



# Gameful Learning for a More Sustainable World

## Measuring the Effect of Design Elements on Long-Term Learning Outcomes in Correct Waste Sorting

Greta Hoffmann · Jella Pfeiffer

Received: 18 May 2021 / Accepted: 15 October 2021  
© The Author(s) 2021

**Abstract** Municipal waste sorting is an important but neglected topic within sustainability-oriented Information Systems research. Most waste management systems depend on the quality of their citizens pre-sorting but lack teaching resources. Thus, it is important to raise awareness and knowledge on correct waste sorting to strengthen current efforts. Having shown promising results in raising learning outcomes and motivation in domains like health and economics, gamification is an auspicious approach to address this problem. The paper explores the effectiveness of gameful design on learning outcomes of waste sorting knowledge with a mobile game app that implements two different learning strategies: repetition and elaboration. In a laboratory experiment, the overall learning outcome of participants who trained with the game was compared to that of participants who trained with standard analogue non-game materials. Furthermore, the effects of two additional, learning-enhancing design elements – repetition and look-up – were analyzed. Learning outcome in terms of long-term retention and knowledge transfer were evaluated through three different testing measures two weeks after the training: in-game, through a multiple-choice test and real-life sorting. The results show that the game significantly enhanced the learning outcome of waste sorting

knowledge for all measures, which is particularly remarkable for the real-life measure, as similar studies were not successful with regard to knowledge transfer to real life. Furthermore, look-up is found to be a promising game design element that is not yet established in IS literature and therefore should be considered more thoroughly in future research and practical implementations alike.

**Keywords** Gameful design · Serious game · Gamification · Game-design elements · Look-up · Repetition · Cognitive learning strategies · Sustainability

### 1 Introduction

In their set of goals for sustainable development, the UN listed targets for different areas of human and environmental wellbeing, one of which concerns waste, sustainable consumption and production (United Nations 2020). Acknowledging the insufficiency of the status quo in terms of waste management, the EU created a plan to raise EU-wide recycling to 55% and decrease landfill use to 10% by 2025 (European Parliament 2018). However, recent studies have shown that global progress is slow, partly due to a lack of appropriate legislation, insufficient financial resources, poor infrastructure, poor environmental attitudes and social norms and a lack of knowledge about what goes into which bin (Schultz et al. 1995; Thomas and Sharp 2013; Filho et al. 2016; Luo et al. 2018). A contributing factor is that many recycling and waste sorting facilities are as yet unable to reach maximum efficiency without pre-sorting measures (Buccioli et al. 2015; Hawlitschek 2020). Countries like Germany, Austria and Switzerland have tackled this issue by making domestic pre-sorting a citizen's responsibility. However, incentivizing citizens to

---

Accepted after three revisions by Stefan Stieglitz.

G. Hoffmann (✉)  
Institute for Information Systems and Marketing (IISM),  
Karlsruhe Institute for Technology, 76135 Karlsruhe, Germany  
e-mail: greta.hoffmann@kit.edu

J. Pfeiffer  
Chair of Business Administration, Spec. Digitalization,  
E-Business and Operations Management, Justus Liebig  
University Giessen, 35394 Gießen, Germany

correctly and consistently dispose of their household waste continues to be a challenge for society, as it is a task requiring individuals to perform for the benefit of society often without rewards for compliance (Abdel-Shafy and Mansour 2018). Furthermore, successful compliance requires citizens to first gain the fundamental knowledge to fulfil the required task. Yet, municipal waste sorting authorities often fail in their education attempts, partly because of outdated measures of communication and information like analogue, paper-based flyers (Luo et al. 2018). Such materials are insufficient for knowledge transmission as they lack incentives to engage mentally, particularly given the amount and depth of information that people need to retain. Of the hundreds of potential waste items, more than 200 are listed on many websites for German waste management organizations as being fundamental to sufficient municipal waste sorting (e.g., Berlin and Hamburg)<sup>1</sup> While citizens do not have to know each item by heart, they need to understand the underlying principles that link different types of waste to their respective bin. To engrain the knowledge in the long term, such extensive amounts of information require adequate training measures.

As stated in an interview on “The Future of Waste Management,” Information Systems (IS) can teach citizens where exactly to dispose of different types of waste (Hawlitshchek 2020). Multi-disciplinary research has shown that games in particular are successful educational tools and supplements (Fileni 1988; Van Eck et al. 2017). By applying gameful design to a real-life context, education can be effectively manipulated, whether as fully conceptualized games or as strategically implemented gamification affordances (Barata et al. 2013; Landers and Landers 2014). In their meta-analysis on digital games and learning, Clark et al. (2016) found significant correlations between quality of design and learning outcomes, highlighting the value of deliberation on specific design decisions.

We created a waste sorting training game based on best practices of game design as well as learning theories with the goal of addressing the prevalent lack of waste sorting knowledge. The game’s release was in 2015 and, as of April 2021, it was downloaded over 31,684 times on the Apple, Microsoft and Windows mobile app stores. As stated by Bellotti et al. (2013), a serious game’s purpose is twofold: to be fun and entertaining as well as educational. Thus, we must assess both aspects. While the field data allowed us to ascertain the game’s success in matters of game fun and engagement with a certain degree of external validity, we could not reliably infer the game’s efficacy in

terms of the intended learning outcome. As this is the game’s primary aim, we prepared a lab experiment to measure the game’s learning outcome under the following research question: *Does gameful design afford learning about correct sorting of waste items into their target bin?*

Gamification/gameful design<sup>2</sup>) is a praxis that consists of designing suitable “service bundles” (Blohm and Leimeister 2013) by adding game design elements to the respective core offer – or core gameplay when the gamified product is a game in itself. The core interaction – core gameplay – of our game is based on a combination of sorting and feedback, the latter being particularly beneficial for knowledge transfer and player engagement (Sicart 2008; Bellotti et al. 2013). However, during the development of the first prototype, our user tests found that the core gameplay by itself did not engage players long enough to benefit from a long-term learning effect. We decided to add optional design elements that would offer players more choices on how to engage with the learning content. We based this decision on motivational theories that highlight autonomy as a fundamental factor of intrinsic motivation (Ryan and Deci 2000). We first chose a repetition element that would allow players to repeat a level – or wave – without penalty. The overall learning benefits of repetition are well-documented across different learning domains (Bygate 1996; Ahmadian 2012). However, as its inherent repetitiveness could interfere with the game fun, we wanted to gain insights into the potential detriments and benefits of including such a design element. We then added an index element, where waste items can be looked up penalty-free during the core gameplay (look-up element). This was inspired by testers frequently asking why certain items were assigned to bins other than expected. We conducted a literature review to find theoretical leads on the expected learning outcome of such an index-based design element. Lacking a related foundational theory, we analyzed research on related contexts: instructional explanations, dictionaries and help tools (Miller and Gildea 1987; Ryan and Shin 2011) finding mixed expectable outcomes.

Thus, we designed our experiment in a way to further answer the second research question: *What effect does a repetition-based and a look-up-based game design element have on the learning outcomes of correct sorting of waste items into their target bin?*

The experiment consisted of five treatment groups to reflect both research questions. The first four learned to correctly sort waste items by playing the game. They

<sup>1</sup> <https://www.bsr.de/die-berliner-stadtreinigung-in-leichter-sprache-24048.php>, <https://www.stadtreinigung.hamburg/privatkunden/abfallabc/>.

<sup>2</sup> In this manuscript, we refer to both practices under the umbrella term gameful design: “affording ludic qualities or gamefulness (the experiential qualities characteristic for gameplay) in nongame contexts” (Deterding 2015). The term encompasses the practice of creating as well as research into the effects of serious games and gamification implementation.

differed with respect to whether participants played only the core gameplay, one or both design elements (repetition, look-up or combined). The fifth group completed the training with common, paper-based teaching materials on waste sorting. As we wanted to ensure long-term retention – long-term memory storage of the learned content – we measured learning outcomes 10 to 12 days after the participants had been trained. Also, while the training was conducted with a game, the learning outcome was supposed to be translated into real life. To test if participants successfully managed this knowledge transfer, we measured the learning outcome in three different ways: first, by testing knowledge retention within the training medium itself in a slightly altered version of the core gameplay designed to test each training item exactly once. Second, we measured knowledge transfer in an abstracted setting via a multiple-choice test featuring only the names (written words) of the trained items. Third, we measured knowledge transfer to real life through a sorting task with real-world waste items.

Our results showed that the treatment trained with the game significantly benefited with regard to learning outcomes of waste sorting knowledge compared to the treatment group given the non-game materials. This is especially remarkable, as, in contrast with other studies (Größler et al. 2000; Luo et al. 2018), we demonstrated that knowledge transfer to real life can be successfully achieved with a gameful application. We further found that the combination of the repetition and look-up game design elements showed significantly higher learning outcomes within the original content domain as well as the reduced setting. Interestingly, this combinatory effect was lost in transfer to real life.

We contribute to the literature and practice in several ways. While growing in numbers, assessments of the effects of specific game design elements are still rare (Bellotti et al. 2013; Kim and Shute 2015). This makes it difficult for both researchers and practitioners to derive informed design decisions from research. We identified a research gap with regard to expectable effects and outcomes of an optional look-up element and found that its implementation contributed to the learning outcome, especially when combined with a repetition-based design element. We also tested learning outcome in terms of long-term retention as well as knowledge transfer to ensure that an actual learning outcome was achieved. While our game successfully achieved the transfer of content to long-term memory, the differences of outcomes between the three different measures highlighted the importance of testing and ensuring successful knowledge transfer in multimedia contexts such as ours.

By constructing and testing a serious game on teaching correct waste sorting and detailing design rationales for

future reproduction, we contributed to the ongoing effort of enhancing sustainable IS (see e.g., Elliot and Webster (2017) and Stanitsas et al. (2019)). Our results showed that successful learning outcomes can be achieved through meticulous gameful design even in less intrinsically motivated and attractive domains such as waste management and even outside a socially mediated learning context.

## 2 Related Work

### 2.1 Empirical Findings on Gameful Design in IS Research

As early as 1991, Duffield showed that computer games provide great learning opportunities for students, as games motivate them to learn, provided that they are adequately designed and that the content of the software, problems presented and instructional methods are carefully aligned (Duffield 1991; Clark et al. 2016). Later, Rieber (2005) and Gee (2003) recommended games as potential learning tools, reasoning that gaming is a complex social practice in which players engage in high-order thinking and where they need to make a complex cognitive effort. Studies have shown that games entertain, instruct, change attitudes and enable skills development. Studies successfully correlating participants' previous gameplay experiences to related real-life skills, e.g., reaction games and driving skills (Vichitvanichphong et al. 2016) and strategy games with management skills (Simons et al. 2020), supported this finding.

In terms of application domains, we found certain topics especially prevalent, such as education (Fotaris et al. 2016; Sanmugam et al. 2016), fitness (Jang et al. 2018; Kappen et al. 2018), health (Allam et al. 2015; El-Hilly et al. 2016; Kurtzman et al. 2018) and the economy (Rodrigues et al. 2016; Hamari 2017). Most sustainability-related studies are connected to broader economic domains like sustainable transport education (Putz et al. 2018) or domestic energy engagement (Gustafsson et al. 2009). Studies solely focused on topics of sustainability are scarce, particularly with regard to sustainable waste management (Elliot and Webster 2017).

### 2.2 Gameful Design and Waste Sorting

In a literature review of serious games in the general domain of sustainability, Stanitsas et al. (2019) found that there has been a radical increase in the development of sustainability-related games since 2010. However, of the 77 listed games (starting in 1990), only two are related to waste management: a board game that teaches about industrial waste management (Jürgen Strohm 2001) and a

role-playing game that educates on irrigation management (Burton 1993). Both games provide a broad perspective on the topic but do not specifically teach about municipal waste sorting. As an extension to Stanitsas et al. (2019), we conducted an additional literature review looking for research studies with a focus on gameful design and waste sorting. In total, we found nine more research studies somewhat related to our topic (Table 1).

Eight of the nine studies did not prove entirely useful to our research efforts, as they either presented their gameful approach without actually evaluating the effectiveness of their design (Bifulco et al. 2011; Berengueres et al. 2013; González-Briones et al. 2018), evaluated their design qualitatively as a whole with a small number of users (Lotfi and Mohammed 2014; Sreelakshmi et al. 2015; Menon et al. 2017; Idrobo et al. 2018) or touched on a relevant but

adjacent topic (Whalen et al. 2018). Except for one (Luo et al. 2018), these studies did not provide insights into the effect of single design choices on learning outcomes but instead looked at their gameful implementations as a whole.

The study closest to our setup was Luo et al. (2018). In their series of experiments, the authors assessed the effect of immediate feedback as a game design element on the learning of accurate recycling and composting. In their lab experiment, they asked 100 students to sort 80 pictures of different waste items into one of the four bins shown on screen. The learning condition received feedback on the correctness of their sorting, while the control condition did not. The learning outcome was tested after one week in the same game to assess long-term retention. Results showed a positive influence of immediate feedback on the learning

**Table 1** Overview of other gameful design-based studies on waste and recycling

Authors	Sub-Domain	Digital/ Analogue	Publishing domain	Country	ql/ qt*	Study design	Study goal
Berengueres et al. (2013)	Gamification	Analogue (digitally enhanced)	Human–Robot Interaction	United Arab Emirates	ql	n not reported (two waste bins installed)	Evaluate effective system to increase usage of recycling bins
Sreelakshmi et al. (2015)	Gamification/ Game-based Learning	Digital (unity 2D)	Computing Communication and Networking Technologies	India	ql	n = 20 participants	Highlight success of game-based learning through a waste sorting game
González-Briones et al. (2018)	Gamification	Analogue (digitally enhanced)	Distributed Computing and Artificial Intelligence	Spain	ql	n not reported (30 waste bins installed)	Generate motivation for citizen participation in the recycling chain
Bifulco et al. (2011)	Serious Game	Digital (3D virtual environment)	Distributed Multimedia Systems	Italy	ql	No report of scientific testing	Present the main concepts of waste collection and garbage recycling to primary school students
Lotfi and Mohammed (2014)	Serious Game	Digital (browser-based)	Computer Applications	Morocco	ql	n = 20 participants	Help instructors and experts improve teaching strategies
Menon et al. (2017)	Serious Game	Digital (Microsoft Kinect)	Serious Games and Applications for Health	India	ql	n = 9 participants	Raise awareness of the importance of trash removal, initiate recycling programs and teach basic hygiene practices
Whalen et al. (2018)	Serious Game	Analogue (board game)	Resources, Conservation and Recycling	Sweden	ql	17, 18, 36 participants (total: n = 71)	Teach the benefits and complexity of the circular economy
Idrobo et al. (2018)	Serious Game	Digital (3D)	Telematics and Computing	Colombia	ql	n = 5 participants	Teach correct waste sorting
Luo et al. (2018)	Digital Sorting Game	Digital (2D)	Environmental Management	Canada	qt	n = 50, n = 100, n = 308	Evaluate the effect of immediate feedback on recycling and composting accuracy
Hoffmann 2021	Serious Game / Gameful Design	Digital (2D / mobile)	Information Systems	Germany	qt	n = 215	Evaluate game / game design elements affording learning of waste sorting knowledge

\*ql = qualitative, qt = quantitative

outcome. Yet, they could not replicate this result in a real-life follow-up experiment. The authors hypothesized that the reason for this null effect could be related to the logistics of accurately measuring changes in real-life waste containers. Evaluating their design artifact – as presented in the manuscript – from a game design perspective, we believe that the lack of game design elements like world-building or colorful aesthetics could also have contributed to this outcome. We came to this conclusion because user-tests of the early iterations of our game indicated that feedback alone lacked incentives to continually engage with the content.

We conclude our literature review with the insight that research on gameful design – particularly with regards to the analysis of game-specific design elements – in the domain of waste management and sustainability can benefit from further research in terms of expectable outcomes on learning. We will discuss this in the following section.

### 3 Hypothesis Development

We based the theoretical foundations of this work on learning theories, particularly on models introduced by instructional design and the didactic method (Wittwer and Renkl 2008; Nitsch et al. 2016). Most learning theories have been developed in and for contexts where social interaction is interwoven into the learning process (e.g., Chi et al. 2001). However, multimedia learning is often designed to work outside of social interaction contexts. We thus built on learning theories and strategies that have proven effective outside of a socially embedded learning context, namely elaborative encoding (for Hypothesis 1) and repetition (for Hypothesis 2). While these strategies afford an empirical learning process, rationalization (the process of sense-making, or understanding “why”) is equally essential for providing meaning and context to learning matter (Wittwer and Renkl 2008). To offer explanations without overwhelming players, we implemented an optional, dictionary-inspired look-up element as a complementary game design element and elaborated on its supposed learning outcome in the development of Hypothesis 3.

Depending on context, the learning outcome can be measured with regard to different facets. In terms of memorization, the learning outcome often differs when tested immediately after the training phase rather than during a post-training transfer phase (Keith and Frese 2008). By measuring the learning outcome in terms of *long-term retention*, we wanted to ensure that the content was memorized in the long term to achieve successful change in real-life waste sorting behavior.

Furthermore, in contexts like ours, where the training medium differed from the application context, it was particularly important to assess the learning outcome with regard to *knowledge transfer* (Barnett and Ceci 2002). In language transfer theory, knowledge transfer happens within the context of translation, where words and meanings are connected between two languages, typically native and second (Mahmood and Murad 2018). In the didactics of mathematics, this transfer is referred to as “conversion,” or the mental merging of different representations – such as graphs and functions – of the same mathematical concept (Dreher et al. 2014). Expanding on the concept of conversion, Nitsch et al. (2016) differentiate two phases of the transfer process: identification (comparison of the incorporated information and identification of similarity with existing schema) and construction (transferability of the incorporated information into new situations). Differentiating the two is important because construction measures whether the content has been understood deeply enough for reapplication in new contexts. Therefore, as later explained in the section on the operationalization of our outcome variable, we measured knowledge transfer with different measurements that capture both identification and construction.

#### 3.1 Elaborative Encoding

The overall story and game design rationale of the game is based on elaboration strategies. Elaborative encoding belongs to the category of learning techniques known as mnemonics. In this learning strategy, loosely adjacent content items are added to the learning matter. Offering more associations that may connect to the learners’ existing knowledge facilitates embedding new information into prevalent mental structures (Bradshaw and Anderson 1982). Examples of mnemonics involve meaning-enhancing additions: constructions or creations that improve one’s memory of what is learnt (Levin 1988). Mnemonics can range from acronyms and rhymes to complex strategies for remembering numbers (Putnam 2015) and character designs (such as the mascot designs commonly found in Japan)<sup>3</sup> Elaborative encoding encompasses the purposeful addition of information, whether visual, semantic, spatial or acoustic, to create more retrieval paths in the mind of the learner from existing knowledge structures to the learning matter (Bradshaw and Anderson 1982). In his meta-analysis of elaboration studies, Mayer (1980) concluded that associative elaboration “increases retention performance as

<sup>3</sup> For example: “[Morio-kun is] a vampire bat who promotes paying taxes by direct debit in Chiba, Japan. He uses direct debit because he’s nocturnal and can’t get to the bank in the daytime.” (<https://twitter.com/mondomascots/status/1020495338644230144>).

compared with control or simple repetition procedures” (p. 771). This technique is particularly valuable in the context of gameful design, as elaboration can occur on several layers at once: the game’s mechanics (rules and systems), aesthetics (visual/auditive representation) and narrative form a multi-sensory context for knowledge transmission (Hunicke et al. 2004; Westera et al. 2008). This is particularly important in serious games like Re-Mission (Hopelab 2006) that translate a specific problem context into gameplay: adolescents are incentivized to take their cancer medication on a regular basis by transferring the game setting to their own bodies and providing them the medication as ammunition against destructive cancer cells. Re-Mission has proven highly successful in improving the health outcomes of its players (Kato et al. 2008). In sum, due to their multimodal elaborative encoding of real-life activities, principles and systems, games have been found to be an effective medium for teaching. Thus, we hypothesize that:

**H1:** Learning waste sorting through a game rather than with state-of-the-art paper-based information on the correct sorting of items increases learning outcome.

### 3.2 Repetition Strategies

However, in terms of the core teaching effort – correctly sorting waste items – studies have shown that a single exposure to new content is not enough for learners to effectively encode that content into memory (Bygate 1996). Repetition is a learning activity in which students repeat individual facts to create firmly anchored connections in their long-term memory. Repetition has long been acknowledged as a powerful learning mechanism: as Horace stated more than two millennia ago, “*repetitio est mater studiorum,*” or repetition is the mother of learning. As a universal principle, it is part of all prevalent learning theories: behaviorism (e.g., in Pavlov (Dunsmoor et al. 2007) and Skinner (1936)), cognitivism (e.g., in schema theory), constructivism (e.g., in Piaget (Greenfield and Savage-Rumbaugh 1993)) and social learning (e.g., Vygotsky (1967) on child development). The underlying theme relates to the formation of memory in the brain. In their research on working memory, Baddeley and Hitch (1974) stated that by repeatedly forming mental connections through reflection and deliberate recall, the stored information gets retrieved more easily and quickly. However, different studies have shown that repeated exposure to the same content does not necessarily lead to improved learning (Crowder and Melton 1965; Nickerson and Adams 1979). Memories are formed more precisely and hold for longer in long-term memory if learners are interested in the content and pay attention (Nickerson and Adams 1979). This is where games might have an additional advantage

compared to learning content presented in a classroom setting.

In their study, Bygate (1996) found that repeating the trained content three days after the initial task led to improvement in fluency and accuracy as well as a marked improvement in repertoire due to growing familiarity with the content. The given reason is that on first contact with the material, learners are primarily concerned with the heuristic planning and understanding of the content matter (Bygate 1996). Ahmadian (2012) corroborated these findings, arguing that it is difficult for learners to focus on form and meaning at the same time. Thus, repetition allows them to gain understanding in both facets. Overall, studies on task-based language learning have reported repetition-based improvements for output factors of accuracy, complexity, repertoire and task success (Lynch and Maclean 2000; Pinter 2005). According to Driskell et al. (1992), there are even benefits to repeating the content beyond perfect retention. In their study on overlearning, they found a significant overall effect: the greater the degree of overlearning, the greater the resulting long-term retention. By raising the number of occurrences in the brain, the significance of the information is enforced and so the content is retained longer. Repetition has been proven to be an effective learning strategy in learning tasks across domains (e.g., education (Johnson 2004), civic knowledge (Ivancic and Hesketh 2000) and games (Clark and Sefton 2001)). Building on the theoretical foundations of repetition, we hypothesize that:

**H2:** Repetition as a game design element increases the learning outcome.

### 3.3 Instructional Design, On-Demand Help and Look-up Strategies

During the teaching process, one of the central functions of the teacher or tutor is to provide context and explanations to learners (Leinhardt and Steele 2005). In non-social contexts, this function must be substituted within the training medium. While there is no dedicated educational or psychological theory for this construct (providing relevant information/answering the “why” question), research on instructional explanations provides foundational insights. Empirical studies show that instructional explanations have often not been successful in terms of raising the learning outcome (Jonassen and Rohrer-Murphy 1999; Leinhardt and Steele 2005). One explanation was that learners merely engage in superficial processing of instructional explanations (Berthold and Renkl 2010) and do not attend to the content of the explanations in a meaningful manner (Roelle et al. 2014). However, the learning outcome was positively influenced through instructional explanations if learners rationally engaged

with the content of the explanations (Wittwer and Renkl 2008) and if there was a meaningful follow-up activity after receiving the explanations (Webb and Farivar 1999). The way the game is designed, each interaction with the look-up element is followed by the sorting of a waste item. Thus, in our case, the instructions can be processed meaningfully.

In self-regulated learning contexts, studies have found help-seeking to be a successful strategy for learning (Ryan and Shin 2011; Webb et al. 2013) if help-seekers were oriented on independent problem-solving (Nelson-Le Gall 1981) and if the process included asking for explanations and hints (Mäkitalo-Siegl et al. 2011). In summary, if learners are invested in the learning process, giving explanations when needed raises the learning outcome. This connection produces positive indications for the success of our look-up element. However, most studies on help-seeking are embedded in a social context: the help is provided by another person. Thus, the expected effects might be weaker outside of a social context. On the other hand, the same studies found that the social context of help-seeking produced a different problem: those learners needing help the most (students with low self-efficacy) were less likely to seek it out, as they feared being perceived as lacking in ability and thus lose social standing (Ryan and Shin 2011). This negative effect could be neutralized in our case, as the game provides social anonymity within the look-up process, potentially resulting in lower inhibitions to use the look-up element and benefit from its content.

Interestingly, the IS literature on help tools (Clarebout and Elen (2009); Größler et al. (2000); Mäkitalo-Siegl et al. (2011)) did not confirm these positive expectations of optional help-seeking tools on learning outcomes. The most common reason provided was that tools were barely used (Aleven, Stahl, Schworm, Fischer and Wallace 2003; Größler et al. 2000; Liu and Reed 1994). The general unwillingness by the participants to accept help partly explained these findings, as has been found in various educational settings (Newman 2000; Ryan et al. 2001; Aleven et al. 2003). One explanation for such usage inhibitions was that the help function was sometimes perceived as cheating (Clarebout and Elen 2009). The factors found to influence how students behaved in open learning environments were the students' self-efficacy, motivation and perception of the task. If they felt the task was performance-oriented, they were less likely to use the help tools than when they perceived it as learning-oriented (Clarebout and Elen 2009). As our game is not only learning-oriented but related to a serious and meaningful task, we believe that such inhibitions regarding the look-up element might be alleviated.

Finally, while looking at the literature on cognitive psychology, we found a dichotomy of two error-related learning strategies: errorful and errorless learning. The former – also referred to as trial-and-error learning – is “the process of making repeated trials or tests, improving the methods used in the light of errors made, until the right result is found” (Webster's 2005). It builds on the repetition-based learning strategy that the repeat element of our artifact is based on. Interestingly, we found that this strategy was juxtaposed with an entirely opposite strategy – errorless learning – which is defined as “an approach whereby the task is manipulated to eliminate/reduce errors. Tasks are executed in such a way that the subject is unlikely to make errors” (Fillingham et al. 2003 p. 339). This was partially fitting for us, as the look-up element would allow players to play the game without error if they chose to use it before every decision. However, when comparing studies that used errorful vs. errorless teaching strategies, neither one was found to be more effectual (e.g., Clare et al. 1999) or the results were inconclusive (e.g., Johnson 2004) (see Table 8 in the appendix; available online via <http://link.springer.com>).

Looking at the volatile nature of instructional explanations and help/look-up tools, we believe that, in particular, the optional and anonymous nature of the look-up element as well as the fact that the game affords a meaningful follow-up activity (sorting the waste after the explanation has been provided) can alleviate some of the negative effects listed in the mentioned theories and studies. Furthermore, given citizens' almost daily interactions with waste, the look-up design element can add meaningful context to already existing knowledge structures. Finally, as our learning element only offers information when the learners reach an impasse and are actively inquiring for context and a solution (as recommended by instructional design theory; Wittwer and Renkl 2008), we hypothesize that:

**H3:** A look-up game design element increases the learning outcome.

## 4 The Design Artifact

### 4.1 General Design Decisions

The downloadable app is a complete and complex game that was released in 2015 for the three mobile platforms Android, iOS and Windows (“Die Müll AG”/ “Trashmonsters”). While playable on a PC, we designed the game with touch interaction in mind, focusing on mobile devices. We embedded the core gameplay into a small and interconnected world that represents the broader cosmos of waste sorting. The full game features an overarching story

narrated through a consecutive quest structure. We aimed to motivate prolonged play through an interplay of unlockable minigames, collectible accessories and an underlying discoverable mystery (see Fig. 1). We added these elements for players to alternate the core gameplay with additional activities connected to the general theme of waste sorting. We made each design decision with metaphorical mapping in mind. The rationale for these design choices (and the decision against other popular game design elements like badges and leaderboards) can be found in the appendix titled “Exclusion of Game Design Elements.”

#### 4.1.1 Player Character

According to Gee (2003), effective learning involves “playing a character.” For example, learning in a science class works best if students “think, act and value like scientists.” This assumption is supported by the findings of a psychological study where participants who were given a virtual body (avatar) communicated as Einstein (signifying super-intelligence), performed significantly better than participants of the control group, considering prior cognitive ability (Banakou et al. 2018). Such studies highlight the weight of design choices concerned with the role players take within the game. For our game, we chose a first-person perspective (the players act in the game as themselves) to keep the attribution of all in-game actions and successes as close to the players as possible to facilitate and suggest reproduction of their in-game actions in real life. Research suggests that players learn best when they are engaged in meaningful, goal-directed activities within the identities of experts (Gee 2003; Shaffer et al. 2005). As such, the role of players is to serve as new and essential members of the workforce, helping the monsters in their struggle to deal with the overwhelming amounts of waste they receive for sorting.

#### 4.1.2 Depiction of Knowledge Items

When deciding on the presentation of knowledge items for the game, we consulted literature on the mental representation of knowledge. During the learning process, different types of memory connections are formed (e.g., typical connections in mathematical didactics are numeric, graphic, situational and algebraic (Nitsch et al. 2016)). Two of the most common items are words (designated representation) and pictures (iconic representation) (Kolers and Brison 1984). According to Mayer’s theory of multimedia learning (2002), active learning entails the coordinated stimulation of both channels of the human information processing system (visual/pictorial and auditory/verbal processing). For our game, we chose to depict our knowledge items (waste items) with a combination of iconic and designated memory connection items through sticker-like pictures and by displaying the name of the waste item when picked up (see Fig. 2). We selected the waste items used for the experiment from a list of the Karlsruhe waste sorting facilities based on the following criteria: (1) relevance (loss of precious resources if sorted incorrectly), (2) frequency of appearance in common households and (3) difficulty (frequency of missorting in real life).

#### 4.2 Core Gameplay

Establishing a fun core gameplay is of great importance before proceeding to the design and implementation of any other design elements (Järvinen 2007; Sicart 2008). We tested the core gameplay extensively with over 20 playtesters. The game went through several iterations before the parameters were finally set. The tests were conducted in the manner of the quiet observer, as is common in user experience testing, with a follow-up session to discuss the highlights and flaws and make suggestions for the gameplay mechanism.



**Fig. 1** Game aesthetics of the unabridged game



**Fig. 2** Metaphorical representation (mapping) of the waste sorting process in the core gameplay of the artifact

#### 4.2.1 Setting

We set the core gameplay within a waste sorting facility. Four waste bins (paper, recycling, bio and residual waste – reflecting the system in Karlsruhe) are placed next to each other behind a conveyor belt. A monster that serves as a visual and charismatic representation of the subsequent process of received waste inhabits each bin, as shown in Fig. 2. For instance, residual waste is represented as a fire-breathing dragon, indicating the subsequent burning of residual waste. We chose a friendly and cartoon-like visual style with a bright color scheme to overcome potential negative associations with the topic of handling waste.

#### 4.2.2 Core Mechanics

As soon as the wave – consisting of 15 waste items – starts, waste items drop onto a conveyor belt that moves them from the right to the left side of the screen, where they then drop off. During this time, players need to pick up each item and sort it into the right bin. If an item drops off, it counts as unsorted and raises the counter of the waste pollution bar, leading to a littering-based Game Over. If it is sorted incorrectly, it is counted towards an air-pollution-based Game Over. The game flow is supposed to represent the ongoing succession of choices we must make with each waste object we encounter in our daily lives as well as the consequences that come with the wrong or negligent dealing of waste.

#### 4.2.3 Feedback System

Feedback should be immediate and comprehensible in terms of the failure or success of the given task (Sicart 2008), with rewards and advancement in the game carefully bound to it (Bellotti et al. 2013), which is an

established rule in games. Thus, we implemented a positive/negative reinforcement system: points (+ 10/-3) for right vs wrong sorting of an item, visual/audio feedback of the monsters (joy/anger), combos (+ 50 points for a correct three-item streak) and combo-breakers (disruptions of the combo counter upon missorting within a streak). A numeric score and a pollution-counter (top left-hand corner in Fig. 2) provide feedback on the overall performance, warning players of an impending Game Over. This counter fills up each time an item is placed in an incorrect bin or drops off the lane and is reduced when an item is placed in the correct waste bin. An appropriate chunking of tasks helps provide a flow experience (Csikszentmihalyi et al. 2005). Inspired by the successful two-minute format of game applications like Angry Birds (Rovio Entertainment 2009), we chunked the learning content into waves that do not exceed playtimes of two minutes so as to encourage shorter but more frequent playtimes. Following advice by Wolfe et al. (1998), we implemented a structure blending the previously learnt items with newly introduced ones.

#### 4.2.4 Tutorial

As is common practice within games (Gee 2003), the first three waves serve as tutorials and differ from regular gameplay. In the first wave, we present the main types of waste (recyclable, bio-degradable, paper and refuse) with an explanation of the underlying attributes with which players can infer the correct bin for each waste item (e.g., inextricably compounded materials go to residual waste). In the second wave, players are supposed to familiarize themselves with the core gameplay through representative waste items for each type. In the third wave, we introduced additional design elements that accompany the core gameplay: the look-up element and the pollution counter, which indicates how close players are to a potential Game Over. In the experiment, we introduced only the pollution counter but not the look-up element to the groups without the look-up element.

