Can't Get You Out of My Head: A Scoping Review on Intrinsic Motivation Techniques and Their Use in Memory-Based Listening Games

Lot M. Wolters

Master Thesis – Applied Cognitive Neuroscience

s3432025
June 2022
Department of Psychology
University of Groningen
Examiner:
Annelieke Roest (RuG)
Daily supervisor:
Henkjan Honing (UvA)

Abstract

What would one do without motivation? According to the self-determination theory (SDT) by Deci & Ryan (1980), one is either intrinsically or extrinsically motivated. For the future design of an online memory-based listening game that investigates demographic differences in catchiness of music, intrinsic motivation is an important element since an external reward for participation seems unrealistic on such large scale. To motivate hundreds of thousands of people from all over the world to participate in the game, effective intrinsic motivation techniques need to be applied. Therefore, the aim of the current study is to find out what techniques are most relevant and useful to intrinsically motivate participants to engage in an online memory-based listening game. This is investigated by means of a scoping review. During the selection of sources of evidence, 22 studies were eligible to be included in the scoping review. Results from studies that implemented augmented reality to those that requested participants to recreate their school building in a videogame, revealed that the techniques that are most relevant and useful are: a) collaboration or competition, b) challenge, c) risk (related to rewards), d) storytelling, as well as e) positive and negative feedback.

Keywords: Games, intrinsic motivation, listening, memory, music perception

Can't Get You Out of My Head: A Scoping Review on Intrinsic Motivation Techniques and Their Use in Memory-Based Listening Games

What would one do without motivation for one's behaviour, actions, or achievements? It is rational to say that motivation drives us to do anything at all. Nevertheless, how a person is motivated, may differ substantially (Deci and Ryan 2008). According to the self-determination theory (SDT) by Deci and Ryan (1980) one might be *intrinsically* or *extrinsically* motivated. The SDT poses that intrinsic motivation is crucial for effective learning. When intrinsically motivated, behaviour is driven by one's motives, goals, values, and interests, causing volitional behaviour. Basic needs for psychological well-being, like relatedness, competence, and autonomy, are important aspects that are facilitated by intrinsic motivation. On the other hand, extrinsic motivation is driven by external values like rewards or competition, rather than the desire for individual development. Rewarded tasks are often perceived as controlling and do therefore not meet the basic psychological need of autonomy (Deci & Ryan, 1985). However, positive feedback may enhance the feeling of competence, therefore it may serve as an intrinsically motivating factor.

Intrinsic versus extrinsic motivation

The reciprocity between extrinsic and intrinsic motivation is known as the overjustification effect. This suggests that "a reward decreases motivation to engage with a previously intrinsically motivated activity when it causes people to attribute their motivation to the reward" (Lepper et al., 1973, as cited in Sansone & Tang, 2021), meaning that an initially intrinsically motivated person will become less intrinsically motivated when a reward is introduced. However, according to Wiersma (1992) the number of studies that support this phenomenon is more or less equal to the number of studies that fail to support this phenomenon. Rewards may increase intrinsic motivation if these rewards facilitate basic psychological needs (Deci & Ryan, 2000). Therefore, it is important to take into account the type of reward that is given.

Based on the goal of one's study, one should consider which motivational technique to adopt when recruiting participants. Extrinsic motivation might be preferred over intrinsic motivation in time-consuming studies or in studies that involve some risk for the participants. Assuming that one would not voluntarily or is less inclined to register for studies like this, rewarding for participation would recruit more participants.

On the contrary, one might prefer intrinsic motivation rather than extrinsic motivation when lots of experimental data are required to draw any conclusions from the data. Intrinsic motivation is an effective technique to recruit a large pool of participants (Honing, 2021). In

this paper, the focus is mainly on intrinsic motivation, since it is written to give an advice on the future design of a study that will investigate the catchiness of music and the demographic differences in the features of catchiness. To be able to draw any conclusions out of the data, large amounts of experimental data are required which cannot be easily obtained in the lab. To get hundreds of thousand participants world-wide, a reward like money is not feasible. Hence intrinsic motivation is an important aspect of the experimental design. With intrinsic motivation, people will be keen to engage in the game and this engagement provides us with data. Therefore, when developing game-like, large scale experiments, it is important to investigate in which ways one can optimise intrinsic motivation.

Examples of intrinsically motivating experiments

One way to intrinsically motivate participants to engage in a psychological experiment and to make the task more pleasant, is by gamification of the experiment. That is, not adding tangible rewards to the non-game context, but rather game-like features to the non-game context (Ferrara, 2013). An example of an experiment that makes use of gamification is the study by Barata et al. (2013). In their five-year experiment, they investigated the impact of gamification on learning experience. They compared the student's behaviour during a master's course (Multimedia Content Production) for different years. The researchers had access to data of attendance to lectures, posts, downloads of course documents, and grades from academic years 2007/2008 to 2011/2012. Of these five years, the course was not gamified for three years, but was gamified for the most recent two years. The differences between these two conditions were analysed. Especially relevant for this paper, is the way in which the course was gamified, thus intrinsically motivating, and what the consequences were.

The course evaluation of the non-gamified course consisted of quizzes, lab evaluations, online participation in the course's forums, a presentation, and a final exam, which resulted in a final grade between 0 and 20. In the gamified course, the final grade was translated to experience points (XP) for the evaluation methods. Next to the XP, progress levels, leader boards, challenges, and badges were added. The students received feedback instantly, by the immediate awarding of XP. Because the environment was uncontrolled, there might be issues with the internal validity and replicability of the results. However, it is reasonable to posit that the external validity might be higher due to the uncontrolled environment that better resembles real-world learning than learning in a lab. After analysing the data, the researchers found no significant increase in final grades and attendance but did find a significant increase in number of downloads in the first year of the gamified course, and an overall increase in

posts on forums in the gamified course. These two increases suggest that students are more engaged in the course when the course is gamified.

Another example of an intrinsically motivating experiment is the study by Butler and Walton (2013). During their experiment, pre-schoolers were instructed to work on a challenging task (i.e., a puzzle). The pre-schoolers were divided in three groups: one experimental condition group that worked physically alone, but was manipulated to feel a psychological collaboration, and two non-collaborative control conditions. The psychological feeling of collaboration was constituted by conveying the children were working on the puzzle together, or that the other child had worked on the puzzle before. This feeling of collaboration was obtained through, for example, instructing the participants that they are "taking turns" of working on the puzzle with other participants in different rooms. The aim of this study was to investigate the role of (a feeling of) collaboration on the motivation of pre-schoolers. A significant increase in motivation for as well as persistence on, and liking of, the puzzle was found, compared to the two non-collaborative control conditions. These results suggest that engaging in collaborative tasks, leads to an increased intrinsic motivation and might therefore be an important aspect to enhance intrinsic motivation.

Thirdly, the online memory-based listening game called 'Hooked on Music' is an app that consists of two tasks: a recognition task, and a verification task (Korsmit et al., 2017). These tasks should enable the researchers to investigate the influence of catchiness of a piece of music on the memory of that piece. The researchers hypothesized that a fragment is catchier when it is better recalled after a long period of time. During the recognition task, the participant hears 15 seconds of the beginning of a verse or chorus and had to report whether they recognize the song or not. If the answer is "Yes", the verification task starts. During this second task, the sound is muted for four seconds, and the participant is instructed to sing along while the sound is muted. During the muting, the sound was either paused or silently continued. When the sound is unmuted, the participant is asked whether the music continued in the correct spot or not. This task verifies whether the participant truly knew the song, by correctly reporting whether the music continued in the correct spot or not. For both tasks, half of the trials the correct answer is "Yes" and for the other half the correct answer is "No".

Within roughly a year, 130.000 participants had played the game, emphasizing the effective intrinsic motivation techniques used. The study especially tackled the challenge and self-efficacy since one might be motivated to participate to proof their competence by answering these questions. In addition, this task is intrinsically motivating due to the dopaminergic effect of the stimulus, i.e., music, and no reward was given for participation

(Salimpoor et al., 2013). The dopaminergic effect of music is explained by one's anticipation to specific moments in one's favourite music. This causes an increase in dopamine production, which in its turn positively affects mood (Schultz, 2010). Since the participants probably will encounter some of their favourite music during the experiment, the experiment itself is intrinsically motivating due to its stimuli. The results revealed that age and music preference or attention to specific musical characteristics play an important role in individual differences in long-term musical memory.

A final example of an intrinsically motivating experiment is the study by Hao and Lee (2019) that investigated the differences between augmented reality (AR) learning and traditional learning of English vocabulary. 147 fifth graders were spread over two conditions: an experimental group (AR learning), and a control group (traditional learning). The control group learned English through slideshows and textbooks, while every 2 participants in the experimental group got an iPad Mini in pairs to learn English with an AR game after they got the instructions from their English teacher. In addition to the expected benefits of gamification, the AR technique made use of computerized adaptive testing (CAT), or challenge. This involves the adjustment of difficulty level of the tasks to the performance level of the player. The experiment continued for 2 hours every week for 4 weeks. The results revealed that for all dimensions (i.e., attention, relevance, confidence, and satisfaction), significant differences were found between the AR games teaching and the traditional teaching condition in favour of the AR games teaching condition. These results suggest that learning with AR games, which include CAT, would be an effective way to improve students' way of learning. However, the learning outcomes did not differ significantly between the two conditions. With respect to motivation, the implementation of peer discussions, challenge and CAT would significantly improve motivation.

Moreover, previous research by Van der Linden & Glas (2000, as cited in Wauters et al., 2010) has already proven the benefits of CAT. An advantage of CAT is that testing time is typically lower, compared to nonadaptive tests. Also, by keeping up the challenge for participants, the participants tend to be more motivated during the experiment, which will be discussed later on.

To sum up the aforementioned techniques to intrinsically motivate participants to engage, one could use gamification, collaboration, competition, and/or CAT.

Online experiments

As well as considering the way in which to motivate participants, it is important to consider the validity and what type of validity is important. Especially in online experiments,

issues with internal validity might arise due to the uncontrolled environment (Bridges et al., 2020). However, in the case of a listening game, it could be argued that the internal validity is indeed lower, but is outweighed by the increased external validity and the more diverse group of participants that is attracted (Honing, 2021). In reality, when one listening to music the environment is often noisy, the quality of the headphones might be low, and one might be distracted by their surroundings. It is questionable if an experiment like this should be executed in the lab since the circumstances in a lab are not comparable to reality. Nevertheless, an uncontrolled environment might result in issues with replicability. However, one could argue that the high internal validity might rather be a cause of than a cure for poor reproducibility (Honing, 2021). Some authors even argue that if an effect is detected despite of all kinds of real-world variance and technological equipment, this is stronger evidence than an effect found in the lab (Honing & Reips, 2008). Specifically in settings of music perception, people do not often listen to music in a fully quiet environment and are often disturbed.

All in all, one should consider what is more important, external, or internal validity. This depends on the goal of the experiment. In the case of an online memory-based listening game, for which lots of experimental data are needed to draw conclusions on a demographical scale, one might be inclined to value external validity more than internal validity. With the focus on external validity, the environment of the experiment does not necessarily need to be controlled, which can be beneficial for the costs of the experiment. In the case of an online experiment, participants are often able to participate on their own devices. Moreover, experiments that focus on external validity are usually better able to replicate real-world situations, which is relevant for the context in which participants listen (to music, but also in conversations).

The current study

In short, the current paper is written in preparation for the construction of a new online memory-based listening game. The goal of this game is to investigate demographical differences in the catchiness of musical fragment and subsequently investigates the characteristics of this catchiness. To be able to draw any conclusions about demographical differences in catchiness of music, lots and lots of experimental data are required. Therefore, extrinsic motivation driven by a reward, for example, seems impossible. Hence, the focus is on intrinsic motivation.

To be able to eventually construct a new online memory-based listening game, the following research question was established: "*How can existing intrinsically motivating*

experiments used in a variety of domains inform us about intrinsic motivation techniques to be used in future designs of online memory-based listening games?". This research question will be answered by means of a scoping review, that discusses experiments that make use of different intrinsic motivation techniques. Both the impact of the different techniques on intrinsic motivation, as well as the impact of the different techniques on learning outcomes are discussed. When confronted with equally effective techniques regarding intrinsic motivation, one might consider looking at the learning outcomes to make a choice between these techniques.

Method

For this scoping review, no review protocol exists and therefore also no registration. The scoping review is written in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement for Scoping Reviews (Tricco et al., 2018). The review only includes experiments that tackles the memory and that intrinsically motivates the participants. Secondly, only peer-reviewed journal articles were included, since these are mostly focused on experiments while other sources may contain reviews or other, for this study, redundant information. The other inclusion criteria are for the language, namely articles that are written in English, the population, namely human participants, and for the participation, namely studies that do not reward participants for participation.

For the identification of possibly relevant documents, the following databases were searched: EBSCO (PsycINFO) and PubMed. Evidently, PsycINFO is most relevant for the search of articles in the field of psychology and memory, so this database was included. In order to also include articles that study the memory in other domains than psychology, PubMed was considered for inclusion. After controlling for duplicates, PubMed was indeed found to be a suitable addition.

The search terms that were used are described in Table 1. As a limit, the checkbox of 'Apply equivalent subjects' was unchecked, so the tightly related papers would not be shown. This would already filter out some irrelevant papers.

Table 1

Key Search Terms

Database	Search strings
PsycINFO	("memory gam*" OR "memory experiment" OR "retrieval gam*" OR "retrieval experiment" OR "learning gam*" OR "learning experiment") AND motivation

PubMed	("memory game" OR "memory experiment" OR "retrieval game" OR "retrieval experiment" OR "learning game" OR "learning experiment") AND motivation

During the selection of sources of evidence, firstly the relevance of the paper was determined based on the title and the abstract. The abstract was not necessarily read, since some papers can be excluded solely based on their title. Next, the method section was read to determine whether the paper describes an experiment and what its characteristics are. Based on this information, the inclusion or exclusion of each paper was decided.

For the data charting, all articles were read, and the relevant information (research question or objectives, participants and study design, intrinsic motivation techniques used, and the outcomes) was selected and processed in Table A in the appendix. Thus, not all information from the articles is discussed in Table A, since not everything discussed in the articles was relevant for the purpose of the current scoping review. The relevant research questions and study design elements were selected, with their corresponding outcomes. To see whether a certain demographic trend is present in the outcomes, Table A contains colours that represent the different demographic regions. The distribution is visible in the appendix, Figure A, and is based on the distribution used in the board game ‘Risk’. Regarding the data items, no specific variables were sought since many types of intrinsically motivating experiments with different variables were useful for this scoping review. Relevant intrinsic motivation elements included, for instance, engagement, motivation, or game attitude of the players.

Table 2 provides an overview of the intrinsic motivation techniques that are discussed in the different studies. The techniques are categorized, meaning that the techniques are not all explicitly mentioned in the articles and may have been described in other words in the articles. For example, the element of peer discussion is grouped under collaboration, and CAT is grouped under challenge.

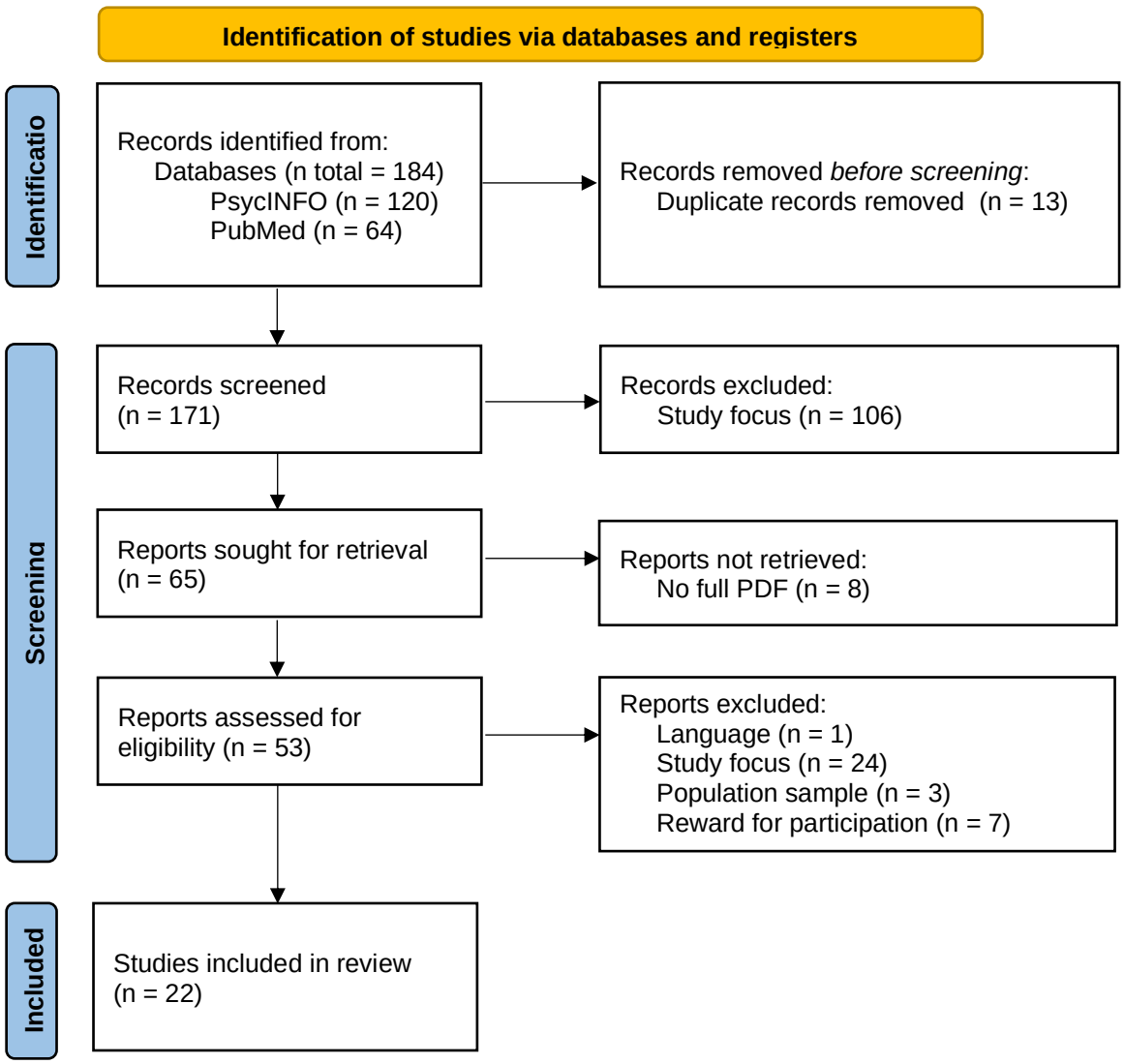
Results

The scoping review yielded 22 studies to include. Figure 1 describes the process of the selection of sources of evidence. The most important findings of the sources of evidence are presented in the appendix, Table A. Most studies investigated different factors of gamification on the engagement, motivation, performance, or game attitude of the players. When thinking of these factors, one can think of collaboration, competition, or uncertainty, for instance. Techniques that were especially often implemented in the design of the

experiment, were CAT (i.e., adapting the level of difficulty to the performance level of the player) and feedback (i.e., giving participants an update on their performance, both positive and negative). Both CAT and feedback were implemented in 7 out of 22 studies. Regarding the demographic trend of intrinsic motivation techniques and their significance, remarkable is that studies from Europe do not report any insignificant results of the association between intrinsic motivation techniques and intrinsic motivation. The number of European studies that investigate these techniques is similar to the number of Asian studies (i.e., 8 studies), but 2 Asian studies do report insignificant associations. Next to this remark, there is no clear demographic trend in effective intrinsic motivation techniques.

Figure 1

PRISMA 2020 Flow Diagram for the Selection of Sources of Evidence.



Note. Adapted from "PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation" by A.C., Tricco et al., 2018, *Annals of Internal Medicine*. *167*(7), 467-473.

In total, 60 statements on associations between 14 different intrinsic motivation techniques and intrinsic motivation or learning outcomes are made. Regarding associations with intrinsic motivation, 33 statements are made of which 27 statements report significant associations. Regarding associations with learning outcomes, 27 statements are made of which 19 statements report significant associations.

Intrinsic motivation

An overview of the intrinsic motivation techniques used and the number of studies it was used in, can be found in Table 2. Motivational techniques that yielded statistically significant associations with intrinsic motivation (include the implementation of augmented reality (AR) (Hao & Lee, 2019), the feeling of competence, (Hong et al., 2017), risk or uncertainty (Howard-Jones & Demetriou, 2009), one's expectations of the usefulness and simplicity of the game (Iten & Petko, 2016), flow in the game (Hong et al., 2017; Iten & Petko, 2016), competition and collaboration (Plass et al., 2013), reward and feedback (Iten & Petko, 2016), as well as perceived competence or self-efficacy (Touati & Baek, 2018). Out of 33 attempts to investigate the association between intrinsic motivation and intrinsic motivation techniques, 81.8% of the results were significant. Of all investigated intrinsic motivation techniques, only stressing benefits/meaning-making got an insignificant association without another study contradicting this with a significant association. The other intrinsic motivation techniques that yielded insignificant associations with intrinsic motivation (i.e., collaboration, challenge, identification/control, competition, and storytelling), were refuted with at least an equal number of studies that found significant associations.

Out of the 22 studies, 10 studies were randomized controlled trials and regarding the analyses, 8 out of 22 studies implemented analyses that were executed by means of an analysis of covariance (ANCOVA).

Learning outcomes

Motivational techniques used by studies that yielded statistically significant associations with learning outcomes, include the implementation of an adventurous story and sense of control (Boeker et al., 2013), a social networking platform (De-Marcos et al., 2014), risk or uncertainty (Howard-Jones & Demetriou, 2009), and complexity-based situation gaming (Yang et al., 2020). Similar to the designs used to investigate intrinsic motivation, most designs were randomized controlled trials and analyses were executed by means of ANCOVA. One study did not include a control group and had a qualitative design with only

9 participants, hence no conclusions on significant effects could be drawn (Bouzid et al., 2016).

Out of the 22 included studies, 27 studies reported associations between intrinsic motivation techniques and learning outcomes. Of these 27 studies, 8 studies reported insignificant differences between the control condition and the experimental (gamified) condition (e.g., Chen et al., 2015; Iten & Petko, 2016). Moreover, one study found a significantly lower learning outcome for the gamified condition than for the control condition. This surprising finding was found by De-Marcos et al. (2014), who discovered that the control group, who learned through three hours of lectures combined with additional online reading material and activities, scored significantly higher on the final written examination than the group who learned through a gamified program ($F(2, 347) = 37.42$, $p < .001$). The researchers created the game with challenges that increase in difficulty. For completion of each level of difficulty, the learner was rewarded with a trophy. Moreover, De-Marcos et al. (2014) found that the gamified approach resulted in significantly lower participation (i.e., engagement) scores than the control group ($F(4, 345) = 60.83$, $p < .001$).

In addition, Iten and Petko (2016) found no significant improvements in learning when using a reward system. Contrarily to the findings of De-Marcos et al. (2014), they did find a significant increase in motivation to learn ($\beta = .26$, $p < .05$).

Relevant techniques for an online memory-based listening game

Just the study by Chen et al. (2015) on collaboration reported insignificant results of the association between collaboration and intrinsic motivation, while all other studies that investigated this association reported significant associations of collaboration with intrinsic motivation (e.g., Plass et al. 2013). In addition, insiginficant results were found for challenge, competition, and storytelling (Yang et al., 2020). However, all these insignificant associations were refuted with at least an equal number of studies that found significant associations. This suggests that all techniques are useful to implement for the promotion of intrinsic motivation, or that a publication bias is present that prevented studies with insignificant results from being published. Especially in Europe, since no insignificant results were published in European studies.

It would be unrealistic to design a new online memory-based listening game that contains all techniques. Therefore, it is important to consider which techniques are relevant for this game. Based on the significance levels of the results, the number of times the technique was used, and the purposes of the future online memory-based listening game, one

might consider the following techniques to be most relevant and useful: collaboration or competition, challenge, risk (related to rewards), storytelling, as well as positive and negative feedback. These techniques were used most often (collaboration: 8; competition: 4; challenge: 7; risk and rewards: 5; storytelling: 3; and feedback 7 times) and all but one study (i.e., Chen et al., 2015) found significant results. For the purpose of the future online memory-based listening game these techniques are relevant, because hundreds of thousands of people should be motivated to engage in the game without any external reward, and these techniques should not be attributed to a certain aspect that is only effective in a certain demographic region. This is important since the purpose of the online memory-based listening game is to investigate world-wide demographic differences in catchiness.

Table 2

Overview of Intrinsic Motivation Techniques

Technique	References	Example of outcomes
Collaboration	Baek and Touati (2017) Chang et al. (2020) Chen et al. (2015) Hao and Lee (2019) Howard-Jones and Demetriou (2009) Plass et al. (2013) Touati and Baek (2018) Vos et al. (2011)	Chen et al. (2015): no significant difference in intrinsic motivation between collaboration vs. individual working. Plass et al. (2013): regarding the enjoyment, competitive and collaborative playing were significantly more enjoyable than playing alone ($p = .03$; $p < .001$).
Feedback	Chang et al. (2020) Chen et al. (2018) Howard-Jones and Demetriou (2009) Iten and Petko (2016) Laine et al. (2016) Plass et al. (2013) Vos et al. (2011)	Iten and Petko (2016): obtaining feedback during gameplay significantly influenced the self-assessed motivational learning gain ($\beta = .26$, $p < .05$).
Challenge	Chang et al. (2020) Hao and Lee (2019) Laine et al. (2016)	Hao and Lee (2019): attention was significantly higher in the AR[a] group with CAT than the traditional learning group ($F(1, 145) = 109.19$, $p < .001$).

	Ruggiero et al. (2013) Savulich et al. (2017) Yang et al. (2020)	
Identification/control	Boeker et al. (2013) Hawkins et al. (2019) Hieftje et al. (2019) Howard-Jones and Demetriou (2009) