CHARLES UNIVERSITY

FACULTY OF MATHEMATICS AND PHYSICS

# HABILITATION THESIS

Mgr. Cyril Brom, Ph.D.

**Learning with Digital Technologies:**
**the Role of Positive Affect and Motivation**

Prague, March 2017

Learning with Digital Technologies: the Role of Positive Affect and Motivation

Habilitation Thesis

Cyril Brom

March, 2017

brom@ksvi.mff.cuni.cz
http://ksvi.mff.cuni.cz/~brom

Charles University
Faculty of Mathematics and Physics
Department of Software and Computer Science Education
Malostranské nám. 25
118 00, Prague 1
Czech Republic

# Contents

# Preface

This thesis presents the author's selected results pertaining to educational research. Specifically, it concerns learning by means of digital technologies. It investigates under what conditions some of these technologies make learning experiences more stimulating and enjoyable and whether this is connected to learning improvements. The research described herein was led by the author and carried out primarily at the Faculty of Mathematics and Physics of the Charles University during the past decade. It was done with the author's many collaborators, a list of which is included in Appendix A. The key collaborating institutions were the Faculty of Arts of the Charles University, the Institute of Physiology of the Academy of Sciences of the Czech Republic, and the Laboratory of Behavioral and Linguistic Studies at the Institute of Psychology of the Academy of Sciences of the Czech Republic. Most of the research was done using digital learning applications co-developed by the author.

The results are presented as five journal papers. Two further studies, unpublished so far, are commented. All the papers focus on specific aspects of the author's overall research agenda. The papers are presented as separate chapters (6 – 10). The commentary introduces the research background (Chapter 1) and the underlying theoretical framework (Chapter 2). Afterwards, the commentary details key instruments pertaining to measuring "stimulation" and "enjoyment", and also learning gains (Chapter 3). Participant samples are discussed in Chapter 4. Finally, the commentary introduces individual studies and their contributions and synthesizes the findings (Chapter 5).

The author was supported primarily by the Czech Science Foundation (also known as the Grant Agency of the Czech Republic, GA CR) and the Faculty of Mathematics and Physics of the Charles University. The key projects were P407/12/P152 (GA CR), 15-14715S (GA CR), and Prvouk P46 (Charles University). Some of the software applications used as research instruments were developed as part of development projects, namely CZ.2.17/3.1.00/31162 (financed by the European Social Fund and the Municipality of Prague) and CZ.04.3.07/3.1.01.3/3213 (financed by the European Social Fund, the state budget of the Czech Republic, and the Municipality of Prague).

Prague, March 2017

Cyril Brom

# Part A Commentary

## 1. Introduction

All of us have participated as learners in plenty of dull and boring, as well as stimulating and likable, educational experiences. The latter experiences arguably enhance learning. Recently, some scholars have argued that harnessing the powers of digital technologies in educational settings is one way to evoke these experiences more often (e.g., Garris, Ahlers, & Driskell, 2002; Prensky, 2001; Squire, 2002; see also D'Mello, 2013; Livingstone, 2012; Passey, Rogers, Machell, & McHugh, 2004; Sung, Chang, & Liu, 2016). Some have even gone so far as to suggest that the educational world needs a "digital revolution", because, without one, kids growing up with digital technologies will not be able to learn effectively (e.g., Oblinger & Oblinger, 2005; Prensky, 2001; Veen, 2007; but see also Cuban, 1986; Curtois et al., 2014; Kirschner & van Merriënboer, 2013, pp. 170-173; Mayer, 2014a, p. 14).

The need to increase interest and motivation of school learners is debated in literature in broader contexts; for instance, as concerns reducing drop-out rates (e.g., Bridgeland, DiLulio, & Morison, 2006) or general decline in intrinsic motivation to learn in schools throughout preadolescence (e.g., Lepper, Corpus, & Iyengar, 2005). There are various ways to address this need (see, e.g., Lazowski & Hulleman, 2016). Do digital technologies offer additional ways? In this thesis, I address some aspects of this question[1].

[1] This thesis describes several studies that have been carried out by me and my colleagues. In the text of the commentary, I use the singular voice when speaking about myself. I use the plural voice when speaking about myself and my collaborators. Appendix A lists the contributions of my colleagues. Also, various bits of text in this thesis were taken from and/or adjusted based on papers written by me and my colleagues and discussed in this thesis.

## 1.1. General Research Agenda

> My overreaching research agenda is whether (and if so, how) certain types of *digital learning technologies* can make educational experiences more *stimulating* and *enjoyable* for *school* learners, thereby increasing their *active cognitive participation* in learning activities and thus eventually improving *learning outcomes*.

To clarify my research agenda, I must explain a couple of points. First, two types of empirical research are important in the present context: so-called value-added and media-comparison intervention studies. *Value-added studies* research whether adding a particular feature to the educational experience enhances learning (Mayer, 2014a, p. 9). Such feature might, for example, be an image of a tutor in a textbook (vs. no tutor; e.g., der Meij, 2013) or graphics with warm and bright colors (vs. black-and-white graphics; e.g., Plass, Heidig, Hayward, Homer, & Um, 2014). *Media-comparison studies* research whether certain new approaches to delivering educational experiences are better than a so-called "traditional" approach (Mayer, 2014a, p. 9; Clark, 2012). It is not always clear what the "traditional approach" should be. However, a media-comparison study can, for instance, compare a learning game ("new") with frontal teaching ("traditional"). Interventions in comparison groups in media-comparison studies differ in many features; unlike in value-added studies. In the present context, *correlational studies*, wherein researchers look for relationships between certain variables (such as enjoyment and learning outcomes), are also tangentially relevant.

Second, educational technologies (digital in the present case) are always used as tools within *educational methods*. In my research agenda, when speaking about learning by means of digital technologies, I always have in mind a technology plus a method (be it an old method or a new, technology-specific one).

Third, by "certain types of digital learning technologies" I refer to *advanced digital multimedia learning materials* (ADMLMs). This term is derived from the notion of *multimedia learning materials*, which are materials that combine written or spoken words with pictures (e.g., illustrations or diagrams) and which are studied in the field of multimedia learning (e.g., Mayer, 2009; Mayer, 2014c; Renkl & Scheiter, 2015). Multimedia learning is to be distinguished from pure text learning (i.e., no pictures; e.g., Schiefele, 1999). Multimedia learning includes learning from teachers who use blackboards, from textbooks with pictures, computerized presentations (i.e., slideshows), brief animations and videos, applications with virtual tutors or educational games. By using the adjective "advanced", I leave out non-animated linear materials (basically textbooks and simple slides) and consider non-interactive animations and brief video clips to be a borderline area between "advanced" and "non-advanced". By using the phrase "advanced digital", I also leave out non-computerized advanced learning materials such as board games. In actuality, this thesis is primarily about computerized games, simulations, and animations.

Fourth, ADMLMs could potentially be used in educational contexts for many purposes. The purpose I am interested in here is emphasized by the words “stimulating” and “enjoyable”. I study whether ADMLMs (including the methods of their usage) offer learners *incentives* in terms of activating stimulation during learning and in terms of induced positive feelings (cf. Schiefele & Rheinberg, 1997, pp. 253-257; Pekrun & Linnenbrink-Garcia, 2012). I also study whether these incentives are able, in turn, to facilitate learning. I focus only on incentives that derive directly from learning experiences (as opposed to from the consequences of learning). Thus, these incentives are of *intrinsic* rather than *extrinsic* nature. For instance, if the educational experience is entertaining in and of itself, I would consider this as an intrinsic incentive. If learners start to perceive the studied topic to be useful for their future carriers, I would consider this to be an extrinsic incentive. Finally, I am interested only in positive incentives and not in negative ones; even though occurrences of, say, confusion can be, in certain situations, beneficial to learning (D’Mello, Lehman, Pekrun, & Graesser, 2014). To conclude, I study in this thesis the ability of ADMLMs to evoke *positive, intrinsic affective-motivational incentives* and thereby enhance learning. For brevity, I will typically use only the term *affective-motivational incentives* in the following text, skipping the words “positive” and “intrinsic”.

Fifth, by referring to “active cognitive participation”, I refer to theoretical frameworks (e.g., Moreno, 2005) that a) use an individual learner as a unit of analysis, b) analyze learners’ mental processes, c) assume that learning is an effortful process and learners must actively cognitively engage in educational activities in order to learn effectively, and d) consider the incentives above as influencing cognitive activity. For instance, I leave out the collaborative learning perspective and behavioristic approaches.

Sixth, by “learning outcomes” I mean primarily *mental models* (Jones et al., 2011), thereby focusing on *mental models acquisition*. I generally study learning in STEM disciplines (science, technology, engineering, and mathematics). Therein, metal models are a typical, though not the only, learning output. Mental models are specific knowledge structures, which are usually viewed as internal representations of the possible behavior of devices and systems and possible unfolding of situations and problems (e.g., Johnson-Laird, 1983; Jones, Ross, Lynam, Perez, & Leitch, 2011). This includes the depiction of causalities and the ability to draw inferences and making predictions about reality; including running “internal simulations” inside one’s mind. The present work is neither about motor learning, nor learning of facts, nor cognitive training, nor learning work ethic, nor promoting teamwork skills, etc.

Finally, by referring to “school” learners, I mean learning in institutional contexts; either in schools directly or for schooling purposes (e.g., doing school homework). The age group is roughly 16 – 30 years of age; especially secondary and tertiary education students.

The complete idea behind my research agenda is summarized in Figure 1.

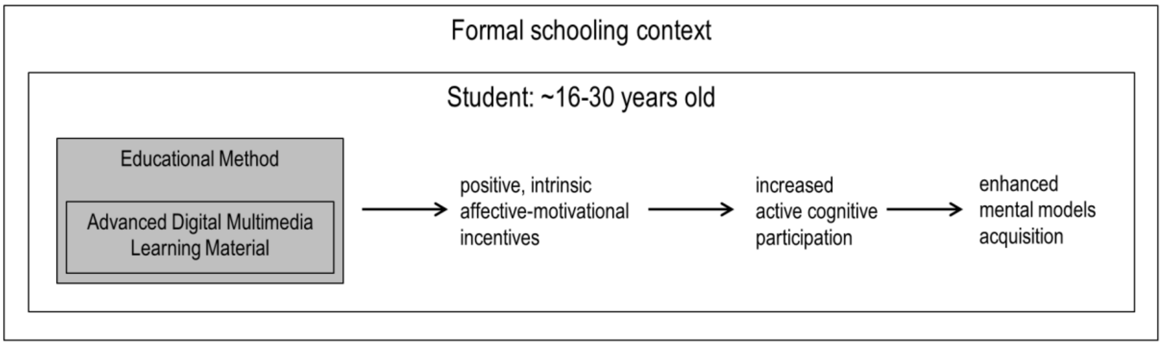


*Figure 1. My research agenda. The links denote my focus, not necessarily causal relationships (specifically, some links may be bidirectional from a causality point of view).*

## 1.2. Emotional Design Research

The ADMLM research field is young, but it partly derives from an older body of knowledge accumulated through research on learning from texts and texts-and-pictures. Research on how to design expository texts and graphics in instructional materials in order to enhance learning via evoking affective-motivational incentives has been around for at least several decades (e.g., Garner, Brown, Sanders, & Menke, 1992; Schiefele, 1999; Snetsinger & Grabowski, 1993). Recently, some called this endeavor (in the context of learning from texts-and-pictures) *emotional design* research (e.g., Um, Plass, Hayward, & Homer, 2012; Plass & Kaplan, 2015; see also Norman, 2004).

This research's ultimate goal can be viewed as the creation of empirically-based *emotional design principles*: a kind of parallel to cognitive principles of multimedia learning (e.g., Mayer, 2009; Mayer, 2014c; Renkl & Scheiter, 2015). Cognitive principles of multimedia learning are guidelines on how to design multimedia learning materials so that they can be processed effectively by learners. These principles tend to derive primarily from what is known about human cognitive architecture and cognitive processes (and from a broad empirical base), but they tend to abstract away from the affective-motivational aspects of learning experiences. The promise of emotional design principles lies in also considering learners' affective-motivational states. To the best of my knowledge, in contrast with cognitive principles of multimedia learning, no "definite" emotional design principles are available (cf. Plass & Kaplan, 2015).

I use the term *emotional design elements* to describe instructional design features that a) can be sources of affective-motivational incentives and b) can be theorized to enhance learning because of their eliciting these incentives. The emotional design principles are thus supposed to list those emotional design elements that actually work in comparison to an emotionally "neutral" baseline.[2]

There are many emotional design elements, but only some of them have been investigated and only a few showed some promise. I review some (but not all) notable examples in this chapter.

A prime example is seductive, also called extraneous, details (e.g., Mayer, 2009). These are typically viewed as additions (textual, pictorial, auditory) to learning materials that a) are aimed at triggering learners' interest and b) provide tangentially relevant (or entirely irrelevant) information not necessary for comprehending the core instructional message. Despite their motivational potential, seductive details have repeatedly failed to improve learning of core information (see, e.g., Garner et al., 1992; Harp & Mayer, 1998; Rey, 2012; but see also Park, Flowerday, & Brünken, 2015; Schneider, Nebel, & Rey, 2016).

A different example is using a conversational, rather than a formal, language style for instructional texts (e.g., changing the text from a third person form of address to first/second person, adding statements directed at the learner or adding personal views). One reason why a conversational style for instructional texts may enhance learning is that it may present an affective-motivational incentive. Other reasons also exist, including those based on cognitive rather than affective-motivational explanations (see Brom, Hannemann, Stárková, Bromová, & Děchtěrenko, 2017). In fact, composing instructional texts in a conversational style is one of the (cognitive) principles of multimedia learning: the so-called *personalization principle* (Mayer, 2009). This "emotional" design element is thus known to work: in certain contexts. It improves deep, conceptual learning (in comparison to a neutral/formal style) in English treatments up to 35 minutes long (meta-analyzed in Ginns et al., 2013). As concerns elevating affective-motivational variables, the findings are inconclusive (Ginns, Martin, & Marsh, 2013). Thus, it is also unknown whether affective-motivational variables mediate the effect of conversational style on learning outcomes. Also, a few studies have been performed with longer treatments and in non-English settings; they reported mixed findings (Ginns et al., 2013; Brom et al., 2017).

Pedagogical agents are static or animated lifelike characters (or torsos, heads) guiding users through multimedia learning environments (the presence of the agent's image is a defining feature) (Heidig & Clarebout, 2011). The promise of these agents is that they can foster motivation and learning outcomes, and they can thus be considered to be emotional design elements. However, a review by Heidig and Clarebout presented a dim picture. As concerns learning outcomes, 9 out of 15 experiments having a no-agent control group reported no difference; one reported a positive result; four mixed, partly positive results; and one a mixed, partly negative result. Only four experiments reported on

[2] The terms emotional design and emotional design elements have no stable definition. I use these terms in relation to any instructional materials and methods and with respect to any positive, intrinsic affective-motivational incentive that has the potential to enhance learning. Some researchers may use the terms slightly differently. Many authors cited in this work did not explicitly use the term emotional design research when describing their work.

motivation; three of them yielded no difference.[3] All 15 experiments, except for two, used university students as participants. Age may be a moderating factor: two studies had partly positive outcomes in the terms of learning gains. Also, in a more recent study, for children 11 years of age, a peer agent-tutor in printed materials improved learning and tended to improve most affective-motivational variables (der Meij, 2013). Several affective-motivational variables all predicted learning gains ($r$ = .26 - .52) (these are the largest correlations I have seen in this context). It is possible that more recently developed, and thus more advanced, agents can also perform better (see Guo & Goh, 2015[4]). Likewise, Mayer (2014b) reported, based on a review of 11 experiments, that replacing a low-embodied on-screen agent with a highly-embodied on-screen agent modestly improves learning (p. 362).

The digital game-based learning (DGBL) subfield studies how people learn from educational videogames. DGBL can be also viewed as an approach to emotional design, with the idea being that games, unlike "traditional" approaches to education, elicit affective-motivational incentives. DGBL media-comparison studies show that, on average, game-based learning modestly enhances learning outcomes (meta-analyzed in Sitzmann, 2011; Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013; Clark, Tanner-Smith, & Killingsworth, 2016). However, both the most recent meta-analyses (Wouters et al., 2013; Clark et al., 2016) reported that the games' benefits attenuated when only studies with randomization were considered, suggesting that part of the "benefits" may be due to uncontrolled differences between comparison groups (this problem is more common in media-comparison studies than in value-added studies; cf. Mayer, 2014a). Much like in the case of pedagogical agents, affective-motivational variables have only been investigated in about one-third of media comparison studies conducted so far (Vogel et al., 2006; Sitzmann, 2011; Wouters et al., 2013); with marginally higher outcomes for games (Wouters et al., 2013[5]). The relationship between affective-motivational variables and learning outcomes has rarely been researched (e.g., van Dijk, 2010); the relationship is not very stable: it is

[3] From the perspective of seductive details literature, this outcome is hardly surprising, because pedagogical agents can be viewed as seductive details (cf. Moreno, Mayer, Spires, & Lester, 2001). This can also be said about some other emotional design elements that I treat separately here such as a narrative (e.g., Adams, Mayer, MacNamara, Koenig, & Wainess, 2012).