#### 4.3 Experiment Version

For the purposes of the experiment, we compiled an abridged version of the game that only included design elements specifically designed to teach correct waste sorting: the core gameplay – including the tutorial – as well as the two additional learning enhancing design elements – the repeat option and the look-up feature. We shortened the core gameplay from 34 levels to 10 and from 201 waste items to 108 (eight were used as exemplary items in the tutorial and the remaining 100 were distributed over the 10 waves, introducing 10 new items and reusing five previously seen ones per wave). To avoid confounding

influences, we stripped the experimental version of all design elements that related to motivation enhancement (narrative elements and unlockable features). We wanted to ensure an isolated observation of the effectiveness of the core gameplay in producing a learning outcome. We kept the underlying worldbuilding and setting (monster design and waste sorting plant (see Fig. 1)) as they are integral to the game feel.

#### 4.4 Repetition-based Design Element

If a level is not completed perfectly, the game shows players how many items they sorted incorrectly and offers them the chance to repeat the level without penalty (see Fig. 3). We strategically placed and colored the “yes” and “no” buttons to favor repetition. If players choose to repeat, their level of pollution is reset to the level when that wave was played for the first time. We were inspired by the quick trial, immediate performance feedback and low inhibition retrieval-loop pattern of games like Cut the Rope (ZeptoLab 2010) and Angry Birds (Rovio Entertainment 2009).

#### 4.5 Look-Up-based Design Element

In his article, Gee (2003) elaborated on the placement of information: that it should be given “on demand” and applied soon after having read it. He based this on people’s poor understanding and retention of information received out of context (Brown et al. 1989; Barsalou 1999; Glenberg and Robertson 1999). The look-up element (see Fig. 4) is an index that can be used to find all previously encountered waste items. For each item, it shows the correct target bin, as well as additional information on why the item belongs there and not in another bin. It can be accessed at any point throughout the game by simply opening it or by pulling an item on top of it (it then scrolls directly to that item). It is introduced in the tutorial and its usage penalty-free. For the



Fig. 3 Repeat element (original game texts were in German)

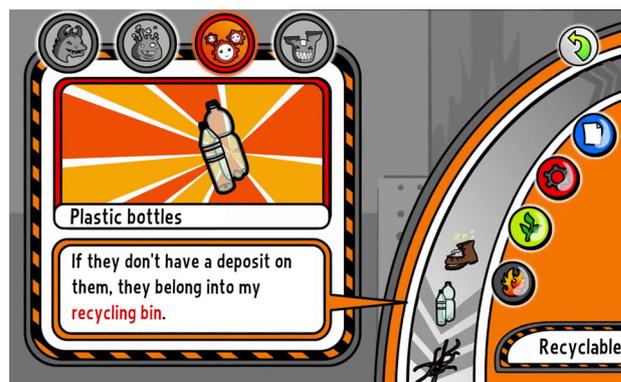


Fig. 4 Look-up Element (original game texts were in German)

mechanics of this look-up design element, two game design elements that serve to offer additional information to the players inspired us. First, we drew insights from the interactive “hint” functions found in puzzle games and point-and-click adventures like Machinarium (Amanita Design 2009). These hints are designed to reduce frustration by guiding the players with incremental tips. They are optional, so players decide for themselves when and if they want to use them. The second inspirational game design element is the *pokédex* used in the Pokémon (Game Freak, 1996) game series: a lexicon-based design element that gradually lists all monsters and their related meta-data that players encounter during the game.

## 5 Experiment

### 5.1 Experimental Design and Independent Variables

We designed the laboratory experiment to test the effect of the game in general as well as two independent variables (look-up and repeat) on the learning outcome. We designed a between-subject experiment in three stages where the 10–12-day duration between Phases 2 and 3 served as the retention period. We designed the experiment with four treatments in a full-factorial design with an additional fifth control group (from now on referred to as non-game material) that was given exemplary teaching material as used by waste management institutions. The used non-game teaching material consisted of the three informative flyers conventionally provided by the city of Karlsruhe to teach citizens correct waste sorting. The first flyer informed on the general categories of waste that go into each of the four bins, the second served to differentiate the general waste categories in combination with the underlying rules of what waste belongs where (see Figs. 11 and 12 in the appendix “Non-Game Materials”) and the third listed exemplary waste items for each bin (see Fig. 5 (excerpt)

and Fig. 10 in the appendix “Non-Game Materials”). An overview of the treatments’ structure is provided in Table 2.

## 5.2 Experimental procedure

We recruited participants from a large German university using the organizing and recruiting software hroot (Bock, Nicklisch, Baetge 2012). Potential participants in the experiment had to meet three requirements to participate: they needed to own a smartphone with an Android-based operating system running on a version higher than 2.3.1 (Gingerbread), be willing to download and install the application on their phone and be fluent in German.

We conducted the experiment in three stages (see Fig. 6): the preparation phase (P1), the training phase including a subsequent passive retention phase (P2) and the testing phase (P3). Participants completed the first two phases remotely. In phase (P2) we instructed the participants on the four game-based treatments to play the game through to the end and then complete the survey. In contrast, we told the control group with the non-interactive materials to attentively read through the teaching materials provided through the link for 25 min (this time was derived from the average playtime of the experimental version of the game during the pre-tests) and to then complete the survey. We conducted the testing phase (P3) in the laboratory to ensure proper supervision of the tests (a detailed description of each phase can be found in the appendix). Participants received a flat payment of €15 for their time.

## 5.3 Operationalization of the Dependent Variable: the Learning Outcome

We measured the learning outcome with special regard to two factors: long-term retention and knowledge transfer. According to cognitive theory, long-term retention can be

tested as soon as two or three days past the training period (Schmidt and Bjork 1992). For our study, we chose an extended retention phase of 10–12 days to ensure success of the transfer to long-term memory (see also Luo et al. (2018); Parkin and Streete (1988)).

In their work on training evaluation, Kraiger et al. (1993) highlighted the importance of conceptually sound measures of learning that ensure training effectiveness with regard to knowledge transfer. We tested knowledge transfer in three ways: first by testing identification (Nitsch et al. 2016) of knowledge by evaluating if players can reproduce the learned content within the training medium. For this, we used a special version of the game (game measure) featuring one wave where all 108 trained items appear one-by-one from the right side of the screen and then have to be sorted into the correct bin before they drop off on the left side (see Table 3). We then tested knowledge transfer via a multiple-choice-based test measure as a power test (number of items answered correctly in an unlimited amount of time (Kraiger et al. 1993). We chose this testing measure because multiple-choice tests are considered best suited for measuring the retention of declarative knowledge (Gagne 1984; Bellotti et al. 2013). In this measure, participants were given the names of the waste items but not images like in the game measure. By offering only one of the two memory connection items, we could differentiate the effectiveness of the representational elements (pictures vs text) (Mayer 2002). Participants were asked to assign the right bin for each of the 108 trained items (the options were residual, recycle, biodegradable and paper waste and separate recycling) (see Table 3).

Finally, we measured knowledge transfer to the final application domain: real-life waste sorting. This measure relates to the construction item introduced by Nitsch et al. (2006), where knowledge is retained and understood in a way so that it can be reapplied to a different context. In this third measure, participants had to sort a selection of real-

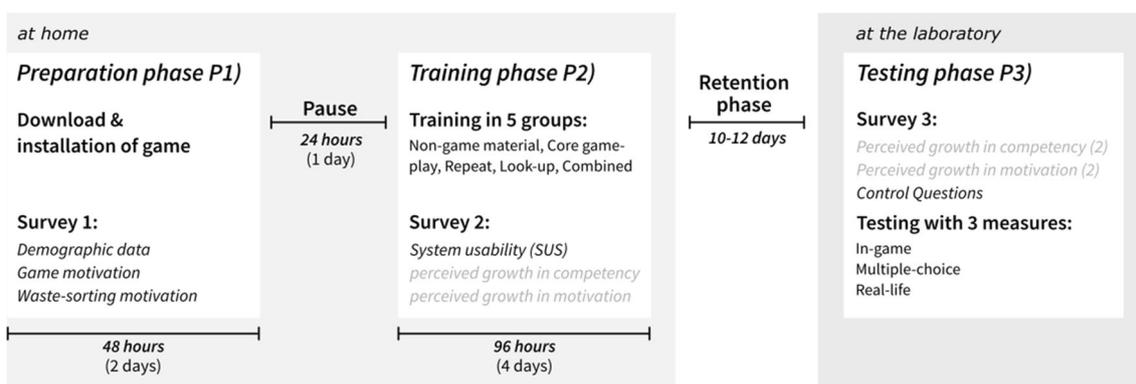
### Four Bins - Examples

Residual waste		Recycleables	Biowaste*	Paper   Cardboard
Ring binder, plastic	Nylon tights	<b>Wood, untreated, like:</b>	Balcony plants	<b>Recycled paper like:</b>
Ashes - packed	Camera lenses	Wooden boards, Fruit crate	Banana peels	Ring binder - cardboard
Baking/grease-proof paper	Paper, very soiled or imbued	<b>Recyclables, like:</b>	Food waste bin liners	Envelopes, with and without viewing panel
Eye glasses, broken	Paper towels & tissues, soiled	CDs	Bread	Brochures
Sanitary pads	Parchment Paper	Bucket - emptied	Eggshells	Books
Photographic slides	Sticky plaster	Plastic crockery	Fish offal	Egg boxes
Floppy disks	Paintbrush	Bottles, canisters	Offal	Wrapping paper, uncoated
Extractor fan filter	Porcelain	Plastic film, plastic bags	Vegetable peel	Notebooks
Bicycle saddle	Dolls	Childrens toys	Hair	Cardboard boxes
Pelts / Skins	Cleaning rags	Mixing bowl	Burlap	Catalogues
Binoculars	Eraser	Styrofoam (sundries in transparent bags)	Coffee grounds	Magazines
Heat-proof glass	Razor blades		Cheese residues	Paper - loose
Lighter, empty			Bones	
Photographic film			Dead parts of plants	

Fig. 5 Flyer on general waste sorting in XX (excerpt) \*translated to English

**Table 2** Treatment overview

Treatment group	Implementation
Control group: Non-game materials	This group received non-interactive learning materials as currently provided by the municipal waste department of Karlsruhe, which consisted of two flyers introducing the general waste assignments and an exemplary list of the items with their correct bins (see Fig. 5)
Game group: Core gameplay	This group was given an instantiation where only the core gameplay was implemented (see Fig. 2)
Game group: Repeat element	On top of the core gameplay, at the end of each wave, the players of this group were given the option to repeat the wave without penalty (see Fig. 3)
Game group: Look-up element	On top of the core gameplay, the players of this group were introduced to and had permanent access to the look-up element, giving them the option to look up the correct bin for any waste item they encountered (see Fig. 4)
Game group: Combined repeat and look-up element	On top of the core gameplay, the players of this group could access the look-up element at any time and after each wave, they were given the option to repeat without penalty

**Fig. 6** Experimental Procedure**Table 3** Dependent variables for measuring the learning outcome

Dependent variable	Range and meaning	Theoretical construct
In-Game Performance	Continuous value between 0 and 1: the percentage of correctly sorted waste items out of 108	Identification
Multiple-Choice Test	Continuous value between 0 and 1: the percentage of correctly sorted waste items out of 108	Knowledge transfer with reduced stimuli
Real-Life Sorting	Continuous value between 0 and 1: the percentage of correctly sorted waste items out of 7	Knowledge transfer to real-life / construction

life waste items into the correct bin (see Figs. 7 and 8). Seven representative waste items were chosen for the real-life sorting according to the participants' performances measured in Phase 2 of the experiment: one from the top five items of best average sorting performance (aluminum), two from the average of their sorting performance (adhesive tape and milk cartons), and four that belonged to the group of the 20 worst-performing items (CDs, thermal paper, empty ring binder and wood shavings).

To increase the comparability of the three measures in consideration of the different number of items, we decided to use percentages of correctly sorted items. Thus, for each person and measure, we divided by the number of items sorted. For example, a measure of 85.71% for the real-life sorting performance meant that the participant sorted five out of the seven items correctly (see Table 3).



**Fig. 7** Representative waste bins: bio, paper, recycle, residual



**Fig. 8** Real-life waste items

#### 5.4 Control Variables

Apart from controlling for demographic factors (age, gender, how long the participants had been living in Germany and the city in which the experiment was conducted), we controlled for the following: **Gaming motivation.** Since gamified systems were previously perceived as less serious than traditional teaching content (Brigham 2015; Hanus and Fox 2015), the acceptance of the medium might influence the willingness to learn. We thus measured user attitude towards the medium in general through self-reporting (the full implementations can be found in Table 9 in the appendix). **General waste sorting motivation.** Since the personal attitude to the topic plays a role in the learning outcome (Garris et al. 2002), we also measured the general attitude towards waste sorting at home through two questions. **System usability.** The usability of the respective information system plays an equally important role as poor user experience can lead to frustration and thus have a negative impact on user interaction (Bangor et al. 2008). We decided to assess user satisfaction with Brooke's (1996) system usability scale (SUS). This decision was based on its widespread usage in IS for such purposes and to allow for comparability between our artifact and similar studies (Bangor et al. 2008).

## 6 Results

The first stage of the experiment was completed by 266 participants. Thirty-one participants did not complete all

three stages of the experiment (17 participants did not start or finish Phase 2 and 14 more did not show up to Phase 3 in the lab). Of the 235 remaining participants, we had to exclude 14 further datasets because of transmission errors (e.g., the game data of the second or third stage of the experiment was missing) and one for failing a crucial control question. Finally, of the remaining 220 data sets, there was a minor data transmission error for 23 participants: not all single item sorts for the in-game performance had been transmitted completely. We decided to exclude the datasets where more than 30% of the item sorts were missing (five out of these 23). This decision was backed by a Kruskal-Wallis test that indicated that the performance of the 18 participants with more than 70% but less than 100% correctly transferred item sorts did not differ significantly from the participants with complete sets of item sorts. We thus decided to include them, leaving us with a total of 215 complete datasets. The average age of the participants was 22.72 years old (one person reported the age of 3, which we set as a missing value because this was either a typo or intentionally misreported), and the gender distribution was 66.05% of participants identifying as male vs 33.49% as female vs one person (0.47%) indicating "other." Table 3 shows the descriptive statistics of the dependent measures for all treatments. For example, in the treatment with non-game materials, participants correctly sorted on average 70.8% of the items in the in-game performance measure, 59% in the multiple-choice test and 70.3% in the real-life sorting task. The pattern of having the lowest performance when measuring with the multiple-choice test compared to the other two learning outcome

**Table 4** Effect of the game in comparison with the non-game material (OLS regression) (Hypothesis 1)

Reference category: Non-game material	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
	coef (bootstr. std. error) [conf. interval]	p (two-tailed)	coef (bootstr. std. error) [conf. interval]	p (two-tailed)	coef (bootstr. std. error) [conf. interval]	p (two-tailed)
Game	.041 (.018) [.012, .070]	.020*	.084 (.020) [.052, .117]	.000**	.069 (.031) [.017, .121]	.027*
Constant	.708 (.016) [.682, .734]	.000**	.590 (.018) [.560, .619]	.000**	.703 (.029) [.656, .751]	.000**
N	215		215		215	
R <sup>2</sup>	.026		.087		.025	
Adj. R <sup>2</sup>	.021		.083		.020	

\*p &lt; 0.1, \*\*p &lt; 0.01

measures is stable over all treatments. The largest value of 78.8% was reached in the combined group for in-game performance. For more details on the descriptive statistics for both the dependent measures as well as the control variables, please see Tables 10 and 11 in the appendix.

For all statistical tests, we computed ordinary least square (OLS) regressions with the three continuous performance measures ranging between 0 (0% correctly sorted) and 1 (100% sorted correctly) as dependent variables. All our hypotheses were directed and therefore a test was significant if p of the two-tailed tests in the presented tables of the statistical tests was below 10%. Robust standard errors were used in all regressions to account for heteroscedasticity, based on the Breusch–Pagan test (Cohen et al. 2014). Furthermore, we bootstrapped the results with a sample size of 5,000 to account for non-normality of residuals (Tibshirani and Efron 1993; Pek et al. 2018).

