to help primary school children to understand fractions. Nevertheless, the learning gains were only marginally better than with the version with an abstract game design. Among the game versions of a game controlled by movement and gestures (embodied learning) tested by Johnson-Glenberg and Megowan-Romanowicz (2017) was one that featured a story; the story did not make a difference.

These findings give a more realistic view of the subject at hand; a story works mainly as an entertaining factor, usually with no other effect. It also seems that in some cases, other approaches are more effective than a narrative serious game (Adams et al., 2012; McQuiggan et al., 2008). However, the limitations of these experiments should be noted as well - except for the experiment by Jimenez (2014), all were built around only one intervention shorter than 2 hours, and the settings were not very natural. For the purposes of this research, it is worth noting that out of all the studies, children younger than 12 years of age participated only in those by Cordova and Lepper (1996) and Jimenez (2014).

Finally, the meta-analysis by Clark et al. (2016) shows that the effect size for learning outcomes was significantly larger for games with no story or a thin story depth relative to those with a thicker story. It also combines multiple aspects, such as story depth and relevance, visual realism, and type viewpoint, into a variable called 'contextualisation'. A meta-regression model showed a small but significant negative relationship, suggesting that increased contextualisation was correlated with smaller learning gains (Clark et al., 2016). These observations are in line with the meta-analysis by Wouters et al. (2013), that notes the superiority of schematic games over cartoon-like or realistic serious games where learning is concerned.

The researchers, of course, wondered as to why a story was not found to be especially beneficial for learning. Apart from possible imperfections in the design of the experiment or the story itself, there is another probable explanation which is in agreement with the *Cognitive-affective theory of learning with media* (Moreno and Mayer, 2007) discussed above. Processing of a narrative might be "eating up" resources needed for learning, forcing the mind to deal with a greater cognitive load. Some elements of the story may also, in some cases, distract the learner from the educational content - disrupting learning processes, such as selection, organisation and integration of the relevant information (Adams et al., 2012; Koenig, 2008; McQuiggan et al., 2008; Pilegard and Mayer, 2016). In multiple cases, the narrative version was more appealing and the pupils might have been more keen on learning with the material, but this may have at best cancelled out the negative effects (suggested by the theory), resulting in negligible differences in the educational outcomes. If a narrative theme is a part of a game's design, Adams et al. (2012) advise to at least align it closely with the instructional goals.

As noted by Adams et al. (2012), the motivational power of (narrative) video games could play a positive role in a situation where learners choose to play the games in their free time. Even if the games are less effective, they could increase the time on task, which would otherwise not happen. Time on task is indeed an important mediating factor for learning (Landers and Landers, 2014). Sitzmann (2011), focused on computer-based simulation games for adults, confirmed that trainees who had unlimited access to the simulation game ($d = .68$) outperformed those who only had limited access ($d = .31$).

It may be that narratives within serious games are not suitable for all educational purposes but might be good for some - for example language skills. Marsh et al. (2011) praise the narrative version of the game they investigated as it "provided a flexible and powerful approach to introduce technical/scientific terms" (Marsh et al., 2011, p. 1). Ryokai et al. (2003) show that preschool children started to use linguistically advanced expressions in their stories after being involved in a storytelling interaction with a virtual peer.

If a narrative game is not performing very well in terms of learning, Pilegard and Mayer (2016) offer a way to improve it. They suggest using additional material - in their case pre-game and in-game worksheets - to focus students' limited cognitive resources on the educational objectives. The researchers also report better performance in students receiving enhanced treatment through written explanation, comprehension tests and transfer tests, none of which affected their enjoyment of the game (Pilegard and Mayer, 2016).

Chapter 2

Research Description & Method

2.1 Basic Research Description

2.1.1 Research Question and Hypotheses

The current body of research might have given rise to curiosity regarding the extent to which it is reasonable to incorporate a story into an educational game for children. The value-added studies with teenage and adult participants showed no statistically significant gains in learning while comparing narrative and non-narrative game versions (Echeverría et al., 2012; Johnson-Glenberg and Megowan-Romanowicz, 2017; Koenig, 2008; Marsh et al., 2011) but the participants only played the game in one session. Jimenez (2014), who let fourth-year children play at school over five sessions, reported only marginally higher learning gains from the story-framed game with characters than from the abstract game version. These findings suggest that narrative may have a small or no effect on learning outcomes.

Cordova and Lepper (1996) praise the use of fantasy to increase motivation, engagement and learning in their value-added study that involved children aged nine to eleven. The story in their game worked like a narrative frame as defined in the previous chapter; the authors called it 'contextualisation'. Others claim that a story has a negative impact as a distraction (Adams et al., 2012; Pilegard and Mayer, 2016). This negative effect appears to have been reported mainly in studies where the educational content was interwoven with a richer story - these two parts therefore competed for the limited cognitive resources. How would these effects (motivation and distraction) play out over multiple user-controlled sessions in the home environment using a game with a storyline distinctly separated from the educational content? This has yet to be determined through research. Johnson-Glenberg and Megowan-Romanowicz (2017) mention that they have not yet found a study focused on the narrative feature and design that incorporated multiple user-controlled sessions.

In our research[1], we used two versions of a commercial serious game for maths training. One version only involves a narrative frame (two heroes on a journey), the other a substantially longer story introduction and a rich story ending in the form of interactive voiced comics. In the second version with a richer story, the players have a chance to learn more about the heroes, their goals and motivations. Both versions are described in detail in section 2.4 Materials: Game & Game Versions.

As participants in general seem to enjoy narrative versions more (Koenig, 2008; Marsh et al., 2011) - get more curious (Wouters et al., 2011) and more engaged (Cordova and Lepper, 1996; McQuiggan et al., 2008) - one of the goals was to discover if the children would play more of the game because of the story. The first hypothesis was formulated as follows:

H1: *The rich story version will be (H1a) more engaging for the players, the rich story group participants will (H1b) advance further in the game and (H1c) solve more tasks than the narrative frame group.*

Continuing this line of thought, it was reasonable to believe that the rich story group would have greater learning gains as solving more tasks would possibly provide more training and increase time on task, which can mediate learning (Landers and Landers, 2014). The CALTM (Moreno and Mayer, 2007) suggests that the learners' potentially higher motivation would also support learning. Of course, according to the theory, the learners' limited cognitive resources would perhaps be drained by the extraneous processing of the narrative, which would likely lead to lower learning gains. Even though the narrative and the learning parts of the rich story game version are separated from each other in our case, and therefore the distraction effect should not be strong, we could not rule the negative effect out. The H2 hypothesis expected either the positive or the negative effect to outweigh the other:

H2: *The learning gains of the rich story group will be different from the narrative frame group.*

Both hypothesis H1 and H2 tend to regard the participants as a homogenous group with similar preferences when it comes to game's narrativity. However, the children taking part

[1] The first person plural is occasionally used when speaking about research activities, as a whole research team was involved in conducting the experiment in the primary schools which took part in it and many details of the experimental design and data analysis were consulted with the supervisor. Nevertheless, I was, as the author if this thesis, in charge and I did most of the activities myself - these were: preparing the tests and questionnaires, implementing changes within the game, communicating with schools, teachers and parents, coordinating administrators at meetings in schools, grading the tests, processing the data including the in-game logs, and later analysing the data using a statistical software.

in the study have various preferences, leading to potentially ambiguous results; hence, we decided to adopt an exploration goal, which was to try to detect how the players' preferences regarding stories would influence their playing and consumption of the narrative parts of the game.

The next section describes shortly how the experiment links hypotheses H1 and H2 with the approaches employed to test them. The methods used to tackle the exploration goal are summarised as well.

2.1.2 Experimental Design

The following paragraphs provide an overview of the experiment; further details can be found in the sections to come - namely sections Participants, Materials: Tests & Questionnaires, Materials: Game & Game Versions, Procedure, and Data Analysis.

Drawing on the value-added study design, the conducted experiment included two experimental groups, which received versions of a game differing only in the aspect of narrativity (described in subsection 2.4.4), and one control group. The control group was given a "placebo" game (described in subsection 2.4.5). Participants were randomly assigned to one of these three groups.

The experiment was initiated at a meeting with both the children and the parents held at each of the participating schools. While the children were completing pre-tests and questionnaires in one classroom, we were speaking with the parents in another classroom; the parents also installed one of the experimental game versions on their (or borrowed) touch devices at the meeting. In a span of 14 days, the children could play in their home (or other preferred locations) as they wished and as their parents let them. The experiment lasted for two weeks and ended at the second meeting where the children and the parents were again separated into two classrooms to complete questionnaires and, in the children's case, a post-test as well. Figure 2.1 shows the experimental schedule.

Two main sources of data were used to assess which of the two narrative versions (narrative frame / rich story) was more engaging. The first source were the answers in the parents questionnaires (see Appendix 4: Parent Questionnaires) describing the way the child behaved while playing or while talking about the game. These were then combined to determine the **child engagement score**, helping to test subhypothesis H1a. The second source were the logs gathered in the game (see subsection 2.4.6) as the child was playing, which made it possible to precisely measure the **progress through game levels** (H1b) and to count the **number of tasks solved** (H1c).

During both the initial and the final session, the participating children had 35 minutes to finish a maths skill test. The pre-post design made possible to calculate the **difference in the**

test scores to test hypothesis H2. In-game measures were employed as well; the **skill scores** achieved by the players were examined to reveal differences between the groups.

Questionnaires for children asking about their opinions and preferences complemented and extended the information mentioned in the paragraphs above. Especially valuable were **self-reported enjoyment** and **self-reported perceived learning gain**. These questionnaires made it possible to also find out how much the children liked stories and liked to tell stories, how much they read, and if they had a favourite book - creating a **story preference score**. Combined with the in-game data about the **time spent on the story parts** and the **number of visits to the story chapters**, it provided the opportunity to explore whether player preferences may have influenced the consumption of the game narrative.

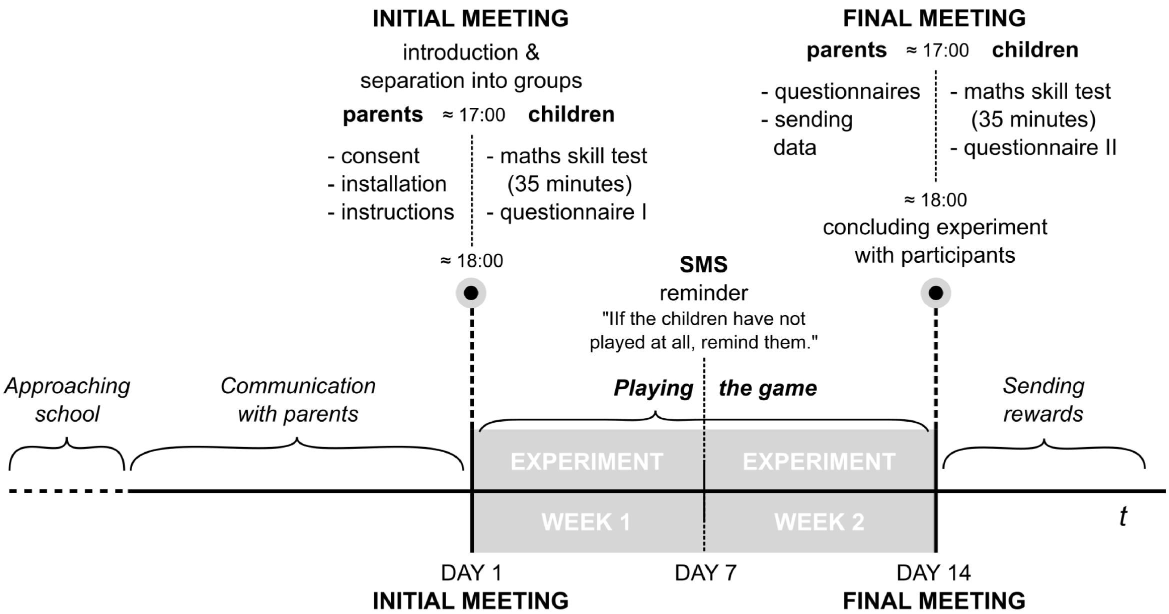


Fig. 2.1 Diagram showing the schedule of the experiment.

2.2 Participants

The research introduced here was conducted in the Czech Republic. The participating families were recruited through six Czech schools which took part in the experiment - two in the Central Bohemian region and four in various districts of the Prague city. The only demand placed on the schools' cooperation was that the maths lessons would be taught 'in the classic way' to ensure that classes using Hejný's constructivist, scheme-oriented approach (Hejný, 2008) would not be included in the experiment as the game itself is influenced by this approach.

Parents of second- or third-year pupils were approached by the school and could voluntarily enter the experiment. The reward (motivation) for their participation was a bundle of four educational maths games by TechSophia s.r.o. worth 25 Euros altogether. Both the parents and the children were separately asked whether the children had been previously exposed to the commercial version of the *Matemág* game at the initial meeting. The participants who had played the game before were allowed to take part in the activities, but their data was excluded from the experiment. From the 90 participants at the beginning of the experiment, six were excluded (or rather their data) due to prior exposure, two were excluded because of contamination between the control group and the condition groups, and 15 did not come to the final meeting at the end of the experiment. Our study is, therefore, based on data from the 67 remaining participants. There were 27 second-year and 40 third-year pupils involved and their average age was 8.67 (SD = 0.4). The number of boys (n = 38) was greater than the number of girls (n = 29). Further characteristics of the experiment and the check we performed to see whether the experiment was balanced are described in chapter 3 that covers the results of the experiment, in subsection 3.1.1.

2.3 Materials: Tests & Questionnaires

2.3.1 Maths Skill Test

The test assessing a subset of maths skills was put together by the research team, consulted with primary school teachers, piloted with children (N =19) and finally refined. It comprised of three types of tasks sharing similarities with the three most frequently occurring puzzles in the game (i.e. near-transfer tasks). One task type was focused on equal division, another used a weighing machine to model simple equations, and the last one examined the children's understanding of arithmetical relations between numbers in a graph-like structure (further referred to as the 'number web' task). An example of a number web task is shown in Figure 2.2, the rest of the test tasks are included in Appendix 2: Near-transfer Maths Skill Test. The grading of the test is described in subsection 2.5.8. Importantly, the dependent variables of interest were the **pretest score**, the **posttest score**, and the **post-pre score** difference.

The test was presented in two parts. The first part was designed to be easier; several tasks of a single type were placed in succession and there was an example solution for the first task of each type. The second part of the test consisted of more difficult tasks and the order of different task types was mixed. The first part of the test contained three ungraded example tasks with an example solution (one for each type) and 12 normal tasks (four for each type).

There were nine tasks (three for each type) in the second part of the test, resulting in the total number of 21 graded tasks for the whole test. The children would receive zero points, half a point or one point for each task depending on the correctness of the solution. Two versions of the test were created for the pre-post design which differed slightly in number values and colours and positions of the graphics. Figure 2.3 gives an example of one task and its modifications for the two test versions.

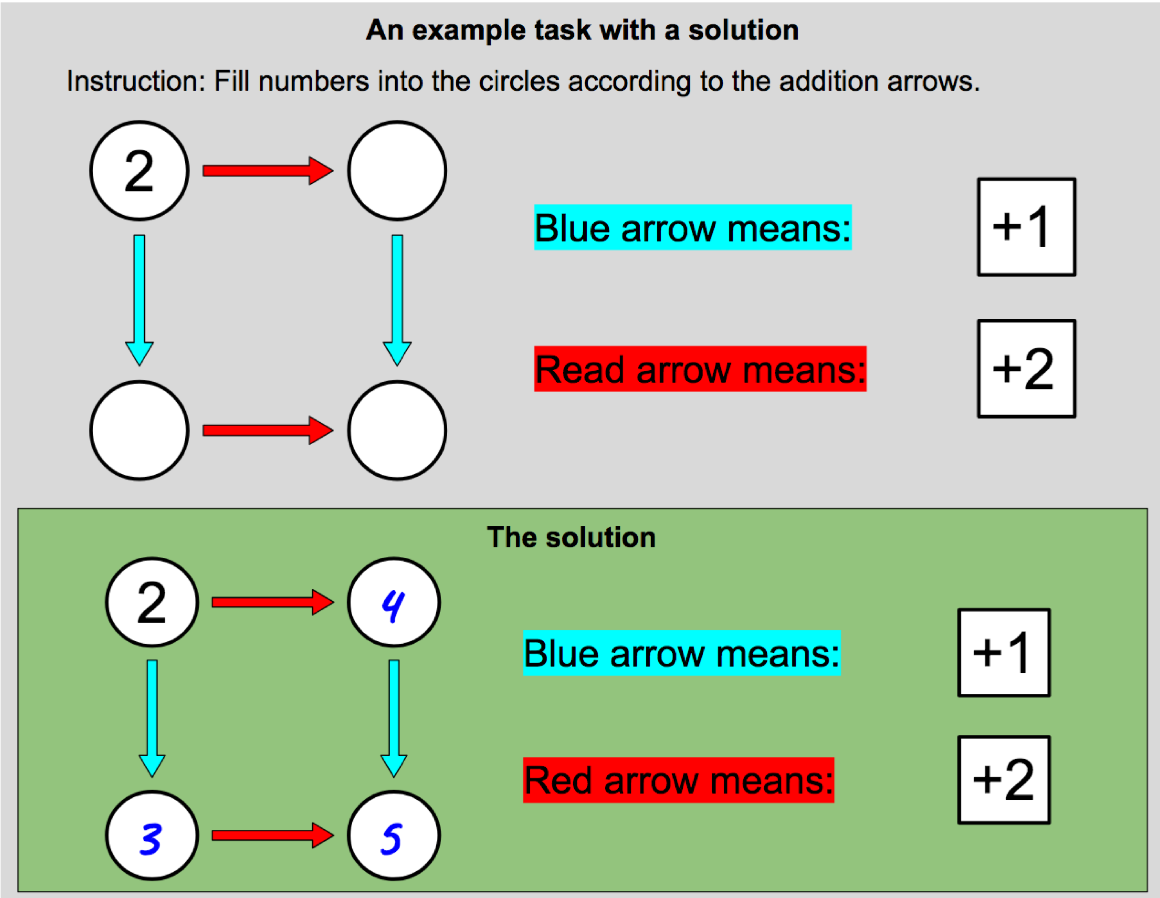


Fig. 2.2 The example number web task from the near-transfer maths skill test. Similar tasks of various difficulty can be found in Czech maths textbooks (Hejný et al., 2008, 2009)

2.3.2 Child Questionnaires

In both of the two sessions (the initial and the final), the participating children completed a questionnaire (a different one in each session) asking primarily about their preferences: preferred school subjects, their reading and storytelling habits etc. The final questionnaire also contained an exploratory section concerning fictional characters, which is irrelevant for the purposes of this research. All of the relevant parts of the questionnaires are attached in

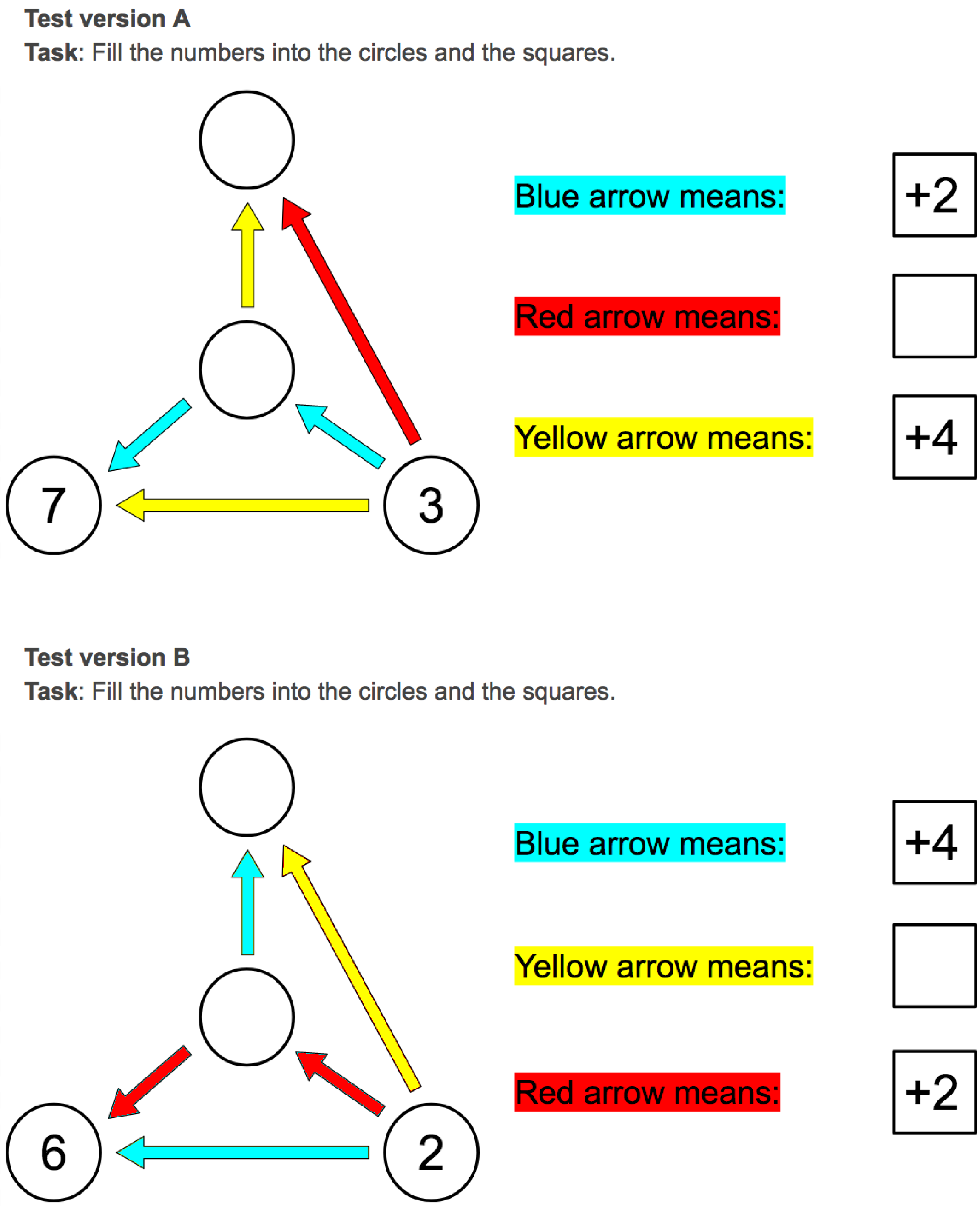


Fig. 2.3 The figure demonstrates a difference between the test versions 'A' and 'B' used in the pre-post design. Each task from version B is a slight modification of its counterpart in version A. The goal was to maintain the same difficulty level.

Appendix 3: Child Questionnaires. The key controlled variables related to the children's preferences were **maths as a favourite subject** (yes/no), **video games as a favourite activity** (yes/no), and the **story preference score** (0-5 points), which was calculated using the following four story-related questions (Cronbach's $\alpha = 0.49$):

- "Do you like stories and fairy tales?" (0-1 points)
- "Do you like to tell stories and fairy tales?" (0-1 points)
- "How often do you read?" (0-2 points)
- "Have you got a favourite book (favourite books)?" (0-1 points)

Other questions for the children were directly related to the experiment itself. Initially we needed to know if the children were familiar with the commercial version of the game (to exclude their data from the experiment). After the two-week experiment, the children were asked to rate how they had enjoyed the game by choosing an emoticon on a scale (from smiling to frowning face), which is how the **child enjoyment score** (0-5 points) was obtained. They also provided answers as to whether they thought that they had learnt something - the more they felt they had learnt, the more stars they were supposed to colour in. Their responses were turned into the **perceived learning gain** variable (0-5 points). Lastly, they gave their opinion on the game tasks - selecting the most interesting task, the most boring one, the easiest and the most difficult task.

2.3.3 Parent Questionnaires

The parents were given two questionnaires at the second (and final) meeting. One of them was centered around the experiment and around how it had worked in their family, the other was more general[2]; it collected demographic data, asked about the parents' opinions about their children, their shared activities, and finally about the parents' approach to the upbringing of their children, especially in relation to technologies. The majority of the questions were closed - either there were options to choose from or the parents were to rate a statement on a scale. Both of the parent questionnaires can be found in Appendix 4: Parent Questionnaires.

To control the balance of the experiment, the focus was put on these demographic data were: **child's age**, **child's gender**, **school year**, **parent's level of education**. The questionnaire also asked about **how often the child plays video games**. Other variables

[2]A note should be made that some questions in the parent questionnaire were too detailed; not all of them are analysed in this thesis. They were included in order to obtain a larger data set for other related research.

were device-related: **device type** (tablet/smartphone), **device OS** (Android/iOS), and whether the device is **used only by the child** (yes/no).

An important part of the questionnaire was concerned with the child's playing of the game. Apart from the **places for playing** and the **social context** (alone/with parents/with siblings/other), the questions explored to which extent the playing was child-driven or parent-driven. We asked about whether the **parent had set time limits** or other rules. A variable named **parent prompted playing** (0-4 points) was constructed from two questions (Cronbach's $\alpha = 0.19$) : one asking if the parent had been reminding his or her child to play the game, the other inquiring if the parent had been asking about the game. Finally, the **child engagement score** (0-12 points) was calculated using the following six questions (Cronbach's $\alpha = 0.76$) :

- "Has your child remembered to play the game himself/herself?" (0-2 points)
- "Has your child been talking about the game?" (0-2 points)
- "Has your child been boasting about his/her achievements while playing?" (0-2 points)
- "Has your child been showing you the game?" (0-2 points)
- "Has your child been forgetting about the time while playing?" (0-2 points)
- "Has your child been wishing he/she would not have to stop playing?" (0-2 points)

2.4 Materials: Game & Game Versions

2.4.1 Devices

The games described below were available for touch devices with the Android or the iOS system. The families were encouraged to bring their own tablets or smartphones; however, there were also spare tablet devices ready for those who wanted to participate but did not have access to these devices. A variable indicating whether a participant under a certain nickname used a **borrowed or own device** was collected and tested for balance between the two narrative condition groups.

2.4.2 Game Description

The video game *Matemág*[3] (TechSophia s.r.o., 2017a) was modified for the purposes of this research, resulting in two versions used for the two experimental conditions. The control

[3]The title "Matemág" could be translated as 'maths magician'.

group received another game introduced below (subsection 2.4.5). As the game versions are the same but for the narrative feature, the following description applies to both the narrative frame and the rich story version. The extra storyline of the rich story version is introduced after all the important aspects shared by the two versions.

Matemág is a single-player, 2D cartoon-style, maths puzzle adventure game intended for children from the first to the third year of primary school (six to nine years old). It was designed for mobile touch devices such as tablets and smartphones and to be played primarily at home. The objectives of the serious game as presented by the creators are: to support the development and training of mathematical skills, to help to establish a positive attitude toward maths, and to bring maths topics outside of school, to the home and family environment. The educational content in the form of maths and logic puzzles was inspired by Hejný's constructivist, scheme-oriented approach (Hejný, 2008). Figure 2.4 represents the structure of the game - the 'Game Core', consisting of 23 levels, is present in both of the experimental game versions, whereas the seven story chapters can be found only in the rich story version.

The core part of the game is framed as a journey through a magical land. Every level represents a certain area of the land, often having a distinct landscape with embedded maths and logic tasks (see Figure 2.5). The player controls the movement of the two main heroes (a girl and a boy), who move on predefined paths. All the maths and logic tasks the player needs to solve are an obstacle standing in the way; Figure 2.6 and Figure 2.7 illustrate unsolved tasks. The correct solution may lead to an event: a bridge is lowered, the way is cleared, a lift starts moving, or the player obtains a key to a gate. The median number of tasks per level is seven. When the player gets to the end of a level, the game switches to the Level Menu scene (see Figure 2.9), where each level is represented by an icon - these unlock one by one as the player progresses through the game. The player may choose to replay a level as well - in that case, new tasks appear within the level (task types remain, but their inner settings change).

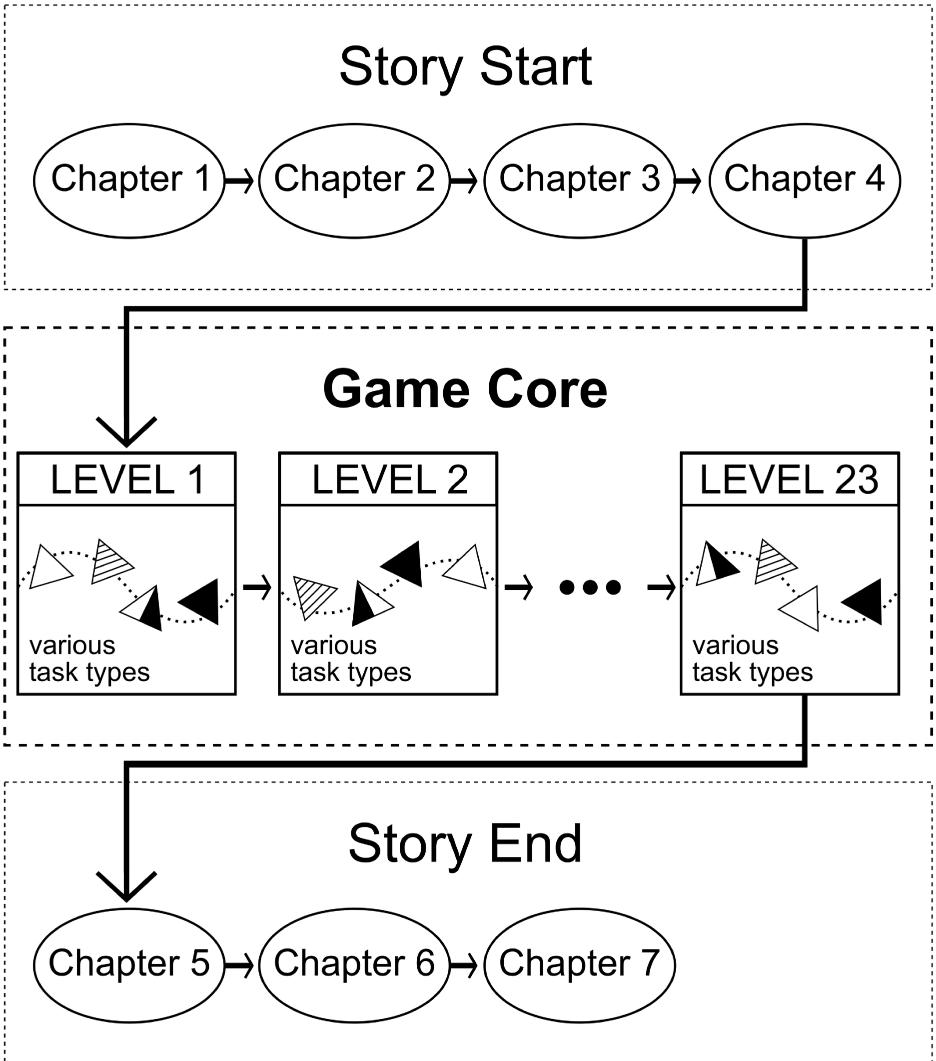


Fig. 2.4 A diagram showing the structure of the experimental game version of *Matemág*. The 'Game Core' is shared by both the rich story and the narrative frame versions used in the research. The four introductory story chapters and three closing story chapters are only included in the rich story version.

Fig. 2.5 A view of a game level. The blue dashed line shows the heroes' path. The view also includes miniatures of game tasks such as the Hungry Plant, the Weighing Machine, or the Music Bridge (TechSophia s.r.o., 2017a).

Fig. 2.6 Before the heroes can continue their journey, they need to feed all the hungry plants equally (TechSophia s.r.o., 2017a).

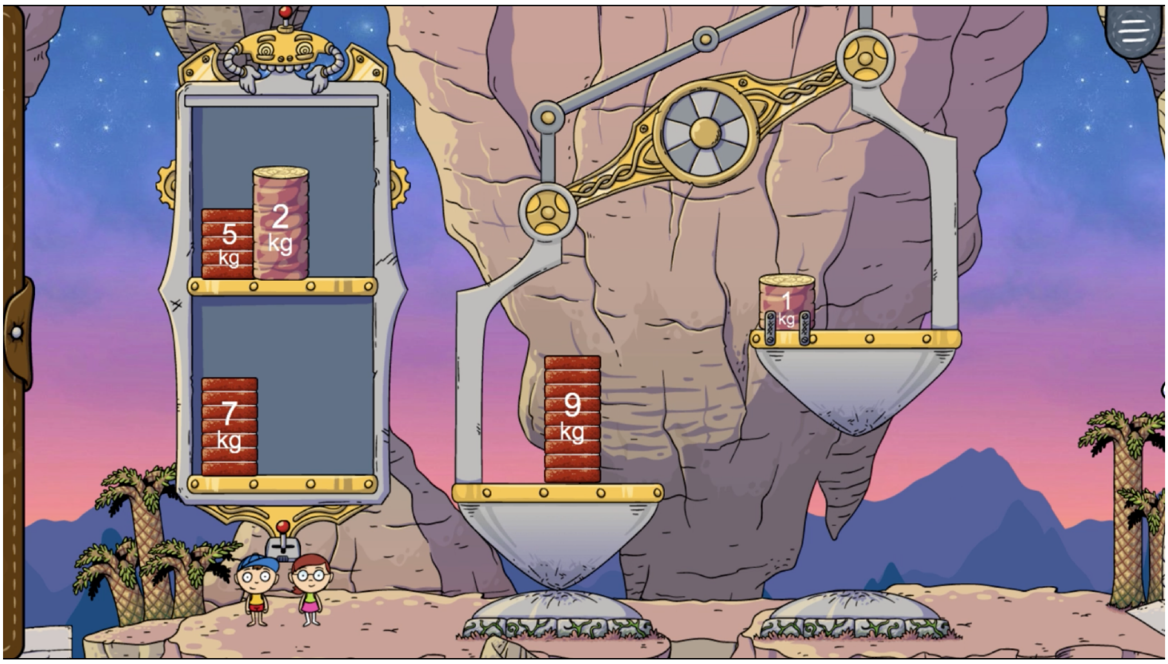


Fig. 2.7 The Weighing Machine task is one of the obstacles the player must overcome in the game. The heroes can pass under the pans only once the machine is balanced (TechSophia s.r.o., 2017a).

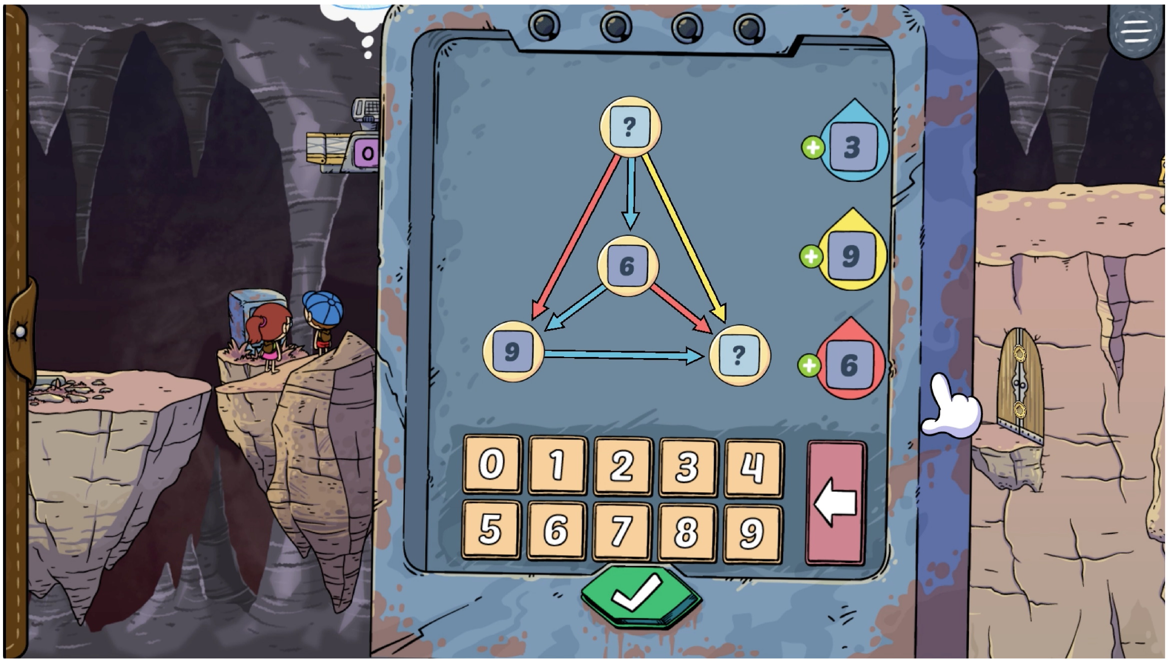


Fig. 2.8 The Number Web task is a kind of lock on a safe in the game. Inside the safe are either keys to locked doors or batteries used to power machines (TechSophia s.r.o., 2017a).

Fig. 2.9 Icons for the first four levels in the game menu (TechSophia s.r.o., 2017a).

Before examining the educational content, this section shortly discusses the game's mechanics, dynamics, and aesthetics within the *MDA* framework by Hunicke et al. (2004). As the target audience of the game is young and might be unskilled in handling mobile devices, both the mechanics and the emerging dynamics are fairly simple. The paths through the magical land are predefined, with little branching within a level. When the heroes arrive to a task, the screen focuses on it. The player may initiate the task instructions, which are then relayed by the heroes. One of the crucial inner (hidden) mechanics is one that chooses the task settings based on the skill score to ensure that the player is facing as appropriate a challenge as possible (see the description of 'adaptivity' below in the list of auxiliary features). In addition to completing the tasks - each of them with its own individual mechanics - the player can also pick up items and put them into a backpack, i.e. keys to gates and batteries, which are sometimes needed to power a flying lift to reach another level. Including the aesthetic aspects, the game strives to catch the player's interest through: *sensation* (high-quality graphics), *fantasy* (an unusual magical land), *challenge* (brain-teasing puzzles) and *discovery* (various landscapes, a gradual introduction of various task types). Last but not least, the *narrative* element might also play a role, depending on the game version.

The designers of the game conceived learning outcomes mainly in the cognitive domain. Throughout the game, the player encounters several mathematical (or maths related) concepts: number comparison, addition, equal division, simple fractions and equations, negative numbers, and patterns (using colour and tone sequences). The concepts appear as the player tries to overcome the puzzles, but the names of the concepts are not visibly highlighted. In the process of solving the tasks, the player mainly practises arithmetics, but also (to a certain extent) trains memory, practises algorithmic thinking, and problem-solving and spatial skills.

To finish all 23 game levels in the experimental versions, the player must surpass 163 puzzles; the game includes the following seven different task types:

1. **Number Web** (61 of 163) - basic arithmetics are incorporated into a simple graph structure. A completed task should comprise of nodes with numbers and oriented edges showing numerical differences between them. The Number Web task type can be found in maths textbooks for the second and third year officially certified by the Czech Ministry of Education (Hejný et al., 2008, 2009). The task exists as a kind of lock on a safe within the game, and two number webs need to be solved to open the door.

2. **Hungry Plant** (22 of 163) - the core of this task is equal division. The magical plant has several “heads”, and the player must give each of the heads an equal amount of food: that can be single objects like apples or carrots, multiple connected parts such as pieces of chocolate, halves, thirds or fourths of a pie, or a wheel of cheese (basic problems with fractions).

3. **Weighing Machine** (21 of 163) - experimenting with the Weighing Machine task should help the player to build the foundations for solving equations. To continue the journey, the player must balance out the machine using various weights, otherwise the heroes cannot pass under the machine.

4. **Flying Platform** (19 of 163) - the player’s goal in this task is to get a flying platform to transport the heroes horizontally from one side of the screen to the other. The platform is controlled through ‘command tiles’, each defining the direction and the number of “steps” the platform will take. By putting the command tiles into a sequence, the player programs the movement of the platform to avoid obstacles and to fly toward the desired destination.

5. **Counting Lift** (19 of 163) - this task has similar mechanics to the Flying Platform task, but the lift only moves vertically. It can stop on different numbered levels - the ground floor is marked zero, the levels above have a positive number and the levels below have a negative number. Depending on the player’s skill, there may be up to three lifts next to each other, which need to be aligned through a sequence of command tiles (a specific lift is controlled by command tiles of a specific colour).

6. **Jigsaw Bridge** (13 of 163) - the task is to put together jigsaw pieces to construct a bridge. More advanced players may also sometimes have to rotate the pieces (always by 90°).

7. **Music Bridge** (8 of 163) - as the heroes follow the path, they can encounter a notestand with a sequence of colourful notes with a discernible pattern. The player must remember the sequence to be able to later play it on a colourful piano keyboard. The correct melody will persuade the bridge guard to lower the drawbridge.

While testing the skills of the participating children, our research team decided to focus primarily on the skills supposedly learnt through these three tasks: the Number Web, the Hungry Plant, and the Weighing Machine. The three chosen tasks represent more than 63 % of the game's content. The main reasons these tasks were given preference over the others were that their educational objectives lay more clearly in the mathematical domain and that they were testable. The test the children took also did not include all the game's task types to limit the the time required to finish so that the children would not be too exhausted and stressed. Although the Flying Platform and the Counting Lift tasks might help to practise simple arithmetics, they were designed to support the development of algorithmic thinking as well. The Jigsaw Bridge and the Music Bridge tasks were not included in the test set as they occur more rarely and their overall benefits for learning measurable maths skills are debatable (the Jigsaw Bridge is more likely to employ spatial skills and the Music Bridge challenges the ability to recognise and remember patterns).

2.4.3 Instructional Support Implemented in the Matemág Game

In addition to the educational content, the game's instructional support can also be identified and classified into groups as defined by Wouters and van Oostendorp (2013) in their meta-analysis. The developers strove to implement instructional support in ways that are further described in the following list of auxiliary features:

- **Adaptivity** - as the player progresses through the game, the difficulty level is adjusted according to previous performance. The player's skill score for each task type is adjusted every time a task is solved.[4]

[4]The outline of the adaptive difficulty algorithm is as follows: The player is given a skill score for each task type. Each specific setting of a task is given a difficulty score. When a task within a level is initiated, the game loads a task setting with a difficulty score close to the player's skill score which has not been solved yet. Once the the task is solved, the game updates the skill score based on the number of tries the player took to finish the task (time is irrelevant for the algorithm): the number of tries is compared with the player's average number of tries for the given task type, and if the player performed better than or equally well as the average, the skill

- **Advice** - puzzle tasks are integrated into the game environment in a meaningful way, helping the player understand their mechanics (contextualisation). The player can also initiate a short spoken task instruction.

- **Feedback** - feedback in the game is simple and provided through the task dynamics. It always helps to determine the correctness/incorrectness of a solution. In some cases, the dynamics hint at what can be done better as well - for example, if the weighing machine leans to one side after a weight is added, it could be that the weight is too big.

- **Interactivity** - the player may experiment freely when solving the puzzles - all actions are reversible until the task is solved, and there is no time limit.

- **Narrativity** - the "journey" frame was designed with tasks as obstacles to reduce the time needed to grasp the newly introduced puzzles throughout the game. The premise is always the same - solve the task to continue your journey. The richer narrative elements included in one of the two versions are discussed later.

- **Variability** - each level contains a different mixture of task types, and the tasks are introduced gradually throughout the game to maintain a feeling of novelty. The settings of the same task type also vary in some aspects (the shape of the Number Web, the types of food for the Hungry Plant, etc.), and the type of landscape changes throughout the levels (e.g. meadow, forest, rocks, cave), keeping the game fresh.

2.4.4 Game Versions: Differences in the Narrative

The simpler version of the game used in the experiment is the narrative frame version. It opens only with a single cartoon image and a sentence that reads: "Theresa and Jacob magically appeared in a wizard's land." The player is given control of the main characters immediately after. When all 23 levels have been completed, another cartoon image appears to close the very simple storyline with a short text: "The children have successfully finished their journey to the castle from which they could easily return back to Earth." These brief introductory and closing parts appear in the narrative frame version so that the menu of both of the game versions looks the same at a glance and also that the player can get to know the main heroes.

The rich story version differs in that the introductory and the final cartoon comics are more extensive and the comic bubbles are voiced. Players can go through the comics at

score is raised; if the performance was worse, the skill score drops. In the special case where the player only needed one try (solved without problems), the skill score raises faster. All the players start with the lowest skill score to ensure that the task settings they encounter at the beginning of the game are not too difficult.

Fig. 2.10 The main heroes, Theresa and Jacob, meeting the wizard, Matemág, for the first time in the story introduction (TechSophia s.r.o., 2017a).

their own pace. Taking into account only the length of the audio in the story scenes - the introduction is approximately 8 minutes long and the final part can be finished in slightly less than 10 minutes - the total time for the story amounts to around 18 minutes[5]. Beyond mere entertainment, the script was written with the intention to illustrate maths as a tool that helps people to think and to overcome problems (see the full story summary in Appendix 1: The Story of Matemág).

The introductory part of the narrative is delivered in four chapters containing multiple cartoon frames. The time needed to go through the chapters is approximately: 2 minutes for chapter 1, 2 minutes 40 seconds for chapter 2, 2 minutes 45 seconds and 35 seconds for chapter 3 and 4 respectively. As mentioned above, the times may vary as the interactive comic is self-paced. The player learns more about a pair of siblings called Jacob and Theresa, the main heroes; how they came to their grandmother's house for holiday, how they discovered a magical book that belonged to a maths magician, how they met him (captured by Figure 2.10) and finally went to visit him in his Land of Abstraction and Imagination. In the comic, the wizard also shows the siblings the power of maths magic, which helps him to invent things and to imagine the possible outcomes of a decision (in that particular case, the unwise decision to eliminate all numbers). In the rich story version, the heroes' goal during their

[5]The time spent on the story scenes can vary from player to player - some might take a longer time admiring the comic frames and some might actually skip to go faster.

search of the magical land is clear: to seek out the wizard again and to learn maths magic. No such motives are mentioned in the narrative frame version of the game.

Fig. 2.11 In one of the final story chapters, Matemág tells the heroes that he will not teach them "maths magic", that they need to discover it for themselves by experimenting. At first the heroes are slightly confused by what he means, but they begin to understand after solving a real-life problem using the power of imagination (TechSophia s.r.o., 2017a).