[4] In their recent meta-analysis, Guo and Goh (2015) reported positive effects of "affective pedagogical agents embodied in computer-based learning environments" (this meta-analysis has around 40% overlap in the primary literature with the review by Heidig and Clarebout (2011)). I have some reservations regarding this meta-analysis and thus interpret it very cautiously. First, inclusion criteria were unclear; for instance, der Meij's work (2013) was included, despite its not seeming to be about "affective pedagogical agents embodied in computer-based learning environments". Second, effect sizes were computed using Pearson's correlation coefficient $r$, which is an atypical approach in the case of studies with between-subject design. Some effect size values were strange, such as $r$ = .89 for retention in the case of der Meij's study, and I was unable to locate in the paper the equation used for converting more traditional effect sizes, such as Cohen's $d$, to $r$. Finally, the authors did not reply to my questions regarding these matters.

[5] Wouters and colleagues (2013) reported that the difference for affective-motivational variables is not significant (Cohen's $d$ = 0.26), without giving the $p$-value, which actually is .076 (Pieter Wouters; email dating from 16 Dec 2013).

generally positive, non-significant or weakly significant, and in a small-to-medium range[6]. The meta-analyses indicate that several factors may (not surprisingly) moderate the effects of games, such as learning domain (Wouters et al., 2013; see also Mayer, 2014a, Ch. 7) or the level of graphical realism (Wouters et al., 2013; Clark et al., 2016).

Value-added DGBL studies investigating the effects of particular game design elements are on the rise. However, with the exception of instructional support features (Wouters & Oostendorp, 2013) and enhanced scaffolding designs (Clark et al., 2016), they have been few so far: the Clark and colleagues' meta-analysis (2016) identified eleven of them. Their results are inconclusive (Clark et al., 2016)[7]. Mayer (2014a; Ch. 5) also reviewed value-added DGBL studies with a somewhat broader definition of "game feature", but the findings were similar (p. 140). For instance, conflicting results were reported regarding the presence of competitive elements (null-positive: Plass et al., 2013; null-negative: Ke, 2008) or narratives (positive: Cordova & Lepper, 1996; null: Adams et al., 2012) in educational games. The effects may be moderated by personal variables such as age (e.g., narrative: Cordova and Lepper researched primary school children but Adams and colleagues worked with university learners).

Adding game design elements (such as points, badges, and leaderboards) into non-game education materials and methods is an approach called gamification (e.g., Deterding, Dixon, Khaled, & Nacke, 2011; Dicheva, Dichev, Agre, & Angelova, 2015). There is a dearth of studies on this approach with a control group. Instead, researchers tend to report user evaluations and/or log analyses (see Dicheva et al., 2015; Hamari, Koivisto, & Sarsa, 2014). Examples of gamification in education pertain mainly to mobile applications or web-based learning platforms; for instance, those used in university courses (Darejeh & Salim, 2016; Dicheva et al., 2015). Gamifying "stand-alone" ADMLMs is rare. Generally, the findings pertaining to learning outcomes from comparative studies, including those on gamified courses, are at best mixed (Domínguez et al., 2013; de-Marcos, Domínguez, Saenz-de-Navarrete, & Pagés, 2014; de-Marcos, Garcia-Lopez, & Garcia-Cabot, 2016; Hanus & Fox, 2015; Katz, Garcia-Lopez, & Garcia-Cabot, 2014; Wang, Zhu, & Sætre, 2016; see also Sandberg, Maris, & Hoogendoorn, 2014).

Some other approaches can be also classified as emotional design research (cf. Mayer, 2014a; Plass & Kaplan, 2015). A notable example is manipulating surface-level graphical representation. Multimedia learning literature suggests that it is beneficial for

---

[6] Correlations are typically expressed here, and in the cited works, as Pearson's correlation coefficient *r*. I use Cohen's (1988) classification of effect sizes (small: $r \sim .1$; medium: $r \sim .3$; large: $r \sim .5$).

[7] Enhanced scaffolding was broadly defined by Clark and colleagues (2016, p.99) as including "personalized scaffolding, intelligent agents, adapting game experiences to student needs or interests, and revised game structuring targeted at emphasizing the learning mechanic". Enhanced scaffolding significantly enhanced learning, which is the only significant finding as concerns game design features (Clark et al., 2016). It is unclear which of the scaffolds were of cognitive origin and which also presented affective-motivational incentives. The meta-analysis by Wouters and Oostendorp (2013) focused primarily on instructional support elements, which, with the exception of narrative and personalization, were of cognitive origin.

college students to anthropomorphize certain graphical elements of the materials; for instance, by adding human-like faces and/or expressive eyes to otherwise non-human entities (Mayer & Estrella, 2014; Park, Knörzer, Plass, & Brünken, 2015; Plass et al., 2014; Um et al., 2012). Other surface representational changes may be less effective; including changes of color alone (Heidig, Müller, & Reichelt, 2015) or change of entities' graphical appearance from a neutral one to a funny one (Snetsinger & Grabowski, 1993; see also Brom, Hannemann, Stárková, Bromová, & Děchtěrenko, 2016). A different example is to manipulate the initial instruction (learning vs. entertainment). Results are ambiguous (e.g., Erhel & Jamet, 2013; Vadercruysse, Vandewaetere, Cornillie, & Clarebout, 2013). The same is true about additions of background music and sounds (e.g., cf. Moreno & Mayer, 2000a and Fassbender, Richards, Bilgin, Thompson, & Heiden, 2012).[8] Also, listening to enthusiastic teachers from videotaped lectures improves likability (Anderson & Withrow, 1981; Tabbers, Reurlings, & de Kievit, 2016). However, learning outcomes may be enhanced only in a home setting, where the learner is likely to get easily distracted by irrelevant stimuli (Tabbers et al.,2016), but not in a classroom (Anderson & Withrow, 1981).

The few studies that have investigated affective-motivational mediation in these contexts report predominantly small-to-medium correlations between affective-motivational variables and learning gains. For instance, correlations in studies on anthropomorphisms are in the range $r$ = .06 – .35) (Um et al., 2012; Plass et al., 2014). Single-group correlational studies in multimedia and digital game-based learning have tended to report negligible correlations between affective-motivational variables and learning outcomes (e.g., der Meij, 2014; Iten & Petko, 2014; Sabourin & Lester, 2014). In the context of learning from texts (i.e., without pictures), Schiefele (1999) reported an average correlation of .33 between ratings of interestingness of narratives or text segments and various measures of text learning (based on a review of 14 text-learning studies); however, these relationships could be partly (though not fully) explained by more interested learners having higher prior knowledge (cf. Tobias, 1994; Schiefele, 1999, p. 271).

Overall, literature that can be classified as pertaining to emotional design makes it clear that the hunt is still on for emotional design elements that would be consistent sources of affective-motivational incentives and, at the same time, facilitate learning (see Table 1). Overall results probably mask the moderating effects of certain variables: such as prior knowledge or age. The relationship between affective-motivational variables and learning outcomes has rarely been investigated, and the limited findings indicate that it tends to be small-to-medium, somewhat fragile, but rarely negative. There are thus ample opportunities for contributions in this research area.

[8] One of the multimedia learning principles is a so-called signaling principle (Mayer, 2009), which states that learning is enhanced "when cues that highlight the organization of the essential material are added" (p. 108). Colors and sounds can be used as such cues, but the studies cited here refer to usage of colors and sounds as triggers for affective-motivational incentives and not as signaling cues.

*Table 1. Overview of some major emotional design approaches and their effectiveness.*

| Emotional Design Approach | Summary of Findings |
|---|---|
| Seductive details | Can trigger interest, but often harms, or does not improve, learning; most likely due to distraction (or "seduction") |
| Personalization principle | Generally beneficial in up to 35-minutes treatments in US context, the reason unclear; insufficient data beyond this context |
| Pedagogical agents | Older agents (up to around 2009) hardly improved learning; unclear if more recent (elaborate and/or better embodied) agents better |
| Digital game-based learning (vs. "traditional" instruction) | Probably slightly beneficial for learning, on average; one, but not the only, reason could be slightly higher motivation |
| Game design elements within games (game with the element vs. game without the element); e.g., competition, narrative | Conflicting results; insufficient data |
| Game design elements in non-game contexts (i.e., gamification) | Not promising so far; insufficient data |
| Anthropomorphisms | Promising so far; replications and extensions needed |

## 1.3. Research Contribution

In this work, I introduce seven educational research studies I have conducted with my colleagues between 2009 – 2016, comprising 12 individual experiments with ADMLMs, with a total sample of 1,271 participants (primarily Czech and Slovak high school and university learners). Five of these studies have been already published and constitute this thesis (Brom, Preuss, & Klement, 2011; Brom, Bromová, Děchtěrenko, Buchtová, & Pergel, 2014; Brom, Levčík, Buchtová, & Klement, 2015; Brom, Šisler, Slussareff, Selmbacherová, & Hlávka, 2016; Brom et al., 2017). The other two studies have been submitted. These studies all investigate questions related to my general research agenda. The interventions include digital educational games, an educational simulation, and educational animations.

The studies use between-subject experimental or quasi-experimental designs. They are either *value-added* (i.e., Brom, Bromová, et al., 2014; Brom et al., 2017; two submitted studies) or *media comparison* (i.e., Brom et al., 2011; Brom et al., 2015; Brom, Šisler, et al., 2016). The value-added studies investigate the following emotional design elements:

a) a conversational language style of instructional texts (as opposed to a formal style); i.e., they investigate the personalization principle in the Czech context (Brom, Bromová, et al., 2014; Brom et al., 2017);

b) a specific form of gamification of a computerized simulation (in the form of a game-like goal, increased freedom of choice, points, virtual currency and praise; as opposed to interaction with the non-gamified simulation – a submitted "gamification" study);

c) a motivating topic (as opposed to a less motivating topic – a submitted “topic” study).

The media-comparison studies focus on:

d) whether digital educational games used for integrating and reinforcing knowledge acquired from an expository text or lecture enhance learning as opposed to integrating and reinforcing this knowledge by a comparable “traditional” instructional approach (Brom et al., 2011; Brom, Šisler, et al., 2016);

e) whether it is better if students play educational games individually, each sitting at one computer, or collectively, i.e., when the teacher plays the game with the whole class while showing it on a projector (Brom et al., 2015).

Overall, the thesis asks (and answers) two general questions:

1) Do the researched emotional design elements improve learning by eliciting affective-motivational incentives?

2) Do learners who are stimulated by and experience enjoyment from the educational experience learn better (across all studies)?

From reading Figure 1, it may seem that the key variable of interest here should be the level of active cognitive participation. However, this variable is difficult to measure (as detailed in Chapter 2). Therefore, the focus is on one of its possible antecedents: variables indexing affective-motivational incentives. The second group of key dependent variables pertains to learning outcomes.

## 1.4. Structure of the Thesis

The thesis proceeds with introducing the theoretical background (Chapter 2). Afterwards, measures (Chapter 3) and samples (Chapter 4) used in my studies are introduced. Then, the main part reviews my work (Chapter 5). That part is followed by a discussion and conclusion, outlining prospects for future research (Chapter 6).

# 2. Theoretical Background

Why should affective-motivational incentives enhance learning and why do they not always do so? One plausible theoretical explanation is offered by the Cognitive-Affective Theory of Learning from Media (CATLM) (Moreno, 2005). It has broad applicability, including learning experiences with ADMLMs. It derives from the Cognitive Theory of Multimedia Learning (Mayer, 2009) and is parallel to the Cognitive Load Theory (Sweller, Ayers, & Kalyuga, 2011; Kalyuga, 2011). The Cognitive Theory of Multimedia Learning further capitalizes on Baddeley's classical memory model (Baddeley, Eysenck, & Anderson, 2009) and Dual Coding Theory (Clark & Paivio, 1991).[9]

CATLM assumes that a learner processes incoming information using his or her visual and verbal channels such that he/she first *selects* information from low-level sensory representations and then *organizes* it in the form of visual and verbal mental models. In the next step, these models are *integrated* – together and also with prior knowledge from long-term memory – resulting in a cross-modal mental model represented in the learner's working memory. Eventually, this model can also be integrated into long-term memory.

The crucial assumptions of the theory are that a) learners must *actively engage* in the selection, organization, and integration processes in order to construct a coherent mental model and b) learners have a *limited capacity* in both channels and in working memory. As concerns the active engagement assumption (a), CATLM posits (apart from other things) that affective-motivational factors influence learning by increasing or decreasing the level of active cognitive processing. These factors can have the form of affective-motivational incentives evoked by the used ADMLM (or any other type of instructional material).

The Cognitive Load Theory is related to the limited cognitive capacity assumption (b). It concerns itself with cognitive load (or working memory load), which is typically defined as the number of information elements that must be simultaneously represented and processed in working memory and the degree of their interconnectedness (Leppink, Paas, Van Gog, van Der Vleuten, & Van Merrienboer, 2014; Sweller, 2010). The theory was recently adjusted by Kalyuga (2011). I favor this adjustment because it is more parsimonious. It assumes two additive types of load that can be imposed on learners during educational experiences: *intrinsic* and *extraneous* (Figure 2).[10] Intrinsic load derives from the complexity of the learning task with respect to the learner's prior knowledge (what is complex for a novice may not be so complex for an expert). This type of load is "useful" in that it is essential for comprehending the learning material. Dealing with this load results in

[9] CATLM is not the only relevant theory for the ADMLM contexts (see, e.g., Mayer, 2014a, Ch. 3; Plass & Kaplan, 2015 for others). However, CATLM is one of the most relevant theories for present purposes. Also, I favor it for its parsimony, and because it is intuitively understandable by non-experts.

[10] Intrinsic and extraneous cognitive load are directly mapped onto the CATLM processes by Mayer (2009, pp. 79-89) and Kalyuga (2011, pp. 5-8).

learning: with no intrinsic load, there is no learning. On the other hand, intrinsic cognitive load should not overwhelm available cognitive resources: that would impede learning. Extraneous load is a “bad” type of load: it arises from suboptimal design of the educational experience, such that the learner must engage in redundant processing that does not help learning but is nevertheless evoked by the experience (e.g., if pieces of information are presented far apart, which complicates their selection, or if a piece of extraneous information is present, which spends the channels’ or the working memory’s capacity). Extraneous load thus typically hinders learning, because its accommodation depletes cognitive resources that could be otherwise devoted to dealing with intrinsic load.

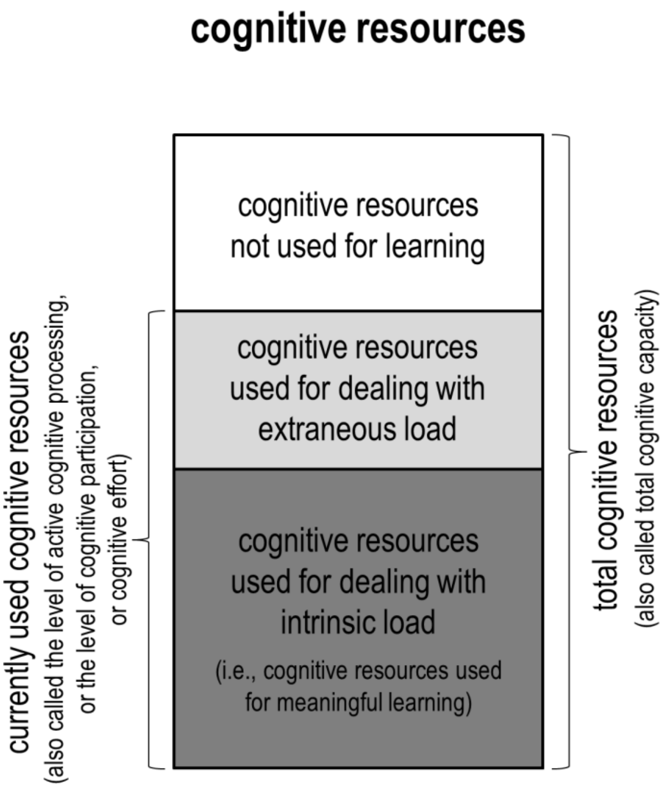


*Figure 2. Cognitive resources versus intrinsic and extraneous cognitive load.*

I have provided above only a rough overview of these two theories, skipping specifics irrelevant for the present purpose (see, e.g., Mayer, 2009; Sweller et al., 2011; Kalyuga, 2011). One thing about these specifics must be said though: most of the cognitive principles of multimedia learning (e.g., Mayer, 2009; Mayer, 2014c; Renkl & Scheiter, 2015) theoretically draw from them; particularly from those specifics that refer to human cognitive processing (rather than affective-motivational factors). I refer to learning materials that follow cognitive principles of multimedia learning as *cognitively optimized*, because these materials have been optimized with respect to what is known about how the human mind works (i.e., designed so that they can be processed effectively by learners; but without much consideration of learners’ affective-motivational states).

The key contribution of CATLM is that it not only offers (capitalizing on the older Cognitive Theory of Multimedia Learning) “cognitive” predictions by which the majority of the cognitive principles of multimedia learning are underlined, but it also offers predictions related to affective-motivational and metacognitive factors. Of these, the key prediction for present purposes is that if the ADMLM evokes affective-motivational incentives, these can increase the level of active cognitive processing and thereby enhance learning. However, these incentives are evoked by specific emotional design elements. These elements must be mentally processed in the first place; that is, selected, and depending on their complexity, possibly also organized and integrated. All these steps increase extraneous cognitive load, reducing cognitive capacity that can be devoted to dealing with intrinsic load and thereby harming learning (Figure 3). It depends on the situation’s specifics whether the positives outweigh the negatives or the other way around. These situation’s specifics include not only the emotional design elements in question, but also the setting (e.g., a lab, home, a classroom), learner characteristics (e.g., prior knowledge, age, personality traits) and interactions of all of these. This also means that one must be very careful when attempting to generalize research findings.

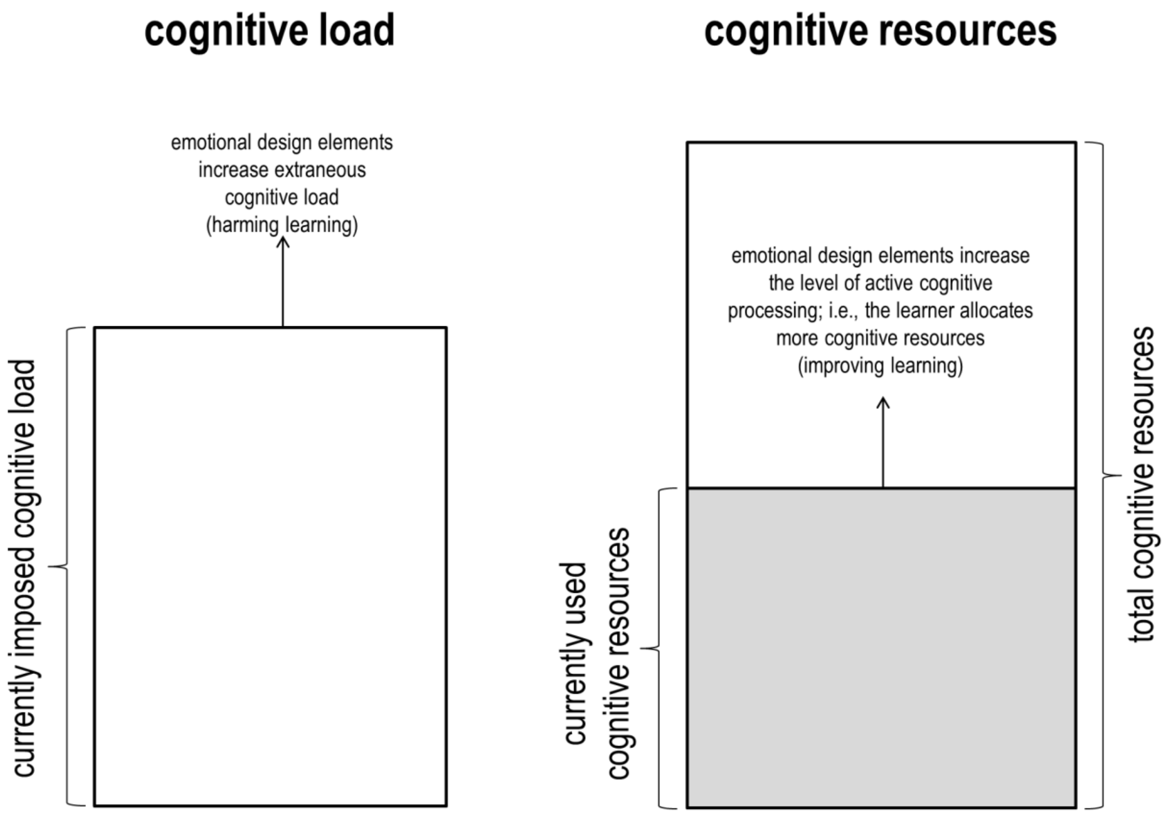


*Figure 3. The effects of affective-motivational incentives on cognitive load and cognitive resources used for learning.*

Throughout this thesis, I assume that instructional materials and methods used for investigating the effects of emotional design elements have been, more or less, cognitively optimized; except for the very presence of the emotional design elements. In other words, without the emotional design elements, the created educational experience should enable effective learning (as opposed to suboptimal learning). For instance, I am not interested here in comparing a poorly-designed instructional animation with anthropomorphic elements to the same poorly-designed animation without these elements.

It is notoriously difficult to measure cognitive load, let alone distinguish between the two types of load (Brünken, Plass, & Leutner, 2003; Brünken, Seufert, & Paas, 2010; de Jong, 2010). Likewise, I am unaware of any simple and reliable measure of the level of active cognitive processing. Therefore, studies rarely attempt to research step by step the whole link depicted on Figure 1.[11] It is possible, as argued below, to measure more reliably certain affective-motivational variables and learning outcome variables. One can thus research the link between these two types of variables and use the skipped concept of active cognitive processing as an explanatory device helping to explain findings and to formulate new research hypotheses.

To summarize:

- I use CATLM as the primary explanatory framework because it makes the following theoretical link clear (even for non-experts): emotional design elements → affective-motivational incentives → increased cognitive effort (but potentially also extraneous cognitive load) → learning outcomes.
- I primarily measure learning outcomes and affective-motivational states.
- The ADMLMs in question should be, more or less, cognitive optimized.

[11] Certain variables, such as "perceived difficulty", have been discredited as proxies to cognitive load (de Jong, 2010). Recently, a brief self-report instrument distinguishing between intrinsic and extrinsic load has been introduced (Leppink et al., 2014) but, based on our experience, some non-psychology students have problems understanding some items, such as "I invested a very high mental effort in the complexity of this activity". Also, certain objective variables can be used as proxies for cognitive load (e.g., based on biofeedback data or dual task paradigm, see Brünken, Plass, & Leutner, 2003; Brünken, Seufert, & Paas, 2010). Measuring these variables, however, complicates the research methodology and makes the study less ecologically valid. For instance, the experiments then look more laboratory-like rather than school-like, and it is problematic to engage learners in longer interventions and hard to collect data from multiple participants at the same time. Also, using devices such as an eye-tracker can influence participants' feelings toward the experiment (Brom, Stárková, Lukavský, Javora, & Bromová, 2016), making it problematic to measure the impacts of an ADMLM's affective-motivational incentives. In a similar vein, certain constructs, such as mind wandering (Smallwood & Schooler, 2015), which can be measured, are related to the notion of active cognitive processing. In brief, if one attempts to use a proxy variable for cognitive load or the level of active cognitive processing measurement, one has to typically alter the research method such that it is more problematic to ask other questions.

# 3. Measures

## 3.1. Measuring Learning Outcomes

The crucial assumption behind this work, and indeed behind the majority of multimedia learning research, is that knowledge can be reliably measured. By knowledge, I now mean mental models (Johnson-Laird, 1983; Jones et al., 2011). In the multimedia learning field, learning outcomes pertaining to mental models acquisition are typically measured by so-called *retention* and *transfer* knowledge tests (Mayer, 2009). Retention tests assess "superficial" learning; basically whether the learner was able to memorize the material without necessarily understanding the core process/model in question. Transfer tests assess "deep" learning; if learners truly understand the point and are thus able to "transfer" and use what they learnt into situations not explicitly mentioned by the learning materials (see Table 2 for examples). Transfer is generally considered the key measure of learning outcome, though both retention and transfer performances should be good in the case of meaningful learning (Mayer, 2009, p. 21). Based on my experience, it is easy to develop retention and transfer knowledge tests for young adult audiences as concerns mental models acquisition.  It is also easy to rate these tests, even though they frequently contain open-ended questions. This is typically done by at least two independent raters, and inter-rater agreements are generally very good. In my work, I usually use both retention and transfer tests and let open-ended questions be graded by two raters.

*Table 2. Retention and transfer tests examples from a biological wastewater treatment domain (Brom et al., 2017).*

| **Test type** | **Question examples** |
|---|---|
| Retention | Based on the animation you just saw, describe **in detail** how biological wastewater treatment works. |
| Transfer | What would happen if a fungus first appeared in the treatment plant and then bacteria?  Write down **all consequences** that come to mind based on the animation you saw today. |
| Transfer | What all does the presence of nutrients in wastewater have to do with the biological wastewater treatment process, i.e., with the functions and functioning of the treatment plant, with bacteria and with fungus? Write down **all possibilities** that occur to you and which relate to any phase of water treatment presented in today's animation.  Write an explanation for each **possibility**. |

Note: emphasis in the original.

Quite often, researchers measure acquired knowledge only immediately after the learning experience (i.e., *immediate* knowledge tests). This way, they assess if the intervention influenced cognitive processes needed for initial knowledge acquisition: such as attention, information processing, retrieval of prior knowledge, and integration of new information with prior knowledge. However, only when learning outcomes are measured also after a longer interval (i.e., *delayed* knowledge tests), can one look into the effects of learning on long-term memory related processes (such as consolidation) and on possible rehearsal.

One of the confounding variables is prior knowledge, and it would thus be useful if it could be measured as well. Here comes the problem. Administering full knowledge tests can be quite lengthy, especially if the topic to be learnt is complex. Based on my experience, the filling in of lengthy tests can be very frustrating for low prior knowledge learners. One usually does not want to frustrate experiment participants at the experiment's beginning. Also, these pretests can cue participants on what they should remember and/or represent a "practice" on its own, thus confounding the outcomes (Fraenkel, Wallen, & Hyun, 2012, p. 171). One can administer pretests, say, a week before the intervention, hoping that the participants will forget the details of this experience during the week (which is itself dubious). However, this poses a risk that learners will improve their knowledge in the period between the tests and the experimental session. A robust solution is using a Solomon Four Group Design, which manipulates the presence/absence of pretests as an additional factor, but this requires more participants. A different way these problems are frequently addressed is by administering brief questionnaires for perceived prior domain knowledge (i.e., learners rate what they think they know) (e.g., Mayer, 2009). I use this approach often. Still another way is recruiting very low prior knowledge learners only and skipping pretests entirely. I also use this approach from time to time. Both approaches have their (obvious) advantages and disadvantages. There is probably no ideal solution.

Having no pretests is sometimes regarded as a poor practice (e.g., All et al., 2016; Sitzmann, 2011). I think differently: in my opinion, it is (often, but not always) methodologically questionable when learning variables are measured at only one point in time (and when next-to-zero prior knowledge cannot be expected in the sample). Specifically, I do think that having immediate tests and delayed tests (but no pretests) is often justifiable, provided one also focuses on the differences between the two measurements. This is the approach I also follow at times.

> In my work, I often measure learning outcomes by retention and transfer tests; usually at two points in time. I also usually measure perceived prior knowledge.

## 3.2. Measuring "Positive Activation"

Learners can experience various affective-motivational states during learning with ADMLM; such as curiosity, delight, engagement, boredom, frustration or disgust (cf. D'Mello, 2013), or relief, pride, anger or sadness (cf. Pekrun & Linnenbrink-Garcia, 2012). One useful way of organizing these states is along three dimensions: activation (activating vs. deactivating), valence (positive vs. negative), and object focus (activity vs. outcome-prospective vs. outcome-retrospective) (Pekrun & Linnenbrink-Garcia, 2012). Here, I am interested in the ADMLMs' alleged ability to elicit positive, intrinsic affective-motivational incentives. Some of these incentives can be viewed as positive, activating activity-related states: such as enjoyment or flow.[12] Other incentives, often interconnected with these states, are more complex affective-motivational constructs: such as situational interest (Hidi & Renninger, 2006) or intrinsic motivation (Ryan & Deci, 2000).

I look at positive, activating affective-motivational incentives and their measurement in three ways. These ways are based on overlaps and differences between the theoretical constructs and the way they are measured. All measures discussed in this thesis are based on self-reports (as opposed to objective methods used as proxies for affective-motivational states; e.g. cortisol levels, cf. Brom, Buchtová, et al., 2014).

I start with *likability*, which I use as an umbrella term for three affective-motivational entities: enjoyment, intrinsic motivation and situational interest. *Enjoyment* can be defined as an activity-related state experienced when the learning activity or materials are positively valued and when the activity is sufficiently controllable by the learner (Pekrun, 2006; p. 323). Enjoyment derives from the activity undertaken per se rather than its instrumental value. *Situational interest* is usually defined as a state of concentration and enjoyment elicited by the features of a specific situation (cf. Hidi & Renninger, 2006; Schiefele, 1999; p. 263); and it usually connects feeling-related (e.g., the object/activity is stimulating or engaging by itself) and value-related (the object/activity is useful or meaningful by itself) components (Schiefele, 1999). Motivation can be viewed as a propensity to start, continue or stop performing a specific activity given current context; and the most basic distinction differentiates between extrinsic and intrinsic motivation (Ryan & Deci, 2000). A person is in an *intrinsically motivating* state when he or she is "doing something because it is inherently interesting or enjoyable" (Ryan & Deci, 2000; p. 55).[13]

Despite conceptual differences (e.g., situational interest is explicitly connected to increased concentration and can be viewed as a precondition to intrinsic motivation; Schiefele, 1999; p. 262), these constructs are actually measured similarly (Table 3). As far as I know, enjoyment is typically measured by a few Likert-type questions; whereas, more robust questionnaires tend to be used (in my opinion) for assessing situational interest and intrinsic motivation. The only notable distinction between these measures seems to be (see Table 3) that the more complex instruments also contain some questions related to attention (e.g., "This activity did not hold my attention at all"), curiosity (e.g., "It stimulates my curiosity") or value ("It was useful."). Notice that the questions relate to the activity being or having been undertaken. I typically measure only enjoyment in my work, i.e. out of all three of these constructs. I use 1-3 questions only because there is a certain limit on the number

[12] A couple of other examples: positive deactivating activity-related: relaxation; positive activating outcome-prospective: hope; positive activating outcome-retrospective: pride; negative activating activity-related: frustration. See Pekrun and Linnenbrink-Garcia (2012) for more examples.

[13] For the sake of brevity, and because I expect a rather general readership, I intentionally do not elaborate on "details"; such as a distinction between triggered and maintained situational interest (Hidi & Renninger, 2006) or different levels of extrinsic motivation (Ryan & Deci, 2000).

of questions one can administer without irritating participants (and the other instruments I use are more complex; as detailed next in this work).

*Table 3. Measures of enjoyment, intrinsic motivation, and situational interest (examples).*

| **Construct** | **Questions**[a] |
|---|---|
| Enjoyment<br>(e.g, Mayer & Estrella, 2014; Giannakos, 2013) | I liked this activity<br>I enjoyed this activity<br>This activity was appealing for me<br>Using this application for learning was fun<br>… |
| Intrinsic Motivation<br>(McAuley, Duncan, & Tammen, 1989)[b] | I enjoyed doing this activity very much<br>This activity was fun to do.<br>This activity did not hold my attention at all. (reverse-coded)<br>I would describe this activity as very interesting.<br>While I was doing this activity, I was thinking about how much I enjoyed it. |
| Intrinsic Motivation<br>(Isen & Reeve, 2005) | It stimulates my curiosity<br>It is interesting<br>It is fun<br>I want to continue investigating it<br>It makes feel curious about it<br>It is enjoyable<br>It makes me want to explore it further<br>I would be willing to come back and participate in a future experiment that used this activity |
| Situational interest[c]<br>(e.g., Magner, Schwonke, Aleven, Popescu, & Renkl, 2014; Schiefele, 1990) | This object/activity was:<br>*(feeling related)*<br>- exciting<br>- entertaining<br>- boring (reverse-coded)<br>*(value related)*<br>- useful<br>- worthless<br>- unimportant |

[a]Questions typically have a Likert scale. [b]Symptomatically, this is a subscale from an "Intrinsic Motivation Inventory", which is called "Interest/Enjoyment". A 7-item version of this subscale also exists. [c]Situational interest can be also assessed prospectively (e.g., Schiefele & Krapp, 1996). More-item versions also exist. The feeling-related and value-related questions are often highly correlated (Schiefele, 1999, p. 262) and thus combined (e.g., Schiefele & Krapp, 1996).

The second way of looking at positively activating affective-motivational incentives is through the lenses of generalized positive and negative affect. This older model, PANA (Watson & Tellegen, 1985), organizes affective states in two dimensions: positive affect, also called positive activation; and negative affect, also called negative activation. These two dimensions are largely orthogonal (Watson, Clark, & Tellegen, 1988) and lay at a 45-degree rotation over axes used more traditionally in two-dimensional models of affect structure: arousal and valence. Therefore, positively activating states can be mapped on high generalized positive affect and low to neutral generalized negative affect.

*Table 4. Positive and Negative Affect Schedule by Watson, Clark, and Tellegen (1988).*

| Construct | Questions[a] |
|---|---|
| Generalized positive affect (Watson, Clark, & Tellegen, 1988) | I feel right now/have felt [time period]:<br>- interested<br>- excited<br>- strong<br>- enthusiastic<br>- proud<br>- alert<br>- inspired<br>- determined<br>- attentive<br>- active |
| Generalized negative affect (Watson, Clark, & Tellegen, 1988) | I feel right now/have felt [time period]:<br>- distressed<br>- upset<br>- guilty<br>- scared<br>- hostile<br>- irritable<br>- ashamed<br>- nervous<br>- jittery<br>- afraid |

[a]In my case, questions have a 5-point Likert scale, which is quite common.

Generalized positive and negative affect can be measured by Positive and Negative Affect Schedule, also called PANAS (Watson, Clark, & Tellegen, 1988; Table 4). Note that the instruction refers to a person's feelings rather than the person's evaluation of an activity or an object, which is a distinction from likability. In all my newer works, I do use PANAS.[14]

[14] As far as I know, there is no standardized version of PANAS in the Czech language. We therefore had to pay specific attention to the translation. Substantial help was provided by Iva Poláčková from the Institute of Psychology of the Czech Academy of Sciences. In pretests, participants' comments on items

*Table 5. Flow Short Scale and Attention-Intensity questions.*

| Construct | Questions[a] |
|---|---|
| Flow | I feel just the right amount of challenge. |
| (Rheinberg, Vollmeyer, & Engeser, 2003)[b] | My thoughts/activities run fluidly and smoothly. |
| | I don't notice time passing. |
| | I have no difficulty concentrating. |
| | My mind is completely clear. |
| | I am totally absorbed in what I am doing. |
| | The right thoughts/movements occur of their own accord. |
| | I know what I have to do each step of the way. |
| | I feel that I have everything under control. |
| | I am completely lost in thought. |
| Attention-Intensity | I was completely caught up in what I was reading [doing]. |
| (Schiefele & Krapp, 1996) | When reading the text [doing X], I was concentrated. |

[a]Questions typically have a Likert scale (in my case, a 7-point Likert scale, as in the original questionnaire). [b]The English translation was taken from (Vollmeyer & Rheinberg, 2006).

The third way of thinking about the notion of positive, activating affective-motivational incentives is through the concept of flow. *Flow* state is usually defined as a pleasant absorption by, and increased attention to, an activity one undergoes (Csikszentmihayli, 1975). Some view flow as a discrete state (rather than a continuum) that refers to extreme intensity which includes total absorption and time distortion; especially in the context of playing videogames (e.g., Cheng, She, & Annetta, 2015, p. 235; Jennett et al., 2008, p. 642). These scholars often use the term immersion to refer to the continuum (but see also Sweetser & Wyeth, 2005; Kiili & Laynema, 2008). Others use parallel constructs, such as attention-intensity, without using the term flow (Schiefele & Krapp, 1996; Table 3). In my work, I use the flow concept as a continuum. Of the numerous ways of measuring flow (see Moneta, 2012), I prefer a pen-and-paper questionnaire, namely the Flow Short Scale (Rheinberg, Vollmeyer, & Engeser, 2003), for its simplicity.[15] Compared to likability and generalized positive affect, the questions are more focused on attention and concentration rather than feelings; particularly on ease of attending to and concentrating on the object/activity (Table 5). Notably, the questions are formulated such that it is possible to

were generally as one would expect, though some items had to be adjusted. In the terms of factor analysis, as concerns my data, positive items load generally well on the positive factor, with the exception of "alert" ("ostražitý"), which tends to load weakly on both factors (though more on the positive one). This could be a translation issue: the term "ostražitý" can be interpreted with negative connotations in the Czech language. Loadings of negative items on the negative factor is generally weaker due to low variability in the data (negative affect tends to be low in the interventions I use; see Table 7). Internal consistency is generally good for the positive affect (Cronbach's alpha > .8) and acceptable for the negative affect (Cronbach's alpha > .7).

[15] Similarly to PANAS, there is no standardized version of Flow Short Scale in the Czech language. Internal consistency of the instrument is usually good (Cronbach's alpha > .8).

measure flow levels during and/or after an activity, but not before (PANAS can be administered also before the intervention). [16]

| |
|---|
| In my work, I often measure enjoyment of the activity by 1-3 Likert items after the activity has been undertaken, learners' generalized positive and negative affect by PANAS at 1-3 moments during the experiment, and flow by Flow Short Scale during or right after the activity. |

[16] There are also different ways to measure positive, activating affective motivational states. For instance, D'Mello and colleagues frequently use "affective grid" (e.g., D'Mello, Olney, Williams, & Hays, 2012).

# 4. Participants

In my studies, I focus on an older high school audience (roughly 16-19 years of age) and university learners (roughly 19-30 years of age). Representative sampling is an ideal that can be rarely achieved in practice in experimental educational research. For instance, even if one has enough money to attempt recruiting a large and representative sample, many would-be-participants that have been chosen refuse to take part in the experiment. This biases the final sample toward those who are willing to participate.

Convenience sampling is thus typical for multimedia learning studies. A notable portion of the university participants are psychology or educational sciences majors, more than half of whom are typically females (cf., e.g., Heidig et al., 2015; Moreno & Mayer, 2000b; Park, Flowerday, & Brünken, 2015; Plass et al., 2014). Also, samples are relatively small: around 30 per cell is relatively common and generally acceptable. This is a very different situation compared to, for instance, online survey studies. This means that care must be taken when attempting to generalize the findings to different audiences (this point is highlighted in my studies on the personalization principle in the Czech context: Chapter 5.1). To ameliorate this general limitation, meta-analyses are typically considered as the key output for this type of research.

My high school participants were generally recruited from schools in the capital. These schools were more often than not above average schools. Sometimes, high school participants were recruited via an online server advertising short-term jobs for students (Brom et al., 2017). Also, part of the research was conducted in direct cooperation with several schools willing to participate (Brom et al., 2011; Brom et al., 2015; Brom, Šisler, et al., 2016). Participants from these samples had diverse backgrounds, as witnessed by the broad spectrum of university studies they were willing to apply for. All in all, the high school samples were relatively heterogeneous, but slightly above average with respect to the yearly cohorts.

My university participants (Brom, Bromová, et al., 2014; Brom, Šisler, et al., 2016; Brom et al., 2017; a "gamification" study; a "topic" study) were recruited from above average universities in the capital. I always strived for a heterogeneous sample with respect to study background. The largest subgroups of participants studied psychology, psychology–special education, computer science, mathematics or physics. However, participants' backgrounds also included new media studies, economics, linguistics, philology, biology, medicine or arts. Again, the samples were relatively heterogeneous, but slightly above average with respect to the yearly cohorts.

Assignment to experimental and control conditions was random (gender balanced). Sometimes, other variables known to correlate with learning gains were also used for balancing the groups, such as prior knowledge (e.g., Brom, Šisler, et al., 2016) or mathematical self-efficacy (e.g., Brom, Bromová, et al., 2014; Brom et al., 2017).

# 5. My Studies

To recapitulate, these are my general questions:

1) Do the researched emotional design elements improve learning by eliciting affective-motivational incentives?
2) Do learners who are stimulated by and experience enjoyment from the educational experience learn better (across all studies)?

I will address these questions in turn. As for the emotional design elements, I will look into the effects of the conversational style of instructional texts (Chapter 5.1), digital game-based learning (Chapter 5.2 and 5.3), gamification (Chapter 5.4), and a motivating topic (Chapter 5.5). I will discuss five published studies constituting this thesis (Brom et al., 2011; Brom, Bromová, et al., 2014; Brom et al., 2015; Brom, Šisler, et al., 2016; Brom et al., 2017) and briefly also two submitted studies (a "gamification" study and a "topic" study).

In all studies, between-group differences are expressed in terms of Cohen's *d* effect size (or an analogical effect size in the case of non-parametric tests). Cohen's *d* was classified as small (*d* ~ 0.2), medium (*d* ~ 0.5), and large (*d* ~ 0.8) (Cohen, 1988).

## 5.1. Personalization Principle

The personalization principle states that people learn better when instructional texts are in a conversational, rather than a formal, style (Mayer, 2009, p. 242). Support for this principle comes primarily from studies using English treatments up to 35 minutes long (see Ginns et al., 2013). These studies repeatedly showed the superiority of learning from instructions in a conversational style, as measured by transfer. Conversational styles can be viewed as emotional design elements, because one of the possible reasons for their superiority is their (alleged) motivational effects (reviewed in Brom et al., 2017).

*Table 6: Summary of the personalization principle experiments.*

| Experiment | Sample | *N* | Treatment | Cohen's *d*[a] for transfer |
|---|---|---|---|---|
| Brom, Bromová, et al., 2014: Exp. 2 | university | 75 | Beer brewing simulation | -0.08 \| -0.16[b] |
| "Gamification" study | university | 77 | Beer brewing simulation | 0.29 \| 0.20[b] |
| Brom et al., 2017: Exp. 1 | university | 57 | Lightning formation animation | -0.45† |
| Brom et al., 2017: Exp. 2 | high school | 73 | Lightning formation animation | 0.48* |
| Brom et al., 2017: Exp. 3 | university | 74 | Wastewater treatment animation | -0.04 |
| Brom et al., 2017: Exp. 4 | high school | 75 | Wastewater treatment animation | 0.22 |

Notes: [a]A positive *d* means higher transfer test scores for participants in the personalized condition. [b]Immediate | delayed tests.

† $p < .10$. * $p < .05$. without a correction for multiple comparisons

We examined whether the personalization principle works also in Czech contexts. Based on six individual experiments with the total sample $N = 431$, I can conclude that, on average, the principle does not seem to work in the Czech context; neither in short nor longer-than-35-minutes treatments (Table 6). Results also showed a slight, unstable, moderating effect of the school level in the case of short treatments such that the conversational style slightly enhanced learning for high school students and slightly hindered learning for college learners. Results from a supplementary experiment ($N = 138$) supported the idea that the limited benefits of conversational instructional texts for Czech learners are probably related to the generally more formal approach to education in the Czech Republic compared to the US schooling system. Language/cultural contexts can thus present a boundary condition for the personalization principle. Having instructional texts in a conversational style, rather than a formal/neutral style, does not seem to be a promising emotional design principle in the Czech context.

In the first study on the personalization principle (Brom, Bromová, et al., 2014), college participants learnt from a 2-hour interactive simulation (Figure 4) how to brew beer either with formal or conversational instructional texts. To justify the usage of the conversational style, the simulation was framed in a simple narrative about a family brewery and the texts were written as if the learner's grandfather were speaking to him or her.

In a small-scale pilot (Brom, Bromová, et al., 2014; Exp. 1, $N = 26$), Czech university learners preferred to study from the simulation with conversation instructional texts when given a choice. However, the main study (Brom, Bromová, et al., 2014; Exp. 2, $N = 75$) did not find significant between-group differences in enjoyment, flow, generalized positive affect, and learning outcomes (besides, learning with conversational instructional texts took longer). That is, no support for the personalization principle was found.

In a subsequent study, unpublished as of January 2017, which primarily investigated the effects of the gamification of the beer brewing simulation (Chapter 5.3), we used two control groups: with conversational and formal instructional texts (as in Brom, Bromová, et al., 2014). Contrasting results from these two groups only ($n = 34 + 33$) can be viewed as a replica of the first study (i.e., Brom, Bromová, et al., 2014). Again, there was no significant effect of the style of instructional texts on learning outcomes, enjoyment, flow, and generalized positive affect.

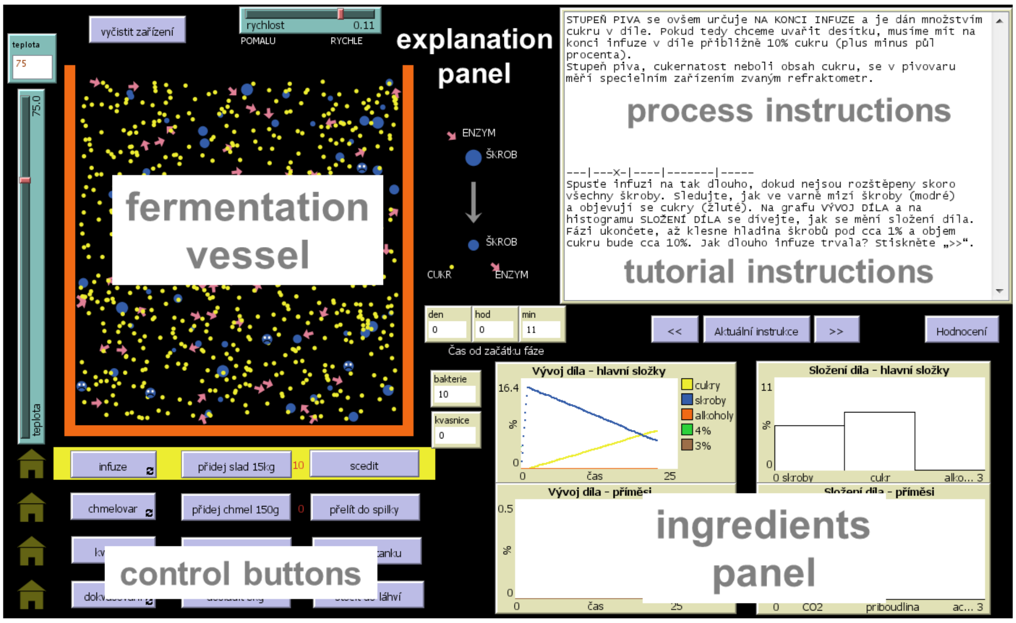


*Figure 4. The simulation screenshot.*

Were these results caused by a longer exposure or by a different language/cultural context? To clarify this issue and, therefore, to support one of the following two ideas – that the personalization principle's boundary condition is the length of exposure or that it is the participants' native language/cultural background – we performed four experiments with self-paced instructional animations roughly five minutes in length in the Czech context (Brom et al., 2017). In Experiments 1 (with university students) and 2 (with high school students), we closely replicated the seminal experiment researching this principle conducted by Moreno and Mayer (2000b; Experiment 2) with an animation on lighting formation (Figure 5). We then replicated the original experiment once again with a slightly different animation (on how a biological wastewater treatment plant functions; Figure 6), which was comparable in terms of length and complexity to the lightning formation animation (Experiment 3: university students; Experiment 4: high school students). In all four experiments, participants received instructional texts either in a conversational style (without a narrative) or in a formal style. Overall, there was no significant effect of the language style on learning outcomes and affective-motivational variables.[17] There was a slight (significant) positive effect of the conversational style for the high school students as opposed to university learners, but it was unstable (cf. Exp. 1 and 2 vs. Exp. 3 and 4) and may have been caused by chance alone.

[17] This study measured knowledge only immediately after the learning session (as in the original experiment by Moreno and Mayer (2000b)).