To compute Hypothesis 1, we had to pool the treatments core gameplay, repeat element, look-up element and combined group into one group because Hypothesis 1 compared the games' performance with the non-game materials. In contrast to the other two hypotheses, it did not focus on the effect of specific game design elements and their related individual treatments. We thus computed a binary variable "Game" that took the value 1 for all observations trained through the game (the pooled group) and the value 0 for the observations in the non-game material treatment. This binary variable was our only independent variable in this main analysis for Hypothesis 1. Table 4 shows the results of the three regressions of this binary variable on each of the three learning outcome measures. We found significant effects on all measures, which supported Hypothesis 1. When tested with the in-game performance measure, the game treatments were estimated to correctly sort 4.1% more items than non-game treatments. For the multiple-choice test, the effects were even larger: the game treatments were estimated to

correctly sort 8.4% more items than non-game treatments. Finally, for the real-life sorting measure, the estimate was 6.9%. To sum up, we could fully support Hypothesis 1 for all three performance measures. The effect for in-game performance was surprisingly the weakest, although this was the measurement for which the medium (the digital game) of training and testing was the same.

In contrast to the analysis of Hypothesis 1, Hypotheses 2 and 3 focused on the effect of the examined design elements. Thus, we did not pool all game treatments but rather compared all five treatments with each other. We coded each treatment with a binary variable that took the value 1 if the observation belonged to the respective treatment. The reference category was the non-game material treatment which meant that all coefficients must be compared to the performance in the non-game material treatment.

Table 5 illustrates the results for Hypotheses 2 and 3. When comparing the in-game performance of the treatments with the non-game material treatment, we found a significantly increased learning outcome for the look-up element treatment (estimated increase of 4% of correct item sorts) and the combined one (8%). An additional Wald-test showed that the effect of the combined treatment was larger than that of the look-up treatment ( $p = 0.04$ ). However, the effect for the look-up element was not significantly larger than for the repeat element treatment (again tested with Wald test,  $p = 0.56$ ). Thus, the look-up treatment performed better than the non-game material treatment but not better than the treatment with only repetition. In sum, for the in-game performance measure, Hypothesis 3 was fully supported: we found better performance for both the groups that only had the look-up element by itself or the look-up element combined with the repetition element. Hypothesis 2 was only partially supported: we did not find a stronger performance when only playing with the repetition element. Hypothesis 2 was only supported if repetition was combined with the look-up

**Table 5** Effect of the design elements in comparison to the non-game material with OLS (Hypotheses 2 and 3)

Reference category: Non-game material	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
	coef	p (two-tailed)	coef	p (two-tailed)	coef	p (two-tailed)
	(bootstr. std. error)		(bootstr. std. error)		(bootstr. std. error)	
	[conf. interval]		[conf. interval]		[conf. interval]	
Repeat element	.028 (.021)	.197	.080 (.024)	.001**	.073 (.039)	.058*
	[-.008, .063]		[.040, .120]		[.010, .137]	
Look-up element	.040 (.022)	.073*	.086 (.024)	.000**	.071 (.038)	.066*
	[.003, .076]		[.047, .125]		[.008, .135]	
Combined	.080 (.021)	.000**	.119 (.023)	.000**	.063 (.040)	.112
	[.046, .114]		[.081, .157]		[-.002, .129]	
Core gameplay	.019 (.021)	.366	.055 (.024)	.024*	.069 (.036)	.054*
	[-.015, .055]		[.015, .095]		[.010, .129]	
Constant	.708 (.016)	.000**	.590 (.018)	.000**	.703 (.029)	.000**
	[.682, .055]		[.560, .619]		[.656, .751]	
N	215		215		215	
R <sup>2</sup>	.070		.121		.025	
Adj. R <sup>2</sup>	.052		.104		.007	

\*p &lt; 0.1, \*\*p &lt; 0.01

**Table 6** Effect of the design elements in comparison to the core gameplay with OLS (Hypotheses 2 and 3)

Reference category: Core gameplay	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
	coef	p (two-tailed)	coef	p (two-tailed)	coef	p (two-tailed)
	(bootstr. std. error)		(bootstr. std. error)		(bootstr. std. error)	
	[conf. interval]		[conf. interval]		[conf. interval]	
Repeat element	.008 (.020)	.681	.025 (.023)	.270	.004 (.034)	.914
	[-.024, .041]		[-.012, .062]		[-.052, .059]	
Look-up element	.021 (.021)	.327	.031 (.023)	.175	.002 (.033)	.954
	[-.014, .055]		[-.007, .068]		[-.052, .056]	
Combined	.061 (.019)	.001**	.064 (.021)	.003**	-.006 (.035)	.860
	[.030, .091]		[.028, .099]		[-.064, .052]	
Non-game material	-.019 (.021)	.366	-.055 (.024)	.024*	-.069 (.036)	.054
	[-.055, .016]		[-.095, -.015]		[-.129, -.010]	
Constant	.727 (.014)	.000**	.645 (.016)	.000**	.773 (.021)	.000**
	[.704, .751]		[.619, .671]		[.738, .808]	
N	215		215		215	
R <sup>2</sup>	.070		.121		.025	
Adj. R <sup>2</sup>	.052		.104		.007	

\*p &lt; 0.1, \*\*p &lt; 0.01

element. For the multiple-choice test, we found even stronger results and could fully support Hypotheses 2 and 3: all four treatments trained through the game performed better in the multiple-choice test than the treatment trained with the non-game materials. The largest effect was

measured for the combined treatment, where on average, participants sorted 11.9% more items correctly than the participants in the control treatment without game materials. For the real-life sorting task, we interestingly found weaker effects for the combined treatment. In detail, we

**Table 7** Descriptive statistics of dependent measures

	n	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
		Mean (min/max)	SD	Mean (min/max)	SD	Mean (min/max)	SD
Non-game material	39	.708 (.463/.889)	.100	.590 (.259/.815)	.114	.703 (.286/1)	.181
Repeat element	46	.736 (.509/.926)	.097	.670 (.444/.861)	.108	.776 (.286/1)	.177
Look-up element	45	.748 (.491/.898)	.107	.676 (.380/.870)	.103	.775 (.429/1)	.171
Combined	41	.788 (.574/.917)	.079	.709 (.528/.870)	.092	.767 (.286/1)	.177
Core gameplay	44	.727 (.544/.870)	.093	.645 (.333/.852)	.107	.773 (.286/1)	.142
Overall	215	.741 (.463/.926)	.098	.659 (.259/.870)	.870	.760 (.286/1)	.171

found that only those treatments that had either one design element or neither of those two elements (the core game), performed significantly better than the treatment that did not play the game. Yet, the coefficients also showed that the effects for all four game treatments were rather similar, ranging between 6.3% for the combined treatment to 7.3% for the repeat element treatment. Thus, when conducting further Wald-tests comparing the coefficients of the game treatments with one another, one cannot claim that one game group performed better than another (all  $p > 0.8$ ). Thus, all in all, we could support Hypotheses 2 and 3 and found that all game treatments did comparably well.

For game or gamification designers, it is interesting to compare the effects of game design elements not only to the non-game material group, but also to the core gameplay group to gain a better understanding of which design elements to include in their gameful applications. Therefore, we want to further focus in detail on the comparison of the different game treatments to the core gameplay group in Table 6.

We found that with the in-game performance measure and the multiple-choice test, the combined treatment achieved a significantly higher learning outcome than the treatment with only the core gameplay available during training, with an increase of 6.1% correctly sorted items with the in-game performance measure and 6.4% with the multiple-choice test. When comparing the groups within the real-life sorting measure, a significantly different learning outcome cannot be discerned. This is a result already highlighted in the analysis above: the game-treatments performed comparably well in the real-life sorting task. Thus, for real-life performance, the overall effect of

the game itself was much stronger than that of adding the single design elements to the core gameplay.

To further assess the robustness of our results, we also computed robust OLS regressions with these control variables: age, gender, how long they lived in Germany (“Living in Germany”), how long they lived in the city the experiment was conducted in (“Living in XX city”), their gaming motivation, their general waste sorting motivation and the SUS (for details, see Tables 12, 13, 14 in the appendix). The results were robust regarding the inclusion of these control variables. However, there was one slight change: for Hypothesis 3, the effect on the repeat treatment became significant. Thus, for the statistics with control variables, we could now fully support Hypothesis 3. Regarding the significance of the control variables, we found that the longer participants lived in Germany, the better they performed in-game and in the multiple-choice test. This control variable can be seen as a proxy for prior knowledge about the participants’ waste sorting. Furthermore, the general waste sorting motivation value showed a tendency to positively affect the performance measures for all three measures ( $p$  ranges between 0.01 and 0.11). The SUS value of the game also had the tendency to positively influence the game performance ( $p = 0.064$  for all three hypotheses).

## 7 Discussion and Conclusions

In terms of our first and overarching research question, we found that the learning outcome for the groups given the game for training was significantly stronger than for the

group given state-of-the-art paper-based information during the training phase. This held true across all three measures. Interestingly, the effect was weakest within the in-game performance measure (with 4.1% more items correctly sorted than by the non-game material group) and strongest in the multiple-choice test (with 8.4% more items correctly sorted). This outcome contrasts with the literature on context reinstatement, which suggests that information encoded in one mindset is more successfully retrieved in the same mindset (Fisher and Kraig 1988). This interesting finding was also apparent in the non-game material treatment, where performance in the in-game measure was significantly higher than in the multiple-choice test (Wilcoxon signed-rank test with  $z = 5.16$ ;  $p < 0.01$ ) although the games' aesthetic and interaction were new to the non-game material treatment.

To gain further insight in this matter, we were interested in whether all participants generally performed better in one measure or the other. We found that performances, when measured in the game (Wilcoxon signed-rank test with  $z = 12.00$ ;  $p < 0.01$ ) and in real life ( $z = 7.85$ ;  $p < 0.01$ ), were significantly higher than when measured with the multiple-choice test. We also found that the game-trained group performed comparably well in the game and in real life ( $z = 1.95$ ;  $p = 0.06$ ) (for the descriptives, see Table 7). A potential explanation for this finding can be linked to the forming of memory connections: the multiple-choice test offered fewer memory connection items (offering only designated connections: words) than the game measure, which presented both iconic and designated connections (words and sticker-like icons) and the real-life measure, which provides real objects. Both the game and real-life objects offered more information items that could connect to existing schemata. This might have helped stimulate memories not activated by the fewer connections offered in the multiple-choice test. This finding is congruent with studies on word and picture learning (Kolers and Brison 1984; Mayer 2002) that found that learners performed better through a combination of words and pictures/objects than with words alone. Similarly, in the domain of mathematical didactics, studies have found that using more mathematical representations (like graphs, numbers, formulas) leads to an increased learning outcome (Ainsworth 2006). Our results showed that this effect works in both directions: learners retrieved formed memories more successfully if we offered more memory connections with their mental schemata.

In terms of Hypothesis 2 – adding repetition as a game design element increases the learning outcome – our results confirmed our conjectures. The group given the additional option to repeat waves showed a significantly higher learning outcome than the non-game material group in two of the three measures (multiple-choice and real-life). This

also held true for the in-game measure when inserting control variables. However, when compared with the core gameplay group, the implementation of a repeat option by itself did not increase the learning outcome significantly. The game design elements enhanced learning potential; however, this manifested within the success of the combined design elements. This suggests that the repeat elements inherently lacking in fun can be compensated for better results. This is underpinned by a study by Kim and Shute (2015), who found that changes in just one design element “significantly impacted players’ interactions with the game by changing players’ mental ‘operational rules’ during play” (p.351). While the use of the repeat element was optional, it was generally well-received, as 63.95% of players who had the repeat element available used it at least once (mean: 3.88, min: 0, max: 24).

For Hypothesis 3 – the increase of the learning outcome through a look-up design element – our results showed that the group given this design element performed significantly better than the non-game material group in all three measures. In terms of usage, the players received it even better than the repeat element, as 67.86% of players who had the look-up element available used it at least once (mean: 14.42, min: 0, max: 85). These are relevant findings given that we found contradictory indicators on the potential outcome in our analysis of related literature (e.g., studies on help tools reported low usage as well as low effects Liu and Reed 1994; Größler et al. 2000; Alevin et al. 2003)). When compared to the core gameplay group, the prevalence of the look-up element by itself did not significantly enhance the learning outcome of the game. However, as mentioned above, in combination with the repeat element, this design element created a significantly stronger effect in the in-game and multiple-choice measures. This showed that look-up features should be considered as important design elements in learning-oriented gameful applications.

Literature on error management training (Chillarege et al. 2003; Keith and Frese 2008) provides a potential explanation for the success of the combination of these two design elements. In contrast to errorful and errorless learning, this method (EMT) consists of helping trainees understand why errors occur, indicating how they can be avoided (as afforded by the look-up element) and then applying that knowledge to solve the problem (as afforded by the repeat element). This offers positive indications that if both affordances are implemented at the same time, they could lead to an especially successful learning outcome. This can be further consolidated within the theory of learning styles. In a study conducted by Liu and Reed (1994), which considered affordance combinations in a hypermedia environment, learning was accomplished by offering a diverse set of tools and aides to groups of students with different learning styles. This suggests that

offering different optional affordances benefits a diverse group of learners and leads to a stronger overall learning outcome. The combined effect could further assist in preventing the perception of cheating that could come with a help or hint-related design element (Clarebout and Elen 2009), as it allows players to test their own abilities in the first iteration of a wave before resorting to looking up the correct solution in the repeated wave.

In summary, the results showed that the core gameplay by itself already performed very well in comparison with the non-game materials. However, for the overall game to be more effective, it can be enhanced successfully by the two design elements that we suggested. Particularly, their combination showed their potential as building blocks for successful learning strategies by combining the mnemonic effect of repetition with easily accessible means for understanding.

When analyzing the control variables, we found that the number of years our participants had been living in Germany positively influenced their learning outcome in terms of the in-game and multiple-choice measures. This connection was expected since this particular control variable was implemented to passively enquire about prior waste sorting knowledge (to prevent priming, we decided against a full pre-measure of waste sorting knowledge – see Limitations Section). General waste sorting motivation also proved to have a significant influence over the learning outcome of the in-game measure alone. However, it is difficult to make sense of the fact that this effect was not replicated in the other measures – especially the real-life waste sorting measure. There could be influences in terms of cognitive dissonance of self-belief and self-actualization, but because of the setup of the experiment, we could not derive any personality-based indicators.

## 8 Contribution

The central goal of our research is to contribute to the rise of sustainable behavior through gameful design, specifically with regard to waste management. This goal stands in line with point 12.8 of the UN catalogue of sustainable goals: “Ensure that people everywhere have the relevant information and awareness for sustainable development and lifestyles in harmony with nature” (United Nations 2020). Our study showed that gameful design can successfully contribute to better municipal waste sorting, even with regard to a transfer of knowledge to real life. To the best of our knowledge, this is the first study to do so.

We further contributed to the ongoing efforts of investigating the potential of serious games and the implementation of gameful design as powerful teaching devices. In particular, the study showed significant positive learning

outcomes within a domain that generally lacks incentives relating to direct personal interest – such as health or fitness-oriented games would offer – and that is hampered by disinterest or even disgust by their target group regarding the general topic. By successfully translating this into more desirable content matter, our research highlighted the benefits of gameful design for teaching under adverse conditions. In terms of theoretical contribution, by conducting a full assessment of design choices with regard to their different learning outcomes, our research added to the ongoing general efforts of methodically assessing learning through gameplay. In this, our study lined up with a growing amount of research dissipating still-existent doubts about the usefulness of game-based learning (Shute et al. 2009).

A factor that contributes to such doubts is that not all studies in gameful learning test the success of their artifact in connection with its transition to real-life knowledge and applicability (e.g., Kim and Shute 2015). This measure, however, is very important, as seen in, for instance, Größler et al. (2000), who found in their study on gamification of business simulators that “participants were not capable of accessing the knowledge gained outside the gaming context” (p. 271). Another example is Ball et al. (2002), who concluded that cognitive training may only improve skills that are specific to the trained cognitive domain. Also, Luo et al. (2018) conducted a study with a similar premise and goal to ours and did not manage to reinstate the learning outcome when measured in real life. In contrast, in our study, we found that our game did overall manage to overcome this hurdle. Despite this success, the difficulty of constructing knowledge could be seen in the differences in learning outcomes between the different testing media. Our study highlighted the importance of measuring in the training medium as well as the true context medium (real life) and proving that the transfer is manageable given good design choices (Van Eck et al. 2017).

We also identified a gap in the IS literature on the effectiveness of look-up/help-based design elements and added to the ongoing discussion by conducting an experimental setup that tested this element in an isolated and a combined treatment. Our results showed that affording an optional, learner-moderated look-up element can be a very promising learning-enhancing design element, especially if added to a repetition-based teaching setup. By intricately testing these specific game mechanics, we contributed to understanding how they function to produce meaningful learning experiences, which is a paradigm suggested by the Games, Learning, and Society initiative (Squire 2007). With regard to the general topic of sustainability in IS, our study was one of few to focus on challenges surrounding

the domain of waste management. We hope to inspire further studies in this seminal area of research.