The final part of the story was also designed in the form of a voiced interactive comic and has three chapters (2 minutes 40 seconds, 2 minutes 10 seconds and 5 minutes long), in which the heroes scale the wizard's tower, meet him again (Figure 2.11) and return to their grandmother's house. The main characters debate as to whether they learnt anything while travelling through the magical land. In the end they use their wits to save a cat stuck on a roof and they conclude that maths skills help them to think problems through. In two of the three final chapters, an interactive scene was included to make them more entertaining; the player's simple task is to build stairs out of predefined blocks, or to attach helium balloons to lift a basket.

2.4.5 "Placebo" Game for the Control Group

It was announced ahead to the participants that they would receive an educational game and so the randomly selected control group was also given a video game similar to the one given to the other groups. The game *Spojovačky*[6] (TechSophia s.r.o., 2017b) was designed in the same cartoon style (Figure 2.13), featuring the same main heroes. It contained only one type of logical puzzle (Figure 2.12) in which the player must connect the nodes of a graph-like structure using one continuous line. The skills potentially acquired by playing the "placebo" game do not overlap with the skills measured by the maths test.

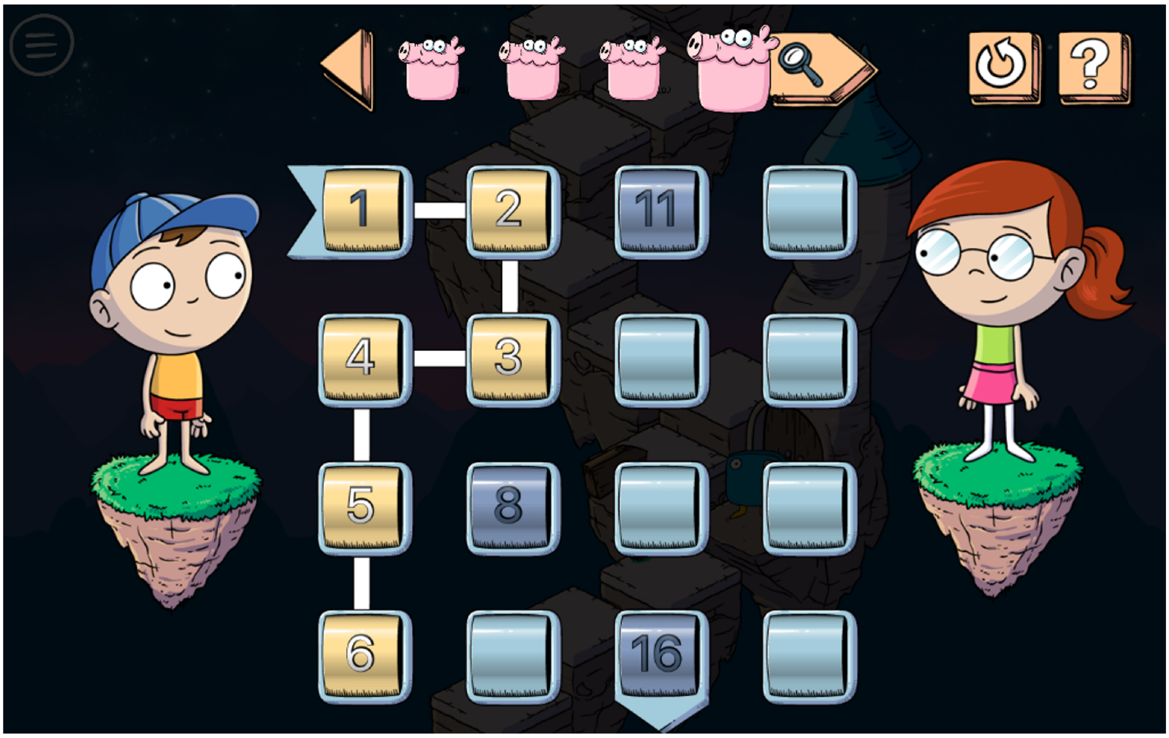


Fig. 2.12 In the "placebo" game (TechSophia s.r.o., 2017b) for the control group, the goal is to connect numbered and not-yet-numbered nodes into a sequence. The main heroes, Jacob and Theresa, are the same characters as the ones in the *Matemág* game.

2.4.6 In-game Measures

The two experimental versions of *Matemág* were logging data as the participants played the game, whereas the "placebo" game given to the control group did not implement any metrics. Table 2.1 provides an overview of the logged data and their uses. Most importantly, the game kept a log of which scene was visited and when, which helped to estimate the **time spent on story**, the **time spent on game levels**, and it showed the **number of sessions**, the **number of days** and the **usual time of day** the game was played. The record of the

[6]The name "Spojovačky" created from the Czech verb 'spojovat', meaning 'to connect' or 'to link', refers to the main game mechanics, which are focused on linking parts of a graph.

Fig. 2.13 The progress through the "placebo" game (TechSophia s.r.o., 2017b) is visualised as a staircase. The player can also choose the difficulty of the task on each stair (feather = easy, carrot = medium, pig = difficulty).

scenes visited revealed information about the player's progress through the game (the **level reached** variable). The **frequency of visits to levels** and the **frequency of visits to story chapters** was also collected; the latter being a valuable tool for shedding light on the player's consumption of the narrative parts.

Each time a task was solved, the game saved several metrics - the details about the task (which task type and which task setting), the time taken to solve the task and the number of tries taken to finish it. In the end, the total **time on tasks** and the **total number of solved tasks** were used for the analysis. As the participants played the game, the players' **skill score** was being recorded for each task type. It was always reevaluated after each task solved (the tasks themselves were given a related 'task difficulty' score). The game uses the skill score to adjust the difficulty of the puzzles as previously mentioned in subsection 2.4.3.

Log type	When logged	Collected data	Use
game scene	at the start of a scene	scene name, timestamp	progress through the game, frequency of visits to story chapters and specific levels, frequency of playing, estimation of gameplay time
task solved	when a task was solved	task type, task difficulty, task id, opened timestamp, number of tries, solved timestamp	number of tasks solved, estimation of time on task, the total gameplay time
skill score (for each task type)	when a task was solved	task type, skill score, timestamp	the player's skill score development, skill reached for a given task type

Table 2.1 An overview of the logged data.

2.5 Procedure

2.5.1 Approaching Schools and Parents

Multiple schools in the Czech Republic from the areas of Prague and the Central Bohemian region were contacted through email and phone and were introduced to the research and its goals. Once a school expressed willingness to take part in the experiment, we proceeded with further communication - we were especially interested in whether they taught maths following Hejný's constructivist scheme-oriented approach (Hejný, 2008). As the game's educational content was influenced by the approach, we had decided to exclude classes where the approach was being used from the experiment as children from such classes may have shown results that would not be comparable with those of children taught using other methods. Moreover, the commercial version of the game was already publicly available at the time and its "evangelisation" was strongly supported by parents and teachers who were in touch with Hejný's approach - the probability of possible contamination of the sample was therefore higher in these school classes.

The contacted schools and teachers then helped contact the parents - they received a one-page handout informing them of the opportunity for them and their children to participate in the study. They were made aware of the general aim of the research - to help to enhance the design of educational games - and the potential reward in the form of a package of four educational games worth 25 Euros after the end of the experiment. They also learnt through the introductory material that the experiment would involve having a maths game installed and available for their children for 14 days. The option to mark down whether they would need to borrow a touch device or had their own was available when enrolling for the experiment as well.

As the last remark in this section, it must be highlighted that two key principles were communicated to both the school officials and the parents: voluntary participation and anonymity. The participants were free to leave the experiment at any point and all their data would be processed anonymously.

2.5.2 General Context for the Meetings with Children and Parents

There were two meetings planned with the children and the parents for each experiment round at a school - the initial and then in 14 days the final meeting. The experiment rounds took place during the school year; the first started on 11 April 2018 and the last round finished on the last school day, on 23 October 2018. Second-year pupils took part in the rounds at the first two schools in the spring of 2018. When the new school year started in September after the summer holiday[7], we decided to target fresh third-year pupils as their age was closer to the children from the spring rounds.

2.5.3 Experiment Administrators

There was one main coordinator for each round of the experiment at a school who communicated with the head of the school, the teachers and the parents. He was responsible for organising the meetings with the children and parents. The main coordinator was also available to answer the parents' and the school's questions throughout the 14 days of the experiment.

A team of administrators was present at both of the school meetings. Usually, one or two played the role of the parent group administrator, the rest of them were children group administrators. All of them were trained before the sessions. It was especially the privacy of the participants that needed to be handled correctly, but the maths skill test administration was also important. The testing, similar in both meetings, is described in subsection 2.5.4 dedicated to the initial meeting. A children group administrator oversaw between seven to ten children. He or she was there to support the children, but the administrators were instructed not to influence the children's performance in any way. If the pupils asked about a particular task, the administrators were allowed only to:

[7]There are two-month holidays in the Czech Republic at the end of the school year starting 1 July and ending 31 August.

- point to the written task instruction (and help with the reading of it if the pupils were struggling with the text),
- remind them that each task type has a corresponding example task,
- remind them that they could skip tasks.

The children group administrators made sure that the children would not cheat or be too loud if they had finished the test earlier than others, and they also ensured that the 35-minute time limit for the test was observed. When some children finished the test faster and managed to complete the questionnaires ahead of others as well, the administrators asked them to draw something nice and to stay quiet. More details are discussed in sections about the meetings.

2.5.4 Initial Meeting

Each initial session was approximately 70 minutes long in total. The participants arrived in the afternoon (around five o'clock) when the teaching hours were over. Two classrooms were ready, one for the children and the other for the parents, but the research team first gathered both groups together for an introduction. Each pair of a child and a parent was randomly assigned to an experimental group - for the participants this meant that they received a paper card with a geometrical symbol; they did not know about the differences between the groups. After welcoming the participants, we received the parents' written consent to the participation in the experiment under the conditions of anonymity and the option to leave the experiment if they decided to do so. The participants were asked - for the sake of anonymity - to invent a nickname under which they would take part. From that moment on, all tests, questionnaires and game statistics were signed only using those nicknames.

The meeting continued by separating the parents and the children into two separate groups. Children entered one classroom to complete a maths skill test and a questionnaire, and the parents entered another to do other tasks, i.e. to install the game. A list of instructions was provided to each parent on a sheet of paper - the specific steps slightly differed depending on the operating system and the experimental group. Nevertheless, the installation steps were similar enough that the participants could go through the instructions without noticing that they were installing a different game (or game version).[8]

Finally, the parents were briefly instructed about what to do during the 14 days of the experiment. They were asked to show the game to their children at least once at the beginning

[8] For example, they were instructed to type in a URL address to download the game. The URLs differed for the experimental groups, but everyone was following their list of instructions on a sheet of paper in front of them and not checking what the others were typing in.

but then behave as they would usually do concerning games and their children's playtime. They were further asked not to install any other educational games during the course of the experiment. We also inquired as to whether the parents had ever installed the commercial version of the *Matemág* game before. If they had, we noted down the participants' nicknames to exclude their data from the experiment. The children were questioned about the game as well - the questionnaire asked whether they knew the game's characters.

What was the procedure in the children's group? One of the children group administrators started by explaining what was going to happen - first there would be a maths test (presented as "interesting math tasks") and then a questionnaire (presented as "a few questions about what you like"). Before the maths skill test, the children were informed that they would have 35 minutes to finish the test, that the test would not be graded and that they could skip a test question if they would rather continue with the next one. We also stressed that they should work alone and that we would make sure that they were not cheating. The children were told to raise their hand if they had any questions during the session and that one of the administrators would come to them. Then the tests were distributed.

There were two versions of the test - version 'A' and 'B' for the purpose of this thesis. These versions differed slightly in number values, and colours and positions of objects within the test questions (minor visual differences). The versions were created mainly for the pre-post experimental design, but they served to prevent cheating as well. Roughly half of the children in the initial session received, at random, version A, the other half received version B. In the final session the children always received the version they had not received at the start of the experiment. Each of the test versions had two parts - set one and set two. The children were given set one first and received the second set before finishing the first one so that they would be able to skip the last tasks of set one if they decided to do so. When children finished the test before the limit was up, they were given the following questionnaire. In 35 minutes the tests were collected whether or not the children had managed to finish them completely.

The rest of the time was dedicated to a questionnaire focused on the children's preferences, expectations about the game, and their possible familiarity with it (see subsection 2.3.2, and Appendix 3: Child Questionnaires). The children went through the questions alone[9], but they always had the option to ask the team if they were uncertain about something. In the end the children group administrators checked if the children recognised the characters from the *Matemág* game and if that was the case, interviewed the children briefly to make sure that the

[9]As the childrens' pace was notably varied, the administrators often asked the faster children to draw a picture on the other side of the questionnaire to keep them occupied and keep them from disturbing the others.

data of those children who had played the game ahead of the experiment would be excluded. Once the children finished the questionnaire, they could rejoin their parents and leave.

2.5.5 The Two Experimental Weeks

It was up to the participating children and parents how often, when, where, and how the games would be played. Both the rich story and the narrative frame version maintained logs from the play sessions and saved them to the devices (the game was played offline). All the parents were given contact information in case they needed to communicate with the main coordinator; that included questions about the final meeting, their apologies, but on rare occasions also technical problems such as a broken device. All cases were dealt with individually.

On day 7 of the experiment, the parents were sent a short message telling them to remind their children about the game if they had not played at all (to ensure that the children would have the opportunity to familiarise themselves with the game). The message also reminded them about the date and the place of the final meeting, which was information they had already received when they enrolled in the experiment.

2.5.6 Final Meeting

The time and the place of the final meeting and some parts of the procedure were similar to those of the initial meeting described in subsection 2.5.4, which is why this section focuses primarily on the differences during the part of the meeting after the parents and the children had been again separated into two groups.

The parents were given two questionnaires commented in subsection 2.3.3 and accessible in Appendix 4: Parent Questionnaires. As they were filling in the answers, a parent group administrator approached each of the parents one by one to upload the logs from their devices to our database. When all the parents were finished (usually in less than 45 minutes), in the rest of the time there was an opportunity for the parents to ask about the experiment in a moderated discussion. Among the topics raised by the administrators was the question if the children had found out about multiple game versions, and if they had played one of the other games (the children were asked as well). Uncovered between-group contamination was noted down for exclusion of the data.

Following the pre-post experimental design, the children were given another maths skill test. Same instructions for the children and instructions for the administrators applied as in the initial meeting. After the test followed another questionnaire prepared for the children, in which they were asked for their opinion about the game and which tried to detect possible between-group contamination. When collecting the questionnaires, the administrators further questioned the children who had provided some information indicating that they had played a different game than they should have done. The contaminated samples that were discovered were excluded from the study.

2.5.7 Concluding an Experimental Round

With the end of the final meeting with the children and parents, the associated experimental round also drew to an end. The participants received promotional codes to download four educational games for free as a reward. The research team gathered all the tests and questionnaires and scanned them before archiving them. The scans and the data logs were processed in a manner described by the following sections.

2.5.8 Tests and Questionnaires Processing

Once scanned, all of the tests were corrected and scored. A solution for each of the test tasks was evaluated and assigned a score of one point, a half-point, or zero points. One point was given for a correct solution, zero points for an incorrect one. If a solution was partially correct, it was given a half-point[10]. The maximum score of the maths skill test was 21 points. The test scores were normalised using z-score normalisation for the two versions of the test.

Once the questionnaires were digitised, the responses were assigned values in plausible cases - many of the questions were answered on some sort of scale. The relevant measures were combined for easier analysis; the variables were already mentioned in section 2.3.

2.5.9 Game Logs Processing

Firstly, the collected data were cleaned, all duplicates were removed, and they were linked with the participants through their nicknames. We also tried to detect errors in the data. In two cases the devices used by the children had a wrong system time, meaning the timestamps were offset by a few years as a result: this was easily corrected as it was possible to shift the first log, which was created at the initial meeting, to the actual time of the meeting and adjust all the other logs using the same difference. There was also an error found in the logs of two task types - the Flying Platform and the Counting Lift; there was no record made when the player opened the task and only the ‘task solved’ timestamp remained. However, the tasks of these two types were not part of the near-transfer maths skill tests and therefore it was not

[10]For example when the task was to equally divide all pieces of a pie between a few people, and in the solution some pieces were left unassigned, but the rest was divided equally.

crucial to have a precise measurement of the time spent on them (if needed, the time on task could have been roughly estimated by taking the timestamp of the closest previous event).

Today's mobile applications often run after the launch for a long time - either in the foreground or in the background - even when the user is not directly using them so the breaks in the playing time had to be identified. By analysing more than 15,000 logged events (scene opened, task started and task solved), we were able to reconstruct the gameplay sessions. It was realisable because of the uncomplicated structure of the game (the player is either reading a story chapter, switching to a level in the menu or solving tasks within the level at any given time) and the similar patterns within the data. Essentially, we compared the time differences between an event with particular features (such as a scene with a specific name) and the event right before and right after; if the time difference was more than five standard deviations, we marked it as a break in the session.

2.6 Data Analysis

The data were analysed using the statistical program R version 3.5.0 (R Core Team, 2018). When internal consistency needed to be tested, e.g. while examining questions from the parent questionnaires, the Cronbach's α (Cronbach, 1951) was calculated. The Pearson's χ^2 test (Pearson, 1900) was used to compare categorical variables between the groups. For comparison of ordinal data and data with non-normal distribution, we made use of Cliff's δ (Cliff, 1996), assessing the magnitude according to the thresholds provided by (Romano et al., 2006): $|\delta| < 0.147$ as "negligible", $|\delta| < 0.33$ as "small", $|\delta| < 0.474$ as "medium", otherwise "large". Normality was evaluated through the Shapiro-Wilk test (Shapiro and Wilk, 1965); t-test and ANOVA were employed only when the test showed that there was no significant departure of the data distribution from the normal distribution. Between-group comparisons for two groups were conducted using independent samples two-sided t-tests (Student, 1908). When comparing samples from all three groups, one-ways ANOVA (Fisher, 2006) was applied; the p values were obtained through the a posteriori Tukey test. Cohen's d was used to estimate the effect sizes, which were classified as suggested by (Cohen, 1988): "small" for $d \sim 0.2$, "medium" for $d \sim 0.5$, or "large" for $d \sim 0.8$. Correlations between variables were evaluated using the Pearson correlation coefficient and the effect sizes were classified according to (Cohen, 1988): $|r| < 0.1$ as "small", $0.1 < |r| < 0.3$ as "medium", and $|r| > 0.5$ as a "large" effect size. The p-values for correlations are not corrected for multiple comparisons.

Chapter 3

Research Results

3.1 Participants

3.1.1 Were the Experimental Groups Balanced?

The number of participants in each of the groups was different (not significantly so - $\chi^2 = 1.1$, $p = .576$); the narrative frame group ($n = 26$) was the largest one, then it was the rich story group ($n = 22$) and finally the control group ($n = 19$). The differences were caused by several factors: some participants did not come to take the posttest and some were excluded due to prior exposure or between-group contamination. As the "placebo" game for the control group had fewer similarities than the two narrative versions of *Matemág*, the children from the control group more frequently noticed that the games were different and played one of the narrative versions. Once these cases were detected, the related data was excluded.

The groups were balanced in age ($F_{2,64} = 0.06$, $p = .942$) and gender ($\chi^2 = 0.97$, $p = .615$). Other variables which may have had influence on the maths skill test score were also tested for balance for all of the three groups. Table 3.1 gives an overview of these variables and shows that there were no significant differences between the groups. Nevertheless, it should be mentioned that the rich story group had the lowest number of university-educated parents ($\chi^2 = 2.81$, $p = .238$). Although the groups had a similar number of third-year pupils, there were more second-year pupils in the narrative frame group ($n = 13$), fewer in the rich story group ($n = 9$) and fewest in the control group ($n = 5$); however, the differences between the numbers were not significant ($\chi^2 = 2.56$, $p = .277$). Although the control group participants claimed that maths was their favourite subject less frequently (not significant, $\chi^2 = 2$, $p = .368$), the control group slightly outperformed the other groups in the maths skill pretest - the difference was not significant, $F_{2,64} = 0.36$, $p = .697$).

Measure		Data for groups			χ^2 (2)
		frame	rich	control	or F (2, 64)
N		26	22	19	χ^2 = 1.1 p = .576
gender	*female*	10	9	10	χ^2 = 0.97
	male	16	13	9	p = .615
age	*mean*	8.69	8.67	8.65	F = 0.06
in years	*SD*	0.42	0.43	0.38	p = .942
school year	*2nd*	13	9	5	χ^2 = 2.56
	3rd	13	13	14	p = .277
maths as	*YES*	19	17	11	χ^2 = 2
favourite subj.	*NO*	7	5	8	p = .368
parent's	*university*	13	6	9	χ^2 = 2.87
education	*other*	13	16	10	p = .238
pretest	*mean*	-0.44	-0.46	-0.25	F = 0.36
z-score	*SD*	0.89	0.89	0.9	p = .697

Table 3.1 An overview of the participants' characteristics that were tested for balance to see whether they may have influenced the test results.

Due to the interest in how the narrative frame and the rich story group compared with regard to playing an educational (story) game, the balance of the two groups in personal preferences, family habits, and technical details had to be examined as well. In general, the children from both groups liked to play video games and had a positive attitude toward stories. Table 3.2 shows that the groups were very much alike in these and the rest of the traits. The table also contains information about the type of device the children used (tablet/smartphone), the device operating system (Android/iOS), whether the device was used solely by the child and whether it was borrowed from us for the research. The narrative condition groups were balanced in terms on these variables.

3.2 Were the Children Really Involved?

At the beginning, it would not have been amiss to ask the question of whether there would actually be anything to analyse, meaning if the children would play the games at all. Based on the in-game logs from the two game versions which differed in their narrative parts, it appears that they did; the average level reached across the two groups was 15.52 out of 23 (SD = 6.51), the average number of days the children played was 6.68 (SD = 2.03), the average number of tasks solved was 138.2 (SD = 56.02) and the average estimated time spent on tasks was 98.58 minutes (SD = 49.97). The players spent an average of 249.29 minutes

Measure		Data for groups frame	rich	χ^2 (1) or Cliff's $\|\delta\|$
playing	*YES*	21	15	χ^2 = 0.45
video games	*NO*	5	7	p = .504
how often play	*mean*	3.69	3.27	δ = .129
video games (0-5)	*SD*	1.52	1.75	(negligible)
story preference	*mean*	4	4.05	δ = .002
score (0-5)	*SD*	1.06	0.94	(negligible)
parent prompted	*mean*	1.46	1.61	δ = .077
playing (0-4)	*SD*	0.62	0.9	(negligible)
parent set time	*YES*	23	19	χ^2 = 0
limit on playing	*NO*	3	3	p = 1
device type	*tablet*	13	10	χ^2 = 0.001
	smartphone	13	12	p = .981
device used	YES	15	13	χ^2 = 0
by child only	NO	11	9	p = 1
borrowed	YES	24	19	χ^2 = 0.04
device	*NO*	2	3	p = .843
device OS	*Android*	22	19	χ^2 = 0.03
	iOS	4	2	p = 0.874

Table 3.2 An overview of gaming-related characteristics relevant mainly for the two narrative condition groups compared in the value-added study.

in the levels (SD = 115.69) - not only solving tasks, but also walking and discovering the land. In addition, the rich story group spent an average of 20.06 minutes on the story parts (SD = 13.88). About 70 % of the participants reached the middle of the game (level 12), and 27.66 % of the participants reached the final level.

3.3 Engagement and Progress Through Game (H1)

3.3.1 Was the Rich Story Version More Engaging? (H1a)

The level of engagement was assessed using the parent and the child questionnaires. The key score, labelled 'child engagement score', was constructed based on responses to six questions in the parent questionnaire (as described in subsection 2.3.3). The possible minimum value for the combined score was 0, which would mean the parent did not observe much engagement in the child, and the maximum was 12 points. The average child engagement score for the control group was the lowest (mean = 4.32, SD = 2.19), and the narrative frame group (mean = 5.96, SD = 2.62) reported lower engagement than the rich story group (mean = 7.14, SD = 2.62). The single-factor ANOVA revealed a significant difference between the three groups: $F_{2,64} = 6.484$, $p < .01$. The *a posteriori* Tukey test showed that the p value is significant only for control vs. rich story ($p < 0.01$), marginally significant for control vs. narrative frame ($p < .083$). The difference between the narrative frame and the rich story group in engagement was not significant ($p > .245$), the effect size could be classified as small to medium ($d = 0.45$).

Measure	Group	Min	Max	Mean	SD
parent-reported child engagement (0-12 points)	frame	2	11	5.96	2.62
	rich	2	11	7.14	2.62
	control	0	9	4.32	2.19
child-reported enjoyment (0-5 points)	frame	2	5	4.65	0.83
	rich	1	5	4.6	1
	control	0	5	3.73	1.49

Table 3.3 The key measures concerned with engagement and enjoyment.

A marginal negative medium effect size correlation found only for the narrative frame group between the parent prompted playing variable and the child engagement score ($r = -0.4$, $p = .054$); both of these variables were obtained from the parent questionnaire. The parent prompted playing score also showed a negative medium effect size linear relationship with total gameplay time ($r = -0.47$, $p = .019$) for the narrative frame group; the correlation was less strong and non-significant for the rich story group ($r = -0.34$, $p = .13$).

Measure	Comparison	F(2, 64)	post hoc p	Cohen's d
parent-reported child engagement (0-12 points)	frame-rich	F = 6.48 p = .003	.245	0.45
	control-rich		.002	1.16
	control-frame		.083	0.67
		Cliff's δ		
child-reported enjoyment (0-5 points)	frame-rich	-0.02	(negligible)	
	control-rich	0.41	(medium)	
	control-frame	0.42	(medium)	

Table 3.4 Effect sizes for engagement & enjoyment.

The information about the children's engagement was complemented by self-reported child enjoyment at the end of the experiment. The games were given up to 5 points. The "placebo" game played by the control group was rated the lowest (mean = 3.73, SD = 1.49, n = 15), and the difference between the ratings of the narrative frame version (mean = 4.65, SD = 0.83, n = 23) and the rich story version (mean = 4.6, SD = 1, n = 20) of the other game was negligible (Cliff's $\delta = 0.02$). The parent prompted playing variable was not, according to correlation, in linear relationship with child enjoyment (narrative frame: r = -0.10, p > .05; rich story: r = -0.11, p > .05).

Correlations	Group	level reached	n. of tasks solved	prompted play
parent-reported	frame	0.54**	0.57**	-0.4
child engagement (0-12 points)	rich	0.18	0.42	-0.15
child-reported	frame	-0.06	-0.07	-0.1
enjoyment (0-5 points)	rich	0.41	0.53*	-0.11

Table 3.5 Correlations for engagement and enjoyment vs. in-game measures.
*Note: p < .0001 '****', p < .001 '***', p < .01 '**', p < .05 '*'*

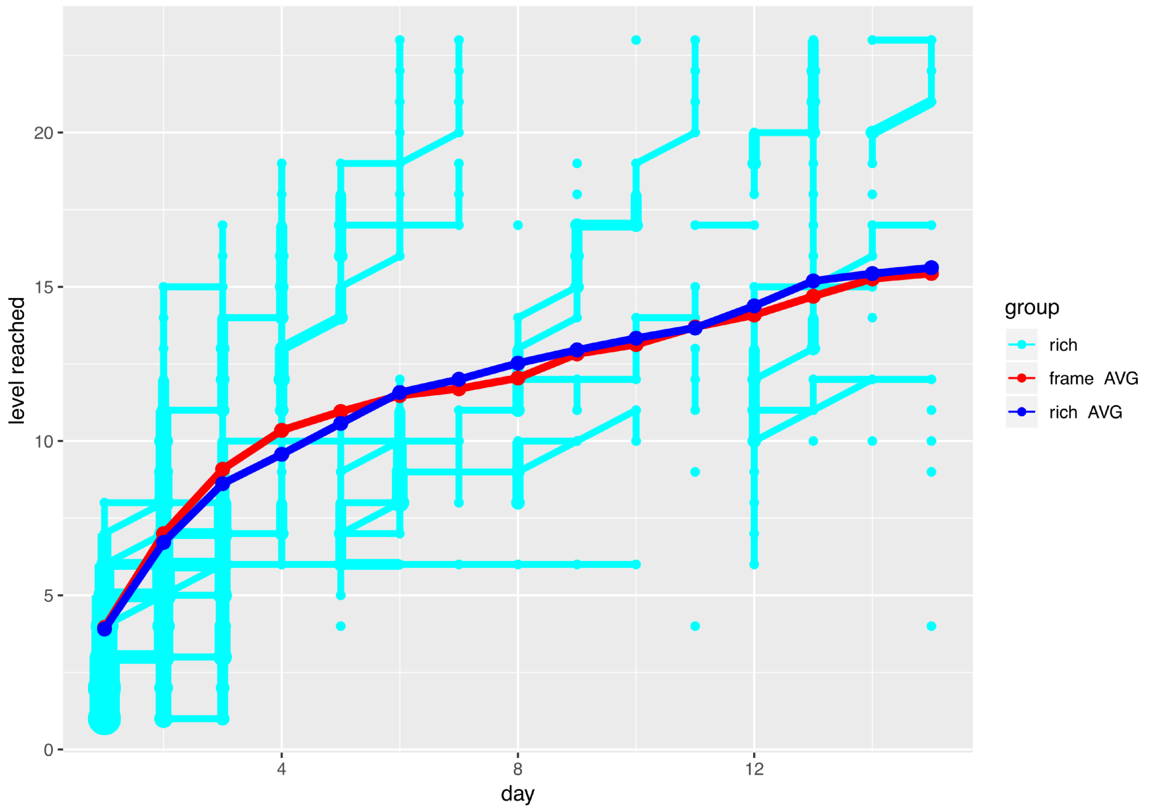


Fig. 3.1 A graph showing how all players from the **rich story group** progressed through the game. If any player had been playing a certain level on a certain day there was a point made in the graph. If the player had been playing also the day before or the next day the two points were connected by a line (the thickness represents that more players reached the same level on the same day). The mean level reached on a given day is highlighted by a line for both narrative condition groups.

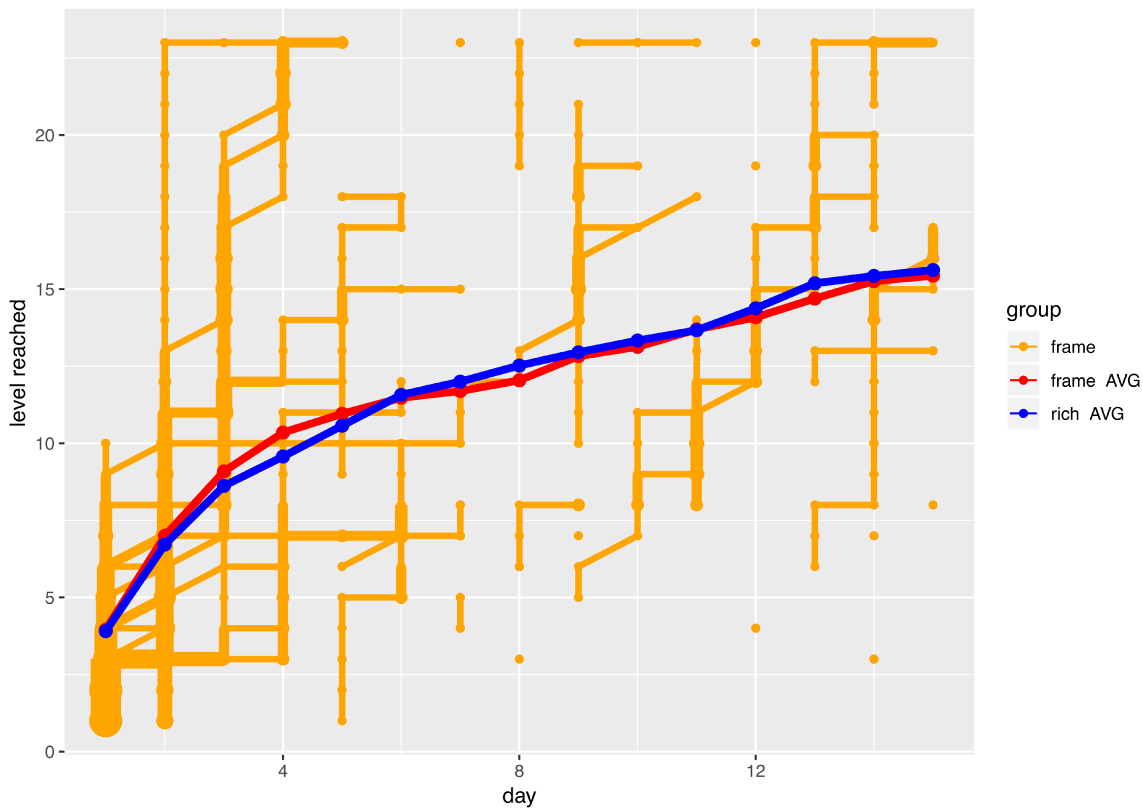


Fig. 3.2 A graph showing how players from the **narrative frame group** progressed through the game. If any player had been playing a certain level on a certain day there was a point made in the graph. If the player had been playing also the day before or the next day the two points were connected by a line (the thickness represents that more players reached the same level on the same day). The mean level reached on a given day is highlighted by a line for both narrative condition groups.

3.3.2 In Which Version Did Participants Progress Further and Solve More Tasks? (H1b & H1c)

Based on the logs from the narrative frame and the rich story game version, this question can be answered with high precision. The children did, in fact, actively play both of the versions: in the narrative frame version, 65 % of the participants reached the middle of the game (level 12) and 30 % reached the last level (23); for the rich story group, it was 71 % and 28.6 % respectively. When measured using the last level reached by the player (out of 23), the difference in progress between the narrative frame (mean = 15.4, SD = 7) and the rich story group (mean = 15.6, SD = 6.1) was negligible (Cliff's $\delta = 0.01$).

As the children had the option to replay any level (the inner settings of the task would be different each time), counting all the tasks solved by each of the players regardless of the type (see the Game Description) can provide further information about their interest in the game. The mean number of tasks solved for the narrative frame group (mean = 137.5, SD = 64.9) compared to the mean for the rich story group (mean = 140.5, SD = 44.2) did not differ significantly (independent samples t test: $t_{40} = -0.182$, $p = .857$).

Based on correlations (Table 3.8) it could be commented that children with better pretest score appear to be progressing further in the game (narrative frame: $r = 0.65$, $p < .001$; rich story: $r = 0.42$, $p > .05$) and solve more tasks (narrative frame: $r = 0.62$, $p < .01$; rich story: $r = 0.43$, $p > .05$). The correlations were stronger and more significant for the narrative frame group.

Measure	Group	Min	Max	Mean	SD	t test	Pr(>|t|)	d
n. of solved tasks	frame	8	243	137.34	64.86	df = 40.72	.857	0.05
	rich	35	210	140	44.23	t = -0.18		
estim. time on task (min.)	frame	0.75	234.87	94.15	51.30	df = 42.79	.547	0.18
	rich	25.53	225.38	103.13	48	t = -0.61		
n. of sessions	frame	2	17	10.71	3.52	df = 38.02	.309	0.31
	rich	4	20	11.95	4.44	t = -1.03		
n. of days	frame	2	9	6.17	1.69	df = 36.92	.08	0.55
	rich	4	11	7.24	2.23	t = -1.8		

Table 3.6 Measures obtained from data logs collected in the game during playing. Only the number of days statistic marginally differed for the two groups compared. The rich story group may have needed extra time to go through the story, which possibly increased the group's number of sessions and days.

3.4 Learning Gains (H2)

3.4.1 Examining the Maths Skill Test and the Test Questions

Learning gains were measured as the difference between the posttest and the pretest scores. The data were firstly used to examine the maths skill test versions and all of the questions (maths tasks): on average, the children scored higher on the posttest (mean z-score = 0.39, SD = 0.95) than on the pretest (mean z-score = -0.39, SD = 0.88). The scores for pretest vs. posttest (all the groups put together) differ significantly with large effect size ($t_{66} = -10.93$, $p < 0.0001$, $d = 1.34$). There was a strong correlation found between the pretest and posttest scores ($r = 0.84$, $p < .001$), whereas the correlation between the pretest score and the post-pre difference in scores was small ($r = -0.15$, $p = .302$).

The correlation of the score for each test question with the total score was always over 0.40 or larger for both test versions, with the exception of the very first questions of the division and weighing task type. The value of r for those tasks ranged from 0.10 to 0.30 depending on the test version. These first tasks were the easiest of the given task type and therefore most of the children solved them correctly, and some children made a mistake by accident but solved the following more difficult questions successfully, which seems to be the reason why the correlation was lower for these two particular questions out of the 21 test tasks.

3.4.2 Experimental Groups Test Results & Game Scores

Measure	Group	Min	Max	Mean	SD	F(2, 64)	post hoc p	d
pretest z-score	frame	-1.82	1.1	-0.44	0.89	F = 0.36 p = .697	frame-rich .997	-0.02
	rich	-1.74	1.88	-0.46	0.89		control-rich .723	-0.24
	control	-1.54	1.47	-0.25	0.89		control-frame .745	-0.22
posttest z-score	frame	-1.35	1.69	0.55	10.1	F = 0.8 p = .456	frame-rich .854	-0.16
	rich	-1.35	1.88	0.4	0.84		control-rich .753	0.23
	control	-1.25	1.88	0.18	1.01		control-frame .422	0.36
post-pre differ.	frame	-0.1	1.8	0.99	0.5	F = 5.98 p = .004	frame-rich .691	-0.25
	rich	-0.13	1.95	0.86	0.56		control-rich .04	0.74
	control	-0.46	1.89	0.43	0.61		control-frame .004	1.00

Table 3.7 Scores showing how the groups performed in the near-transfer maths skill test.

The results for the three experimental groups show that the control group performed better on the pretest (mean z-score = -0.25, SD = 0.9) than the narrative frame group (mean z-score = -0.44, SD = 0.89) and the rich story group (mean z-score = -0.46, SD = 0.89); however, as has been mentioned before, when the experiment was examined to see whether it was balanced (see subsection 3.1.1), the difference was not proven to be significant (the p values obtained from the Tukey post hoc test were .745 for control-narrativeFrame, and .723 for control-richStory). In the posttest the control group scored the lowest (mean z-score = 0.18, SD = 1.01), having been outperformed by the rich story group (mean z-score = 0.4, SD = 0.84) and even more so by the narrative frame group (mean z-score = 0.55, SD = 1.01).

The comparison of the post-pre score differences should indicate whether the games had any effect on learning. The control group had the lowest learning gains (mean = 0.43, SD = 0.61), followed by the rich story group (mean = 0.86, SD = 0.56) and finally by the narrative frame group (mean = 0.99, SD = 0.5). The single-factor ANOVA showed a significant difference between the learning gains of the three groups: $F_{2,64} = 5.982$, $p < .01$. The *a posteriori* Tukey test showed that only the control group was significantly different both from the narrative frame group (p = .004) and from the rich story group (p = .04) in

Correlations		pretest z-score	posttest z-score	post-pre difference	level reached	n. of tasks solved
posttest	frame	0.86****				
z-score	rich	0.83****				
post-pre	frame	-0.07	0.44*			
difference	rich	-0.33	0.26			
level	frame	0.65***	0.69***	0.23		
reached	rich	0.42	0.37	-0.12		
n. of tasks	frame	0.62**	0.7***	0.27	0.97****	
solved	rich	0.43	0.44*	-0.01	0.90****	
time	frame	0.42*	0.49*	0.23	0.83****	0.88****
on task	rich	0.03	0.20	0.27	0.81****	0.83****

Table 3.8 A correlation table for the test scores and the in-game measures of progress and task-solving. Interestingly, the pretest and posttest scores correlated significantly with all of these in-game measures only for the narrative frame group.
*Note: p < .0001 '****', p < .001 '***', p < .01 '**', p < .05 '*'*

terms of gains. The two narrative groups, whose game versions were similar but for the extent of the game's story, did not differ significantly as to their learning gains (p = .691). Cohen's d calculated for control vs. each of the other groups separately was larger for control-narrativeFrame (d = 1.00) than for control-richStory (d = 0.74).

Skill score for task type	Group	Min	Max	Med	Mean	SD	t-test	Pr(>|t|)
Hungry	frame	10.21	98.3	63.83	62.61	21.39	df = 42.1	.379
Plant (0-100)	rich	24.26	84.26	60	57.63	16.08	t = 0.89	
Number	frame	0	100	64.78	60.56	28.64	df = 42.43	.758
Web (0-100)	rich	18.81	91.94	60.6	58.21	22.21	t = 0.31	
Weighing	frame	0	97.6	44.8	42.67	28.48	df = 43	.774
Mach. (0-100)	rich	0	81.6	46.4	44.95	24.59	t = -0.29	

Table 3.9 In-game skill scores for each of the task types which were covered by the maths skill test. As the Weighing Machine task is introduced slightly later in the game, some children only received the initial zero skill score, resulting in a mean value lower than for the other two task types.

The results of the comparison of the narrative frame and the rich story game versions in terms of learning gains (similar learning gains) using the near-transfer maths skill test correspond with the results of the comparison of the participants using their in-game skill scores, which showed no significant difference for the three tasks types which were covered by the maths test; the details can be found in Table 3.9.

Even the self-reported perceived learning gain reported in the final questionnaire, that the children could evaluate using 0 to 5 points according to how much they thought they had learnt, was remarkably similar (Cliff's δ = 0.03) for the narrative frame (mean = 4.09, SD = 1.14, n = 22) and the rich story group (mean = 4.05, SD = 1.16, n = 20). When comparing the self-reported perceived learning gain of the control group (mean = 2.67, SD = 1.36, n = 15) with the narrative condition groups it is possible to classify the effect size for both the narrative frame (Cliff's δ = 0.55) and the rich story group (Cliff's δ = 0.56) as large.

Correlations		Hungry Plant skill reached	Number Web skill reached	Weighing Machine skill reached
pretest	frame	0.26	0.73****	0.58*
z-score	rich	0.23	0.43	0.35
posttest	frame	0.41	0.77****	0.68***
z-score	rich	0.24	0.44*	0.33
post-pre	frame	0.37	0.21	0.28
difference	rich	-0.02	0.04	-0.06
Hungry Plant	frame		0.39	0.40
skill reached	rich		0.11	0.19
Number Web	frame			0.77****
skill reached	rich			0.51*

Table 3.10 A correlation table for the test scores and the in-game skill scores for the three main task types. All of the skill scores correlated positively with the posttest score, but the linear relationship was significant for both groups only in the Number Web task skill score. *Note: p < .0001 '****', p < .001 '***', p < .01 '**', p < .05 '*'*

3.5 Players & the Story (Exploration Goal)

Examination of the logs from the introductory comic in the rich story version was the best starting point in trying to uncover interesting information about story consumption as all the children from the rich story group (n = 22)[1] had a chance to go through the initial narrative. The game allows the player to revisit already unlocked scenes (chapters and levels) so the number of visits was clearly relevant.

Adding up all the visits to the four introductory chapters that were longer than 20 seconds for each player, it was found that about half of the children went through the introduction only once; the median number of visits was 4 (equal to the number of introductory story

[1] The number of samples for the data analysis differed by 1 or 2 data points due to technical issues that occurred in some cases; it never dropped below 20 data points.

scenes), the average was higher (mean = 6.57, SD = 4.39). About 30 % of the players paid at least 8 visits to one of the introductory scenes. Further, it was worth to explore the estimated time spent on the story scenes, which strongly correlated with the number of visits ($r = 0.89$, $p < .001$).

Intro-story-related measure	Min	Max	1st Qu	Med	3rd Qu.	Mean	SD
n. of visits to intro scenes	2	19	4	4	8	6.57	4.39
n. of diff. days intro was visited	1	4	1	1	3	1.81	1.12
time on intro scenes (minutes)	2.32	38.1	8.92	11.33	19.47	15.64	10.63

Table 3.11 Measures gathered to examine the consumption of the game's story introduction by the rich story group. Each of the four introductory story scenes was usually visited at least once. The audio of the narration is approximately 8 minutes long - a median of 11.33 minutes suggests that at least half of the children went through the story only once.

The audio in the introductory story comic is 8 minutes in total, but the player can speed up[2] or look at the pictures longer and spent more time with it. The median number of minutes spent in total on the introduction (11.33 min) did not exceed the length of the audio by a lot, but the mean (15.64 min) and especially the standard deviation (10.63 min) show that the time spent on the story chapters varied. About 30 % of participants spent at least twice more time (more than 16 min) on the introduction chapters than the actual length of the audio, mostly by revisiting the scenes. In three cases the children spent more than 35 minutes on the initial part of the story throughout all their visits (12, 14, and 19 visits)[3]. It would be interesting, with more data, to see whether this was a distinct group of children who like to come back to the story in general. Only one player spent less than six minutes on the introduction (skipping the comic frames extensively).

Examining correlations helped to see whether there might be any relationship between estimated time spent on the introductory story chapters (further referred as 'intro time') and

[2] A swipe stops the current line of the script, allowing the player to skip to the next frame.

[3] To ensure that the number was not a mistake in the estimation of time spent on the story from the game logs, the logs were checked manually. One of the estimates may have seemed suspicious as the player once spent 12 minutes on a scene for which the median was 2 minutes, but the player continued playing afterwards - it was not an end to a session - and even after subtracting these 12 minutes, the player still spent more than 23 minutes on the introductory scenes (twice more than the group median of 11.33). The rest of the time estimates for these three players seem to be correct - the values accumulated during multiple visits to the story scenes over the course of three or four different days. These three players did not spend more than five minutes on each of the scenes with the exception of the one aforementioned case.

the player's preferences, maths skill, way of playing or learning gains. The pretest score, representing a subset of maths skills at the start of the experiment, correlated neither strongly nor significantly with the intro time (r = -0.15, p = .523). Surprisingly, the intro time did not correlate significantly with the story preference score either (r = 0.18, p = 0.451). It also does not seem to have had a strong relationship with maths being the participant's favourite subject (r = -0.27, p = .24). However, there was a significant negative linear relationship with a medium to large effect size found between intro time and playing video games being the participant's favourite activity (r = -0.49, p = .023). The top four "story consumers" stated in their questionnaire that they do not play video games despite having access to a device (two of them having one solely to themselves).