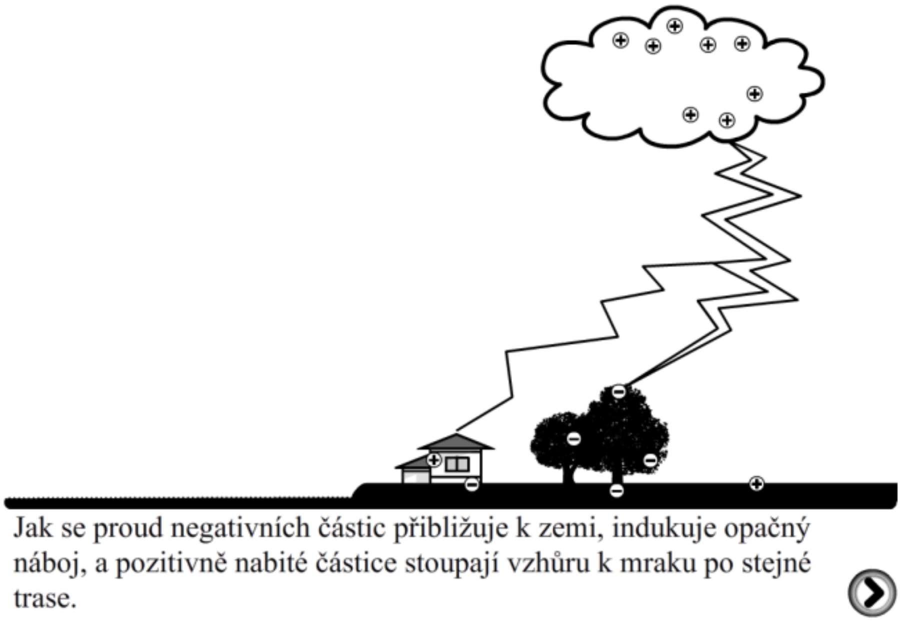


*Figure 5. A screenshot from the lightning formation animation (instructions in Czech). Based on work by Moreno and Mayer (2000b).*

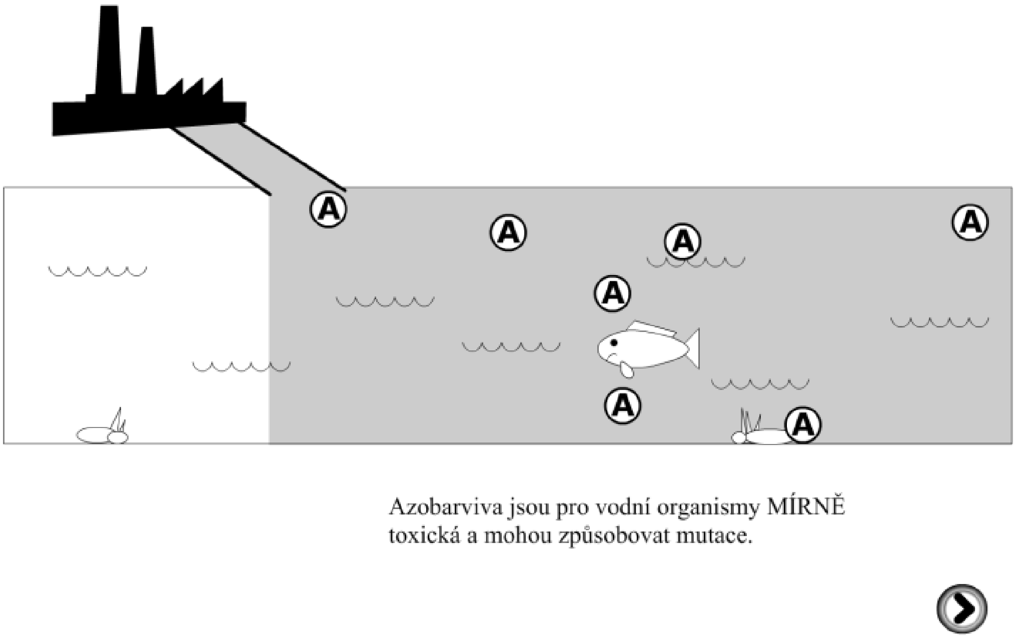


*Figure 6. A screenshot from the biological wastewater treatment plant animation (instructions in Czech).*

The levels of flow and generalized positive affect were relatively high for the beer brewing simulation and both animations compared to other treatments I used (Table 7). In the case of the wastewater treatment animation, generalized positive affect was measured also at the beginning of the experiments, and it was lower before the participants viewed the animation than after. This means that the learning experience increased generalized positive affect. In all cases, i.e., immediately before, during or after learning in any of the experiments measuring this variable, generalized positive affect was substantially higher than during the mere filling in of questionnaires (Table 7), which supports the validity of this measure. Generalized negative affect was consistently low in all experiments (Table 7).

The relationship between affective-motivational variables and learning gains was generally in the medium range for flow; small to medium for generalized positive affect; and, with the exception of the first beer brewing study (i.e., Brom, Bromová, et al., 2014, Exp. 2), negligible for enjoyment (Table 8). Therefore, the positive relationship between affective-motivational and learning outcome variables was partially supported.

Taken together, these findings (Table 6) indicate that the participants' cultural/language background (Czech in this case) presents a boundary condition for the personalization principle. I believe that these findings can be explained by the generally more formal approach to education in the Czech Republic compared to the US schooling system (see Brom et al., 2017 for details). This explanation is also corroborated by the results of the fifth experiment in my study (Brom et al., 2017). Therein, Czech high school and university participants ($N$ = 138) had to rate preferences on a computer tutor's printed statements. These statements had already been evaluated by US learners in terms of politeness (Mayer, Johnson, Shaw, & Sandhu, 2006) (politeness can be considered one way of personalizing formal instructions). Direct rather than polite statements were preferred; suggesting that learners' preferences for statements made by computer tutors, with differing levels of politeness, may differ across language/cultural contexts. This result seemingly contradicts the result of Experiment 1 from the beer brewing study (Brom, Bromová, et al., 2014), wherein participants preferred the simulation with conversational instructions. However, in the beer brewing study, the conversational simulation's version was framed in a narrative, which – together with the prospect of a 2-hour learning session – might have swayed the participants. In the experiment with rating preferences (Brom et al., 2017; Exp. 5), participants' task was only to judge individual tutor's statements. This might explain the difference.

In conclusion, these results warn us that certain emotional design elements may work only in specific language/cultural contexts. Generalization of the results is not a given. At the same, there is provisional support for the idea that elevated affective-motivational states enhance learning from ADMLM; particularly as concerns flow and, to a lesser extent, generalized positive affect.

*Table 7. Means and SDs for generalized positive affect, flow and generalized negative affect from the contexts discussed in this thesis.*

| Experiment | Characteristics | *n* | Age group | Generalized positive affect [10-50] | Flow [21-74] | Generalized negative affect [10-50] |
|---|---|---|---|---|---|---|
| **Beer Brewing** | 2-hour simulation, personalized version | 36 | university | 32.89 (6.73) | 55.43 (7.06) | 14.24 (4.00) |
| (Brom, Bromová, et al., 2014) | 2-hour simulation, direct version | 39 | university | 31.26 (7.28) | 55.71 (8.12) | 13.86 (3.89) |
| **Europe 2045** | 5-hour digital game | 103 | high school + university | 30.95 (6.34) | 50.86 (8.28) | 17.86 (6.11) |
| (Brom, Šisler, et al., 2016) | 5-hour non-digital game | 96 | high school + university | 30.84 (7.21) | 49.65 (8.36) | 18.00 (6.57) |
| | 5-hour discussion without gaming elements | 126 | high school + university | 26.00 (6.77) | 46.18 (7.77) | 18.06 (6.15) |
| **Beer Brewing – Gamified** | 2-hour simulation, gamified | 31 | university | 32.26 (7.49) | 57.23 (7.33) | 13.45 (4.22) |
| (submitted) | 2-hour simulation, personalized version | 34 | university | 32.82 (7.03) | 56.27 (8.28) | 14.53 (5.28) |
| | 2-hour simulation, direct version | 33 | university | 30.20 (5.88) | 54.11 (8.28) | 13.71 (3.76) |
| **Wastewater** | 6-minute animation, personalized version | 37 | university | 31.86 (7.42) | 57.32 (6.22) | 12.30 (2.92) |
| (Brom et al., 2017) | 6-minute animation, direct version | 37 | university | 30.49 (6.03) | 55.32 (8.21) | 12.19 (3.69) |
| | 6-minute animation, personalized version | 37 | high school | 30.91 (7.15) | 55.26 (8.17) | 13.33 (3.33) |
| | 6-minute animation, direct version | 37 | high school | 30.89 (6.48) | 54.47 (8.02) | 14.59 (5.96) |
| **Beer vs. Citrate** (submitted) | 90-minute simulation, direct version, citrate substrate production | 35 | university | 28.71 (7.21) | 53.43 (6.83) | 13.51 (3.91) |
| | 90-minute simulation, direct version, beer brewing | 30 | university | 32.10 (6.48) | 57.28 (7.86) | 12.87 (3.46) |
| **First-aid training course; actors** (Brom, Buchtová, et al., 2014) | 15-minute life action training simulation; actors | 12 | young adults | 34.08 (8.74) | 58.58 (7.39) | 14.25 (4.96) |
| **Filling in of questionnaires at a delayed testing session** | At the beginning a 30-minute long testing session | 165 | adults 18-34 years of age | 23.27 (6.74) | - | 13.89 (4.81) |
| (preliminary data) | At the end of a 30-minute long testing session | 165 | adults 18-34 years of age | 22.19 (7.17) | - | 13.41 (5.04) |
| **Beginning of an experiment** | 6-minute animation (Wastewater, both conditions) | 37+37 | high school | 28.60 (6.14) | - | 16.97 (5.63) |
| (Brom et al., 2017) | 6-minute animation (Wastewater, both conditions) | 37+37 | university | 27.74 (6.28) | - | 14.55 (3.69) |

*Table 8. Correlations between affective-motivational and learning outcome variables (Pearson correlation coefficients).*

<table>
<tr><th rowspan="3">Experiment</th><th rowspan="3">Type</th><th rowspan="3">N[a]</th><th rowspan="3">No. of groups</th><th rowspan="3">Mean Age (SD)</th><th colspan="2">Enjoyment</th><th colspan="2">Generalized positive affect</th><th colspan="2">Flow</th></tr>
<tr><th>Immediate</th><th>Delayed</th><th>Immediate</th><th>Delayed</th><th>Immediate</th><th>Delayed</th></tr>
<tr><th>Reten. | Trans.</th><th>Reten. | Trans.</th><th>Reten. | Trans.</th><th>Reten. | Trans.</th><th>Reten. | Trans.</th><th>Reten. | Trans.</th></tr>
<tr><td>Beer Brewing<br>(Brom, Bromová, et al., 2014)</td><td>Value-added, personalization principle</td><td>75 (70)</td><td>2</td><td>22.1 (2.3)[b]</td><td>.45*** | .37 ***</td><td>.40*** | .22†</td><td>.27* | .33**</td><td>.32** | .15</td><td>.41*** | .45 ***</td><td>.40*** | .37***</td></tr>
<tr><td>Beer Brewing – gamified<br>(submitted)</td><td>Value-added, gamification</td><td>98 (97)</td><td>3</td><td>23.1 (2.53)</td><td>.10 | .13</td><td>.11 | .05</td><td>.23* | .18†</td><td>.25* | .18†</td><td>.31** | .27 **</td><td>.40*** | .30**</td></tr>
<tr><td>Beer vs. Citrate<br>(submitted)</td><td>Value-added, topic interest</td><td>65 (64)</td><td>2</td><td>23.6 (3.75)</td><td>-.09 | -.06</td><td>-.00 | -.04</td><td>-.06 | .01</td><td>.05 | .05</td><td>.31* | .33**</td><td>.31* | .30 *</td></tr>
<tr><td>Animations[c] (Brom et al., 2017)</td><td>Value-added, personalization principle</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>- Lightning formation (university)</td><td></td><td>57</td><td>2</td><td>22.2 (2.7)</td><td>.03 | -.20</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>- Lightning formation (high school)</td><td></td><td>73</td><td>2</td><td>17.3 (0.7)</td><td>.19 | .12</td><td>-</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>- Wastewater treatment (university)</td><td></td><td>74</td><td>2</td><td>22.1 (2.5)</td><td>.16 | .08</td><td>-</td><td>-.14 | .04<br>[.19 | .00][d]</td><td>-</td><td>.25* | .36*</td><td>-</td></tr>
<tr><td>- Wastewater treatment (high school)</td><td></td><td>74</td><td>2</td><td>17.1 (0.9)</td><td>.14 | .11</td><td>-</td><td>.05 | .18<br>[.28* | .10][d]</td><td>-</td><td>.06 | .27*</td><td>-</td></tr>
<tr><td>Europe 2045 (Brom, Šisler, et al., 2016)[e]</td><td>Media comparison, DGBL</td><td>325 (287)</td><td>3</td><td>16.8 (2.1)</td><td>.16**</td><td>.30***</td><td>.26***</td><td>.42***</td><td>.21***</td><td>.30***</td></tr>
<tr><td>- Europe digital game</td><td></td><td>n = 103 (93)</td><td>subgroup</td><td>-</td><td>.01</td><td>.18</td><td>.04</td><td>.37***</td><td>.12</td><td>.30**</td></tr>
<tr><td>- Europe non-digital game</td><td></td><td>n = 96 (84)</td><td>subgroup</td><td>-</td><td>.13</td><td>.43***</td><td>.32**</td><td>.48***</td><td>.13</td><td>.23*</td></tr>
<tr><td>- Europe discussion</td><td></td><td>n = 126 (110)</td><td>subgroup</td><td>-</td><td>.18†</td><td>.21*</td><td>.30**</td><td>.32**</td><td>.26**</td><td>.25*</td></tr>
<tr><td>Animal Training<br>(Brom et al., 2011)</td><td>Media comparison, DGBL</td><td>100 (100)</td><td>2</td><td>16.0 (0.9)</td><td>.24*, .05[f, g]</td><td>0.18†, .12[f, g]</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Animal Training, Genetics (Brom et al., 2015)</td><td>Media comparison, DGBL</td><td>166 (166)</td><td></td><td>16.8 (0.6)</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>- Animal Training</td><td></td><td>n = 93 (93)</td><td>2</td><td>-</td><td>.08 | .06</td><td>.14 | .13</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>- Genetics</td><td></td><td>n = 73 (73)</td><td>2</td><td>-</td><td>.28* | .03</td><td>.23† | .13</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
</table>

Notes: The purpose of this table is only to provide a provisional look at the studied relationships; the table is not intended to be a meta-analytical synthesis. [a]Number of participants attending the delayed testing session in brackets. [b]Also one 40-year-old outlier (not counted in the average) was included. [c]No delayed testing session. [d]Correlations with the difference between generalized positive experiment prior to experiment and immediately after the experiment…

*(the notes continue on the next page)*

*Table 8 (continuation). Correlations between affective-motivational and learning outcome variables (Pearson correlation coefficients).*

*(Notes continuation)*

…in square brackets. [e]The Europe2045 experiment did not distinguish between retention and transfer tests but between four different types of tests. Here, correlations with the overall score are presented. [f]The animal training experiment featured two types of test questions based on the following topics: (a) general animal training knowledge, b) positive reinforcement knowledge, practiced in the game; the correlations are given in this order). Both types were a mixture of retention and transfer test questions. [g]Spearman correlation coefficients. The majority of learners selected "3" or "4" (4-point Likert scale).

†$p < .10$ *$p < .05$ **$p < .01$ ***$p < .001$ (not corrected for multiple comparisons)

## 5.2. Digital Game-Based Learning

Digital games can be used for many purposes in training activities. One of them is for reinforcing and integrating knowledge (cf. Thomas & Hooper, 1991) acquired from an expository lecture or a text. Does usage of games for this purpose enhance learning, when compared to reinforcing and integrating knowledge acquired from the same expositions through a more "traditional" learning activity? The argument for games is that a "traditional" instructional format, whatever it is, would incur lower active cognitive participation (because it offers fewer affective-motivational incentives compared to games). The argument against games is that, albeit more motivating, they can direct learners' attention away from learning: reducing the amount of cognitive activity devoted to meaningful learning. This is the classical emotional design trade-off. As said (in Chapter 2), it is difficult enough to measure the level of active cognitive processing, let alone distinguish between cognitive activity devoted to meaningful learning vs. that used to accommodate extraneous cognitive load. Still, the question can be approached by measuring affective-motivational and learning gain variables. Positive findings would imply that the following statement is a promising emotional design principle: "people integrate and reinforce knowledge gained from an exposition better from games than from less stimulating and less enjoyable non-game formats".

We examined whether game-based learning improves learning outcomes when used for the purpose of reinforcing and integrating knowledge acquired from an expository lecture or a text. Based on two media-comparison experiments with mainly high school samples ($N$ = 100, 325), which compared game-based learning to a "traditional" type of learning, I can conclude that games used for this purpose indeed modestly enhance learning as concerns long-term learning outcomes (though the immediate learning benefits may not be detectable). In the larger of these experiments, the effect of the game on learning outcomes was partly mediated by affective-motivational variables. This means that the principal game mechanics used in that game (i.e., a mild competition and a lightweight, team role playing) are promising emotional design elements.

In the first study (Brom et al., 2011), high school students were exposed, in a classroom, to a lecture on the topic of animal training. Afterwards, part of the knowledge learnt from that lecture was reinforced and integrated either by playing an animal training mini-game over 20 minutes or by receiving a comparable 20-minute lecture with the same objective. In the game, students trained a virtual dog to wave one front paw based on a verbal stimulus (Figure 7). To this end, they had to combine the various animal training methods they were introduced to in the lecture. In the control condition, the teacher (i.e., myself) explained how to train the dog to do the same task using slides that contained primarily the in-game graphics and a supplementary video. As concerns knowledge practiced in the game, no statistically significant between-group difference was detected in immediate learning outcomes[18], but a moderate-to-large significant effect of the game on delayed learning outcomes was observed. Enjoyment was measured with one 4-point Likert item and no significant between-group difference was detected. Flow and generalized positive affect were not measured in this study.


*Figure 7. A screenshot from the Animal Training mini-game.*

This study started in 2009 and it is my oldest educational research project. It had a couple of issues. Most notably, I served as a lecturer, which could cause confirmation bias. Although it is hard to imagine how I could have influenced long-term, but not immediate, learning outcomes, this is definitely a limitation of the study. Also, trying to use one 4-point Likert item as a proxy to the broad spectrum of affective-motivational incentives was not the best idea. Indeed, the item's sensitivity was limited: over 90% of learners selected "3" or "4" (i.e., "good" or "very good"). On the positive side, the study probably succeeded in creating a reasonably comparable game-replacement for the control condition, which is a thorny issue for media-comparison studies. The groups were at least equalized with respect to time-on-task and the core information given during these 20 minutes. It can be argued that what the study actually manipulated was the level of active cognitive processing during the last 20 minutes (presumed to be higher in the game-group).

[18] Both studies introduced in this chapter used tests that combined retention and transfer type questions.