In terms of its practical contribution, we believe that if implemented into the teaching curriculum of sustainability classes, our artifact can have a beneficial impact on the topic of correct waste sorting. Our research aims to support the process of research informing practice and aide designers in optimizing their design decisions, as they have to make efficient decisions under time pressure (Stacey and Nandhakumar 2009). Furthermore, as stated by Clow (2013), educators need to be given insights about additional tools, as well as their strengths and limitations, which we provide in this manuscript. By affording detailed insights into the rationales behind the design decisions that went into the creation of our game and the design elements used, we facilitated easy means of reproduction for practitioners and researchers. While the mechanisms we looked at are embedded in the framework of a game, any learning or training context can serve as the foundation for the design mechanisms we analyzed in our study (Deterding 2016). Thus, we argued that in a playful setting that allows a certain degree of make-believe, a broad variety of teaching tasks (e.g., vocabulary, geography training, digital management training and onboarding) can benefit from applying the findings of our study.

## 9 Limitations and Future Work

One potential limitation concerns the fact that we omitted assessment of prior knowledge on waste sorting. Due to the three-phase setup of the experiment, we consciously decided against this assessment because of concerns about priming the participants and thus skewing the results. While it is common in the assessment of serious games to use pre- and post-testing, “the main problem with the pre- and post-test experimental design is that it is impossible to determine whether the act of pre-testing has influenced any of the results.” (Bellotti et al. 2013 p.3). By conducting a prior assessment like completing a survey-based multiple-choice test, we were concerned that participants would influence the actual results by looking up certain items they were unsure of before the first task. Instead, we measured “living in Germany” as a proxy indicator for prior knowledge, which turned out to have a significant influence on the learning outcome. Because we conducted an experiment by randomly assigning participants to treatments, we trust the internal validity of our results. Thus, the effect should be independent of confounders such as prior knowledge.

We believe the exclusion of prior knowledge as a predictor in our models, as well as the omission of measuring other variables that might influence waste sorting knowledge – e.g., participants’ exposure to the topic in school or

other contexts or their families’ attitudes towards sustainability and eco-friendliness – are the main reasons for the rather low  $R^2$  of our main models that included only the treatment variables. However, a low  $R^2$  is not unusual for experimental research and does not harm the interpretation of the effect of the treatment variables. Our further analyses in the appendix also show that the inclusion of the control variables – e.g., the number of years that the participants have lived in German and general waste sorting motivation – helps reduce the unexplained variance substantially, yielding  $R^2$  values around 0.2.

We further see that there could be an underlying cultural bias given the generally high range of results. Frese et al. (1991 p. 90) noted that errors may be perceived as especially stressful in German culture, “where perfectionism is highly valued.” For transferability to other cultures with different prior mentalities regarding correct waste sorting, future studies will be necessary to assess mentality as a moderating factor. Another important point is that we assessed the real-life measure with only seven items. While this arguably weakens comparability with the other two measures, practical concerns in terms of implementing a much larger number of items (limited setup and timeframe, participants’ resistance to interacting with certain items) limited our options for this measure. It should also be noted that within the non-game materials used during the training phase, one of the three flyers (the one showing examples of waste items for each bin in Fig. 10 in the appendix) featured more items than were presented to the game groups during the experimental task. When we designed the experiment, we wanted to approximate a real-life scenario and thus chose to use an unabridged set of standard materials provided by the local waste management (see in appendix “Non-Game Materials”). In hindsight, the overall experimental design would have been cleaner if we had reworked the flyer to feature the exact number of items that were trained in the game. However, it is important to note that the goal of the game was not to teach the specific relationship of each featured waste item to their bin but to help players understand and train them on the rules of the underlying waste systems. As all objects will eventually turn into a waste item, citizens need to learn how to correctly sort any object they encounter to the respective system by understanding and internalizing the underlying principles.

We wanted to test the learning outcome in a rigorous and controlled manner to obtain clear and interpretable results, so we decided to conduct a laboratory experiment to provide high internal validity. However, as our findings were based on an experimental setting including mostly students, any gained insights were only applicable for the tested age group (17–41). A future step would be to test whether the effects found are replicable in the field. Another facet of this relates

to knowledge transfer. Even though we found that knowledge transfer to real life (construction) was successfully achieved in the game, we believe that this effect might be enhanced by transferring the game to a virtual/augmented reality environment by bringing the medium of training closer to the actual application context.

Finally, while we chose to separate learning from motivation to isolate our findings, this approach might have omitted an important influence on learning. On this basis and because of our overall goal to teach correct waste sorting and to boost the motivation to act upon that knowledge, we want to design and conduct another motivation-focused experiment to build on our findings and enhance gameful design-based learning even further.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s12599-021-00731-x>.

**Acknowledgements** Many thanks to Raphael Martin, Brice Clocher and Anke Greif-Winzrieth for their practical and invaluable contributions and the reviewer team for their insightful feedback and constructive criticism.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

#### Declarations

**Conflict of interest** The authors declared that there is no conflict of interest

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

#### References

- Abdel-Shafy HI, Mansour MSM (2018) Solid waste issue: sources, composition, disposal, recycling, and valorization. *Egypt J Pet* 27:1275–1290. <https://doi.org/10.1016/j.ejpe.2018.07.003>
- Ahmadian MJ (2012) Task repetition in ELT. *English Lang Teach J* 66:380–382. <https://doi.org/10.1093/elt/ccs020>
- Ainsworth S (2006) DeFT: a conceptual framework for considering learning with multiple representations. *Learn Instr* 16:183–198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>
- Aleven V, Stahl E, Schworm S et al (2003) Help seeking and help design in interactive learning environments. *Rev Educ Res* 73:277–320. <https://doi.org/10.3102/00346543073003277>
- Allam A, Kostova Z, Nakamoto K, Schulz PJ (2015) The effect of social support features and gamification on a web-based

- intervention for rheumatoid arthritis patients: randomized controlled trial. *J Med Internet Res* 17(1):e14. <https://doi.org/10.2196/jmir.3510>
- Baddeley A, Hitch G (1974) Working memory. *Psychol. Learn Motiv* 8:47–89. <https://doi.org/10.1016/j.cub.2009.12.014>
- Baddeley A, Wilson BA (1994) When implicit learning fails: amnesia and the problem of error elimination. *Neuropsychologia* 32:53–68. [https://doi.org/10.1016/0028-3932\(94\)90068-X](https://doi.org/10.1016/0028-3932(94)90068-X)
- Ball K, Berch DB, Helmers KF et al (2002) Effects of cognitive training interventions with older adults: a randomized controlled trial. *J Am Med Assoc* 288(18):2271–2281. <https://doi.org/10.1001/jama.288.18.2271>
- Banakou D, Kishore S, Slater M (2018) Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Front Psychol* 9:917. <https://doi.org/10.3389/fpsyg.2018.00917>
- Bangor A, Kortum PT, Miller JT (2008) An empirical evaluation of the system usability scale. *Intl J Human-Computer Interact* 24:574–594. <https://doi.org/10.1080/10447310802205776>
- Barata G, Gama S, Jorge J, Gonçalves D (2013) Improving participation and learning with gamification. In: *Gamification '13: Proceedings of the First International Conference on Gameful Design, Research, and Applications*. pp 10–17
- Barnett SM, Ceci SJ (2002) When and where do we apply what we learn? a taxonomy for far transfer. *Psychol Bull* 128(4):612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Barsalou LW (1999) Language comprehension: archival memory or preparation for situated action? *Discourse Process* 28(1):61–80. <https://doi.org/10.1080/01638539909545069>
- Bellotti F, Kapralos B, Lee K, et al (2013) Assessment in and of serious games: an overview. *Adv. Human-Computer Interact*.
- Berengueres J, Alsuwairi F, Zaki N, Ng T (2013) Gamification of a recycle bin with emoticons. In: *2013 8th ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, pp 83–84
- Berthold K, Renkl A (2010) How to foster active processing of explanations in instructional communication. *Educ Psychol Rev* 22:25–40
- Bifulco I, Francese R, Lettieri M, et al (2011) The TIE project: agile development of a virtual world serious game on waste disposal. In: *Distributed Multimedia Systems*. pp 204–209
- Blohm I, Leimeister JM (2013) Gamification: design of IT-based enhancing services for motivational support and behavioral change. *Bus Inf Syst Eng* 5:275–278. <https://doi.org/10.1007/s12599-013-0273-5>
- Bradshaw GL, Anderson JR (1982) Elaborative encoding as an explanation of levels of processing. *J Verb Learn Verb Behav*. [https://doi.org/10.1016/S0022-5371\(82\)90531-X](https://doi.org/10.1016/S0022-5371(82)90531-X)
- Brigham TJ (2015) An introduction to gamification: adding game elements for engagement. *Med Ref Serv Q* 34:471–480. <https://doi.org/10.1080/02763869.2015.1082385>
- Brooke J (1996) SUS-A “quick and dirty” usability scale. *Usability Eval Ind* 189:4–7. <https://doi.org/10.1201/9781498710411-35>
- Brown JS, Collins A, Duguid P (1989) Situated cognition and the culture of learning. *Educ Res* 18:32–42. <https://doi.org/10.3102/0013189X018001032>
- Buccioli A, Montinari N, Piovesan M (2015) Do not trash the incentive! Monetary incentives and waste sorting. *Scand J Econ* 117:1204–1229. <https://doi.org/10.1111/sjoe.12122>
- Burton MA (1993) The irrigation management game: a role playing exercise for training in irrigation management. *Irrig Drain Syst* 7:305–318. <https://doi.org/10.1007/BF00881557>
- Bygate M (1996) Effects of task repetition: appraising the developing language of learners. In: Willis J, Willis D (eds) *Challenge and change in language teaching*. Heinemann, Oxford, pp 136–146

- Chi MTH, Siler SA, Jeong H et al (2001) Learning from human tutoring. *Cogn Sci* 25:471–533. [https://doi.org/10.1207/s15516709cog2504\\_1](https://doi.org/10.1207/s15516709cog2504_1)
- Chillarege KA, Nordstrom CR, Williams KB (2003) Learning from our mistakes: error management training for mature learners. *J Bus Psychol* 17:369–385. <https://doi.org/10.1023/A:1022864324988>
- Clare L, Jones RSP (2008) Errorless learning in the rehabilitation of memory impairment: a critical review. *Neuropsychol Rev* 18:1–23. <https://doi.org/10.1007/s11065-008-9051-4>
- Clare L, Wilson BA, Breen K, Hodges JR (1999) Errorless learning of face-name associations in early Alzheimer's disease. *Neurocase* 5:37–46. <https://doi.org/10.1093/neucas/5.1.37>
- Clarebout G, Elen J (2009) The complexity of tool use in computer-based learning environments. *Instr Sci* 37:475–486. <https://doi.org/10.1007/s11251-008-9068-3>
- Clark K, Sefton M (2001) Repetition and signalling: experimental evidence from games with efficient equilibria. *Econ Lett*. [https://doi.org/10.1016/S0165-1765\(00\)00381-5](https://doi.org/10.1016/S0165-1765(00)00381-5)
- Clark DB, Tanner-Smith EE, Killingsworth SS (2016) Digital games, design, and learning: a systematic review and meta-analysis. *Rev Educ Res* 86:79–122. <https://doi.org/10.3102/0034654315582065>
- Clow D (2013) An overview of learning analytics. *Teach High Educ* 18:683–695. <https://doi.org/10.1080/13562517.2013.827653>
- Cohen P, West SG, Aiken LS (2014) Applied multiple regression/correlation analysis for the behavioral sciences. Psychology Press
- Crowder RG, Melton AW (1965) The Ranschburg phenomenon: failures of immediate recall correlated with repetition of elements within a stimulus. *Psychon Sci* 2:295–296. <https://doi.org/10.3758/BF03343462>
- Csikszentmihalyi M, Abuhamdeh S, Nakamura J (2005) Flow. In: Elliot AJ, et al (eds) *Handbook of competence and motivation*. Guilford Publications, pp 598–608
- Amanita Design (2009) *Machinarium*. <https://amanita-design.net/games/machinarium.html> Accessed 14 Nov 2021
- Deterding S (2016) Make-believe in gameful and playful design. In: Turner P, Harviainen JT (eds) *Digital make-belief*. Springer, pp 101–124
- Donaghey CL, McMillan TM, O'Neill B (2010) Errorless learning is superior to trial and error when learning a practical skill in rehabilitation: a randomized controlled trial. *Clin Rehabil* 24:195–201. <https://doi.org/10.1177/0269215509353270>
- Dreher A, Kuntze S, Winkel K (2014) Empirical study of a competence structure model regarding conversions of representations—the case of fractions. In: Nicol C, et al. (eds) *Proceedings of the Joint Meeting of PME 38 and PME-NA 36*. ERIC, Vancouver, pp 425–432
- Driskell JE, Willis RP, Copper C (1992) Effect of overlearning on retention. *J Appl Psychol* 77:615–622. <https://doi.org/10.1037/0021-9010.77.5.615>
- Duffield JA (1991) Designing computer software for problem-solving instruction. *Educ Technol Res Dev* 39:50–62. <https://doi.org/10.1007/BF02298106>
- Dunn J, Clare L (2007) Learning face-name associations in early-stage dementia: comparing the effects of errorless learning and effortful processing. *Neuropsychol Rehabil* 17:735–754. <https://doi.org/10.1080/09602010701218317>
- Dunsmoor JE, Bandettini PA, Knight DC (2007) Impact of continuous versus intermittent CS-UCS pairing on human brain activation during Pavlovian fear conditioning. *Behav Neurosci* 121:635–642. <https://doi.org/10.1037/0735-7044.121.4.635>
- Van Eck RN, Shute VJ, Rieber LP (2017) Leveling up: game design research and practice for instructional designers. In: Reiser R, Dempsey J (eds) *Trends and issues in instructional design and technology*. Pearson, pp 227–285
- El-Hilly AA, Iqbal SS, Ahmed M et al (2016) Game on? Smoking cessation through the gamification of mHealth: a longitudinal qualitative study. *JMIR Serious Games* 4(2):e18. <https://doi.org/10.2196/games.5678>
- Elliot S, Webster J (2017) Editorial: special issue on empirical research on information systems addressing the challenges of environmental sustainability: an imperative for urgent action. *Inf Syst J* 27:367–378. <https://doi.org/10.1111/isj.12150>
- Rovio Entertainment (2009) *Angry Birds*. <https://www.angrybirds.com/> Accessed 14 Nov 2021
- Evans JJ, Wilson BA, Schuri U et al (2000) A comparison of “errorless” and “trial-and-error” learning methods for teaching individuals with acquired memory deficits. *Neuropsychol Rehabil* 10:67–101. <https://doi.org/10.1080/096020100389309>
- Fileti F (1988) Educational and cognitive aspects of videogames. *Simulation-gaming Educ Train* 252–260
- Filho WL, Brandli L, Moora H et al (2016) Benchmarking approaches and methods in the field of urban waste management. *J Clean Prod* 112:4377–4386. <https://doi.org/10.1016/j.jclepro.2015.09.065>
- Fillingham JK, Hodgson C, Sage K, Lambon Ralph MA (2003) The application of errorless learning to aphasic disorders: a review of theory and practice. *Neuropsychol Rehabil* 13:337–363. <https://doi.org/10.1080/09602010343000020>
- Fotaris P, Mastoras T, Leinfellner R, Rosunally Y (2016) Climbing up the leaderboard: an empirical study of applying gamification techniques to a computer programming class. *Electron J e-Learning* 14(2):94–110
- Frese M, Brodbeck F, Heinbokel T et al (1991) Errors in training computer skills: on the positive function of errors. *Human-Comput Interact* 6:77–93. [https://doi.org/10.1207/s15327051hci0601\\_3](https://doi.org/10.1207/s15327051hci0601_3)
- Gagne RM (1984) Learning outcomes and their effects: useful categories of human performance. *Am Psychol* 39:377–385. <https://doi.org/10.1037/0003-066X.39.4.377>
- GameFreak (1996) *Pokémon*. <https://www.pokemon.com/us/> Accessed 14 Nov 2021
- Garris R, Ahlers R, Driskell JE (2002) Games, motivation, and learning: a research and practice model. *Simul Gaming* 33:441–467. <https://doi.org/10.1177/1046878102238607>
- Gee JP (2003) What video games have to teach us about learning and literacy. *Comput Entertain* 1(1):20. <https://doi.org/10.1145/950566.950595>
- Glenberg AM, Robertson DA (1999) Indexical understanding of instructions. *Discourse Process* 28:1–26. <https://doi.org/10.1080/01638539909545067>
- González-Briones A, Valdeolmillos D, Casado-Vara R et al (2018) GarbMAS: simulation of the Application of gamification techniques to increase the amount of recycled waste through a multi-agent system. In: De La Prieta F et al (eds) *Distributed Computing and Artificial Intelligence, 15th International Conference*. Springer, Cham, pp 332–343
- Greenfield PM, Savage-Rumbaugh ES (1993) Comparing communicative competence in child and chimp the pragmatics of repetition. *J Child Lang*. <https://doi.org/10.1017/S0305000900009090>
- Größler A, Maier FH, Milling PM (2000) Enhancing learning capabilities by providing transparency in business simulators. *Simul Gaming* 31:257–278. <https://doi.org/10.1177/104687810003100209>
- Gustafsson A, Katzeff C, Bang M (2009) Evaluation of a pervasive game for domestic energy engagement among teenagers. *Comput Entertain* 7:1–19. <https://doi.org/10.1145/1658866.1658873>