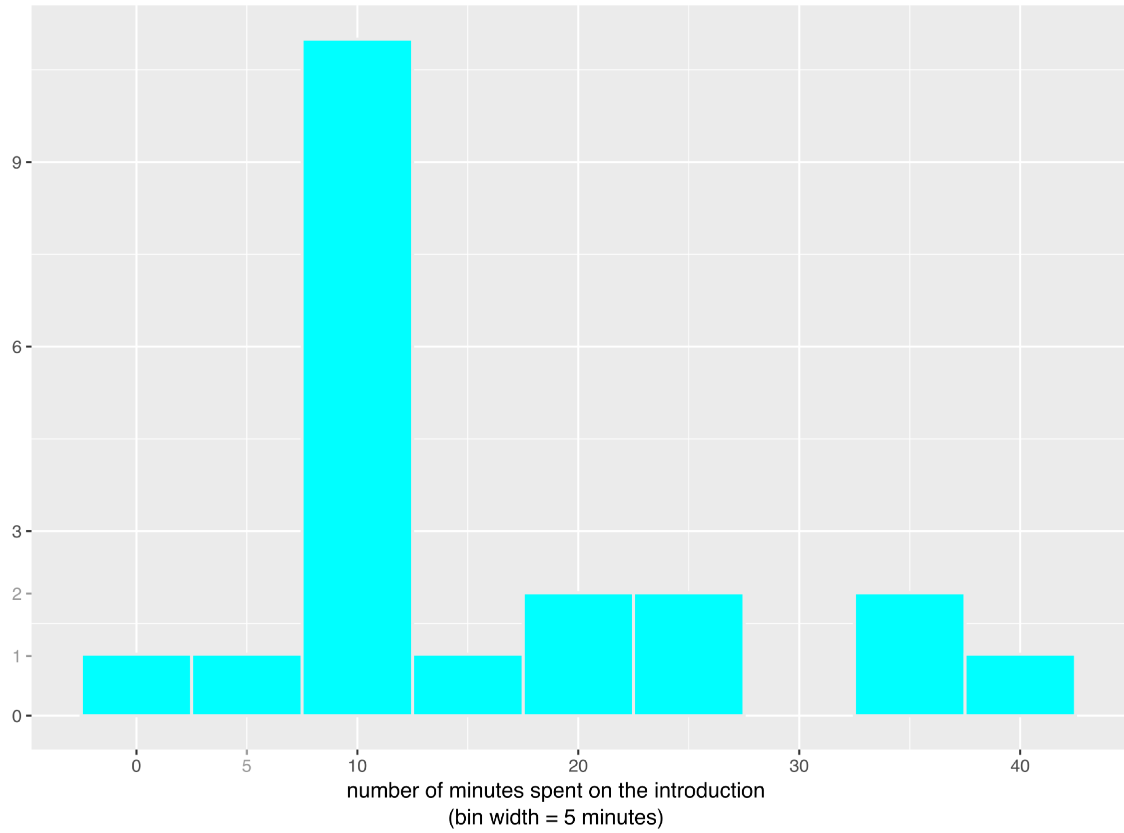


Fig. 3.3 A histogram showing how many players spent a certain amount of time on the story introduction. Although a clear majority spent between 5 to 15 minutes on the story, there is a group of players who spent from 20 up to about 40 minutes on the story (always over multiple visits).

Correlations with the collected in-game data revealed that those players who spent more time on the introductory story chapters were more likely to revisit and replay already completed game levels - the linear relationship was significant with a large effect size (r = 0.72, p < 0.001). It seems that some children did not only like to go through the story again but also prefered to revisit some game levels. The players who spent more time on the introductory story were also often revealed to be the ones whose parents reported higher engagement (r = 0.45, p = .043); significant with a medium effect size. No significant relationship was found between intro time and the level reached (r = -0.22, p = .341) or the number of tasks solved (r = 0.03, p = .898). Last but not least, there was a positive significant linear relationship with a medium effect size found between the estimated time spent on the introductory story parts and the learning gains measured by the difference between the posttest and the pretest (r = 0.46, p = .037).

Correlations	estimated time spent on story intro scenes
number of visits to intro story scenes	0.89****
child-reported story preference score	-0.04
child's favourite subject is maths	-0.27
child likes to play video games	-0.49*
parent-reported child engagement score	0.45
number of revisits to game levels	0.72***
level reached	-0.22
number of tasks solved	0.03
pretest z-score	-0.15
posttest z-score	0.14
post-pre test score difference	0.46*

Table 3.12 The analysis of correlations between the estimated time spent on the story introduction with other variables revealed a very strong and significant linear relationship with the number of visits to the story scenes and also the number of revisits to the game's levels. This suggests that the players were not only coming back to the story, they were returning to replay other parts of the game as well.
*Note: p < .0001 '****'; p < .001 '***', p < .01 '**', p < .05 '*'*

Six players from the rich story group (28.6 %) finished the game, reaching the final parts of the story, and all but one of them went through all the three final chapters. It is worth mentioning that after reaching the end of the story, they did not tend to return to the levels to solve more tasks, unlike some players from the narrative frame group. The data set is small, but the six players from the rich story group solved altogether only two tasks after finishing the game's storyline; on the other hand, out of the seven players from the narrative frame group, four solved more than ten tasks (two even more than 40) after opening the very short story ending in their version. It may be interesting to examine, with a larger data set, whether a more conclusive story ending (which potentially creates an impression that the game is truly finished) demotivates players and therefore stops them from returning back to the game[4].

3.6 Further Exploration

3.6.1 How the Game Was Played Within the Families

Data from the game logs and the questionnaires from both narrative condition groups (n = 48) provide material for a brief description of where, when, and how the game was played. All the parents reported that their children played at home, and 1/4 of the parents reported that their children also played while travelling (the majority in a car). The children usually played in the afternoon from around five o'clock till eight o'clock on workdays, but the span was longer at the weekends - from eleven o'clock in the morning till half past seven in the evening. An average playing session lasted 23 minutes (SD = 21 min), the first quartile was 8 minutes and the third quartile 33 minutes. Only about 12.5 % of the parents set a strict time limit (from 20 to 60 minutes), others limited the playing session "when it got too long". The children mostly played alone, but 1/5 of the parents did report that they played with their children. 33 % of the children approached their parents for advice often, 60 % did so from time to time, and about 6 % were reported not to have asked for advice at all. There was no big difference between the narrative frame and the rich story group in terms of asking for advice (Cliff's δ = 0.15).

3.6.2 The Children's Opinions About the Game Tasks

Out of the 48 children, 41 answered questions about the game tasks, indicating how they perceived them. There were all together seven task types (listed in subsection 2.4.2). According to roughly 34 % of the children, the easiest task was the Hungry Plant (simple division), which was a task also included in the near-transfer maths skill test. About 24 % rated the Weighing Machine (simple equations), another task from the maths skill test, as the easiest. The third task type covered by the maths skill test - the Number Web - was rated as the most difficult task by the majority of the children (61 %). Interestingly, even though the Number Web task occured in the game three times more frequently than the Hungry Plant or the Weighing Machine and it was rated as the most difficult task, the children did not find it to be the most boring task (only 15 % did). There was no remarkable agreement about the most boring task - the "first place" was shared by the Hungry Plant and the Flying Platform

[4]The experiment was limited to two weeks - the question of whether the participants return to the game even after they have completed it should ideally be studied over a longer timespan.

(algorithmic thinking) with 17 % of votes. The majority of the 61 % participants chose the Music Bridge (memory and patterns training), the scarcest task type in the game (8 out of 163), as their favourite task. The Weighing Machine was voted the second most favourite task by 22 % of the children.

Chapter 4

Discussion

4.1 Initial Notes

This RCT value-added user-controlled study was conducted mainly in order to examine whether a richer in-game story has more influence on engagement and learning gains than a simple narrative frame. The consumption of the narrative in the rich story group was also explored in an attempt to uncover differences among player approaches. There is a section in this chapter dedicated to each of these key topics, and the limitations of the study are discussed immediately after. Before proceeding to the discussion, it should be noted that the experimental groups were not found to be significantly different from each other in terms of participant characteristics and background. Another crucial fact is that even though the children, under the supervision of their parents, were free to decide where and when to play the game themselves, they did indeed play it over the course of the two experimental weeks. The participants from the two narrative condition groups played, on average, for more than six hours (not including the story parts) and solved around 140 maths and logic tasks.

4.2 Discussing Engagement and Progress Through Game

The first hypothesis, which was based on previous studies focused on narrative educational games (Cordova and Lepper, 1996; Jimenez, 2014; Koenig, 2008; Marsh et al., 2011; McQuiggan et al., 2008; Wouters et al., 2011), viewed in-game stories as a way to captivate the player's attention, suggesting that the rich story version would be (H1a) more engaging for the players, and the participants in the rich story group would (H1b) advance further in the game and (H1c) solve more tasks than those in the narrative frame group.

The parent-reported child engagement score seemed promising in terms of the hypothetical positive effects of the narrative as the mean score for the rich story group was higher than for the narrative frame group. However, the difference was not significant ($p = .245$), and the effect size would be classified as small to medium ($d = 0.45$). Moreover, child-reported enjoyment was practically the same for both of the narrative condition groups (Cliff's $\delta = 0.02$). These results do not appear to be sufficient enough evidence to rule out the null hypothesis and to accept subhypothesis H1a. Some of the previous studies which reported higher engagement often compared a narrative version of a game to a non-narrative version (Cordova and Lepper, 1996; Koenig, 2008), or to a different instructional method, (McQuiggan et al., 2008), which may have been the reason why the effect of the narrative on engagement was shown to be greater in these cases; interestingly, other studies comparing narrative/non-narrative versions of games, by Echeverría et al. (2012) and Johnson-Glenberg and Megowan-Romanowicz (2017), did not find results that would support a significant difference in engagement. A study by Marsh et al. (2011) which involved multiple narrative versions reported that versions with more story elements were consistently rated better on average in terms of fun and excitement, but the differences were always less than 1 standard deviation. In light of these studies, the fact that the richer story was not proven to have had a substantially greater effect on engagement than the narrative frame in our experiment is not entirely surprising.

The similar engagement results for the two narrative condition groups are further supported by the measures for progress through the game and number of tasks solved in that they were not found to be significantly different. Therefore, it must be concluded that the experiment failed to disprove the null hypothesis in the case of both H1b and H1c. According to these in-game statistics, the narrative frame version performed similarly well in terms of motivating children to play as the rich story version.

The correlations between the participant-reported measures and the in-game measures may raise a question as to why the child engagement score was found to significantly correlate with the level reached ($r = 0.54$, $p < .01$) and the number of tasks solved ($r = 0.57$, $p < .01$) only for the narrative frame group. A possible explanation could be that the correlations for the rich story group were found to be weaker (but still positive) because when the children were interacting with the game, i.e. the parents perceived them as "engaged", some of the activity was dedicated to the story parts rather than only to task-solving and making further progress.

4.3 Discussing Learning Gains

In opposition to the the null hypothesis that the story would not make any difference for learning, the possibility that the outcomes would be tipped in one or the other direction was proposed - that the story may either promote learning through engagement or hinder it due to distraction. These possibilities were suggested based on the CATLM by Moreno and Mayer (2007). That is why hypothesis H2 (alternative to the null) assumed there would be some differences in learning gains between the narrative frame and the rich story group.

To make sure that learning could be assessed, a control group was involved in the experiment that received a "placebo" game instead of the maths educational game. As the ANOVA and the *a posteriori* Tukey test showed, the control group indeed had significantly different learning gains compared to the narrative condition groups; the estimated effect size was classified as large for both of the narrative condition groups compared to the control, and Cohen's d was larger for the narrative frame group (1 > 0.74). The difference in learning gains between the narrative frame and the rich story group indicated by the difference in effect size was not, however, significantly large (post hoc p .691). It is quite possible that the sample size was not big enough to capture the effect. A negative influence of the richer story would be aligned with meta-analytical findings reported by Clark et al. (2016), but no significant difference in learning gains was reported by several previous value-added studies comparing narrative and non-narrative game versions (Echeverría et al., 2012; Johnson-Glenberg and Megowan-Romanowicz, 2017; Koenig, 2008), or multiple narrative versions (Marsh et al., 2011). Jimenez (2014) reported marginally higher learning gains, but only when comparing a version with more story elements with an abstract one; no significant difference was found comparing two narrative versions in his case either. At any rate, the extent of the narrative seems not to play a very important role in learning, particularly in the cases mentioned above and even in the case of our study, where the "richer" story was still quite limited in depth and scope. Excessive generalisation should therefore be avoided.

Other measurements consistently show a small or neutral effect of the richer story on learning gains. Both of the narrative condition groups achieved similar in-game skill scores for the tasks covered by the near-transfer test, and the children's self-reported perceived learning gain was almost equal in the two groups (Cliff's $\delta = 0.03$). No significant difference was found in estimated time on tasks (df = 42.79, t = -0.61, p = .547).

Considering all the measures, the narrative frame version appears to work well as an educational tool for use at home, and the rich story version did not perform significantly worse. Some designers of educational maths games may prefer a narrative frame as the choice reduces the risk of distraction and the time required to complete the game (small or no time spent on a story). Others may choose to include a richer story - probably not to influence the learning of mathematics (lacking the evidence supporting the choice), but for other potential reasons, e.g. in order to fulfill educational objectives in areas like literacy and language learning (Ryokai et al., 2003) in addition to maths, or even for marketing purposes in the case of commercial games[1].

4.4 Exploring the Narrative Feature

Another question asked was how the story was consumed and whether the players' preferences and characteristics played a role. There are some remarks to be made regarding this exploration goal that may be of interest; the data of about 20 participants from the narrative group could indicate more common practices, but a more nuanced and definitive conclusion would require more research.

Most of the players went through the introductory story chapters only once without much skipping: the median time spent on the introduction (11.33) was about 3.5 minutes longer than the length of the audio; a subgroup consisting of 30 % of the players returned to the introductory chapters at least once - some revisited them multiple times, spending up to around 35 minutes solely on the introduction. It seems the story captured their attention. They either revisited it on their own, or perhaps showed it to their family or friends. Those players who spent more time on the introductory story chapters were more likely to revisit and replay already completed game levels - the linear relationship was significant with a large effect size ($r = 0.72$, $p < 0.001$). It is unlikely that this indicates their preference for the story, but rather a different way of consuming the game as a whole as opposed to focusing only on progressing further and further. Another noteworthy detail is the significant negative linear relationship with a medium to large effect size found between the time spent on the introduction and playing video games being the participant's favourite activity ($r = -0.49$, $p = .023$). The top four story consumers were non-gamers as well. They may have returned to the game because of the story, or perhaps they were, as previously suggested, just playing the game differently. Other characteristics, such as a preference for stories (expressed in the questionnaire by most of the children) or maths being the participant's favourite subject (the majority of participants liked maths) did not seem to influence the time spent on the story. In conclusion, the way players consumed the narrative parts was not uniform and the matter is something that may be interesting to investigate further in the future.

The exploration goal focused on the rich story group, but some attention should also be given to the narrative frame version. A speculation worth investigating in future studies is

[1]One of the marketing claims for *Matemág* was that it includes a rich story and an interactive voiced comic (TechSophia s.r.o., 2018).

that the narrative frame version may have actually been more intriguing because it created more tension by not providing enough information to the players; the participants may have kept wondering about why the heroes went to the magical land and what was going to happen at the end of their journey as they played. Moreover, the players' imaginations may have been filling in the blanks of the simple narrative frame the whole way through the game - children are good at inventing own stories after all.

4.5 Limitations

Multiple limitations of the study have been identified that need to be taken into account when interpreting the results. The limitations concern aspects of the experimental design and method, procedure imperfections, limited sample size, technical issues and the use of particular games.

A good starting point for assessment of the experimental design and method would be asking a question if two weeks was a long enough time to identify differences between the conditions. In fact, the study was able to find a difference in learning gains between the control group and the narrative condition groups through the near-transfer maths skill test. The effects of the richer story were probably not remarkably large as we failed to measure them through this particular method. Finally, about 30 % of the participants finished the game, in one case even in only two days, which is one matter that would suggest that a span of 14 days is appropriate.

The fact that the experiment involved classes where participants from different groups were able to meet could certainly be criticised. Unfortunately, ways of recruitment other than approaching public schools were not available due to a limited budget. One possible problem with between-group contamination could be that children from the narrative frame group may have come in contact with the richer story and been influenced by it. After much scrutiny, we believe that total contamination (see the rich story) occurred only in a minority of cases; the children played mostly at home on their own, and the logs from the rich story group serve as evidence that the majority of them went through the story only once; if the general practice had been to show the story to their friends from the narrative frame group, the logs would have shown a greater number of revisits to the introductory story parts. The worst case scenario would have been if all the 30 % of participants who had revisited all the introductory story chapters had done it to show the story to their friends. However, we believe that this is not likely as we paid careful attention to detect contamination between the groups and only two children from the control group (the "placebo" game was clearly different from the other maths game) had to be excluded due to between-group contamination.

Partial contamination may have occurred within classes as some story details may have leaked between the groups. Moreover, one might argue that the effect of multiple classmates playing one game at the same time could have influenced the way the game was played. Admittedly, this is not entirely unlikely, but this "social" influence would have been the same for all of the conditions in any case so any substantial effect caused by the story would have been noticeable.

As has been mentioned before, learning gains were measured by the difference between pretest and posttest scores. The test itself may have certainly been demanding for the second- and third-year pupils; not only in terms of maths, which was the focus, but also in terms of reading and text comprehension as the children were not yet experienced readers. All the children from all the conditions had the option to notify the administrators if they were experiencing problems during the test, including reading difficulties, in which case an administrator would help with the reading (not with the solution in any way). That the children tackled the written instruction well is illustrated by the fact that in the pretest, 99 % of the participants solved at least one equal division task correctly, 81 % solved at least one number web task correctly and 89 % solved at least one weighing machine task correctly. As the three task types repeated throughout the test, the need to read new instructions gradually decreased.

As demonstrated in the results section, the experimental groups were balanced in the key characteristics. Unfortunately, the number of participants in each group differed as some were absent during the final meeting and some were excluded due to contamination. It would have also been more suitable if the experiment had involved a greater total number of participants ($N = 67$); our situation did not allow detection of smaller effect size and moderators for the examined effects of the narrative feature. The participants are unlikely to represent the whole population of the Czech Republic well, considering that the research was conducted in cooperation with schools in the richer regions around the capital city of Prague. Moreover, the invitation to participate in the research contained the information that the children would be playing an educational maths game on a touch device, which could have potentially attracted or discouraged certain participants.

Through the experimental procedure, the research team strove to keep high standards regarding similar treatment of participants from all the schools during the meetings. However, the initial meeting at the first school was, admittedly, imperfect as children from the different groups were not assigned to three administrators at random, but rather according to their group. To avoid making the sample size even smaller, it was decided that the data from the first school would be kept as the administrators were, in general, instructed not to influence the test solutions in any way (therefore reducing the impact of the flaw in the meeting to an acceptable amount).

There were also a few technical issues. Several parents forgot their devices at home or did not remember their password so the installation, which should have been done on the day of the initial meeting, was completed a day later. This minor difference in the number of days which occured in a few cases was not deemed disruptive to the experiment. Three parents also reported that their children got stuck at level 17 of the game due to a bug in the software. As the children had already completed about 3/4 of the game, we partially used the data in some analyses. These data were excluded whenever they would have compromised the analysis (such as in the analysis of the progress of the groups, etc.). Among the technical limitations was also the collection of logs, a wider range of which would have been suitable for a more detailed analysis - for example to log all the touch input. We collected only the most important measures mainly due to time and budget constraints. An error in the code was found in the logs concerning task solutions, which rendered the timestamps for the Flying Platform and the Counting Lift tasks irrelevant. The time on task from these tasks was not included in the analysis.

The choice of this particular game for research and the consequences of it may be subject to scrutiny as well. The commercial version of the game had been published before the start of the experiment; a situation which demanded that prior exposure be checked and data associated with such exposure be excluded from the analysis. Another point could be made that as there was no significant difference found between the narrative frame and the rich story group, the story (mainly the story introduction to which all the rich story group children were exposed) may have been too short; the mean number of minutes spent on the introduction was 15.64 and the shortest time it would take to finish it without skipping was 8 minutes. The response to this issue is that the narrative in the experimental rich story version had the same form as in the commercial version of the educational game so the gaming experience of the participants from the rich story group was very similar to the real-life experience of a common user. A lot of information is conveyed in those 8 minutes of the audio in the introductory interactive comics: it introduces the characters, highlights the usefulness of numbers and maths and explains what motivates the heroes to begin their journey. To put the matter into a perspective - TV series for children often have a similar running time: for example, the infamous animated series *SpongeBob SquarePants* (Hillenburg et al., 1999) has a regular running time of around 11 minutes (Wikipedia contributors, 2018).

The last point to be made as to the limitations of this study is that there is a wide variety of serious games; the findings presented in this thesis are to be associated with a specific game with a specific design of both the educational content and the narrative. Any generalisations are to be made with caution.

Conclusion

At the beginning a question was posed as to whether the narrative feature in educational digital games for young children has positive, neutral, or negative effects where learning is concerned. An objective was further established to assess when and to which extent it is reasonable to use the narrative feature. Studies involving participants older than 12 years of age have reported mostly non-significant differences in educational outcomes in the comparison of two groups where one group received the narrative and the other the non-narrative version of a serious game (Echeverría et al., 2012; Johnson-Glenberg and Megowan-Romanowicz, 2017; Marsh et al., 2011; Wouters et al., 2011). However, two studies with children younger than 12 years of age associated contextualisation provided by a story with a positive influence on learning (Cordova and Lepper, 1996; Jimenez, 2014). Results from our empirical value-added RCT user-controlled study with second- and third-year schoolchildren (N = 67, 8.67 years) who played a maths game over the course of two weeks suggest that the richer story in one game version had a similar effect to the simple narrative frame in the other version. The following paragraphs present the conclusions reached based both on the findings from the experiment (see chapter 3) and the studies reviewed in chapter 1.

For clarity's sake, the conclusions should be prefaced with the overall objective: the general goal is to support the educational process. The following reflections assume that the decision to use a digital game to achieve the goal has been made[2], the text further focuses on the narrative as a possible game feature.

Should a story be employed in an educational game for primary school children? The total majority (95 %) of the children participating in our research who answered the question

[2]It should be clarified that another instructional method may sometimes perform better than a game in terms of fulfilling educational purposes (Adams et al., 2012; McQuiggan et al., 2008). As previously mentioned in the Theoretical Background, Mayer (2016) recommends using games for targeted learning of well-specified objectives in those situations for which there is some evidence grounded in GBL research. Literature reviews may help to identify the promising domains where games have been shown to perform well, such as science or second-language learning (Mayer, 2014).

about their relationship with stories (n = 57)[3] expressed that they like them in general. Marsh et al. (2011) reported that participants (13-14 years old) from their study consistently rated narrative versions of the game used in their study as “more fun, more exciting, and more engaging” (Marsh et al., 2011, p. 18). These examples suggest that using a story in a game may (to some extent) raise attractiveness for the end users. Games outside the educational domain often do well enough with a simple story (in this thesis referred to as the ‘narrative frame’), as illustrated by legendary titles like *Space Invaders* (Taito, 1978) or *Mario Bros* (Nintendo, 1983). Two studies, one by Cordova and Lepper (1996) and the other by Jimenez (2014), in which primary school children took part in playing a maths game speak in favour of the narrative frame. Cordova and Lepper (1996) mention ‘contextualisation’ (which is analogical to the narrative frame) among experimental conditions which produced “dramatic increases, not only in students’ motivation but also in their depth of engagement in learning, the amount they learned in a fixed time period” (Cordova and Lepper, 1996, p. 1). In addition, Jimenez (2014) reported a strong correlation between fun and learning for an experimental group which played a game version with story elements such as characters and interesting events.