The second study (Brom, Šisler, et al., 2016) extended this finding into a different domain and addressed the most notable limitations. During a half-day workshop in a lab, students (in teams of 6-10 members) learnt from expository texts about certain European Union topics. After reading the texts, this knowledge was integrated and reinforced via a debate-based educational method, which was derived from an educationally successful method called academic controversies (Johnson, Johnson, & Smith, 1996). The reading-debating procedure was repeated several times. The whole method was used either within a) the *Europe 2045* computer game (Brom, Šisler, & Slavík, 2010; Figure 8) I had co-developed in the past, b) a very similar game played without computers, and c) a non-game workshop. The principle game mechanics were a light-weight, team role playing and a mild competition. We ran 16 different workshops. We used a pool of eight teachers, three of whom were randomly assigned to groups at the beginning of each workshop (or two when not enough students arrived to form three subgroups). Also, teachers served mainly as discussion moderators. The learning content was delivered via expository texts and by debating with peer learners. We also measured learners' flow and generalized positive affect (and enjoyment with a 6-point Likert item[19]).


*Figure 8. A screenshot from the Europe 2045 game.*

The results were as follows. First, no notable differences between the two game conditions were detected, suggesting that the delivery media (i.e., presence/absence of computers) made no difference. This is in fact an unsurprising finding (cf. Clark, 2012). The

[19] The paper (Brom, Šisler, et al., 2016) does not report on enjoyment; only on generalized positive affect, flow, and learning outcomes.

letter "D" in the ADMLM acronym make no difference, provided the educational materials and methods are the same (i.e., when no extra affordances of computers are used).

Second, both games induced comparably higher generalized positive affect and flow, and they were liked more by learners. Immediate learning outcomes were slightly better in game groups and delayed learning outcomes were moderately better. Generalized positive affect, flow and enjoyment correlated with learning outcomes; with the strongest correlation having been observed for generalized positive affect and the weakest for enjoyment (Table 8). Generalized positive affect was found to mediate partly the effect of games on learning outcomes. These findings are consistent with the idea that the benefits of increased active cognitive participation (due to better affective-motivational incentives in game conditions) counterbalanced the possible negative effects of a higher extraneous cognitive load (which could be comparably higher in the game conditions). This interpretation must be treated cautiously, because the crucial measures – of the level of active cognitive processing and extraneous cognitive load – were lacking. Still, the second study's results better support this interpretation compared to the animal training game study (i.e., Brom et al., 2011) because of better affective-motivational measures.

When looking at correlations within the three groups[20], they were not always aligned with the net correlation (i.e., with respect to the whole sample). This suggests that the net correlation could have been caused partly by between-groups differences in correlated variables and/or that the treatment type moderated the affect-motivational–learning relationship. This experiment, out of all the experiments reported here, also triggered the highest generalized negative affect, one of the lowest generalized positive affect levels, and the lowest flow (Table 7). This makes sense because the experience required team-based interaction, which we showed to be somewhat stressful for male participants with higher social interaction anxiety (Brom, Buchtová, et al., 2014).

To conclude, on the practical level, the findings indicate that the principal game mechanics worked well as emotional design elements. In the animal training mini-game, the key game elements were actually just the presence of a clear game goal within a simulation environment and visibility of progress in terms of the increased performance of the virtual dog. In the *Europe 2045* game, the key mechanics were more elaborate: the light-weight, team role playing and the mild competition (which implicitly included a goal and the visibility of progress). This interpretation is provisional though. As the next step, the benefits of the indicated game mechanics should be tested using a value-added approach.

On a theoretical level, one may wonder why these mechanics worked. In the case of the animal training minimalistic approach, it is quite plausible that the game triggered a comparably higher level of active cognitive processing because the control group was just passively exposed to a lecture. I will return to this idea in Chapter 5.3 referencing another

[20] In my opinion, this is the only study described here in which it makes sense to look at correlations within individual groups (due to the sample size).

experiment with the animal training game. The control condition in the *Europe 2045* experiment was organized around debates and thus likely demanded higher active cognitive participation from the learners compared to a passive lecture from the animal training experiment, but the game mechanics of *Europe 2045* were more elaborate than in the animal training mini-game and could thus have triggered even higher levels of active cognitive processing (compared to the control condition in the *Europe 2045* experiment). Some participants might have been nervous due to the team-based nature of the game and increased level of interaction. However, on average, this limitation did not counterbalance the positive effect of elevated affective-motivational states.

On a methodological level, the findings from these two studies further support the idea that measuring flow levels and generalized positive affect with 10-item instruments is a better idea than measuring enjoyment with one item. Methodologically, the strength of the studies also lies in their measurement of delayed knowledge outcomes. Wouters and colleagues reported in their DGBL meta-analysis (2013) total sample size $N$ = 5,547, but only $n$ = 499 learners received delayed knowledge tests (including 100 from Brom et al., 2011). As for Sitzmann's meta-analysis (2011), the total sample size was $N$ = 6,476; whereas, only $n$ = 824 learners received delayed knowledge tests (the overlap between these two meta-analyses was only $n$ = 60 as concerns delayed knowledge assessments). Even though there are newer DGBL studies also assessing knowledge after a delay (e.g., McLaren et al., 2017), at the time when my studies started, there was quite limited information about the impacts of digital games on delayed learning outcomes. A point requiring further research is that, unlike our studies, none of the meta-analyses reported higher overall effect of games on delayed learning outcomes compared to immediate learning outcomes.

## 5.3. Active Watching of Game-Playing

One might wonder if learning by game playing would be enhanced, should the game be compared to an activity that can be theorized to trigger a high level of active cognitive processing.

> We examined whether the minimalistic approach to game playing (used in the animal training mini-game) also outperforms an activity presupposed to trigger a higher level of active cognitive processing (i.e., higher than listening to an extra-lecture triggers). After an expository lecture, high school students (N = 166) either played a mini-game individually at computers (as in the animal training experiment: Brom et al., 2011), or the teacher played it while showing it to the class on a projector and prompting the students on how to proceed with the game. The results indicated that the two methods of play were nearly comparable as concerns immediate and one-month-delayed learning gains.

Participants in this study (Brom et al., 2015) were exposed, after a lecture, to one of two mini-games; either the animal training game or a game on Mendelian genetics, in which learners explored this subject by practicing breeding of animals (Novak & Wilensky, 2007; Figure 9). Half of the students then played the game themselves as in the previous study (i.e., Brom et al., 2011), each sitting at one computer; whereas, the other half remained in the classroom, where the teacher played the game for them, demonstrated the outcomes on the projector, and prompted students on how to proceed with the game (with the same time allotment). The theoretical idea behind this design was that, although individual play seemed to be more engaging than passively looking at the game play, not all learners playing individually would devote their cognitive capacity to learning; whereas, the teacher's prompts could cognitively engage a fraction of the learners during the teacher-led game play. This idea was based on our observation from the previous study (i.e., Brom et al., 2011) and also supported by our informal observations from the present experiment (e.g., several students playing the game individually always clicked their way through the game, arranged animals on the screen, etc.). Also, when the teacher addressed the whole class in the individual play condition, he typically distracted learners away from game playing, leading to attention switching and possibly increasing extraneous cognitive load (unlike in the teacher-led play condition). On average, we reasoned that the learners' cognitive activity devoted to meaningful learning might be similar in the two conditions.

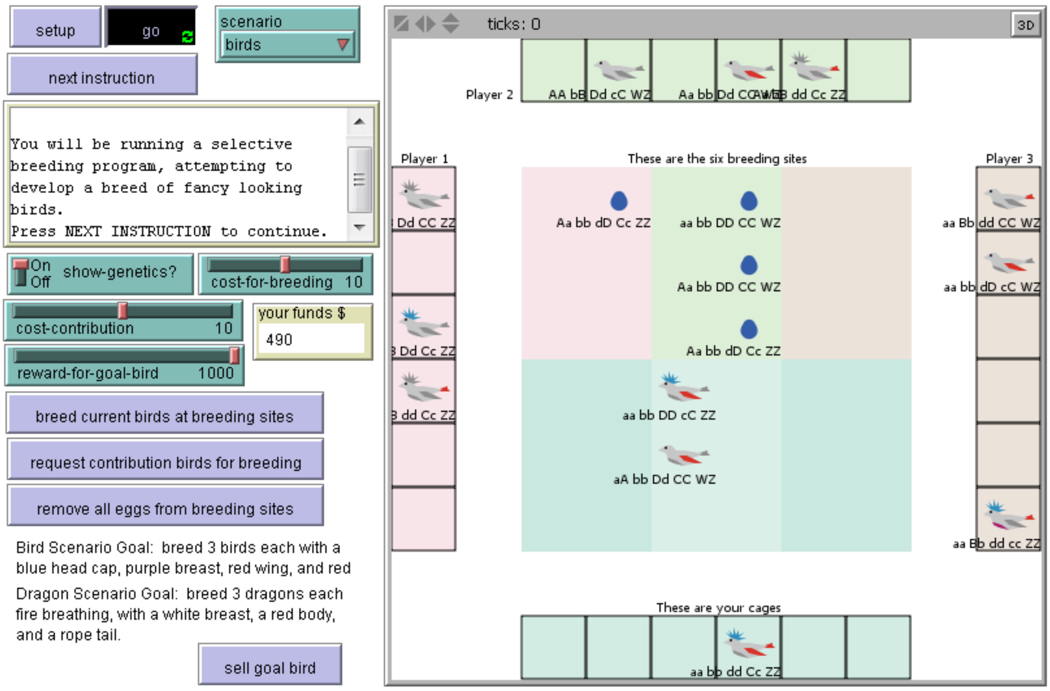


*Figure 9. A screenshot from the Bird Breeder mini-game (with permission of Uri Wilensky).*

The results were as follows. No notable between-group differences in learning outcomes were found: out of eight comparisons of learning outcome variables, only one was marginally significant with medium effect size (in favour of individual game play in the case of the bird breeding game; for immediate retention test). This was an older study, so again it used only one enjoyment question (now with a 6-point Likert scale). The enjoyment–learning correlations were positive, but generally negligible: one out of eight was significant (Table 8). For the bird breeding game, enjoyment was significantly higher for the individual play with small-to-medium effect size. No significant difference in enjoyment was detected for the animal training game. I again served as the sole teacher, which means that confirmation bias was possible; but then again, it would not be easy to influence the study unconsciously so that null results were produced.

This study's practical implication is that it may be reasonably efficient to implement mini-games in schools using the teacher-led play mode, which is substantially easier to realize than moving the whole class to the computer lab. On the theoretical level, the study points at the importance of cognitive activity devoted to meaningful learning. On the methodological level, this study indicates that one really must be careful to what he or she compares the experimental treatment (and one should avoid overgeneralization of the findings). This study also serves as a useful reminder that one cannot say if an emotional design element works in general, but only in comparison to something else.

## 5.4. Gamification

Gamification refers to the usage of game design elements in non-game contexts (Deterding et al., 2011). One of these contexts is education (Dicheva et al., 2015). Unlike the DGBL approach, which focuses on the usage of "entire" games, the gamification approach puts emphasis only on individual game elements.

Results described in Chapter 5.2 and 5.3 are consistent with the idea that learning by playing educational games can be more appealing and, therefore, promote better learning outcomes than a "traditional" instructional approach, but only if the "traditional" approach cognitively engages learners less (and the game play is not overly distracting from the learning process). One could start to wonder if some *non-game* instructional approaches might not offer comparable affective-motivational incentives as games do and thereby cognitively stimulate learners to the same extent. Such approaches could also be promising from the emotional design perspective.

One can notice that the beer brewing simulation and both instructional animations (described in Chapter 5.1) triggered higher generalized positive affect and flow compared to the *Europe 2045* game (Table 7). This could be an artifact: the measurement instruments may work just a bit differently in these contexts. However, this could also be because learning from the *Europe 2045* game actually was less stimulating than learning from the simulation or the animations (e.g., because of the team-based nature of *Europe 2045*; see

Chapter 5.3). This brings us to the following question: if we add some game-based elements to these simulation/animation treatments, i.e., if we gamify them, would that further enhance learning?

| |
|---|
| I examined whether adding several game design elements to the beer brewing simulation (particularly a game goal, increased freedom of choice, points, virtual currency and praise – all combined) enhances learning via eliciting more affective-motivational incentives (*N* = 98; university learners). No difference between the gamified simulation and the two non-gamified versions was detected as concerns affective-motivational variables and learning outcomes. |

The study compared the gamified simulation to its two non-gamified versions from the personalization principle experiment (Chapter 5.1; Brom, Bromová, et al., 2014). One of these non-gamified versions used instructional texts in a conversational style and the other in a formal style. The set of employed game design elements was a minimal one in the sense that removing any single element would undermine the believability of the treatment (in practice, it is not always useful to add just one game design element, because the resulting simulation could look artificial).

This study has not yet been published and I will thus not discuss the results here in detail. The key points include the following: the gamified simulation was perceived to be significantly easier than the non-gamified versions ($\eta_p^2 = 0.10$), but no significant difference was found regarding enjoyment, flow, generalized positive affect, and learning outcomes ($\eta_p^2 \leq 0.05$). As for the affective-motivational–learning link, the pattern with the weakest correlations for enjoyment and strongest for flow once again emerged (Table 8).

What caused the null results? It is possible that the combination of the game design elements just did not work well in this context. However, it is also possible, as I speculated above, that the simulation worked well even without the game-based embellishments (i.e., these embellishments had nothing to add). Both interpretations are compatible with other null/mixed results regarding use of gamification elements in educational contexts (see Chapter 1.2). One can also object that our affective-motivational variables exhibited a ceiling effect. I also have data pertaining to flow levels and generalized positive affect reported by actors participating in a 15-minute experiential simulation of a car accident (Table 7; Brom, Buchtová, et al., 2014)[21]. One the one hand, these actors reported the highest levels of both variables out of all my experiments. One the other hand, differences between the actors and beer brewing learners were negligible to modest. Thus, the ceiling effect may be partly to blame, but it is likely not the whole story.

[21] This non-computerized simulation took place as part of first-aid training courses organized by ZDrSEM (http://www.zdrsem.cz). The ZDrSEM group stems from the Vacation School Lipnice (VSL), a civic organization promoting experiential pedagogy in the Czech Republic.

## 5.5. Motivating Topic

So far, the beer brewing simulation seems to be an effective learning tool. Can its effectiveness somehow be compromised; both in affective-motivational and learning outcome terms? If so, the missing element could be viewed as a promising emotional design element.

It is known that topic interest affects learning from instructional texts (e.g., Fulmer D'Mello, Strain, & Graesser, 2015; Schiefele, 1999; Tobias, 1994). There is also considerable evidence that topic interest, and related task-value and relevance beliefs, play an important role in achievement settings (e.g., Eccles & Wigfield, 2002; Keller, 2010). The beer brewing topic is probably interesting for Czech learners on average. Would learning be compromised with a different, less interesting topic, particularly citrate substrate preparation[22]?

If so, this would mean that the statement "people learn better from instructional simulations when the topic is inherently motivational" can indeed be viewed as a promising emotional design principle. Given substantial support from literature on topic importance, this principle may be viewed as kind of trivial, and it would actually be surprising if we would not find the difference. However, the key point of this study is that if the difference is found (as expected), this would bring about a methodological advantage: we would have at our disposal a research method that can be used for investigating how affective-motivational variables mediate the effects of the manipulation on learning outcomes. So far, with the exception of the *Europe 2045* game (which is a quite complex treatment), none of the instruments I used could also be used for this purpose (and very few instruments in general actually generated affective-motivational incentives and, and the same time, enhanced learning – see Chapter 1.2).

To enable the drawing of meaningful conclusions, the study must manipulate just the presence of the topic: not the learning content. This might seem to be an oxymoron, but there is way to tackle this.

[22] Citrate substrate is a specific bacterial growth medium.

We examined whether learning outcomes and affective-motivational variables were both enhanced when university learners from the Czech Republic, a country where beer brewing is a source of national pride, studied how to brew beer (high intrinsic motivation) compared to studying citrate substrate preparation (low intrinsic motivation). The simulation environment was about beer brewing in both cases, with only superficial changes made to instructions and graphics to disguise topic manipulation. Learning outcomes and affective-motivational variables were higher in the beer brewing condition. Yet, only flow positively mediated the influence of the topic manipulation on immediate learning outcomes. There was no mediation by any of the variables for the delayed tests. The results suggest that affective-motivational states can be differentially related to learning. Also, there can be other mechanisms by which a topic-based intrinsic motivation manipulation influences learning.

The point of this study was that it tried to minimize between-group differences in extraneous cognitive load (which, as said in Chapter 2, is difficult to measure) so that the differences in learning outcomes could be attributed to affective-motivational differences rather than extraneous load differences. For the citrate substrate version, only the title, certain key words in the instructions (e.g., yeast – fungal culture, acetone – toxin), and two graphic icons were replaced. The structure of the to-be-learnt process, the simulation's graphics (with two minor exceptions), how learners interacted with the simulation, the amount of extraneous information, and the language style of instructional texts remained unaltered.

This study has not yet been published and I will not discuss it here in detail. I mention the core findings though, so that they can be compared to findings from my other studies (Table 7, 8). Learning outcomes (controlling for perceived prior knowledge) and affective-motivational variables were higher in the beer brewing condition with medium to large effect sizes (some effects were only marginally significant, probably because the citrate substrate simulation was not that boring after all). Crucially, the meditational analysis corroborated, yet again, the idea that flow levels are a better predictor of immediate learning outcomes than enjoyment (measured with three items in this study). This was despite the fact that the between-group difference in enjoyment ($d$ = 0.87) was higher compared to the difference in flow ($d$ = 0.55) and generalized positive affect ($d$ = 0.44, marginally significant). Yet, enjoyment was unrelated to learning outcomes ($r$ = -.09 – -.00). In fact, in this study, generalized positive affect was also unrelated to learning outcomes ($r$ = -.06 – .05). Finally, not even flow levels mediated the influence of topic on delayed learning outcomes when the immediate learning outcomes were co-varied out.

The main points of this study are a) that affective-motivational states can be differentially related to learning and b) that there can be other mechanisms by which topic-based intrinsic motivation manipulation influences learning. These points should not be forgotten in future emotional design studies.

## 5.6. Summary

In this thesis, I asked two general questions:

1) Do the researched emotional design elements improve learning by eliciting affective-motivational incentives?
2) Do learners who are stimulated by and experience enjoyment from the educational activity learn better (across all studies)?

For the first question, the answer is:

- the personalization principle does not seem to work in the Czech context, especially not for university learners (Brom, Bromová, et al., 2014; Brom et al., 2017; "gamification" study);
- using games for integrating and reinforcing knowledge acquired from an expository lecture or a text is a promising emotional design approach (Brom et al., 2011; Brom, Šisler, et al., 2016);
- but mini-games can be equally effective at schools no matter whether played by learners individually or by the teacher along with the whole class "collectively" (Brom et al., 2015);
- a cognitively-optimized, computerized simulation, gamified by adding a game goal, increased freedom of choice, points, virtual currency and praise (all combined), may not be better than the non-gamified, but still cognitively optimized, simulation alone ("gamification" study);
- manipulating interestingness of topics ("topic" study) works in the context of computerized simulations as expected by motivational theories; our findings corroborate earlier results on topic interestingness from different domains (e.g., learning from texts).

As for the second question: what does it mean "learners who are stimulated and experience enjoyment"? If "stimulation" is considered to be measured primarily by flow, the answer is *yes, modestly* (Table 8; median $r$ = .30; minimal $r$ = .06; maximal $r$ = .45). If "enjoyment" is considered to be measured primarily by my 1-3 Likert-type enjoyment questions, the answer is less clear (Table 8; median $r$ = .13; minimal $r$ = -.20; maximal $r$ = .45). Because the enjoyment measure comprises so few items, it can be criticized for insufficient sensitivity (as apparent in the experiment by Brom et al. (2011), but not in the submitted "topic" study or *Europe 2045* study (Brom, Šisler, et al., 2016)). It can be that the likability–learning relationship would be more robust and positive if measured differently, such as by intrinsic motivation or a situational interest questionnaire (see Table 3). However, one argument supports the idea that the likability–learning link may be generally weaker than the flow–learning link; notwithstanding the measure: generalized positive affect can be viewed as a concept in between enjoyment and flow and the strength of its relationship to learning outcomes was in the middle (Table 8; median $r$ = .205; minimal $r$ = -.14; maximal $r$ = .48).

This answer to the second question should be treated as highly provisional though. For instance, despite the fact that only flow mediated the influence of topic on immediate learning gains in the "topic" study, flow was not that good a predictor of learning as generalized positive affect was in the *Europe 2045* study (Brom, Šisler, et al., 2016). This difference could be due to background noise, but it could also have real substance: affective-motivational variables can be differentially related to performance variables in different contexts. Other variables (for instance, the level of extrinsic motivation) can play a moderating role. More light could be shed on this question if the provisional report (from Table 8) were expanded to incorporate studies from different labs and used a proper meta-analytical approach.

# Part B Articles

## 6. Personalized Messages in a Brewery Educational Simulation: Is the Personalization Principle Less Robust than Previously Thought?

Cyril Brom
Edita Bromová
Filip Děchtěrenko
Michaela Slussareff – Buchtová
Martin Pergel

Published in the *Computers & Education*, Vol. 72, 2014, pp. 339-366
DOI: 10.1016/j.compedu.2013.11.013.
Accepted 29 November 2013

Impact Factor (2014): 2.556
Impact Factor Without Journal Self Cites (2014): 2.005

# 7. The Role of Cultural Background in the Personalization Principle: Five Experiments with Czech Learners

(Preprint)

Cyril Brom
Tereza Hannemann – Selmbacherová
Tereza Stárková
Edita Bromová
Filip Děchtěrenko

Accepted for publication in *Computers & Education* (2017)
DOI: 10.1016/j.compedu.2017.01.001.
Accepted 1 January 2017

Impact Factor (2015): 2.881
Impact Factor Without Journal Self Cites (2015): 2.269

# 8. Are Educational Computer Micro-Games Engaging and Effective for Knowledge Acquisition at High-Schools? A Quasi-Experimental Study

Cyril Brom
Michal Preuss
Daniel Klement

Published in the *Computers & Education*, Vol. 57, 2011, pp. 1971-1988
DOI: 10.1016/j.compedu.2011.04.007.
Accepted 19 April 2011

Impact Factor (2015): 2.881
Impact Factor Without Journal Self Cites (2015): 2.269

# 9. You Like It, You Learn It: Affectivity and Learning in Competitive Social Role Play Gaming

Cyril Brom
Vít Šisler
Michaela Slussareff – Buchtová
Tereza Hannemann – Selmbacherová
Zdeněk Hlávka

Published in the *International Journal of Computer-Supported Collaborative Learning*, Vol. 11, Issue 3, 2016, pp. 313-348
DOI: 10.1007/s11412-016-9237-3
Accepted 27 June 2016

Impact Factor (2015): 2.200
Impact Factor Without Journal Self Cites (2015): 1.000

# 10. Playing Educational Micro-games at High Schools: Individually or Collectively?

Cyril Brom
David Levčík
Michaela Slussareff – Buchtová
Daniel Klement

Published in the *Computers in Human Behavior*, Vol. 48, 2015, pp. 682-694
DOI: 10.1016/j.chb.2015.02.025
Accepted 9 February 2015

Impact Factor (2015): 2.880
Impact Factor Without Journal Self Cites (2015): 1.914

# Part C Conclusion

---

## 11. Conclusion

The seven studies introduced in this thesis provided provisional support for the idea that the more stimulating and enjoyable ADMLM experiences are, the better learning outcomes tend to be. This verdict may seem trivial because it merely confirms what can be viewed as intuitively obvious, but one should be cautious: the effect sizes are generally small to medium only. Also, cognitive stimulation is likely more important for enhancing learning than mere enjoyment. Effect sizes are not much different from what has been reported in the context of learning from texts (Schiefele, 1999) and emotional design studies (e.g., Um et al., 2012; der Meij, 2013; Plass et al., 2014). Lower correlations for likability variables are also not entirely a new thing (e.g., Plass et al., 2014).

This work also suggested that using games (that impose low extraneous cognitive load) for integrating and reinforcing knowledge acquired from an expository lecture or a text can be considered to be a promising emotional design approach. The overall results also indicated that a cognitively optimized educational simulation or animation can offer in and of itself (without additional emotional design elements) substantial affective-motivational incentives and support learning well. However, gamifying a simulation turned out not to be promising and the Czech context presented a boundary condition for the personalization principle.

These studies highlight several points that must be reflected in further research.

1. Cultural contexts should be considered; cross-cultural multimedia learning studies (let alone emotional design studies) are rarely done, yet they are much needed.

2. Related to that, present findings concern university and high school students from the capital city and above average schools; these learners were probably above average with respect to the yearly cohorts. It is not only unclear if present findings would generalize for different age groups, e.g., children, but also for less able learners of the same age. Participants' characteristics, such as prior level of scientific knowledge or age, should be considered as possible moderators.
3. The issue not always considered in emotional design research, but quite apparent in our "topic" study, is that manipulations eliciting affective-motivational incentives may influence learning through alternate mechanisms unrelated to affective-motivational states (for instance, anthropomorphisms can have attention-capturing effects). These must also be considered and studied.
4. A proper meta-analytical approach should be employed to analyze the affective-motivational–learning relationship. However, affective-motivational constructs must be treated as distinct entities, since they may be differentially related to learning outcomes. Too often, they are merged into one variable in meta-analyses.
5. The thorny issue is the lack of a measure of extraneous cognitive load and the level of active cognitive participation. Such measures are much needed.
6. Enjoyment appeared to be the weakest predictor of learning outcomes. This is theoretically justifiable, but before strong conclusions are made, the issue of the likability–learning link should be addressed with better measures than 1-3 Likert items.
7. Related to the previous two points, flow can be theorized to be more closely related to active cognitive processing than enjoyment and generalized positive affect. Present data are compatible with this idea. Yet flow did not originate as a measure of learning effort. One might wonder if it would not be possible to measure learning effort more directly; such measure could be a better proxy to the level of active cognitive processing.
8. Correlations reported here were between-subject, not within-subject. correlations. Studies with within-subject design may complement present data.
9. Emotional design principles are still few and provisional; more emotional design principles should be proposed and more studies conducted.

In conclusion; the present work shows that there are empirical reasons to believe that certain affective-motivational incentives offered by ADMLM learning experiences are connected to better learning outcomes. The hunt for emotional design principles, in the ADMLM context and beyond, should continue. Not only new emotional design elements should be tested, but also new research methods and new research audiences should be considered. I will continue along some of the suggested lines in my research.

# References


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

All, A., Castellar, E. P. N., & Van Looy, J. (2016). Assessing the effectiveness of digital game-based learning: Best practices. *Computers & Education, 92*, 90-103.

Andersen, J. F., & Withrow, J. G. (1981). The impact of lecturer nonverbal expressiveness on improving mediated instruction. *Communication Education, 30*(4), 342-353.

Baddeley, A. D., Eysenck, M., & Anderson, M. C. (2009). *Memory*: Hove: Psychology Press.

Bridgeland, J. M., DiIulio, J. J., & Morison, K. B. (2006). *The silent epidemic: Perspectives of high school dropouts. A report by Civic Enterprises in association with Peter D. Hart Research Associates for the Bill & Melinda Gates Foundation.* Retrieved from http://files.eric.ed.gov/fulltext/ED513444.pdf

Brom, C., Bromová, E., Děchtěrenko, F., Buchtová, M., & Pergel, M. (2014). Personalized messages in a brewery educational simulation: Is the personalization principle less robust than previously thought? *Computers & Education, 72*, 339-366.

Brom, C., Buchtová, M., Šisler, V., Děchtěrenko, F., Palme, R., & Glenk, L. M. (2014). Flow, social interaction anxiety and salivary cortisol responses in serious games: A quasi-experimental study. *Computers & Education, 79*, 69-100.

Brom, C., Hannemann, T., Stárková, T., Bromova, E., & Dechterenko, F. (2016). Anthropomorfic faces and funny graphics in an instructional animation may improve superficial rather than deep learning: A quasi-experimental study. *Proceedings of European Conference on e-Learning* (pp. 89 - 97).

Brom, C., Hannemann, T., Stárková, T., Bromová, E., Děchtěrenko, F. (2017) The role of cultural background in the personalization principle: Five experiments with Czech learners. *Computers & Education* (in press; DOI: 10.1016/j.compedu.2017.01.001)

Brom, C., Levčík, D., Buchtová, M., & Klement, D. (2015). Playing educational micro-games at high schools: Individually or collectively? *Computers in Human behavior, 48*, 682-694.

Brom, C., Preuss, M., & Klement, D. (2011). Are educational computer micro-games engaging and effective for knowledge acquisition at high-schools? A quasi-experimental study. *Computers & Education, 57*(3), 1971-1988. doi:10.1016/j.compedu.2011.04.007

Brom, C., Stárková, T., Lukavský, J., Javora, O., & Bromová, E. (2016). Eye Tracking in Emotional Design Research: What are its Limitations? *Proceedings of the 9th Nordic Conference on Human-Computer Interaction* (pp. 114): ACM.

Brom, C., Šisler, V., & Slavík, R. (2010). Implementing digital game-based learning in schools: Augmented learning environment of 'Europe 2045'. *Multimedia Systems, 16*(1), 23-41. doi:10.1007/s00530-009-0174-0

Brom, C., Šisler, V., Slussareff, M., Selmbacherová, T., & Hlávka, Z. (2016). You like it, you learn it: affectivity and learning in competitive social role play gaming. *International Journal of Computer-Supported Collaborative Learning, 11*(3), 313-348.

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 53-61.

Brünken, R., Seufert, T., & Paas, F. (2010). Measuring cognitive load *Cognitive load theory* (pp. 181-202).

Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning a systematic review and meta-analysis. *Review of Educational Research, 86*(1), 79-122.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational psychology review, 3*(3), 149-210.

Clark, R. E. (Ed.) (2012). *Learning from Media: Arguments, Analysis, and Evidence, Second Edition* (2nd. ed.): Information Age Publishing.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of educational psychology, 88*(4), 715-730.

Courtois, C., Montrieux, H., De Grove, F., Raes, A., De Marez, L., & Schellens, T. (2014). Student acceptance of tablet devices in secondary education: A three-wave longitudinal cross-lagged case study. *Computers in Human behavior, 35*, 278-286. doi:10.1016/j.chb.2014.03.017

Csikszentmihalyi, M. (1975). *Beyond Boredom and Anxiety*: Jossey–Bass, San Francisco, CA.

Cuban, L. (1986). *Teachers and machines: The classroom use of technology since 1920*: Teachers College Press.

D'Mello, S. (2013). A Selective Meta-Analysis on the Relative Incidence of Discrete Affective States During Learning With Technology. *Journal of educational psychology, 105*(4), 1082-1099. doi:10.1037/a0032674

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170. doi:10.1016/j.learninstruc.2012.05.003

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human-Computer Studies, 70*(5), 377-398.

Darejeh, A., & Salim, S. S. (2016). Gamification Solutions to Enhance Software User Engagement—A Systematic Review. *International Journal of Human–Computer Interaction, 32*(8), 613-642. doi:10.1080/10447318.2016.1183330

de-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education, 75*, 82-91.

de-Marcos, L., Garcia-Lopez, E., & Garcia-Cabot, A. (2016). On the effectiveness of game-like and social approaches in learning: Comparing educational gaming, gamification & social networking. *Computers & Education, 95*, 99-113.

De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: some food for thought. *Instructional science, 38*(2), 105-134.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification" *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9-15): ACM.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: a systematic mapping study. *Educational Technology & Society, 18*(3), 1-14.

Domínguez, A., Saenz-De-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380-392.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology, 53*(1), 109-132.

Erhel, S., & Jamet, E. (2013). Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education, 67*, 156-167.

Fassbender, E., Richards, D., Bilgin, A., Thompson, W. F., & Heiden, W. (2012). VirSchool: The effect of background music and immersive display systems on memory for facts learned in an educational virtual environment. *Computers & Education, 58*(1), 490-500.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.): McGraw-Hill New York.

Fulmer, S. M., D'Mello, S. K., Strain, A., & Graesser, A. C. (2015). Interest-based text preference moderates the effect of text difficulty on engagement and learning. *Contemporary educational psychology, 41*, 98-110.

Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). "Seductive Details" and Learning from Text *The role of interest in learning and development* (pp. 239-254).

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming, 33*(4), 441-467.

Giannakos, M. N. (2013). Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education, 68*, 429-439.

Ginns, P., Martin, A. J., & Marsh, H. W. (2013). Designing instructional text in a conversational style: a meta-analysis. *Educational psychology review, 25*(4), 445-472.

Guo, Y. R., & Goh, D. H. L. (2015). Affect in Embodied Pedagogical Agents: Meta-Analytic Review. *Journal of Educational Computing Research, 53*(1), 124-149. doi:10.1177/0735633115588774

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? - A Literature Review of Empirical Studies on Gamification *2014 47th Hawaii International Conference on System Sciences* (pp. 3025-3034): IEEE.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80*, 152-161.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of educational psychology, 90*(3), 414-434.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? *Educational Research Review, 6*(1), 27-54.

Heidig, S., Müller, J., & Reichelt, M. (2015). Emotional design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human behavior, 44*, 81-95.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127.

Cheng, M. T., She, H. C., & Annetta, L. A. (2015). Game immersion experience: its hierarchical structure and impact on game-based science learning. *Journal of Computer Assisted Learning, 31*(3), 232-253.

Isen, A. M., & Reeve, J. (2005). The influence of positive affect on intrinsic and extrinsic motivation: Facilitating enjoyment of play, responsible work behavior, and self-control. *Motivation and emotion, 29*(4), 295-323.

Iten, N., & Petko, D. (2014). Learning with serious games: is fun playing the game a predictor of learning success? *British Journal of Educational Technology, 47*(1), 151-163.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International Journal of Human-Computer Studies, 66*(9), 641-661.

Johnson-Laird, P. N. (1983). *Mental Models*: Hardvard University Press.

Johnson, D. W., Johnson, R. T., & Smith, K. A. (1996). *Academic Controversy. Enriching College Instruction through Intellectual Conflict.*: ASHE-ERIC Higher Education Report Volume 25, No. 3. Washington, D. C: The George Washington University, Graduate School of Education and Human Development.

Jones, N. A., Ross, H., Lynam, T., Perez, P., & Leitch, A. (2011). Mental Models: An Interdisciplinary Synthesis of Theory and Methods. *Ecology and Society, 16*(1).

Kalyuga, S. (2011). Cognitive load theory: How many types of load does it really need? *Educational psychology review, 23*(1), 1-19.

Katz, B., Jaeggi, S., Buschkuehl, M., Stegman, A., & Shah, P. (2014). Differential effect of motivational features on training improvements in school-based cognitive training. *Frontiers in human neuroscience, 8*, 242. doi:10.3389/fnhum.2014.00242

Ke, F. (2008). Computer games application within alternative classroom goal structures: cognitive, metacognitive, and affective evaluation. *Educational Technology Research and Development, 56*(5-6), 539-556.

Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*: New York: Springer.

Kiili, K., & Lainema, T. (2008). Foundation for measuring engagement in educational games. *Journal of Interactive Learning Research, 19*(3), 469-488.

Kirschner, P. A., & van Merrienboer, J. J. G. (2013). Do Learners Really Know Best? Urban Legends in Education. *Educational Psychologist, 48*(3), 169-183. doi:10.1080/00461520.2013.804395

Lazowski, R. A., & Hulleman, C. S. (2016). Motivation Interventions in Education A Meta-Analytic Review. *Review of Educational Research, 86*(2), 602-640.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of educational psychology, 97*(2), 184-196. doi:10.1037/0022-0663.97.2.184

Leppink, J., Paas, F., Van Gog, T., van Der Vleuten, C. P., & Van Merrienboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction, 30*, 32-42.

Livingstone, S. (2012). Critical reflections on the benefits of ICT in education. *Oxford Review of Education, 38*(1), 9-24. doi:10.1080/03054985.2011.577938

Magner, U. I., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction, 29*, 141-152.

Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.): Cambridge University Press.

Mayer, R. E. (2014a). *Computer Games for Learning: An Evidence-Based Approach.*: The MIT Press.

Mayer, R. E. (2014b). Principles based on social cues in multimedia learning: Personalization, voice, image, and embodiment principles *The Cambridge handbook of multimedia learning* (pp. 345-368).

Mayer, R. E. (Ed.) (2014c). *The Cambridge handbook of multimedia learning* (2nd ed.): Cambridge University Press.

Mayer, R. E., & Estrella, G. (2014). Benefits of emotional design in multimedia instruction. *Learning and Instruction, 33*, 12-18.

Mayer, R. E., Johnson, W. L., Shaw, E., & Sandhu, S. (2006). Constructing computer-based tutors that are socially sensitive: Politeness in educational software. *International Journal of Human-Computer Studies, 64*(1), 36-42.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research quarterly for exercise and sport, 60*(1), 48-58.

McLaren, B. M., Adams, D. M., Mayer, R. E., & Forlizzi, J. (2017). A Computer-based Game that Promotes Mathematics Learning More than a Conventional Approach. *International Journal of Game-Based Learning (IJGBL), 7*(1), 36-56.

Moneta, G. B. (2012). On the measurement and conceptualization of flow. *Advances in flow research* (pp. 23-50): Springer.

Moreno, R. (2005). Instructional technology: Promise and pitfalls. *Technology-based education: Bringing researchers and practitioners together* (pp. 1-19): Information Age Publishing.

Moreno, R., & Mayer, R. E. (2000a). A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of educational psychology, 92*(1), 117.

Moreno, R., & Mayer, R. E. (2000b). Engaging students in active learning: The case for personalized multimedia messages. *Journal of educational psychology, 92*(4), 724-733.

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19*(2), 177-213.

Norman, A. D. (2004). *Emotional design: Why we love (or hate) everyday things*: Basic Books.

Novak, M. and Wilensky, U. (2007). NetLogo Bird Breeder model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL, available: http://ccl.northwestern.edu/netlogo/models/BirdBreeder [accessed 25 Feb 2017].

Oblinger, D., & Oblinger, J. L. (2005). *Educating the Net Generation*: EDUCAUSE.

Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human behavior, 44*, 267-278.

Park, B., Knörzer, L., Plass, J. L., & Brünken, R. (2015). Emotional design and positive emotions in multimedia learning: An eyetracking study on the use of anthropomorphisms. *Computers & Education, 86*, 30-42.

Passey, D., Rogers, C., Machell, J., & McHugh, G. (2004). *The motivational effect of ICT on pupils (Research Report RR523)*: Department of Educational Research, Lancaster University.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational psychology review, 18*(4), 315-341.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic Emotions and Student Engagement *Handbook of Research on Student Engagement* (pp. 259-282): Springer Science+Business Media.

Plass, J. L., Heidig, S., Hayward, E. O., Homer, B. D., & Um, E. (2014). Emotional design in multimedia learning: Effects of shape and color on affect and learning. *Learning and Instruction, 29*, 128-140. doi:10.1016/j.learninstruc.2013.02.006

Plass, J. L., & Kaplan, U. (2015). Emotional design in digital media for learning *Emotions, technology, design, and learning* (pp. 131-162): Academic Press.

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The Impact of Individual, Competitive, and Collaborative Mathematics Game Play on Learning, Performance, and Motivation. *Journal of educational psychology, 105*(4), 1050-1066. doi:10.1037/a0032688

Prensky, M. (2001). *Digital Game-Based Learning*: McGraw-Hill.

Renkl, A., & Scheiter, K. (2015). Studying visual displays: How to instructionally support learning. *Educational psychology review*, 1-23.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review, 7*(3), 216-237.

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). Die Erfassung des Flow-Erlebens [in German]. In J. Steinsmeier-Pelster & F. Rheinberg (Eds.), *Diagnostik von Motivation und Selbstkonzept* (pp. 261-279): Hogrefe.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology, 25*(1), 54-67.

Sabourin, J. L., & Lester, J. C. (2014). Affect and Engagement in Game-BasedLearning Environments. *IEEE Transactions on Affective Computing, 5*(1), 45-56.

Sandberg, J., Maris, M., & Hoogendoorn, P. (2014). The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education, 76*, 119-130.

Schiefele, U. (1999). Interest and learning from text. *Scientific studies of reading, 3*(3), 257-279.

Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and individual differences, 8*(2), 141-160.

Schiefele, U., & Rheinberg, F. (1997). Motivation and knowledge acquisition: Searching for mediating processes. *Advances in motivation and achievement, 10*, 251-301.

Schneider, S., Nebel, S., & Rey, G. D. (2016). Decorative pictures and emotional design in multimedia learning. *Learning and Instruction, 44*, 65-73.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology, 64*(2), 489-528.

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: empirically navigating the stream of consciousness. *Annual review of psychology, 66*, 487-518.

Snetsinger, W., & Grabowski, B. (1993). Use of Humorous Visuals To Enhance Computer-Based-Instruction. *Visual Literacy in the Digital Age: Selected Readings from the Annual Conference of the International Visual Literacy Association* (pp. 262-270). Rochester, New York.

Squire, K. (2002). Cultural framing of computer/video games. *Game studies, 2*(1), 1-13.

Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education, 94*, 252-275. doi:10.1016/j.compedu.2015.11.008

Sweetser, P., & Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. *ACM Computers in Entertainment, 3*(3), Art. No. 3A.

Sweller, J. (2010). Cognitive load theory: Recent theoretical advances *Cognitive load theory* (pp. 29-47): Cambridge University Press.

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

Tabbers, H., Reurlings, R., & de Kievit, J. (2016). *Getting the message across: The learning benefit of enthusiasm in online lectures*. Paper presented at the EARLI SIG 2 meeting, Geneva.

Thomas, R., & Hooper, E. (1991). Simulations: An opportunity we are missing. *Journal of research on computing in education, 23*(4), 497-513.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37-54.

Um, E. R., Plass, J. L., Hayward, E. O., & Homer, B. D. (2012). Emotional Design in Multimedia Learning. *Journal of educational psychology, 104*(2), 485-498. doi:10.1037/a0026609

van der Meij, H. (2013). Motivating agents in software tutorials. *Computers in Human behavior, 29*(3), 845-857.

van der Meij, H. (2014). Developing and testing a video tutorial for software training. *Technical Communication, 61*(2), 110-122.

van Dijk, V. (2010). *Learning the triage procedure. Serious gaming based on guided discovery learning versus studying worked examples.* (Master thesis), University Utrecht, the Netherlands.

Vandercruysse, S., Vandewaetere, M., Cornillie, F., & Clarebout, G. (2013). Competition and students' perceptions in a game-based language learning environment. *Educational Technology Research and Development, 61*(6), 927-950.

Veen, W. (2007). *Homo Zappiens and the need for new education systems.* Paper presented at the OECD seminar on Digital Natives and Education, Florence, Italy.

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. *Journal of Educational Computing Research, 34*(3), 229-243.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational psychology review, 18*(3), 239-253.

Wang, A. I., Zhu, M., & Sætre, R. (2016). The Effect of Digitizing and Gamifying Quizzing in Classrooms *10th European Conference on Games Based Learning: ECGBL 2016* (pp. 729-737): Academic Conferences and Publishing International.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology, 54*(6), 1063-1070.

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological bulletin, 98*(2), 219-235.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & 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. doi:10.1037/a0031311

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

# Appendix A

## My Colleagues' Contributions

This work would not be possible without the contributions of my many collaborators (Table A1). I am also thankful to Tomáš Holan, Pavel Töpfer, and Edita Bromová for commenting on this text and to Filip Děchtěrnko, Zdeněk Hlávka, and Daniel Klement for conducting supplementary data analyses not presented in the papers included in this thesis. I also thank Brady Clough for proof reading my manuscripts and this commentary.

*Table A1. Contributions of individual researchers (other than myself).*

| Paper | Contribution | Names |
|---|---|---|
| **Animal Training** (Brom et al., 2011; Chapter 5.2) | Study design and preparation | Daniel Klement, Michal Preuss, František Šusta |
| | Intervention preparation | Petr Jakubíček, Daniel Klement, Veronika Mašková, Michal Preuss, František Šusta |
| | Research administration | Daniel Klement, Michal Preuss, František Šusta |
| | Data analysis | Daniel Klement, Hana Voňková |
| | Paper writing | Daniel Klement, Michal Preuss |
| | Consultations | Daniel Frynta, Jiří Lukavský |
| | Special thanks | Schools and teachers enabling us to conduct the experiment |
| | | Generation Europe CZ (http://www.generationeurope.cz/) |
| **Animal Training, Genetics** (Brom et al., 2015; Chapter 5.3) | Study design and preparation | Michaela Slussareff – Buchtová, Daniel Klement, David Levčík, Tereza Nekovářová, František Šusta |
| | Intervention preparation | *Animal Training:* Petr Jakubíček, Daniel Klement, Veronika Mašková, Michal Preuss, František Šusta |
| | | *Bird Breeder:* Novak and Wilensky (2007) |
| | Research administration | Michaela Slussareff – Buchtová, Monika Hampacherová, Daniel Klement, David Levčík, Markéta Popelová – Tomková, Tereza Hannemann – Selmbacherová |
| | Data analysis | Daniel Klement |
| | Paper writing | Michaela Slussareff – Buchtová, Daniel Klement, David Levčík |
| | Consultations | František Šusta |
| | Special thanks | Schools and teachers enabling us to conduct the experiment |
| | | Generation Europe CZ (http://www.generationeurope.cz/) |

*(the table continues on the next page)*

*Table A1 (continuation). Contributions of individual researchers (other than myself).*

| Paper | Contribution | Names |
|---|---|---|
| **Europe 2045** (Brom, Šisler, et al., 2016, Chapter 5.2) | Study design and preparation | Michaela Slussareff – Buchtová, Lisa M. Glenk, Jan L. Plass, Ivo Šebek, Vít Šisler |
| | Intervention preparation | Europe 2045's team (http://www.evropa2045.cz/kdo_jsme.php) |
| | Research administration | Miroslava Abrahámová, Lucie Bauerová, Eva Bednaříková, Michaela Slussareff – Buchtová, Martina Denemarková, Viktor Dobrovolný, Sandra Feyglová, Jakub Fiala, Marie Grecká, Monika Hampacherová, Tomáš Holan, Jakub Janovský, Lukáš Kolek, Jiří Lacík, Kamil Pavelka, Ivana Pecháčková, Tomáš Pospíšil, Radka Římanová, Tereza Hannemann – Selmbacherová, Ondřej Smíšek, Tereza Stárková, Vlaďka Šnoblová, Ondřej Šíp, Václav Šálený, Ivo Šebek, Vít Šisler, Jana Tupá, Kateřina Vávrová, Jakub Vlasák, Jiří Volák, David Wagner, Veronika Zemanová |
| | Data analysis | Zdeněk Hlávka |
| | Paper writing | Michaela Slussareff – Buchtová, Zdeněk Hlávka, Tereza Hannemann – Selmbacherová, Vít Šisler |
| | Consultations | Sidney D'Mello, Lisa M. Glenk, Jiří Lukavský, Rupert Palme, Jan L. Plass |
| | Special thanks | Schools and teachers enabling us to conduct the experiment |
| | | Institut certifikovaného vzdělávání (http://www.icv.cz) |
| | | ZDrSEM (http://www.zdrsem.cz) |
| | | Generation Europe CZ (http://www.generationeurope.cz/) |
| **Beer Brewing** (Brom, Bromová, et al., 2014; Chapter 5.1) | Study design and preparation | Edita Bromová, Michaela Slussareff – Buchtová, Martin Pergel |
| | Intervention preparation | Edita Bromová, Martin Pergel |
| | Research administration | Edita Bromová, Michaela Slussareff – Buchtová, Filip Děchtěrenko, Tereza Stárková |
| | Data analysis | Filip Děchtěrenko |
| | Paper writing | Edita Bromová, Michaela Slussareff – Buchtová, Filip Děchtěrenko, Martin Pergel |
| | Consultations | Jiří Lukavský, Jan L. Plass, Bruce M. McLaren |
| | Special thanks | Laboratory of Behavioral and Linguistic Studies (http://labels.ff.cuni.cz/) |
| **Animations** (Brom et al., 2017; Chapter 5.1) | Study design and preparation | Edita Bromová |
| | Intervention preparation | *Lightning formation:* Edita Bromová, based on Moreno and Mayer (2000b) |
| | | *Wastewater treatment:* Edita Bromová, Kateřina Svobodová |
| | Research administration | Nela Bendová, Viktor Dobrovolný, Lukáš Kolek, Tereza Hannemann – Selmbacherová, Tereza Stárková, Ondřej Smíšek, Martina Stejskalová |
| | Data analysis | Filip Děchtěrenko |
| | Paper writing | Edita Bromová, Filip Děchtěrenko, Tereza Hannemann – Selmbacherová, Tereza Stárková |
| | Consultations | Slava Kalyuga, Richard Mayer, Kateřina Svobodová |
| | Special thanks | Laboratory of Behavioral and Linguistic Studies (http://labels.ff.cuni.cz/) |

*(the table continues on the next page)*

*Table A1 (continuation). Contributions of individual researchers (other than myself).*

| Paper | Contribution | Names |
|---|---|---|
| **Gamification Study** (submitted; Chapter 5.1, 5.4) | Study design and preparation | Edita Bromová, Tereza Stárková |
| | Intervention preparation | Edita Bromová, Martin Pergel |
| | Research administration | Edita Bromová, Karolína Brožová, Viktor Dobrovolný, Filip Děchtěrenko, Tereza Hannemann – Selmbacherová, Tereza Stárková, Martina Stejskalová |
| | Data analysis | Filip Děchtěrenko |
| | Paper writing | Edita Bromová, Filip Děchtěrenko, Tereza Stárková |
| | Consultations | -- |
| | Special thanks | Laboratory of Behavioral and Linguistic Studies (http://labels.ff.cuni.cz/) |
| **Topic Study** (submitted; Chapter 5.5) | Study design and preparation | Edita Bromová, Sidney D'Mello |
| | Intervention preparation | Edita Bromová, Martin Pergel, Kateřina Svobodová |
| | Research administration | Edita Bromová, Filip Děchtěrenko, Nikola Frollová, Kateřina Koppová, Ondřej Smíšek, Tereza Stárková, Petra Šustová |
| | Data analysis | Filip Děchtěrenko |
| | Paper writing | Edita Bromová, Filip Děchtěrenko, Sidney D'Mello, Nikola Frollová, Tereza Stárková |
| | Consultations | Kateřina Svobodová |
| | Special thanks | Laboratory of Behavioral and Linguistic Studies (http://labels.ff.cuni.cz/) |