- Hamari J (2017) Do badges increase user activity? A field experiment on the effects of gamification. *Comput Human Behav* 71:469–478. <https://doi.org/10.1016/j.chb.2015.03.036>
- Hanus MD, Fox J (2015) Assessing the effects of gamification in the classroom: a longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Comput Educ* 80:152–161. <https://doi.org/10.1016/j.compedu.2014.08.019>
- Hawlitschek F (2020) Interview with Benjamin Scheffler on “The future of waste management.” *Bus Inf Syst Eng* 63:207–211. <https://doi.org/10.1007/s12599-020-00671-y>
- Hopelab (2006) Re-Mission. <https://hopelab.org/product/re-mission/> Accessed 14 Nov 2021
- Horaz (Quintus Horatius Flaccus) (19 BC according to The Oxford Companion to Classical Literature, p. 74) *De Arte Poetica*
- Hunicke R, LeBlanc M, Zubek R (2004) MDA: a formal approach to game design and game research. In: *Proceedings of the AAAI Workshop on Challenges in Game AI*, p 1722
- Hunkin NM, Squires EJ, Parkin AJ, Tidy JA (1998) Are the benefits of errorless learning dependent on implicit memory? *Neuropsychologia* 36:25–36. [https://doi.org/10.1016/S0028-3932\(97\)00106-1](https://doi.org/10.1016/S0028-3932(97)00106-1)
- Idrobo ML, Saenz MF, Márceles K et al (2018) Recycling: a serious game focused on the classification of waste. In: Mata-Rivera M et al (eds) *International Congress of Telematics and Computing*. Springer, Cham, pp 234–245
- Ivancic K, Hesketh B (2000) Learning from errors in a driving simulation: effects on driving skill and self-confidence. *Ergonomics* 43:1966–1984. <https://doi.org/10.1080/00140130050201427>
- Jang S, Kitchen PJ, Kim J (2018) The effects of gamified customer benefits and characteristics on behavioral engagement and purchase: evidence from mobile exercise application uses. *J Bus Res* 92:250–259. <https://doi.org/10.1016/j.jbusres.2018.07.056>
- Järvinen A (2007) Introducing applied ludology: hands-on methods for game studies. In: *Digital Games Research Association Conference*. pp 1–11
- Johnson A (2004) We learn from our mistakes—don’t we? *Ergon Des* 12:24–27. <https://doi.org/10.1177/106480460401200206>
- Jonassen DH, Rohrer-Murphy L (1999) Activity theory as a framework for designing constructivist learning environments. *Educ Technol Res Dev* 47:61–79. <https://doi.org/10.1007/BF02299477>
- Jones RSP, Eayrs CB (1992) The Use of errorless learning procedures in teaching people with a learning disability: a critical review. *Ment Handicap Res* 5:204–212. <https://doi.org/10.1111/j.1468-3148.1992.tb00045.x>
- Kappen D, Mirza-Babaei P, Nacke L (2018) Gamification of older adults’ physical activity: an eight-week study. In: *Proceedings of the 51st Hawaii International Conference on System Sciences*. pp 1207–1216
- Kato PM, Cole SW, Bradlyn AS, Pollock BH (2008) A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. *Pediatrics* 122:e305–e317. <https://doi.org/10.1542/peds.2007-3134>
- Keith N, Frese M (2008) Effectiveness of error management training: a meta-analysis. *J Appl Psychol* 93:59–69. <https://doi.org/10.1037/0021-9010.93.1.59>
- Kessels RPC, Haan EHF (2003) Implicit learning in memory rehabilitation: a meta-analysis on errorless learning and vanishing cues methods. *J Clin Exp Neuropsychol* 25:805–814. <https://doi.org/10.1076/jcen.25.6.805.16474>
- Kessels RPC, van Loon E, Wester AJ (2007) Route learning in amnesia: a comparison of trial-and-error and errorless learning in patients with the Korsakoff syndrome. *Clin Rehabil* 21:905–911. <https://doi.org/10.1177/0269215507077309>
- Kim YJ, Shute VJ (2015) The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Comput Educ* 87:340–356. <https://doi.org/10.1016/j.compedu.2015.07.009>
- Kolers PA, Brison SJ (1984) Commentary: on pictures, words, and their mental representation. *J Verb Learn Verb Behav* 23:105–113. [https://doi.org/10.1016/S0022-5371\(84\)90539-5](https://doi.org/10.1016/S0022-5371(84)90539-5)
- Kraiger K, Ford JK, Salas E (1993) Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *J Appl Psychol* 78:311–328. <https://doi.org/10.1037/0021-9010.78.2.311>
- Kurtzman GW, Day SC, Small DS et al (2018) Social incentives and gamification to promote weight loss: the LOSE IT randomized, controlled trial. *J Gen Intern Med* 33(10):1669–1675. <https://doi.org/10.1007/s11606-018-4552-1>
- Landers RN, Landers AK (2014) An empirical test of the theory of gamified learning: the effect of leaderboards on time-on-task and academic performance. *Simul Gaming* 45:769–785. <https://doi.org/10.1177/1046878114563662>
- Leinhardt G, Steele MD (2005) Seeing the complexity of standing to the side: Instructional dialogues. *Cogn Instr* 23:87–163. [https://doi.org/10.1207/s1532690xci2301\\_4](https://doi.org/10.1207/s1532690xci2301_4)
- Levin JR (1988) Elaboration-based learning strategies: powerful theory = powerful application. *Contemp Educ Psychol* 13:191–205. [https://doi.org/10.1016/0361-476X\(88\)90020-3](https://doi.org/10.1016/0361-476X(88)90020-3)
- Liu M, Reed WM (1994) The relationship between the learning strategies and learning styles in a hypermedia environment. *Comput Human Behav* 10:419–434. [https://doi.org/10.1016/0747-5632\(94\)90038-8](https://doi.org/10.1016/0747-5632(94)90038-8)
- Lotfi E, Mohammed B (2014) Players performances analysis based on educational data mining case of study: interactive waste sorting serious game. *Int J Comput Appl* 108:13–18. <https://doi.org/10.5120/18954-0217>
- Luo Y, Zelenika I, Zhao J (2018) Providing immediate feedback improves recycling and composting accuracy. *J Environ Manage* 232:445–454. <https://doi.org/10.1016/j.jenvman.2018.11.061>
- Lynch T, Maclean J (2000) Exploring the benefits of task repetition and recycling for classroom language learning. *Lang Teach Res* 4:221–250. <https://doi.org/10.1177/13621688000400303>
- Mahmood AH, Murad IMA (2018) Approaching the language of the second language learner: interlanguage and the models before. *English Lang Teach* 11:95–108. <https://doi.org/10.5539/elt.v11n10p95>
- Mäkitalo-Siegl K, Kohnle C, Fischer F (2011) Computer-supported collaborative inquiry learning and classroom scripts: effects on help-seeking processes and learning outcomes. *Learn Instr* 21:257–266. <https://doi.org/10.1016/j.learninstruc.2010.07.001>
- Mayer RE (1980) Elaboration techniques that increase the meaningfulness of technical text: an experimental test of the learning strategy hypothesis. *J Educ Psychol* 72:770–784. <https://doi.org/10.1037/0022-0663.72.6.770>
- Mayer RE (2002) Multimedia learning. In: *Psychology of learning and motivation*. Academic Press, pp 85–139
- Menon BM, Unnikrishnan R, Muir A, Bhavani RR (2017) Serious game on recognizing categories of waste, to support a zero waste recycling program. In: *2017 IEEE 5th International Conference on Serious Games and Applications for Health*. IEEE, pp 1–8
- Miller GA, Gildea PM (1987) How children learn words. *Sci Am* 257:94–99. <https://doi.org/10.1038/scientificamerican0987-94>
- Nelson-Le Gall S (1981) Help-seeking: an understudied problem-solving skill in children. *Dev Rev* 1:224–246. [https://doi.org/10.1016/0273-2297\(81\)90019-8](https://doi.org/10.1016/0273-2297(81)90019-8)
- Newman RS (2000) Social influences on the development of children’s adaptive help seeking: the role of parents, teachers,

- and peers. *Dev Rev* 20:350–404. <https://doi.org/10.1006/drev.1999.0502>
- Nickerson RS, Adams MJ (1979) Long-term memory for a common object. *Cogn Psychol* 11:287–307. [https://doi.org/10.1016/0010-0285\(79\)90013-6](https://doi.org/10.1016/0010-0285(79)90013-6)
- Nitsch R, Bruder R, Kelava A (2016) Schülerhandlungen als Elemente fachdidaktisch motivierter Kompetenzmodellierungen. *J Fur Math* 37:289–317. <https://doi.org/10.1007/s13138-015-0084-y>
- Ohlsson S (1996) Learning from performance errors. *Psychol Rev* 103:241. <https://doi.org/10.1037/0033-295X.103.2.241>
- Olaf B, Andreas N, Ingmar B (2012) hroot: Hamburg registration and organization online tool. WiSo-HH Working Paper Series No. 1, 2012
- Page M, Wilson BA, Shiel A et al (2006) What is the locus of the errorless-learning advantage? *Neuropsychologia* 44:90–100. <https://doi.org/10.1016/j.neuropsychologia.2005.04.004>
- Parkin AJ, Streete S (1988) Implicit and explicit memory in young children and adults. *Br J Psychol* 79:361–369. <https://doi.org/10.1111/j.2044-8295.1988.tb02295.x>
- European Parliament (2018) The circular economy package: new EU targets for recycling. In: News. <http://www.europarl.europa.eu/news/en/headlines/priorities/circular-economy/20170120STO59356/the-circular-economy-package-new-eu-targets-for-recycling>. Accessed 20 Feb 2019
- Pek J, Wong O, Wong A (2018) How to address non-normality: a taxonomy of approaches, reviewed, and illustrated. *Front Psychol* 9:2104. <https://doi.org/10.3389/fpsyg.2018.02104>
- Pinter A (2005) Task Repetition with 10-year-old Children. In: Edwards C, Willis J (eds) *Teachers Exploring Tasks in English Language Teaching*. Palgrave Macmillan, London. [https://doi.org/10.1057/9780230522961\\_11](https://doi.org/10.1057/9780230522961_11)
- Prather DC (1971) Trial-and-error versus errorless learning: training, transfer, and stress. *Am J Psychol* 84:377–386. <https://doi.org/10.2307/1420469>
- Putnam AL (2015) Mnemonics in education: current research and applications. *Transl Issues Psychol Sci* 1:130–139. <https://doi.org/10.1037/tps0000023>
- Putz LM, Treiblmaier H, Pfoser S (2018) Field trips for sustainable transport education: impact on knowledge, attitude and behavioral intention. *Int J Logist Manag* 29:1424–1450. <https://doi.org/10.1108/IJLM-05-2017-0138>
- Rieber LP (1996) Seriously considering play: designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educ Technol Res Dev* 44:43–58. <https://doi.org/10.1007/BF02300540>
- Rieber LP (2005) Multimedia learning in games, simulations, and microworlds. In: Mayer R (ed) *The Cambridge handbook of multimedia learning*. Cambridge University Press, pp 549–567
- Rodrigues LF, Costa CJ, Oliveira A (2016) Gamification: a framework for designing software in e-banking. *Comput Human Behav*. <https://doi.org/10.1016/j.chb.2016.04.035>
- Roelle J, Berthold K, Renkl A (2014) Two instructional aids to optimise processing and learning from instructional explanations. *Instr Sci* 42:207–228. <https://doi.org/10.1007/s11251-013-9277-2>
- Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation. *Am Psychol* 55:68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan AM, Shin H (2011) Help-seeking tendencies during early adolescence: an examination of motivational correlates and consequences for achievement. *Learn Instr* 21:247–256. <https://doi.org/10.1016/j.learninstruc.2010.07.003>
- Ryan AM, Pintrich PR, Midgley C (2001) Avoiding seeking help in the classroom: who and why? *Educ Psychol Rev* 13:93–114. <https://doi.org/10.1023/A:1009013420053>
- Sanmugam M, Abdullah Z, Mohamed H, et al (2016) The impacts of infusing game elements and gamification in learning. In: 2016 IEEE 8th International Conference on Engineering Education: Enhancing Engineering Education Through Academia-Industry Collaboration, pp 131–136
- Schmidt RA, Bjork RA (1992) New conceptualizations of practice: common principles in three paradigms suggest new concepts for training. *Psychol Sci* 3:207–218. <https://doi.org/10.1111/j.1467-9280.1992.tb00029.x>
- Schultz PW, Oskamp S, Mainieri T (1995) Who recycles and when? A review of personal and situational factors. *J Environ Psychol* 15:105–121. [https://doi.org/10.1016/0272-4944\(95\)90019-5](https://doi.org/10.1016/0272-4944(95)90019-5)
- Shaffer DW, Squire KR, Halverson R, Gee JP (2005) Video games and the future of learning. *Phi Delta Kappan* 87:105–111. <https://doi.org/10.1177/003172170508700205>
- Shute VJ, Ventura M, Bauer M, Zapata-Rivera D (2009) Melding the power of serious games and embedded assessment to monitor and foster learning: flow and grow. *Serious Games Mech Eff* 2:295–321. <https://doi.org/10.4324/9780203891650>
- Sicart M (2008) Defining game mechanics. *Int J Comput Game Res* 8(2)
- Simons A, Wohlgenannt I, Weinmann M, Fleischer S (2020) Good gamers, good managers? a proof-of-concept study with Sid Meier's Civilization. *Rev Manag Sci* 15(4):957–990. <https://doi.org/10.1007/s11846-020-00378-0>
- Skinner BF (1936) The effect on the amount of conditioning of an interval of time before reinforcement. *J Gen Psychol* 14:279–295. <https://doi.org/10.1080/00221309.1936.9713155>
- Squire K (2007) Games, learning, and society: building a field. *Educ Technol* 47:51–55
- Sreelakshmi R, McLain ML, Rajeshwaran A, et al (2015) Gamification to enhance Learning using Gagne's learning model. In: 2015 6th International Conference on Computing, Communication and Networking Technologies. IEEE, pp 1–6
- Stacey P, Nandhakumar J (2009) A temporal perspective of the computer game development process. *Inf Syst J* 19:479–497. <https://doi.org/10.1111/j.1365-2575.2007.00273.x>
- Stanitsas M, Kirytopoulos K, Vareilles E (2019) Facilitating sustainability transition through serious games: a systematic literature review. *J Clean Prod* 208:924–936. <https://doi.org/10.1016/j.jclepro.2018.10.157f>
- Jürgen Strohm (2001) Industrial waste. <https://boardgamegeek.com/boardgame/2476/industrial-waste>
- Halfbrick Studios Pty. Ltd. (2010) Fruit Ninja. <https://www.halfbrick.com/games/fruit-ninja>
- Tailby R, Haslam C (2003) An investigation of errorless learning in memory-impaired patients: improving the technique and clarifying theory. *Neuropsychologia* 41:1230–1240. [https://doi.org/10.1016/S0028-3932\(03\)00036-8](https://doi.org/10.1016/S0028-3932(03)00036-8)
- Thomas C, Sharp V (2013) Understanding the normalisation of recycling behaviour and its implications for other pro-environmental behaviours: a review of social norms and recycling. *Resour Conserv Recycl* 79:11–20. <https://doi.org/10.1016/j.resconrec.2013.04.010>
- Tibshirani RJ, Efron B (1993) An introduction to the bootstrap. Springer
- United Nations (2020) SDG–Goal 12: ensure sustainable consumption and production. <https://www.un.org/sustainabledevelopment/sustainable-consumption-production/>. Accessed 30 Mar 2020
- Vichitvanichphong S, Talaei-Khoei A, Kerr D et al (2016) Good old gamers, good drivers: results from a correlational experiment among older drivers. *Australas J Inf Syst* 20:1–21. <https://doi.org/10.3127/ajis.v20i0.1110>
- Vygotsky LS (1967) Play and its role in the mental development of the child. *Sov Psychol* 5:6–18. <https://doi.org/10.2753/RPO1061-040505036>