The narrative frame seems to have some advantages in educational games for children, but what about a richer story? The *Cognitive-affective theory of learning with media* (Moreno and Mayer, 2007) stresses the limited cognitive capacities of the learner. These capacities may be depleted by the processing of a game’s narrative, which could lead to less-effective learning. The more complex the story, the greater the potential risk. This claim, based on the theory, is also supported by research - the meta-analysis by Clark et al. (2016) reports that the effect size for learning outcomes was significantly larger for games with no story or a thin story depth relative to those with a thicker story.

However, in our research comparing two narrative versions which differed in the depth of the story the difference in learning gains between the narrative frame and the rich story group was not found to be significant ($d = -0.25$, $p = .691$). Both narrative condition groups performed significantly better than the control group. When comparing the effect size for the narrative frame and the rich story game versions to the control "placebo" game it can be commented that Cohen’s d was larger for the narrative frame ($d = 1.00$, $p = 0.004$) than for the rich story ($d = 0.74$, $p = 0.04$).

The study did not find the richer story to be a significant distraction in learning, the negative effect might have been too small to be measured as each of the experimental groups had only about 20 participants. When interpreting the outcome one has to take into account

[3] Seven children did not complete the second questionnaire due to their time limitations. Three children left this particular question unanswered.

several other factors, be it the specific game used, the length of the story compared to the length of the whole game, the length of the whole experiment (two weeks) and other nuances. One of the suggested explanations for a weak distraction effect is that the story parts were not interwoven with the core educational gameplay. Learning gains aside, it could be noted that the richer narrative had small to medium non-significant effect on engagement ($d = 0.45$, $p = .245$).

It seems - considering both our research and the previous studies - that there are potential risks (distraction) and potential benefits (higher engagement) to using the narrative feature. Drawing conclusions from the evidence at hand, it appears that the use of a narrative frame in a maths educational game for children is a safer option than the use of a richer story as the choice reduces the risk of distraction while still maintaining the game's potential to be engaging. The decision is, in the end, primarily the responsibility of the game designer, who should ideally be aware of its consequences. To provide more possibilities for both designers and educators, Pilegard and Mayer (2016) presented a research-proven way to reduce distraction during learning with a narrative serious game. Their findings suggest using additional material such as pre-game and in-game worksheets to focus students' limited cognitive resources on the educational objectives.

The very last remark to be made as to the use of deeper stories within educational games in general involves special cases in which it is more reasonable to consider a richer narrative from the start. These cases occur when the story itself is central to the learning and therefore would not be considered a "distraction". One example would be literacy and language learning as indicated by a study by Ryokai et al. (2003) in which young children improved their skills through free play and interaction with a virtual storyteller. Another such case would be history as demonstrated by the *Attentat 1942* (Charles University and Academy of Sciences of the Czech Republic, 2017) game, in which players explore World War II from the perspective of common citizens through stories, memories and other materials.

References

Abbott, H. P. (2002). *The Cambridge introduction to narrative*. Cambridge University Press.

Abt, C. C. (1987). *Serious games*. University press of America.

Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., and Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology*, 104(1):235–249.

Apple Inc. (2018). iTunes Education Category. https://itunes.apple.com/pw/genre/ios-education/id6017?mt=8. Accessed: 2018-11-29.

Bage, G. (2012). *Narrative matters: Teaching history through story*. Routledge.

Barab, S. A., Sadler, T. D., Heiselt, C., Hickey, D., and Zuiker, S. (2007). Relating Narrative, Inquiry, and Inscriptions: Supporting Consequential Play. *Journal of Science Education and Technology*, 16(1):59–82.

Berkley-Patton, J., Goggin, K., Liston, R., Bradley-Ewing, A., and Neville, S. (2009). Adapting effective narrative-based HIV-prevention interventions to increase minorities' engagement in HIV/AIDS services. *Health communication*, 24(3):199–209.

Bower, G. H. and Clark, M. C. (1969). Narrative stories as mediators for serial learning. *Psychonomic Science*, 14(4):181–182.

Broderbund (1985). Where in the World Is Carmen Sandiego? [PC].

Brom, C. (2017). Learning with Digital Technologies: the Role of Positive Affect and Motivation.

Casey, B., Erkut, S., Ceder, I., and Young, J. M. (2008). Use of a storytelling context to improve girls' and boys' geometry skills in kindergarten. *Journal of Applied Developmental Psychology*, 29(1):29–48.

Charles University and Academy of Sciences of the Czech Republic (2017). Attentat 1942. [Steam].

Clark, D. B., Tanner-Smith, E. E., and Killingsworth, S. S. (2016). Digital Games, Design, and Learning: A Systematic Review and Meta-Analysis. *Review of Educational Research*, 86(1):79–122.

Cliff, N. (1996). Ordinal methods for behavioral data analysis.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences*. Hillsdale, NJ: erlbaum.

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences.* Taylor and Francis, Hoboken. OCLC: 845258623.

Cordova, D. I. and Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of educational psychology*, 88(4):715.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, 16(3):297–334.

Duveskog, M., Kemppainen, K., Bednarik, R., and Sutinen, E. (2009). Designing a story-based platform for HIV and AIDS counseling with Tanzanian children. In *Proceedings of the 8th International Conference on Interaction Design and Children*, pages 27–35. ACM.

Echeverría, A., Barrios, E., Nussbaum, M., Améstica, M., and Leclerc, S. (2012). The atomic intrinsic integration approach: A structured methodology for the design of games for the conceptual understanding of physics. *Computers & Education*, 59(2):806–816.

Feitelson, D., Goldstein, Z., Iraqi, J., and Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, pages 71–79.

Fisher, R. A. (2006). *Statistical methods for research workers*. Genesis Publishing Pvt Ltd.

Fullerton, T. (2014). *Game design workshop: a playcentric approach to creating innovative games*. AK Peters/CRC Press.

Genette, G., Sheridan, A., and Logan, M.-R. (1982). *Figures of literary discourse*. Columbia University Press New York.

Gillam, S. L., Gillam, R. B., and Reece, K. (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, speech, and hearing services in schools*, 43(3):276–291.

Google LLC (2018). Google Play Educational Games. https://play.google.com/store/apps/category/GAME_EDUCATIONAL. Accessed: 2018-11-29.

Graesser, A. C. and Ottati, V. (1995). Why stories? Some evidence, questions, and challenges. *Knowledge and memory: The real story*, 8:121–132.

Hainey, T., Connolly, T. M., Boyle, E. A., Wilson, A., and Razak, A. (2016). A systematic literature review of games-based learning empirical evidence in primary education. *Computers & Education*, 102:202–223.

Halpern, D., Graesser, A., and Hakel, M. (2007). Learning principles to guide pedagogy and the design of learning environments. *Washington, DC: Association of Psychological Science Taskforce on Lifelong Learning at Work and at Home*.

Harp, S. F. and Mayer, R. E. (1998). How Seductive Details Do Their Damage: A Theory of Cognitive Interest in Science Learning. page 21.

Hejný, M. (2008). Scheme–oriented educational strategy in mathematics. *Supporting Independent Thinking Through Mathematical Education, Wy. Univ. Rzeszow*, pages 40–48.

Hejný, M., Jirotková, D., and Slezáková-Kratochvílová, J. (2008). *Matematika pro 2. ročník základní školy*. Fraus.

Hejný, M., Jirotková, D., Slezáková-Kratochvílová, J., and Michnová, J. (2009). *Matematika pro 3. ročník základní školy*. Fraus.

Hillenburg, S., Drymon, D., Hill, T., and LeFrancis, L. (1999). Help wanted.

Hudson, M. E. and Test, D. W. (2011). Evaluating the evidence base of shared story reading to promote literacy for students with extensive support needs. *Research and Practice for Persons with Severe Disabilities*, 36(1-2):34–45.

Hunicke, R., LeBlanc, M., and Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, volume 4, page 1722.

Jimenez, O. (2014). Reflecting on educational game design principles via empirical methods. Boulder, CO: International Society of the Learning Sciences.

Johnson-Glenberg, M. C. and Megowan-Romanowicz, C. (2017). Embodied science and mixed reality: How gesture and motion capture affect physics education. *Cognitive Research: Principles and Implications*, 2(1).

Juul, J. (1999). A clash between game and narrative. *Danish literature*.

Koenig, A. D. (2008). *Exploring effective educational video game design: The interplay between narrative and game-schema construction*. Arizona State University.

Lacey, S., Stilla, R., and Sathian, K. (2012). Metaphorically feeling: comprehending textural metaphors activates somatosensory cortex. *Brain and language*, 120(3):416–421.

Landers, R. N. and Landers, A. K. (2014). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, 45(6):769–785.

Lee, K. M., Park, N., and Jin, S.-A. (2006). Narrative and Interactivity in Computer Games.

Magnavox (1972). Magnavox Odyssey. [Magnavox Game Card].

Magnavox (1973). Magnavox Odyssey TV Commercial. [TV].

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual review of psychology*, 62:103–134.

Mares, M.-L. and Pan, Z. (2013). Effects of Sesame Street: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology*, 34(3):140–151.

Marsh, T., Nickole, L. Z., Klopfer, E., Xuejin, C., Osterweil, S., and Haas, J. (2011). Fun and Learning: Blending Design and Development Dimensions in Serious Games through Narrative and Characters. In Ma, M., Oikonomou, A., and Jain, L. C., editors, *Serious Games and Edutainment Applications*, pages 273–288. Springer London, London.

Mayer, R. E. (2011). Multimedia learning and games.

Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. MIT Press.

Mayer, R. E. (2016). What Should Be the Role of Computer Games in Education? *Policy Insights from the Behavioral and Brain Sciences*, 3(1):20–26.

McKillup, S. (2011). *Statistics explained: an introductory guide for life scientists*. Cambridge University Press.

McQuiggan, S. W., Rowe, J. P., Lee, S., and Lester, J. C. (2008). Story-Based Learning: The Impact of Narrative on Learning Experiences and Outcomes. In Woolf, B. P., Aïmeur, E., Nkambou, R., and Lajoie, S., editors, *Intelligent Tutoring Systems*, volume 5091, pages 530–539. Springer Berlin Heidelberg, Berlin, Heidelberg.

Microsoft Corporation (2018). Microsoft Store Most Popular Educational Games. https://www.microsoft.com/en-us/store/most-popular/games/mobile?category=Educational. Accessed: 2018-11-29.

Moreno, R. and Mayer, R. (2007). Interactive Multimodal Learning Environments: Special Issue on Interactive Learning Environments: Contemporary Issues and Trends. *Educational Psychology Review*, 19(3):309–326.

Morrow, L. M., O'Connor, E. M., and Smith, J. K. (1990). Effects of a story reading program on the literacy development of at-risk kindergarten children. *Journal of Reading Behavior*, 22(3):255–275.

Nintendo (1983). Mario Bros. [Arcade].

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302):157–175.

Pilegard, C. and Mayer, R. E. (2016). Improving academic learning from computer-based narrative games. *Contemporary Educational Psychology*, 44-45:12–20.

Prince, G. (1982). *Narratology: the form and functioning of narrative*. Number 108 in Janua linguarum. Mouton, Berlin ; New York.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Richardson, B. (2000). Recent concepts of narrative and the narratives of narrative theory. *Style*, 34(2):168–175.

Romano, J., Kromrey, J. D., Coraggio, J., and Skowronek, J. (2006). Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen'sd for evaluating group differences on the NSSE and other surveys. In *annual meeting of the Florida Association of Institutional Research*, pages 1–33.

Rowe, J. P., Shores, L. R., Mott, B. W., and Lester, J. C. (2011). Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments. *International Journal of Artificial Intelligence in Education*, (1-2):115–133.

Rubin, D. C. and others (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press on Demand.

Ryan, M.-L. (2007). Toward a definition of narrative. In Herman, D., editor, *The Cambridge Companion to Narrative*, pages 22–36. Cambridge University Press, Cambridge.

Ryokai, K., Vaucelle, C., and Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of computer assisted learning*, 19(2):195–208.

Sawilowsky, S. S. (2009). New effect size rules of thumb.

Schrader, C. and Bastiaens, T. J. (2012). The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. *Computers in Human Behavior*, 28(2):648–658.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.

Sitzmann, T. (2011). A META-ANALYTIC EXAMINATION OF THE INSTRUCTIONAL EFFECTIVENESS OF COMPUTER-BASED SIMULATION GAMES. *Personnel Psychology*, 64(2):489–528.

Stephens, G. J., Silbert, L. J., and Hasson, U. (2010). Speaker–listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences*, 107(32):14425–14430.

Student (1908). The probable error of a mean. *Biometrika*, pages 1–25.

Taito (1978). Space Invaders. [Arcade].

Tappan, M. and Brown, L. M. (1989). Stories told and lessons learned: Toward a narrative approach to moral development and moral education. *Harvard Educational Review*, 59(2):182–206.

TechSophia s.r.o. (2017a). Matemág: Dobrodružství matemagie. [Mobile].

TechSophia s.r.o. (2017b). Spojovačky. [Mobile].

TechSophia s.r.o. (2018). Matemág - edukativní hra pro děti.

The Learning Company (2015). Carmen Sandiego Returns. [PC, Mobile].

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., and Wright, M. (2006). Computer Gaming and Interactive Simulations for Learning: A Meta-Analysis. *Journal of Educational Computing Research*, 34(3):229–243.

Wikipedia contributors (2018). Spongebob squarepants — Wikipedia, the free encyclopedia. [Online; accessed 16-December-2018].

Wouters, P., Oostendorp, H. v., Boonekamp, R., and Spek, E. v. d. (2011). The role of Game Discourse Analysis and curiosity in creating engaging and effective serious games by implementing a back story and foreshadowing. *Interacting with Computers*, 23(4):329–336.

Wouters, P., Van der Spek, E. D., and Van Oostendorp, H. (2009). Current practices in serious game research: A review from a learning outcomes perspective. In *Games-based learning advancements for multi-sensory human computer interfaces: techniques and effective practices*, pages 232–250. IGI Global.

Wouters, P., van Nimwegen, C., van Oostendorp, H., and van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2):249–265.

Wouters, P. and van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, 60(1):412–425.

Šisler, V. (2016). Contested Memories of War in Czechoslovakia 38-89: Assassination: Designing a Serious Game on Contemporary History. *Game Studies*, 16(2).

Appendix 1: The Story of Matemág

First Story Chapter

The siblings Jacob and Theresa (about 10 years old) and their parents are arriving at their grandparents' summer house, they are on summer vacation. When saying goodbye, their dad reminds the children to study, whereas their mother tells them to mainly learn by playing. When the parents are gone, Jacob and Theresa are meeting their grandmother's cats Micka and Packal. Then the children start to unpack in their room in the attic. Theresa is examining a bookshelf, taking a closer look at different books. She discovers an old, heavy book about a mighty magician called "Matemág" and reads out loud about him and his magnificent land. Jacob is thrilled by the magician's powers and proclaims that it would be great if such a magician could make all numbers disappear so the kids would not have to study math. Suddenly, the book starts to glow and strange things are about to happen.

Second Story Chapter

A wizard appears in a pillar of light shining from the book, the children are flabbergasted. Matemág introduces himself and suggests to first try to imagine what might happen if all numbers disappeared. He creates an illusion in which Jacob and Theresa see their grandma in a discussion with the postman. The adults are complaining about their miseries - missing numbers on postcards, houses, the oven, in shops. Furthermore, they vividly describe how the world outside is in chaos - cars, trains, airplanes, phones and computers either don't work at all or are malfunctioning. That is enough of an illustration for Jacob and his sister and they understand that the wish to make the numbers disappear was foolish.

Third Story Chapter

Matemág agrees that it would be better if the numbers stayed. He invites the kids into his Land of Abstraction and Imagination (Abima) in case they would like to learn "mathmagic" which would allow them to imagine possible futures and invent things. Then the wizard jumps into the book and disappears. Theresa and Jacob decide to follow him as they have some time left before lunch. As they prepare to jump as well, Packal, one of the cats, makes a leap into the book before them. The children follow him an instant later.

Fourth Story Chapter

The heroes appear on a monolith surrounded by a beautiful landscape of floating islands with rocks, trees, meadows and funny animals. Matemág's castle is visible in far distance and the children conclude that they first need to find a way to the castle; so their journey begins.

Game Levels

The heroes are on their journey through the magical land solving a diversity of tasks which allow them to overcome obstacles and continue on their path.

Final Part

Fifth Story Chapter

Finally, after a long journey (23 levels, 163 puzzles), Jacob and Theresa come to a tall tower in which the wizard dwells. But they are facing a problem: There is no entrance, no way how to get up into Matemág's laboratory. They get an advice from a magic hat they had found on the way. It reveals that in Abima, land of fantasy, whenever someone can imagine something in perfect detail, he or she may create the object with a thought. The heroes start to brainstorm ideas how to scale the tower. At first they suggest an elevator but they do not know how it works. Then they are coming up with either a simple staircase or balloons attached to a basket. The player can choose between those options and needs to solve a final task.

Sixth Story Chapter

Matemág is awaiting Theresa and Jacob in his tower offering them a cup of tea. When the kids bring up that they came to learn about mathmagic, the wizard replies: "You already learnt some mathmagic, otherwise you could not have made it here." He suggests that they can stay for some experiments, discover the laboratory as long as they want to. But the siblings suddenly realize that their granny might be searching for them already in the real world. They promise to return and Matemág provides them with a book which transports them back to their grandparents' house.

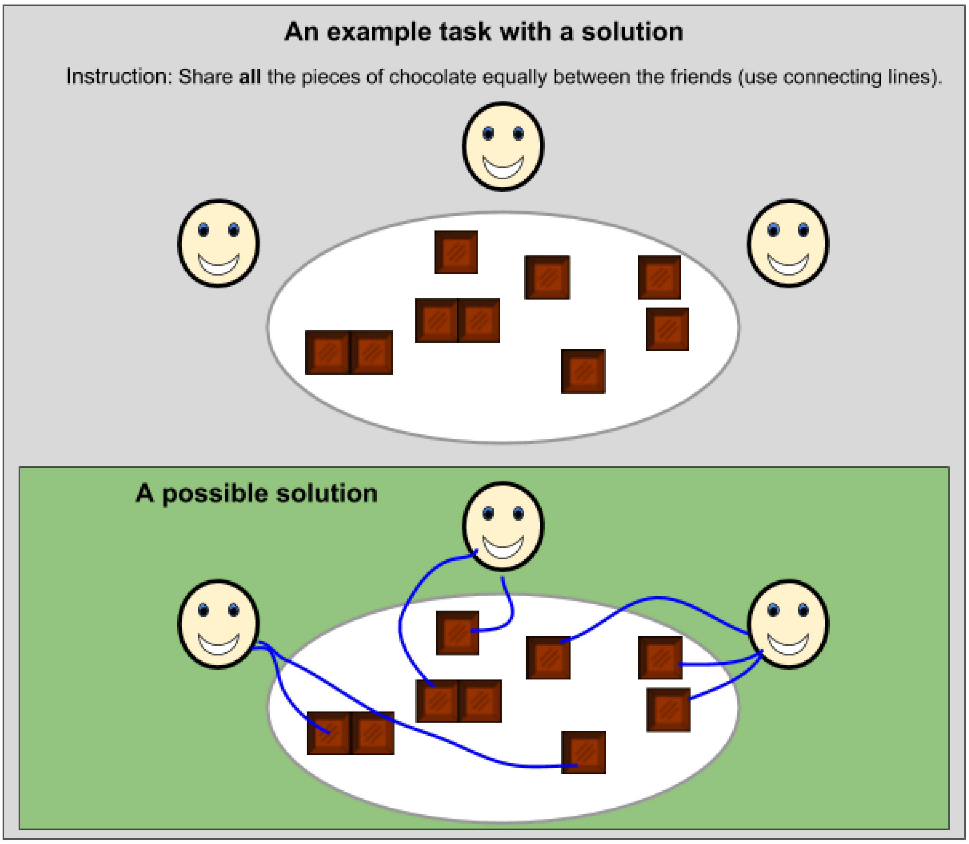
Seventh Story Chapter

After returning from Abima, the two heroes are wondering if they actually learnt anything at all. Their discussion is interrupted by the grandmother who presents them with a real-life problem: Micka the cat was looking for her friend Packal who went to Abima with Jacob and Theresa. As she jumped from a tree onto the roof of a shed, she is now stuck up there and too scared to jump down again. The kids' grandfather is away with the ladder. Jacob and Theresa accept the challenge and after analysing the situation, they decide to build some sort of stairs out of boxes and other things they find in the shed. The children rescue Micka and realize that their whole journey was about training of problem-solving and this is what they learnt.

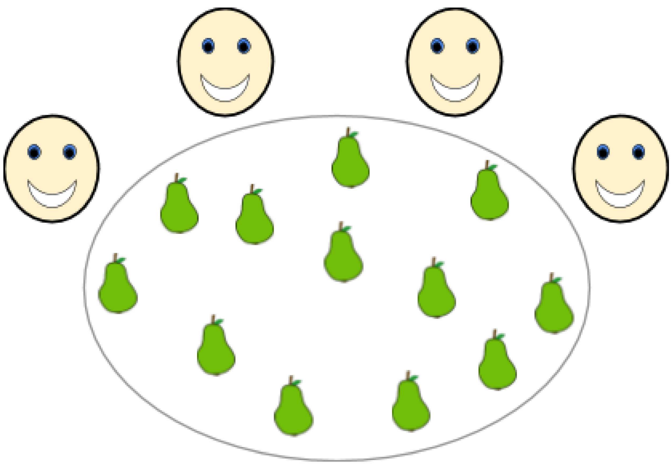
Appendix 2: Near-transfer Maths Skill Test

Test Set 1 A

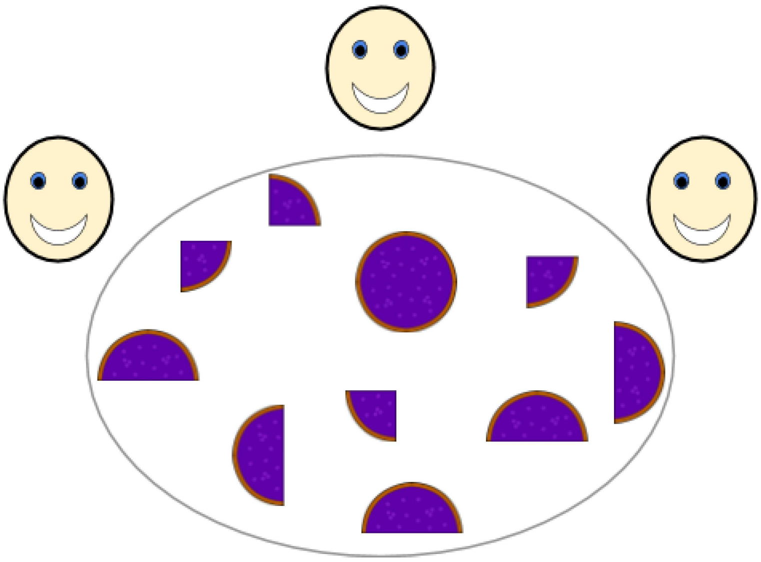
Friends at a table



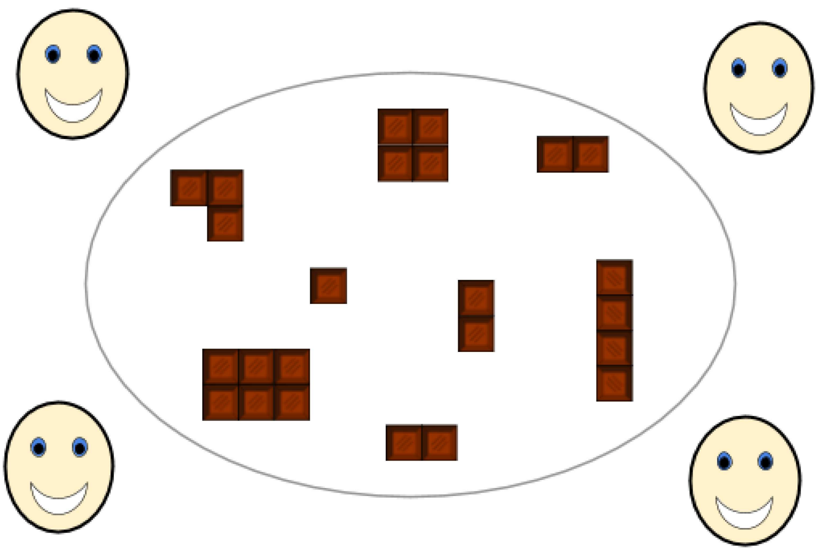
1) **Task**: Share **all** pears equally between the friends.



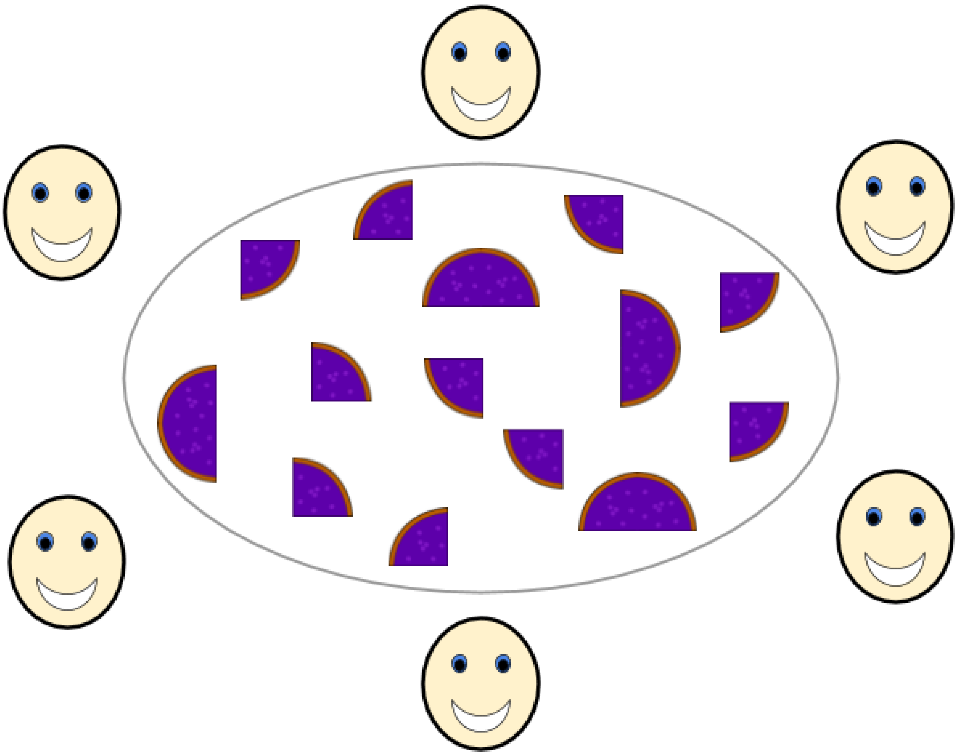
2) **Task**: Share **all** the slices of pie equally between the friends.



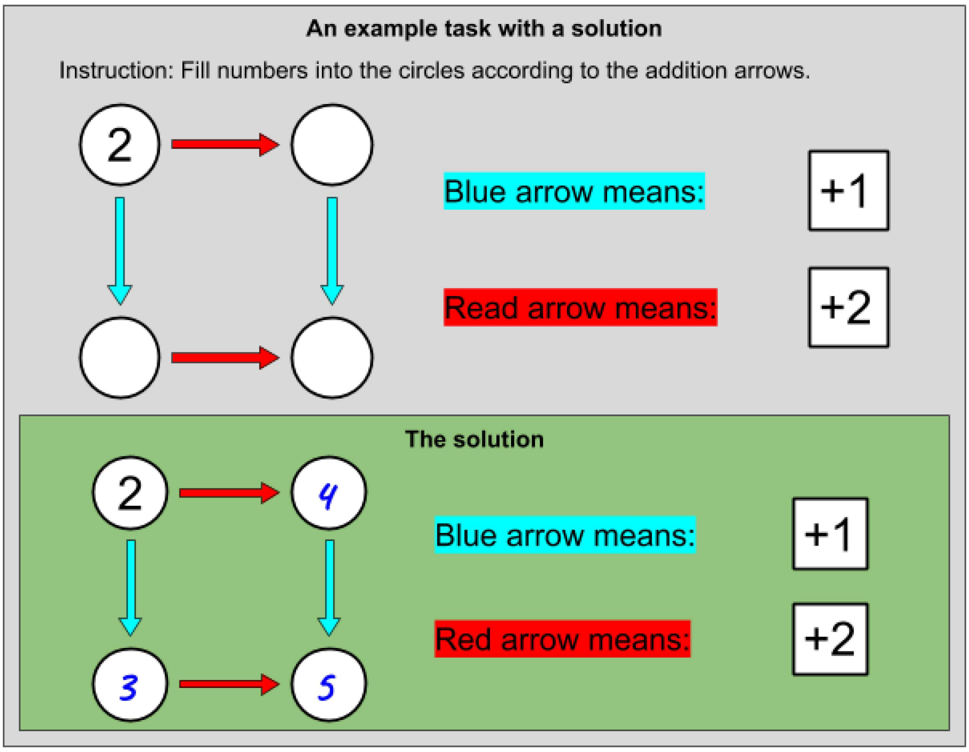
3) **Task**: Share **all** the pieces of chocolate equally between the friends.



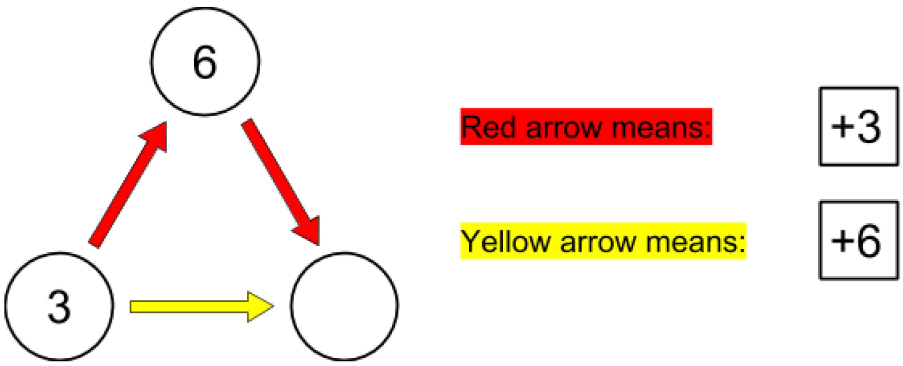
4) **Task**: Share **all** the slices of pie equally between the friends.



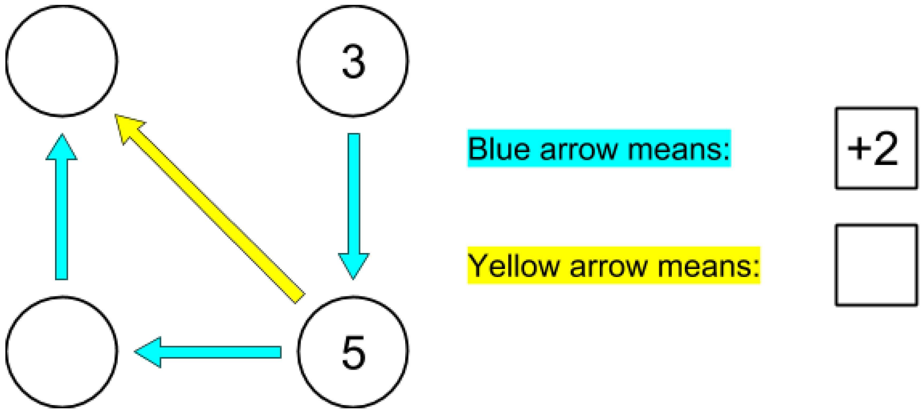
Addition arrows



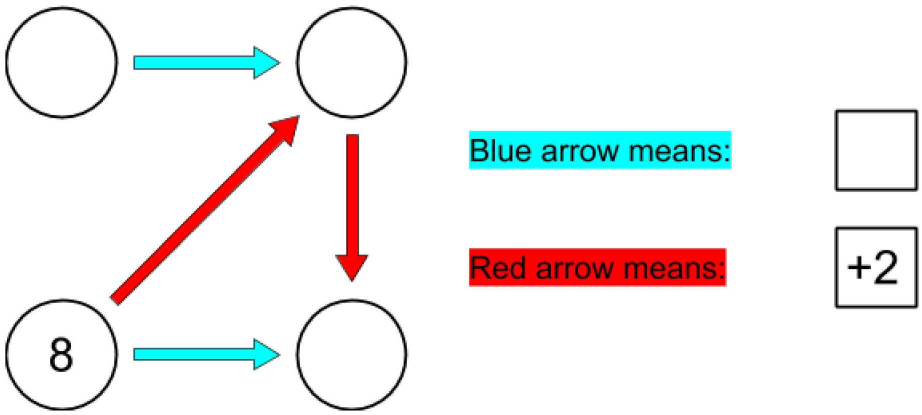
1) **Task**: Fill numbers into the circles according to the addition arrows.



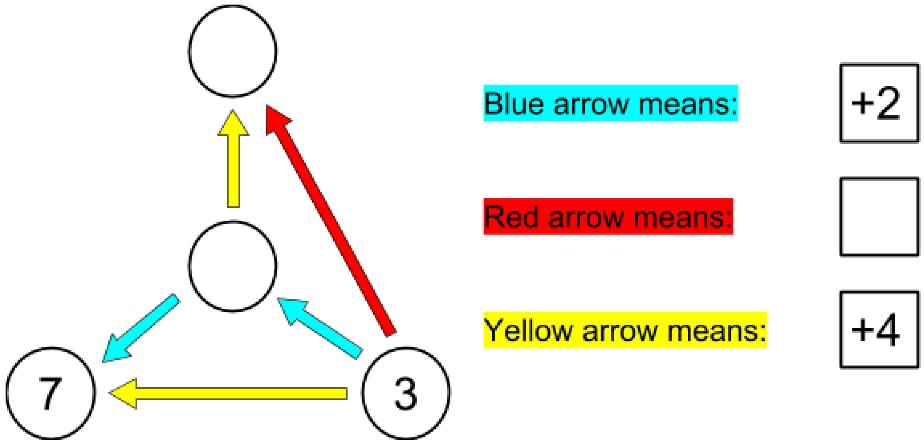
2) **Task**: Fill numbers into the circles according to the addition arrows. Fill as well the value of the yellow arrow into the square.



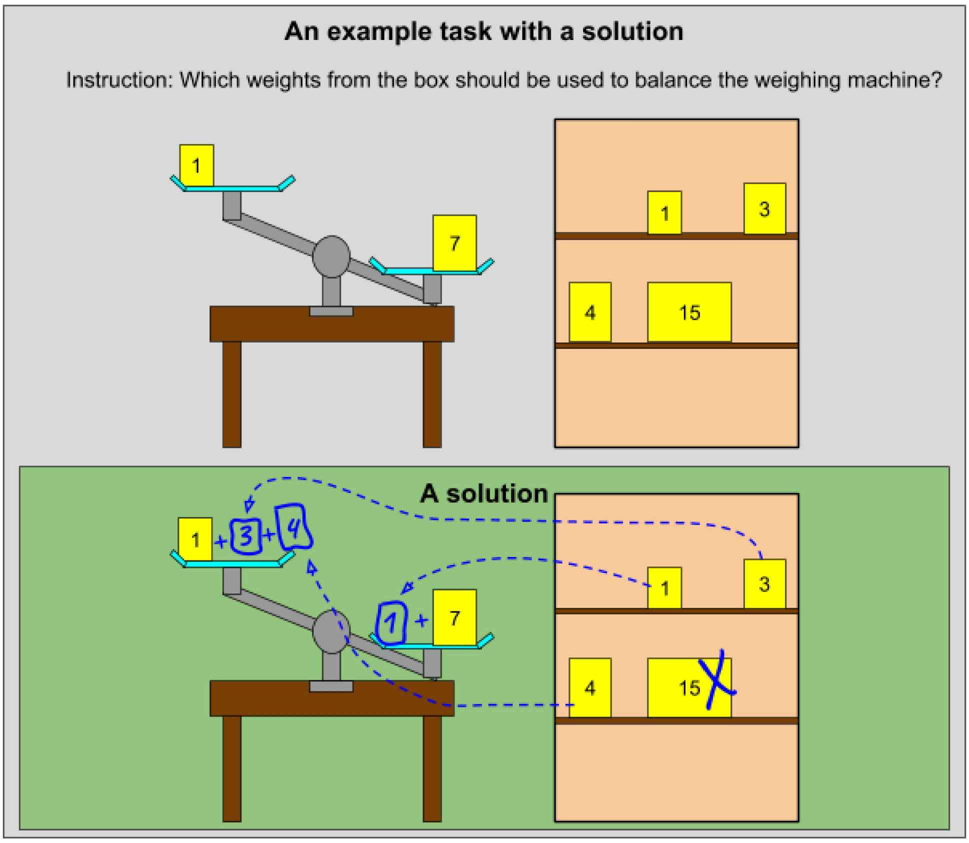
3) **Task**: Fill the numbers into the circles and square.



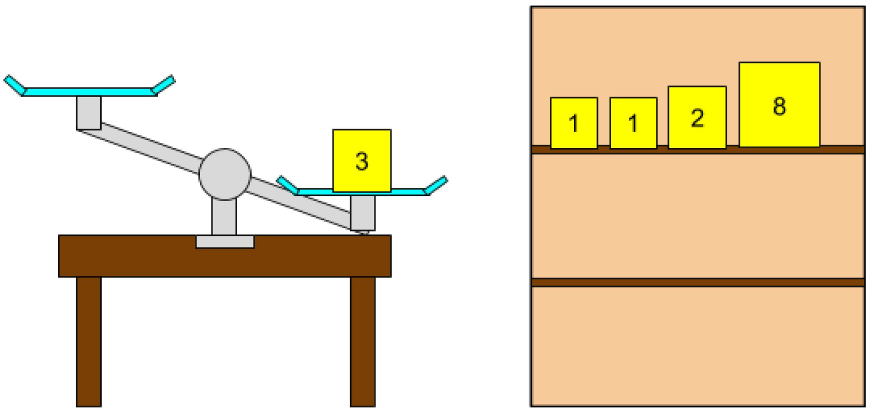
4) **Task**: Fill the numbers into the circles and the squares.



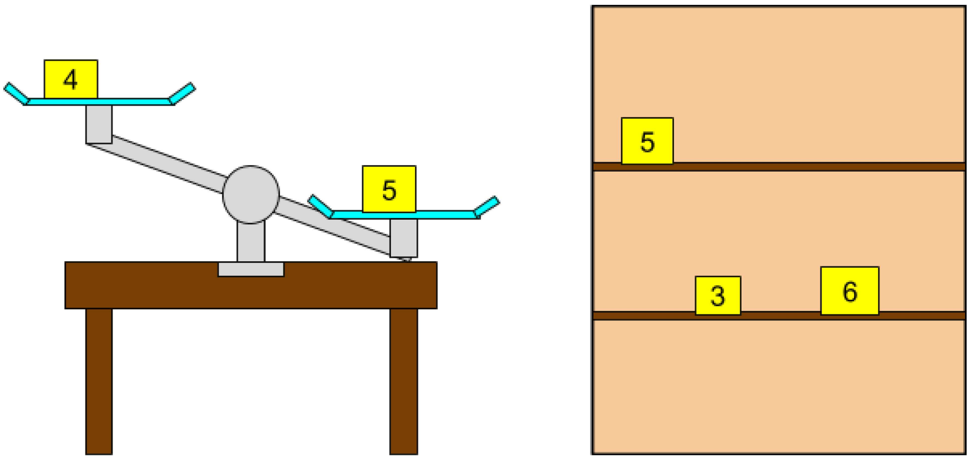
Weighing machine



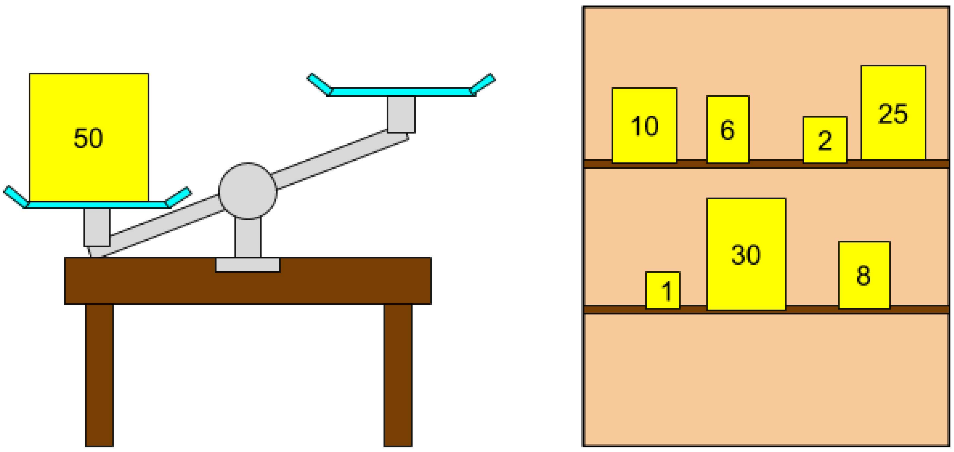
1) **Task**: Which weights from the box should be used to balance the weighing machine?



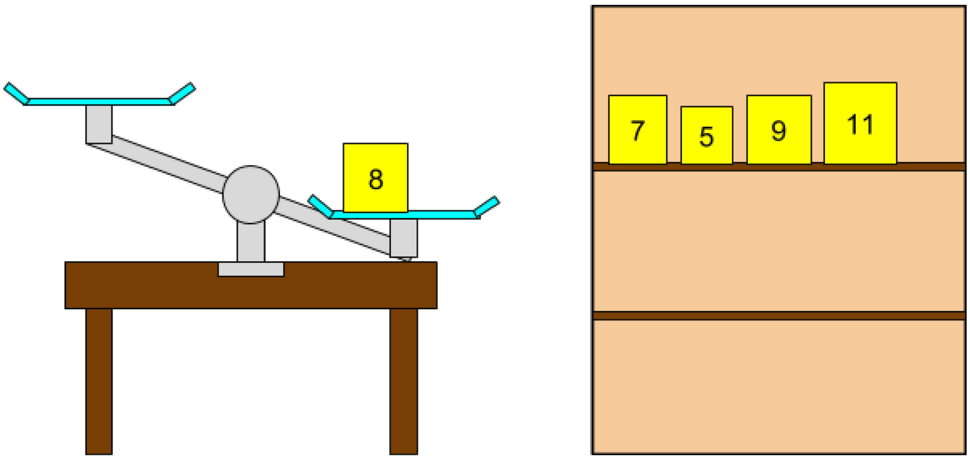
2) **Task**: Which weights from the box should be used to balance the weighing machine?



3) **Task**: Which weights from the box should be used to balance the weighing machine?

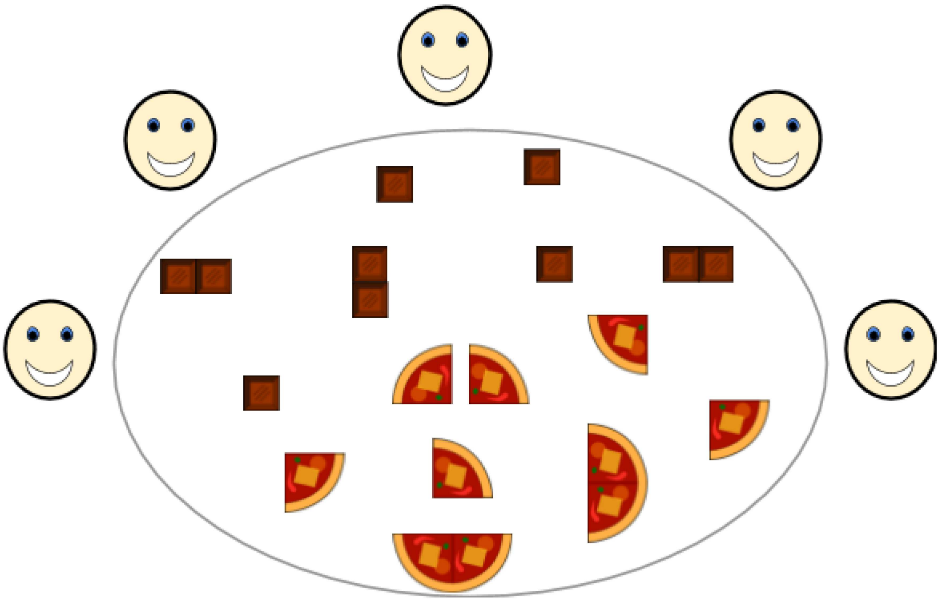


4) **Task**: Which weights from the box should be used to balance the weighing machine?

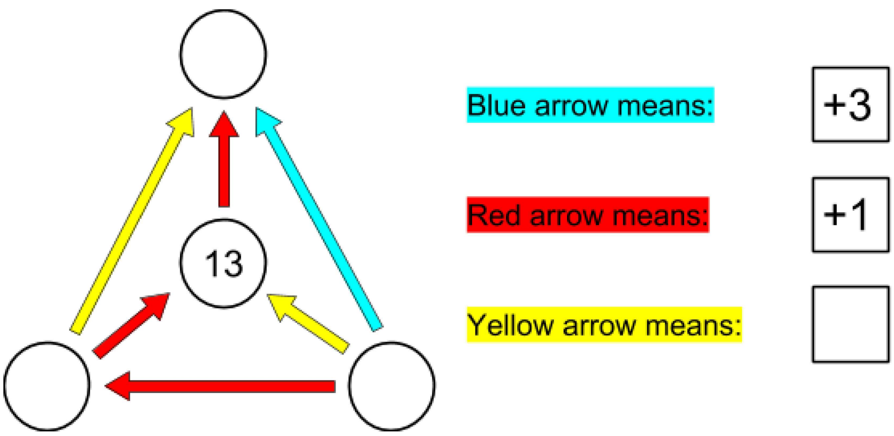


Test Set 2 A

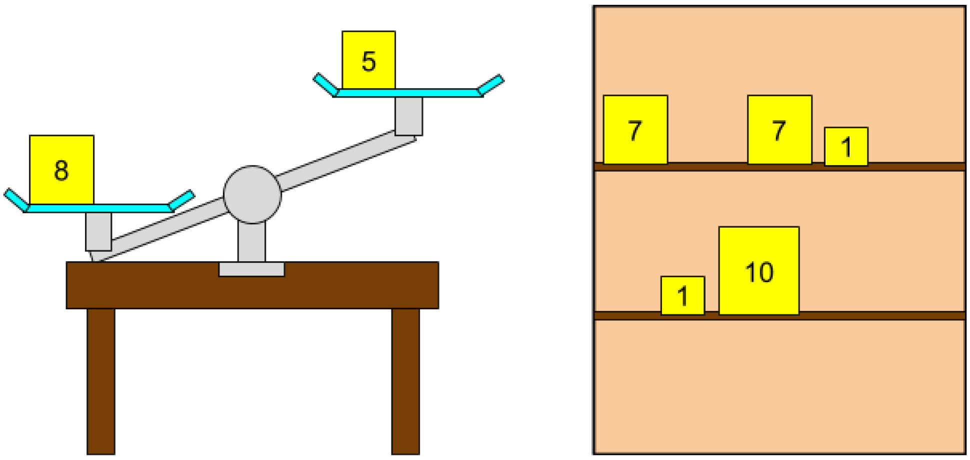
1) **Friends at a table**: Share **all** the slices of pizza and **all** the pieces of chocolate equally between the friends.



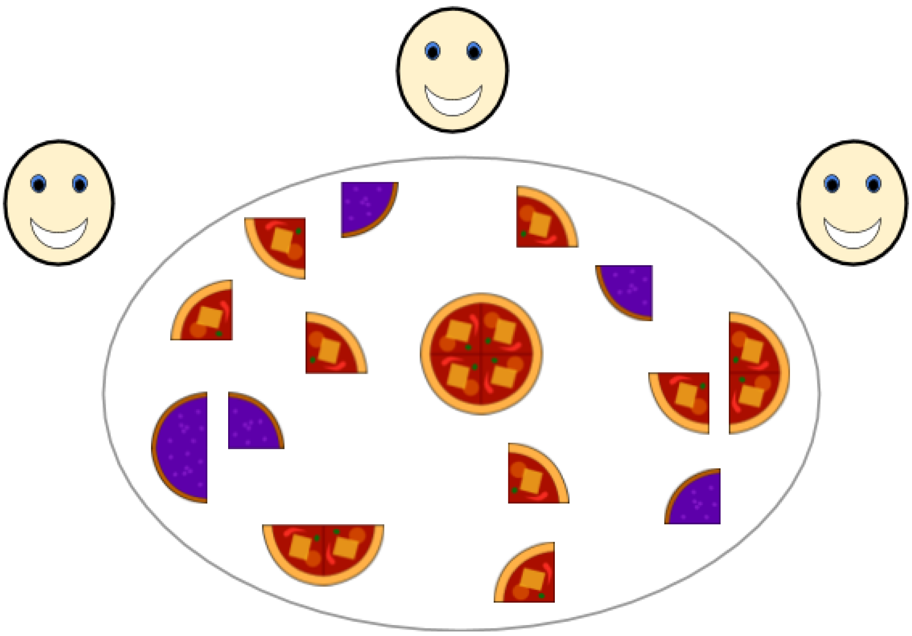
2) **Addition arrows**: Fill the numbers into the circles and square.



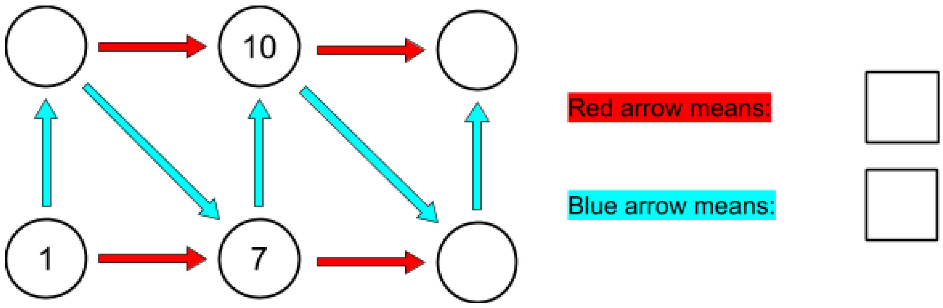
3) **Weighing machine**: Which weights from the box should be used to balance the weighing machine?



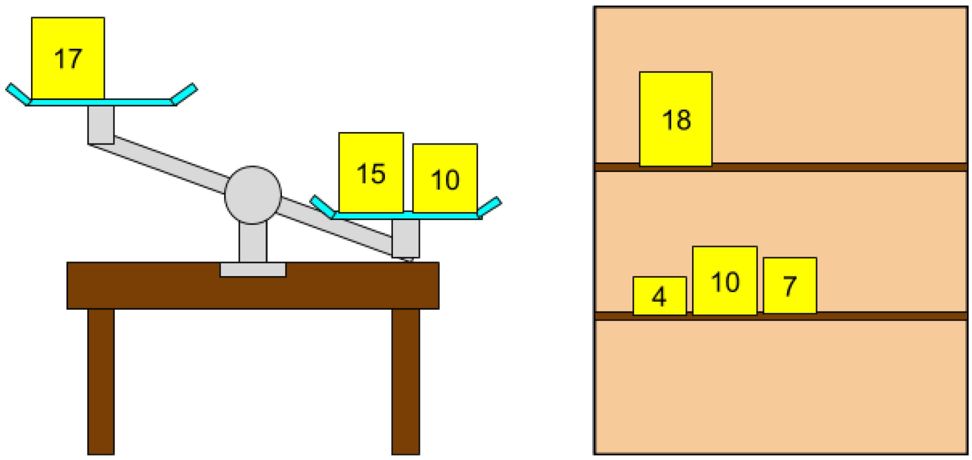
4) **Friends at a table**: Share **all** the slices of pizza and pie equally between the friends.



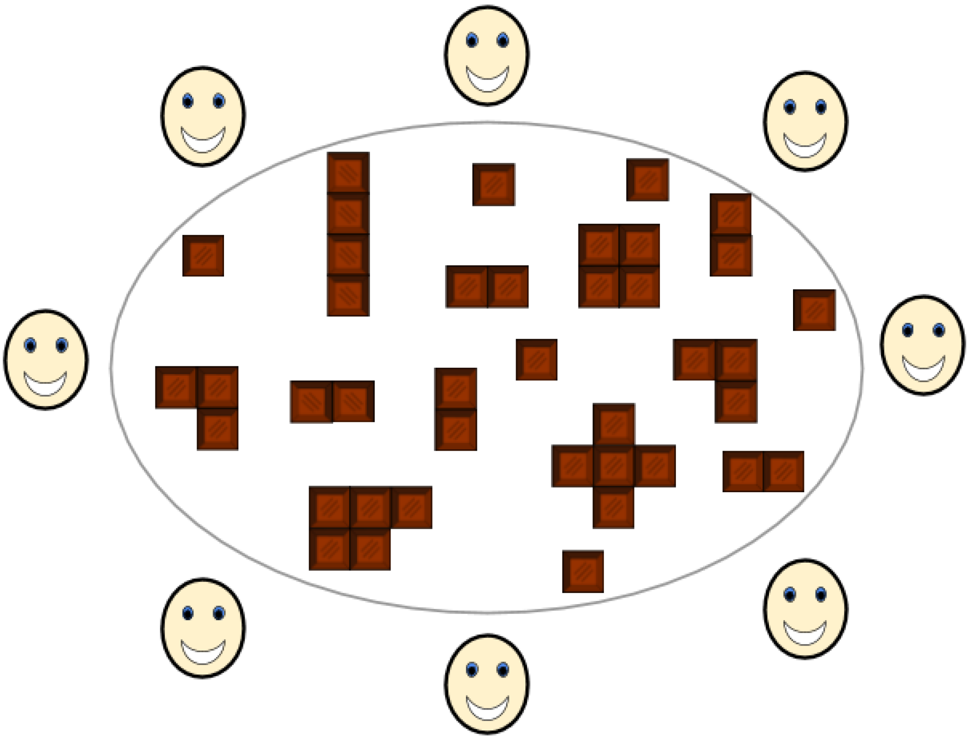
5) **Addition arrows**: Fill the numbers into the circles and square.



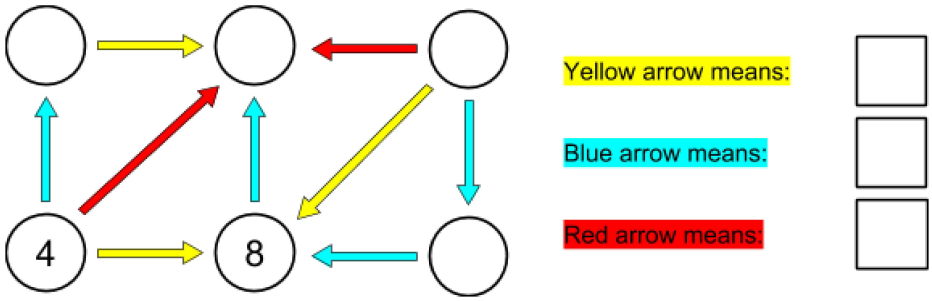
6) **Weighing machine**: Which weights from the box should be used to balance the weighing machine?



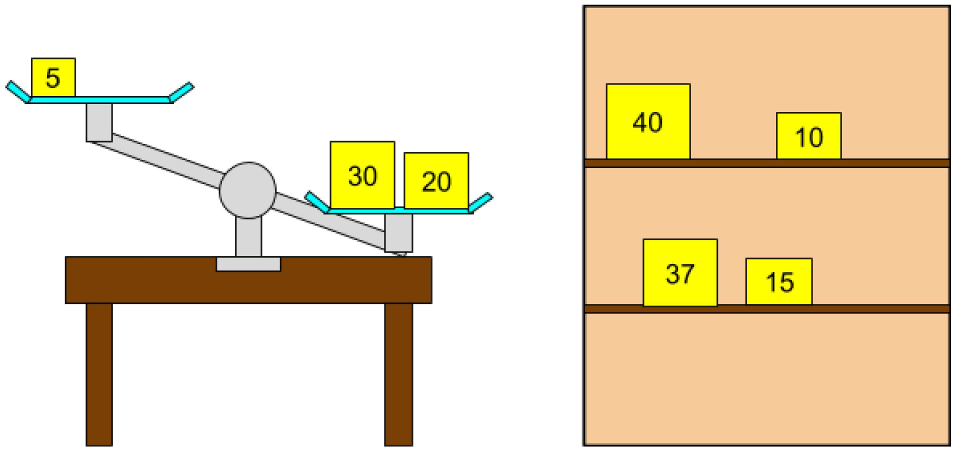
7) **Friends at a table**: Share **all** the pieces of chocolate equally between the friends.



8) **Addition arrows**: Fill the numbers into the circles and square.



9) **Weighing machine**: Which weights from the box should be used to balance the weighing machine?



Appendix 3: Child Questionnaires

Initial Questionnaire

1) Encircle • like this all your favourite school subjects:

- Czech language
- Math
- English language
- Basic science
- Music
- Art
- Physical education
- Other? Please write down which other subject: ____________________

2) How often do you read (books or magazines)? Encircle your answer:

- I read often
- Sometimes I read, but not often
- I don't read books or magazines at all

3) Encircle what is true:

- I read alone
- Sometimes I read alone, sometimes I read with my parents
- I read with my parents
- I don't read too much, neither myself nor with my parents

4) Have you got a favourite book?

- No, I don't
- Yes, but I don't remember the name
- Yes, it is: ______________________________
- Yes, but there are more than one. For example:

____

5) Do you play games on the smartphone or tablet or notebook?

- No I don't, I rather like to: __________________
- Yes
    - Write down which games you play:

      ____

6) How good do you think you will be at playing the game?
(The more stars you color, the better you think you will be)

7) How much do you think you will enjoy the game?

8) Do you know these characters?
(Encircle the ones you know. And write down their names if you know them.)

___________ ___________ ___________

Final Questionnaire - Narrative Groups

1) Do you **like** to listen to stories and fairytales?

- Yes
- No
- Other:

2) Do you **like** telling stories and fairytales?

- Yes
- No
- Other:

3) How much did you enjoy the game?

4) How good do you think you were at playing the game?
(The more stars you color, the better you think you were)

5) Did you learn anything while playing?
(The more stars you color, the more you think you learnt)

6) Do you know where Jacob and Theresa met the wizard Matemág for the first time?
(If you don't know, make a cross)

…………………………………….

7) What do you think about the tasks in the game?

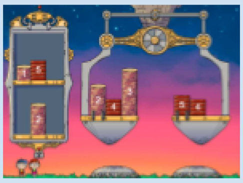
I enjoyed the most: …………………………………..

I was bored by: …………………………………..

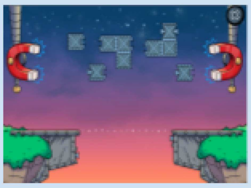
The most difficult was: …………………………………..

The easiest was: …………………………………..

Here all the tasks are listed:



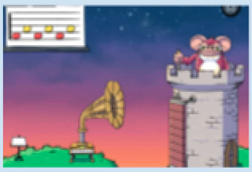
A) WEIGHING MACHINE



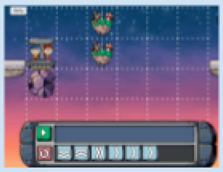
B) JIGSAW BRIDGE

C) HUNGRY PLANT

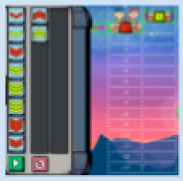
D) NUMBER WEB



E) MUSIC BRIDGE



F) FLYING PLATFORM



G) COUNTING LIFT

Final Questionnaire - Control Group

1) Do you **like** stories and fairy tales?
- Yes
- No
- Other:

2) Do you **like** to tell stories and fairy tales?
- Yes
- No
- Other:

3) How much did you enjoy the game?

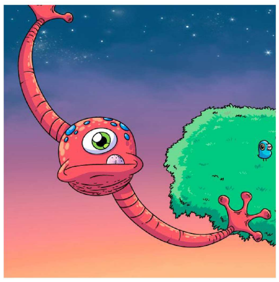
4) How good do you think you were at playing the game?
(The more stars you color, the better you think you were)

5) Did you learn anything while playing?
(The more stars you color, the more you think you learnt)

6) Do you know these characters?

(Encircle the ones you know. And write down their names if you know them.)

______________ ______________ ______________



______________ ______________ ______________

Appendix 4: Parent Questionnaires

Questionnaire About Playing

1. **The nickname your child chose:** ……………………………………….

2. **How often do you think your child played the game?**

In the first week	Daily	Every other day	1-2 times a week	Never	I do not know
In the second week	Daily	Every other day	1-2 times a week	Never	I do not know

3. **Did your child play on a device that is used by him/her alone?**

   Yes No, the device is also used by: ……………………….

4. **What type of device did your child use?**

   Tablet Mobile phone Other: ………………………………..

5. **Which operating system does the device used for the game have?**

   Android iOS (Apple device)

6. **On which place has your child been playing?**

   a. At home
   b. With friends
   c. While travelling by car
   d. While travelling by train
   e. On other occasions: ………………………………………………………….

7. **Select and, if necessary, add:**

   a. your child played on his/her own
   b. your child played with: ……………………………………..……………..

8. **Do you think your child showed the game to his/her friends?**

   Yes No

9. **Please rate, how did you like the game as a parent?**

   (as in school - 1 the best, 5 the worst)

1	2	3	4	5

10. **Which attributes or features of the game did you like (if any)?**

……………………………………………………………………………………………………

……………………………………………………………………………………………………

11. **Would you recommend this game to other parents?**

    Definitely yes Rather yes Rather not Definitely not I don't know

For the questions below, encircle the answer that corresponds most to reality:

12. **Has your child remembered to play the game himself/herself?**

    Yes, frequently Yes, sometimes Not at all I don't know

13. **Have you been explicitly reminding your child about the game?**

    Yes, frequently Yes, sometimes Only when the sms came Not at all

14. **Have you been asking your child about the game?**

    Yes, frequently Yes, sometimes Not at all I don't know

15. **Has your child been talking about the game?**

    Yes, frequently Yes, sometimes Not at all I don't know

16. **Has your child been mentioning some phrases from the game?**

    Yes, frequently Yes, sometimes Not at all I don't know

    Possibly **which**?

    ………………………………………………………………..

17. **Has your child been boasting about his/her achievements while playing?**

    Yes, frequently Yes, sometimes Not at all I don't know

18. **Has your child been showing you the game?**

    Yes, frequently Yes, sometimes Not at all I don't know

19. **Has your child ever needed your help while playing?**

    Yes, frequently Yes, sometimes Not at all I don't know

20. **Has your child been forgetting about the time while playing?**

    Yes, frequently Yes, sometimes Not at all I don't know

21. **Has your child been wishing he/she would not have to stop playing??**

    Yes, frequently Yes, sometimes Not at all I don't know

22. **Have you ever interrupted the gameplay when it was too long?**

    Yes, frequently Yes, sometimes Not at all I don't know

23. **Did your child have a time limit or other rules about playing?**

……………………………………………………………………………………………

……………………………………………………………………………………………

24. **Have you been communicating with other parents who participated in the research?** (Only communication about the research - e. g. about the game)

    Yes, frequently Yes, sometimes Not at all

25. **Were there any technical problems with the game?**

……………………………………………………………………………………………

26. **This is the opportunity for any further comments about the game or about the research:**

……………………………………………………………………………………………

……………………………………………………………………………………………

Thank you for your cooperation.

General Questionnaire

1. **The nickname your child chose:** ……………………………………

2. **Year and month of your child's birth:** ……………….

3. **Gender:** girl - boy - ……………….

4. **Which class does your child attend?** ……………………………………

   (Including class number and letter)

5. **Does your child have siblings?** (Please write down how many, sex, age)

………………………………………………………………………………………

6. **Do you think your child is happy going to school?**

Definitely yes   Rather yes   Rather not   Definitely not   I don´t know

7. **Do you think your child enjoys math?**

Definitely yes   Rather yes   Rather not   Definitely not   I don´t know

8. **Which of the following activities do you think your child enjoys?**

	Extremely fun	Definitely fun	Rather fun	Rather not fun	Definitely not fun	Extremely not fun	Not applicable
1. Studying	1	2	3	4	5	6	7
2.Reading	1	2	3	4	5	6	7
3.Playing at home (with toys, board games...)	1	2	3	4	5	6	7
4.Playing outside (in the garden, on the playground…)	1	2	3	4	5	6	7
5.Going out into nature	1	2	3	4	5	6	7
6.Going to the cinema/the theatre	1	2	3	4	5	6	7
7.Doing sports	1	2	3	4	5	6	7
8.Hobbies (e.g.: ceramics, singing…)	1	2	3	4	5	6	7
9.Being and communication with friends	1	2	3	4	5	6	7
10.Playing computer games	1	2	3	4	5	6	7
11.Playing games on tablet/mobile phone	1	2	3	4	5	6	7
12.Watching TV or videos	1	2	3	4	5	6	7

9. **What do you do with your child?**

	Daily	Weekly	Monthly	Less Often	Not at all	Not applicable
1.Studying	1	2	3	4	5	6
2.Working in the household	1	2	3	4	5	6
3.Playing at home (with toys, board games…)	1	2	3	4	5	6
4.Playing outside (in the garden, on the playground …)	1	2	3	4	5	6
5.Playing computer games/games on tablet or mobile phone	1	2	3	4	5	6
6.Watching TV and videos	1	2	3	4	5	6
7.Reading	1	2	3	4	5	6
8.Hobbies (e.g.: ceramics, singing…)	1	2	3	4	5	6
9.Playing computer games	1	2	3	4	5	6
10.Going out into nature	1	2	3	4	5	6
11.Going to the cinema/theatre	1	2	3	4	5	6
12.Doing sports	1	2	3	4	5	6

10. **How much do you check on your child in the following areas?**

	I check a lot	Rather check	Rather do not check	Do not check at all	Not applicable
1.Who your child's friends are	1	2	3	4	5
2.How your child spends his/her money	1	2	3	4	5
3.Where your child stays after school	1	2	3	4	5
4.What your child does in his/her free time	1	2	3	4	5
5.In which environment the child is outside of the family	1	2	3	4	5
6.Which websites your child visits	1	2	3	4	5
7.Whom your child is "talking" to online	1	2	3	4	5

11. **Rate how strongly you agree with the following statements:**

	Definitely agree	Rather agree	Rather disagree	Definitely disagree
1.The internet should be used by all family members	1	2	3	4
2.The internet harms children's learning process	1	2	3	4
3.The internet helps with the development of the child	1	2	3	4
4.Computers harm children in their cognitive development	1	2	3	4
5.Children need to work with computers now to be successful in the future	1	2	3	4

12. **Please encircle how often your child plays video games:**
(On the tablet/mobile phone/computer)

a) Daily
b) Several times a week
c) Once a week
d) Several times a month
e) Less often
f) Not at all

13. **How much time does your child spend by playing video games on average?** (During one "gaming session")

…………………………………………………………………………………………

14. **How often do you personally use the internet?** (E. g.: e-mail, watching online videos, browsing etc.)

a) Daily
b) Several times a week
c) Once a week
d) Several times a month
e) Less often
f) Not at all

15. **Does your child have his/her own device with internet access?**

a) Yes
b) No

16. **Please encircle which of these devices there are in your home:**

a) TV
b) Computer
c) Notebook/Laptop
d) Tablet
e) Smartphone
f) Other
g) I don't know

17. **If your child has his/her own device with internet access, encircle which:**

a) TV
b) Computer
c) Notebook/Laptop
d) Tablet
e) Smartphone
f) Other
g) I don't know

15. **Which of the following activities is your child allowed to do online?**

	Allowed	Allowed with supervision	Not allowed	I don't know
1.Playing games	1	2	3	99
2.Using communication applications (e. g.: e-mail, Messenger, WhatsApp, Skype etc.)	1	2	3	99
3.Downloading music or movies from the Internet	1	2	3	99
4.Watching videos (e. g. on YouTube)	1	2	3	99
5.Using their own profile on a social network?	1	2	3	99
6.Sharing personal information with others on the Internet (e. g.: full name, address or telephone number)	1	2	3	99
7.Uploading photos, videos or music and sharing it with others	1	2	3	99

16. **Can you describe in which ways "computers" are helping your child?**

Technologies help with:	Definitely yes	Rather yes	Rather no	Definitely no
1.School work	1	2	3	4
2.Having fun	1	2	3	4
3.Communication with friends	1	2	3	4
4.Communication with parents	1	2	3	4
5.Communication with other family members (e. g.: grandparents)	1	2	3	4
6.Finding information about children's interests	1	2	3	4

Finally, we will ask you for some information about you. The information is used for the statistical evaluation of the collected data.

17. **How old are you? ………………………**

18. **I am: mother, father** or **other: ………………………**

19. **Your education:**

a) Primary school

b) High school

c) Vocational colleges

d) University

20. **What is your nationality? ………………………**

Thank you for your cooperation.