- Webb NM, Ing M, Kersting N, Nemer KM (2013) Help seeking in cooperative learning groups. In: Help seeking in academic settings. Taylor and Francis, pp 45–88
- Webb NM, Farivar S (1999) Developing productive group interaction in middle school mathematics. In: King A, O'Donnell AM (eds) Cognitive perspectives on peer learning. Erlbaum, pp 117–149
- Webster's (2005) Webster's New World College Dictionary, 4th edn. Ref Rev 19:35–36. <https://doi.org/10.1108/09504120510632705>
- Westera W, Nadolski RJ, Hummel HGK, Wopereis IGJH (2008) Serious games for higher education: a framework for reducing design complexity. *J Comput Assist Learn* 24:420–432. <https://doi.org/10.1111/j.1365-2729.2008.00279.x>
- Whalen KA, Berlin C, Ekberg J et al (2018) 'All they do is win': lessons learned from use of a serious game for circular economy education. *Resour Conserv Recycl* 135:335–345. <https://doi.org/10.1016/j.resconrec.2017.06.021>
- Wittwer J, Renkl A (2008) Why instructional explanations often do not work: a framework for understanding the effectiveness of instructional explanations. *Educ Psychol* 43:49–64. <https://doi.org/10.1080/00461520701756420>
- Wolfe MBW, Schreiner ME, Rehder B et al (1998) Learning from text: matching readers and texts by latent semantic analysis. *Discourse Process* 25:309–336. <https://doi.org/10.1080/01638539809545030>
- ZeptoLab (2010) Cut the Rope. <https://www.cuttherope.net/> Accessed 14 Nov 2021

# **Gameful Learning for a More Sustainable World – Measuring the Effect of Design Elements on Long-Term Learning Outcomes in Correct Waste Sorting**

**Greta Hoffmann, Jella Pfeiffer**

Business & Information Systems Engineering (2021)

**Appendix (available online via <http://link.springer.com>)**

## **Experimental Procedure – Detailed**

**P1:** After closing the registration for participation in the experiment, we randomly assigned participants to one of the five experimental groups in a between-subject design. We sent instructions via email for participants to fill out the first survey and download the app (as an apk-file). The game within was locked by a server to prevent premature play and could only be accessed once the first phase officially started. In the survey, we assessed demographic information (age, gender, how long they had been living in Germany, how long they had been living in the city in which the experiment was conducted), participants' game motivation (how much they were involved in and how they felt about these games in general) and their general waste sorting motivation (how they felt about municipal waste sorting). We also included several controls checking language proficiency and conscientiousness in answering the questions. To ensure absolute anonymity in the datasets when linking the game data to the survey entries, each app showed a unique code that participants had to report in each respective survey. For this phase, we set a 48-hour timeframe followed by a pause of 24 hours that allowed for troubleshooting.

**P2:** In the second phase, we sent the next set of instructions as well as another survey link via email. We instructed the participants on the four game-based treatments to open the application and to play it through to the end and then complete the survey. In contrast, we told the control group with the non-interactive materials to attentively read through the teaching materials provided through the link for 25 minutes (this time was derived from the average playtime of the experimental version of the game during the pre-tests) and to then complete the survey. The last part of the survey was the same for all treatments: we measured the perceived usability of the application—or the materials in the case of the non-game material treatment—with the system usability scale (Brooke 1996) as well as self-stated perceived growth in competency and growth in motivation. To adapt the 30 minutes of focused attention to the survey and training, we gave participants a four-day timeframe—including a weekend—to finish the

task. We scheduled the final sessions 10-12 days after the deadline for the second phase, depending on the day of the assigned session.

**P3:** The experiment took place in a laboratory in 19 experimental sessions. Each participant was seated in a cabin where they were guided through the first part of the experiment with the final survey. We first asked participants about their perceived growth in competency and growth in motivation, and there was a final control question on any prior knowledge about the project. Next, we tested the learning outcome in three different performance measures. First, the participants completed a multiple-choice test in which they had to match all 108 trained waste items. Second, we asked all participants to take their phones and start the game application, where they had to sort all 108 items in a special version of the game. Here, each item appeared only once in one big game wave without the two additional design elements. Third, we called the participants into a separate room, where we asked them to sort a selection of real-life waste items. The design of the experimental procedure was pre-tested with seven participants.

## **Design Elaborations on the General Game**

The experiment as described in this manuscript is based on a reduced version of the serious game we built. As a complete game, it features a full set of additional game mechanics.

It can be downloaded here (Apple: <https://apps.apple.com/de/app/die-m%C3%BCll-ag/id1046221391>, Android:

<https://play.google.com/store/apps/details?id=com.bunnyandgnome.mullag>, Windows -Ger:

<https://www.microsoft.com/de-de/p/die-mull-ag/9nblggh6bvny>, Windows-Eng:

<https://www.microsoft.com/en-us/p/trash-monsters/9nblggh6bvny>). As it might be interesting to

readers to know which design decisions we made to inspire long-term interactions with the game, the following sections give further insights into some of the mechanics that were excluded from the experimental version of the game.

## **General Setting**

We designed the overall setting as a cartoon world on a waste sorting planet that serves as the main hub of the game. This world represents a metaphorical holistic view of the waste management process. The waste planet is inhabited by monsters that represent the different waste recycling processes and they all live and work together on the planet as it is their job to take care of the city planets' waste. Their homes can be visited by the player, which we designed to achieve higher emotional involvement and commitment to the topic through social interactions with the monsters. The planet overview screen connects the games' different locations. These locations are i) the waste sorting facility, where the core gameplay takes place, ii) the monsters' living spaces, where players accept quests and different minigames can be played, and iii) the info centre, where players get information on the current state of the game (pollution, sorting correctness and unlocked quests). Adjacent to the waste planet is another smaller planet that represents the respective waste supplier (in the current version, the waste system of Karlsruhe city). As each region in Germany has autonomy in its choice of waste management system, the game is designed to switch systems according to the city planet to which it is connected.

## **Story**

We designed each monster with a different type of personality that is linked to the type of waste they represent. For example, the monster representing the residual waste bin is a dragon that burns incoming residual waste. With this, we wanted to make it transparent what currently happens to objects that are thrown into the residual waste bin. We also made its personality cynical and grumpy as it understands the necessity of its job within the waste management system but at the same time hates the inevitable waste of resources that comes with its assigned job. Players can visit each monster in their home and explore their personalities through conversation. A questline is connected to each monster, resulting in the unlocking of mini quests or additional areas within the game. The questline and story progress are regulated through the game waves and new content is unlocked after each wave. Apart from the

consecutive quest structure, an additional story point is introduced in the middle of the game: the volcano starts to be active again. After a few warning earthquakes, it erupts toward the end of the game. Instead of emitting lava, it erupts into a fountain of waste that had been accumulating within the core of the planet over many years. This initiates the final waves of the game, where players have to sort double and triple amounts of waste at maximum speed to get on top of the emergency. Related to the main story, there is also an underlying mystery surrounding the planet and its history that curious and meticulous players can explore.

### **Unlockable Content (Minigames, Accessories, Mystery)**

We embedded three minigames into different locations that impart additional waste sorting information to the players. The first minigame represents the inner workings of a composting plant and is inspired by the mobile game Fruit Ninja (Halfbrick Studios pty. ltd. 2010). The second minigame represents the process of glass separation at the glass container and the third the operating principles of a battery recycling process. Every time players successfully complete a minigame, they are awarded one of nine accessories that they can present to any monster on the planet as wearables. There are also three upgrades that players can unlock to enhance the core gameplay: a lever that is unlocked in two parts and allows players to either slow down or speed up the conveyor belt, as well as a second conveyor belt that transports waste that would otherwise have fallen off the first one back across the screen. We included these items to give players the control to readjust the difficulty of the main game, to allow them to explore “the outer edges of their competency” (Gee 2003) or take away some pressure from the core gameplay. Finally, the game features a guide where players learn how to recycle paper by themselves and there is a dog that represents bulky waste management and can be trained to pick up bulky waste that occasionally blocks the main game.

### **Exclusion of Game Design Elements**

We designed the game as single player due to limitations in development resources. Thus, game incentives that build upon multiplayer interactions such as leaderboards were not included. Also, while leaderboards can be a strong incentive for certain player types, we

designed the game with a strong focus on collaboration and shared responsibility. For this, we particularly focused our efforts on the design of the relationship that players develop with the monsters. We reasoned that especially young players might be deterred from irresponsibly misorting waste items in real life if they feel an empathic connection to the monsters they encounter within the game.

Apart from the adventure mode, where the core gameplay is embedded into a narrative structure, the app also offers an endless mode of the core gameplay, where the waste items are randomly dropped and players can train and test their sorting skills outside the story mode. This mode would be suitable for the inclusion of competitive elements like leaderboards, which is planned for, if further development of the game becomes possible.

However, badges can particularly be easily perceived as a low-cost/effort playing motivator that could even lead to the opposite of the desired effect (Hamari 2017). Thus, from a design perspective, we believe that such game design elements should only be included in a game or gameful application if there is neither the time nor the budget to design a fully realized experience.

# Non-Game Materials

## Four Bins - Examples

Residual waste		Recycleables	Biowaste*	Paper   Cardboard
Ring binder, plastic	Nylon tights	<b>Wood, untreated, like:</b>	Balcony plants	<b>Recycled paper like:</b>
Ashes - packed	Camera lenses	Wooden boards, Fruit crate	Banana peels	Ring binder - cardboard
Baking/grease-proof paper	Paper, very soiled or imbued	<b>Recycleables, like:</b>	Food waste bin liners	Envelopes, with and without viewing panel
Eye glasses, broken	Paper towels & tissues, soiled	CDs	Bread	Brochures
Sanitary pads	Parchment Paper	Bucket - emptied	Eggshells	Books
Photographic slides	Sticky plaster	Plastic crockery	Fish offal	Egg boxes
Floppy disks	Paintbrush	Bottles, canisters	Vegetable peel	Wrapping paper, uncoated
Extractor fan filter	Porcelain	Plastic film, plastic bags	Hair	Notebooks
Bicycle saddle	Dolls	Childrens toys	Burlap	Cardboard boxes
Pelts / Skins	Cleaning rags	Mixing bowl	Coffee grounds	Catalogues
Binoculars	Eraser	Styrofoam (sundries in transparent bags)	Cheese residues	Magazines
Heat-proof glass	Razor blades	<b>Metals, like:</b>	Bones	Paper - loose
Lighter, empty	Roller skates	Antenna	Dead parts of plants	Papertowels - if only slightly moist
Photographic film	Soot - packed	Baking Dish, metal	Nutshells	Paper packaging
Felt-tip markers, dried out	Shoes - unusable	Sheet metal	Fruit waste	Cardboard
Photographs	Napkins - used	Cans - emptied	Orange peel	Posters
Fountain pen, empty	Skateboard	Electric cable	Seeds	Brochures
Doormat	Mirror glass	Crown cap	Cut flowers	Writing paper
Garden hose	Syringes, safely packed	Brass keys	Left-overs, raw (no soups and sauces)	Packaging from paper, cardboard, carton
Gift wrap, coated	Vacuum cleaner bag	Pans, pots	Tea bags	Advertisements - printed
Light bulbs	Tampons	Tool parts	Potted plants	Journals
Rubber materials	Wallpaper leftovers	<b>Aluminium, like:</b>	Rotting food, without packaging	Newspapers
Suspenders	Electric torch, without batteries	Cling film	Sausage and processed meat leftovers	
Inline skates	Pieces of carpet, chopped up	Yoghurt pot lids	Lemon peel	<b>Recycled paper can also be donated to the wastepaper collection.</b>
Cassettes (audio   video)	Thermal paper	Chocolate foil	<b>Biowaste that does not emerge from households has to be disposed of commercially.</b>	<b>Tips:</b> Remove the tape from your shipping boxes, to fold them in a space-saving manner. Carton packaging can equally be folded flatly. Avoid unnecessary paper waste e.g. by using an advertising ban sticker on your mailbox.
Chewing gum	Thermos flask	Tubes, no contaminants, emptied		
Sweepings	Animal bedding	<b>Composite packaging like:</b>		
Ceramics, no sanitary ceramics	Clocks - no batteries	Blister packs		
Candle stubs	Wound dressing	Milk carton - emptied		
Sticky tape	Packaging, strongly soiled	Paper bag with synthetic padding		
Sticky labels	Hot-water bottle - gummi	Juice cartons - emptied		
Carbon paper	Absorbent cotton	Vacuum packaging		
Ball pen refill	Wicker basket			
Pleather	Diapers	<b>Packaging, scraped clean, from:</b>		
Cuddly toys	Cigarette ends	Wood, plastic, metal		
Leatherbags, -belts	Ignition plugs			
Left-over linoleum				
Airbed				
Rags				
Crayons, solvent-free				
Left-over bits of food				

Some waste types are collected via alternative waste disposal facilities like recycled glass-, used textiles-, organic waste containers, composting facilities, recycling stations, bulk waste collections as well as contaminant collection facilities.

Electrical and electronic appliances up to 50 centimeters edge length can be dropped free of charge at all recycling stations.

Large appliances can also be dropped free of charge at all recycling stations. Large electrical appliances (ovens, stoves, cooling- and freezing appliances, dryers, washing machines) will be picked up on demand. Small equipment can be placed adjacently free of charge.

Figure 1. Flyer on General Waste Sorting in Karlsruhe, translated



**RESTMÜLL**  
 Lumpen, Gummi, Windeln, Hygieneartikel, stark Verschmutztes, Ton, Staubsaugerbeutel, Kippen, Porzellan, Glühbirnen

**WERTSTOFF**  
 Kunststoff, Metall, unbehandeltes Holz und Verpackungen aus diesen Materialien. Alufolie, Getränkeverpackungen, Styropor

**BIOABFALL**  
 Gemüse- und Obstreste, gekochte und ungekochte Speisereste, Eierschalen, Kaffeefilter, Fleischreste, Blumen und Topfpflanzen

**PAPIER/PAPPE**  
 Papier, Pappe, Karton und Verpackungen aus diesen Materialien, Papiertüten, Zeitungen, Schreibpapier, Bücher, Kataloge

Figure 3. Flyer on Bins and Representative Waste Items

**RESTMÜLL**

**WAS?**  
 Staub, Klebricht, Winceln, Asche (alle verpackt), Keramik, Porzellan, stark verschmutztes Material (wie Hygieneartikel, Glühbirnen), zerstückte Teppichreste, Tapetenreste, nicht tragfähige Kleidung und Schuhe

**WOHIN?**  
 in die Restmülltonne oder Anlieferung an der Wertstoffstation.  
 Für vorübergehend anfallende Spitzenmengen können Sie Ihr Handelsspezial-Abfallsäcke verwenden.

**ORGANISCHE ABFÄLLE**

**WAS?**  
**Küchenabfälle** wie Obst- und Gemüseabfälle, Teebeutel, Kaffeesatz, Eierschalen, Hefe und gekochte Speisereste, verwelkte Blumen  
**Grünabfälle** wie Laub, Rasen und Baumanschnitt, verwelkte Pflanzen

**WOHIN?**  
**Küchenabfälle** in die Biotonne oder selbst kompostieren.  
**Grünabfälle** zum Grünabfallcontainer, zu den Wertstoffstationen und zu den Kompostierungsanlagen.

**GLAS**

**WAS?**  
**Hohlglas** wie Einwegflaschen, Marmeladen- und Obstgläser  
**Flachglas** Fensterglas ohne Rahmen, Aquarienglas, Glasregalböden

**WOHIN?**  
**Hohlglas** getrennt nach Farben in die Altglasstationen. Asen, besser Pfandflaschen benutzen.  
**Flachglas** nur zu den Wertstoffstationen.

**WERTSTOFFE UND PAPIER**

**WAS?**  
 Metall, Kunststoff, Styropor, unbehandeltes Holz und Verpackungen aus diesen Materialien  
 Papier, Pappe, Karton und Verpackungen aus diesen Materialien

**WOHIN?**  
**Wertstofftonne** für Metall, unbehandeltes Holz und Kunststoff.  
**Papiertonne** für Papier, Pappe und Kartonen. Wertstoffe vorsortiert zu den Wertstoffstationen.  
 Papier zur Glasabholung.

**SCHADSTOFFE**

**WAS?**  
 Energiesparlampen, Leuchtstoffröhren, Perle, Kleber, Lacke, Acrylnitrat, Kosmetika, Netze, Öle, Lötlötzinn, sonstige Chemikalien, Medikamente, Pflanzen- und Schädlingsbekämpfungsmittel, dazugehörig

**WOHIN?**  
 Mobile Schadstoffsammlung oder Schadstoffannahmestellen  
**Energiesparlampen** auch zu allen Wertstoffstationen.  
 Feuerlöscher (Leeren zurück zum Händler) oder zu den Batteriesammelbehältern im Stadtgebiet. Autoblech (Leeren zum Händler (Pfand) oder zur Schadstoffsammlung).

**TEXTILIEN**

**WAS?**  
 Kleinststücke, die noch gut erhalten sind  
 Schulleinweisung zusammengebunden

**WOHIN?**  
 Alttextiltrommel bei einem Altglascontainer und auf den Wertstoffstationen oder karitative Organisationen, Tausch- und Verkaufsmärkte, Second-Hand-Läden, Flohmärkte.  
 Nicht mehr tragbares in die Restmülltonne, Teppiche zum Sperrmüll.

Figure 2. Flyer on Waste Categories

## Additional Literature Overviews

**Table 1.** Literature Comparison between Errorful (EF) and Errorless (EL) Learning

Authors	Context	Subjects	Conclusion
Baddeley and Wilson (1994)	Clinical	16 people with brain injuries and memory impairment, and 16 young and older controls each	EL is better than EF
Clare et al. (1999)	Clinical	One participant with Alzheimer's disease	EL is effective and useful for memory problems
Clare and Jones (2008)	Clinical	Six participants with early-stage DAT	EL is effective and useful for memory problems
Donaghey et al. (2010)	Clinical	30 people with an amputated limb, randomly assigned to either the experiment or control group	EL is better than EF
Dunn and Clare (2007)	Clinical	10 people with different conditions	No difference
Evans et al. (2000)	Clinical	Phase 1: 18 people with brain injuries and memory impairment. Phase 2: 16 people with brain injuries and memory impairment. Phase 3: 34 people with brain injuries and memory impairment	Mixed results but overall better performance with EL
Hunkin et al. (1998)	Clinical	Eight people with memory impairment	EL is better than EF
K. Ivancic and Hesketh (2000)	Driving Education	Experiment 1: 44 people in two equal groups Experiment 2: 32 people in two equal groups	EF is better than EL
Johnson (2004)	Learning Strategies	Evidence aggregation of different studies	EF is better than EL
Jones and Eayrs (1992)	Teaching Strategies	Literature synopsis	Inconclusive
Kessels and Haan (2003)	Natural Ageing	18 elderly and 16 young controls	EL is better than EF
Kessels et al. (2007)	Clinical	10 people with Korsakoff Syndrome	No difference
Ohlsson (1996)	Learning Strategies	Tests on the evaluation of own performance errors—more theoretical	Inconclusive
Prather (1971)	Airforce Education	96 people	EF and EL are similarly effective
Page et al. (2006)	Clinical	Experiment 1: 23 people with memory impairment and 20 controls Experiment 2: 20 people with memory impairment	EL is better than EF
Tailby and Haslam (2003)	Clinical	24 people in three groups of eight each with different severity of memory impairment	EL is better than EF

## Control and Additional Variables

**Table 2.** Operationalization of Control and Additional Variables

Controls	English	German	Tested with
Age	“Please tell us your age”	„Bitte teile uns Dein Alter mit.“	Integer value
Gender	“Which gender do you identify with?”	„Welchem Geschlecht ordnest Du Dich zu?“	Male/female/other <i>Männlich/ Weiblich/ Sonstiges:</i>

Living in Germany	How long have you been living in Germany? Please answer with number of full years.	„Wie lange wohnst Du schon in Deutschland? Bitte antworte in ganzen Jahren.“	Integer value
Living in XX City	How long have you been living in XX? Please answer with number of full years.	„Wie lange wohnst Du schon in XX? (Bitte antworte in ganzen Jahren)“	Integer value
Game motivation (medium acceptance)	Please tell us about your attitude towards games.	„Bitte teile uns Deine Einstellung gegenüber Gaming mit.“	(sub-headline)
	I play videogames (computer games, smartphone games, console games, ...) in my free time.	„Ich spiele in meiner Freizeit Videospiele (Computerspiele, Handygames, Konsolenspiele,...).“	Likert (five-point): Strongly disagree, rather disagree, neither agree nor disagree, rather agree, strongly agree <i>Stimme gar nicht zu, stimme eher nicht zu, teils-teils, stimme eher zu, stimme voll und ganz zu</i>
	I am prejudiced towards grown-ups who play videogames. (r)	„Ich habe Vorurteile gegenüber erwachsenen Menschen, die Videospiele spielen.“(r)	
	I wish videogames were more accepted in society.	„Ich wünschte, Videospiele würden eine höhere Akzeptanz in der Gesellschaft genießen.“	
	I think videogames are a waste of time. (r)	„Ich denke, dass Videospiele eine Form der Zeitverschwendung sind.“(r)	
	Videogames are my hobby.	„Videospiele sind mein Hobby.“	
	I feel that too much attention is spent on videogames. (r)	„Ich finde, dass man Videospiele zu viel Aufmerksamkeit schenkt.“(r)	
General waste sorting motivation (general interest in the topic)	What is your attitude towards waste sorting at home? Please answer honestly.	„Wie ist Deine Einstellung zu Mülltrennung? Bitte antworte ehrlich.“	Likert (five-point) Fully applicable, rather applicable, partly applicable, rather not applicable, not applicable <i>trifft voll zu, trifft eher zu, teils-teils, trifft eher nicht zu, trifft nicht zu</i>
	I have never given any thought to waste sorting.	„Ich habe mir noch nie über Mülltrennung Gedanken gemacht.“	
	Waste sorting at home is very important to me.	„Mir ist Mülltrennung im Haushalt sehr wichtig.“	
Waste sorting motivation and competency	Please let us know to what extent you agree with the following statements.	„Bitte teile uns mit, inwiefern Du den folgenden Aussagen zustimmst“	Likert (five-point) Strongly disagree, rather disagree, neither agree nor disagree, rather agree, strongly agree <i>Stimme gar nicht zu, stimme eher nicht zu, teils-teils, stimme eher zu, stimme voll und ganz zu</i>
Waste sorting motivation: last two weeks	Since part 2 of the experiment, have you been more motivated to correctly sort your waste?	Warst Du seit Teil 2 des Experimentes motivierter, Deinen Müll korrekt zu trennen?	
Waste sorting motivation: from now on	Since part 2 of the experiment, have you felt more skilled at correctly sort your waste?	Hast Du Dich seit Teil 2 des Experimentes kompetenter darin gefühlt, Deinen Müll richtig zu trennen?	
Waste sorting competency: last two weeks	After participating in this experiment, do you feel more motivated to correctly sort your waste from now on?	Bist Du nach Abschluss dieses Experiments motivierter, ab jetzt Deinen Müll korrekt zu trennen?	
Waste sorting competency: from now on	After participating in this experiment, do you feel more skilled at correctly sort your waste from now on?	Fühlst Du Dich nach Abschluss dieses Experiments kompetenter darin, Deinen Müll ab jetzt richtig zu trennen?	
SUS	See Brooke (1996).	See Brooke (1996)	

(r) refers to the questions being reverse-coded

**Table 3.** Control Variables - Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Scale/Type of Measure
Age	22.72	3.01	17	41	Age in years (integer values)
Living in Germany	20.73	5.90	0	30	Number of years (integer values)
Living in XX City	4.28	5.46	0	28	Number of years (integer values)
Gaming motivation	3.12	.90	1.17	5	Likert five-point (six items, three reverse-coded)
General waste sorting motivation	4.23	.80	1.5	5	Likert five-point (two items)
SUS	78.79	12.93	30	100	SUS score: map answers (Likert five-point) from 0 (lowest) to 4 (highest), add the values of all 10 items and multiply by 2.5

**Table 4.** Control Variables – Descriptive Statistics per Treatment

	Non-game material		Repeat element		Look-up element		Combined		Core gameplay	
	mean (min/max) / percent for gender	std. dev.	mean (min/max)	std. dev.	mean (min/max)	std. dev.	mean (min/max)	std. dev.	mean (min/max)	std. dev.
Age	23.28 (19/30)	3.28	22.6 (18/41)	3.76	22.42 (17/28)	2.46	23.34 (18/32)	2.83	22.09 (19/30)	2.47
Gender (male)	71.8%		65.2%		75.6%		63.4%		54.5%	
Gender (female)	28.2%		32.6%		24.4%		36.6%		45.4%	
Gender (diverse)			2.2%							
Living in Germany	21.85 (3/30)	5.46	20.22 (2/28)	5.33	21.13 (1/28)	5.48	21.34 (3/28)	5.64	19.30 (0/30)	7.27
Living in XX City	4.08 (0/28)	5.60	4.82 (0/23)	6.13	3.77 (0/22)	4.00	4.46 (0/27)	5.74	4.26 (0/27)	5.77
Gaming motivation	3.25 (1.17/5)	1.05	3.04 (1.17/5)	.89	3.21 (2/4.83)	.80	3.03 (1.33/4.67)	.88	3.07 (1.17/4.67)	.88
General waste sorting motivation	4.13 (1.5/5)	.92	4.23 (2/5)	.74	4.13 (2/5)	.84	4.44 (2.5/5)	.64	4.23 (2/5)	.84
SUS	76.73 (32.5/95)	14.13	78.91 (47.5/95)	11.91	75.44 (45/97.5)	13.86	81.59 (42.5/100)	10.81	81.31 (30/100)	13.13

## Additional Analyses

**Table 5.** Effect of the Game in Comparison with the Non-Game Material Group with Control Variables

	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
<b>Reference category: Non-game material</b>	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)

Game (all 4 game treatments)	.045 (.016) [.019, .072]	.005**	.090 (.019) [.058, .121]	.000**	.068 (.031) [.018, .119]	.025*
<b>Control Variables</b>						
Age	.000 (.003) [-.005, .005]	.961	-.001 (.003) [-.007, .005]	.714	-.007 (.004) [-.015, .000]	.060
Gender	-.020 (.014) [-.048, .007]	.145	-.017 (.016) [-.049, .015]	.306	-.049 (.027) [-.102, .005]	.073
Living in Germany	.005 (.001) [.002, .007]	.000**	.005 (.002) [.002, .008]	.002*	.003 (.003) [-.002, .009]	.229
Living in XX City	.001 (.001) [-.001, .003]	.341	.002 (.001) [-.001, .004]	.170	.002 (.002) [-.002, .005]	.406
Gaming motivation	.009 (.007) [-.005, .023]	.210	.009 (.008) [-.007, .025]	.290	.010 (.013) [-.016, .036]	.446
General waste sorting motivation	.020 (.008) [.004, .035]	.012*	.016 (.009) [-.002, .033]	.085	.027 (.016) [-.005, .059]	.100
SUS	.001 (.000) [-.001, .002]	.064	.001 (.001) [-.000, .002]	.262	.001 (.001) [-.001, .003]	.181
Constant	.426 (.066) [.318, .534]	.000**	.361 (.079) [.231, .492]	.000**	.570 (.108) [.392, .748]	.000**
N		213		213		213
R <sup>2</sup>		.193		.212		.091
Adj. R <sup>2</sup>		.161		.181		.055

For the treatment groups, we used an alpha-error level of 10% (\*p<0.1, \*\*p<0.01).

For the other controls that did not have directed hypotheses, we set the alpha-error level to 5% (\*p<0.05, \*\*p<0.01).

Male was coded as 0, female as 1 and diverse as 2.

**Table 6.** Effects of the Design Elements in Comparison with the Non-Game Material Group with Control Variables

Reference category: Non-game material	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)
Repeat element	.033 (.020) [.001, .066]	.094*	.086 (.023) [.048, .124]	.000**	.073 (.038) [.010, .135]	.056*
Look-up element	.044 (.021) [.009, .078]	.037*	.090 (.023) [.052, .127]	.000**	.072 (.037) [.012, .132]	.050*
Combined	.076 (.019) [.044, .107]	.000**	.117 (.023) [.079, .154]	.000**	.056 (.040) [-.010, .123]	.163
Core gameplay	.029 (.020) [-.004, .062]	.144	.065 (.023) [.027, .104]	.005**	.071 (.035) [.013, .129]	.045*
<b>Control Variables</b>						
Age	-.000 (.002) [-.005, .004]	.841	-.002 (.003) [-.007, .004]	.542	-.007 (.004) [-.015, .001]	.074
Gender	-.019 (.014) [-.046, .008]	.175	-.014 (.016) [-.046, .018]	.389	-.049 (.028) [-.103, .005]	.078
Living in Germany	.004 (.001) [.002, .007]	.000**	.005 (.002) [.002, .008]	.002**	.003 (.003) [-.002, .009]	.239
Living in XX City	.001 (.001) [-.001, .003]	.265	.002 (.001) [-.001, .004]	.140	.002 (.002) [-.002, .005]	.433
Gaming motivation	.009 (.007) [-.005, .023]	.195	.009 (.008) [-.007, .026]	.267	.010 (.013) [-.016, .036]	.461

General waste sorting motivation	.018 (.008) [.003, .033]	.019*	.014 (.009) [-.003, .032]	.114	.028 (.017) [-.005, .060]	.095
SUS	.001 (.000) [-.000, .002]	.064	.001 (.001) [-.000, .002]	.255	.001 (.001) [-.001, .003]	.170
Constant	.449 (.067) [.318, .580]	.000**	.382 (.079) [.228, .536]	.000**	.561 (.108) [.347, .774]	.000**
N		213		213		213
R <sup>2</sup>		.219		.233		.092
Adj. R <sup>2</sup>		.176		.191		.042

For the treatment groups, we used an alpha-error level of 10% (\*p<0.1, \*\* p<0.01).

For the other controls that did not have directed hypotheses, we set the alpha-error level to 5% (\* p<0.05, \*\* p<0.01).

**Table 7.** Effects of the Design Elements in Comparison with the Core Gameplay

Reference category: Core gameplay	In-Game Performance		Multiple-Choice Test		Real-Life Sorting	
	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)	coef. (bootstr. std. error) [conf. interval]	p (two-tailed)
Repeat element	.004 (.019) [-.027, .035]	.831	.021 (.022) [-.015, .056]	.337	.002 (.033) [-.052, .056]	.958
Look-up element	.015 (.021) [-.019, .049]	.470	.024 (.021) [-.011, .059]	.256	.001 (.033) [-.023, .055]	.978
Combined	.047 (.019) [.016, .077]	.012*	.052 (.022) [.017, .086]	.015*	-.015 (.036) [-.073, .044]	.681
Non-game material	-.029 (.020) [-.062, .004]	.144	-.065 (.023) [-.104, -.027]	.005*	-.071 (.035) [-.129, -.013]	.045*
<b>Control Variables</b>						
Age	-.000 (.002) [-.005, .004]	.841	-.002 (.003) [-.007, .004]	.542	-.007 (.004) [-.015, .001]	.074
Gender	-.019 (.014) [-.046, .008]	.175	-.014 (.016) [-.046, .018]	.389	-.049 (.028) [-.103, .005]	.078
Living in Germany	.004 (.001) [.002, .007]	.000**	.005 (.002) [.002, .008]	.002**	.003 (.003) [.002, .009]	.239
Living in XX City	.001 (.001) [-.001, .003]	.265	.002 (.001) [-.001, .004]	.140	.002 (.002) [-.002, .005]	.433
Gaming motivation	.009 (.007) [-.005, .023]	.195	.009 (.008) [-.007, .026]	.267	.010 (.013) [-.016, .036]	.461
General waste sorting motivation	.018 (.008) [.003, .033]	.019*	.014 (.009) [-.003, .032]	.114	.028 (.017) [-.005, .060]	.095
SUS	.001 (.000) [.000, .002]	.064	.001 (.001) [-.000, .002]	.255	.001 (.001) [-.001, .003]	.170
Constant	.478 (.065) [.350, .606]	.000**	.447 (.078) [.294, .601]	.000**	.632 (.107) [.421, .842]	.000**
N		213		213		213
R <sup>2</sup>		.219		.233		.092
Adj. R <sup>2</sup>		.176		.191		.042

For the treatment groups, we used an alpha-error level of 10% (\*p<0.1, \*\* p<0.01).

For the other controls that did not have directed hypotheses, we set the alpha-error level to 5% (\* p<0.05, \*\* p<0.01).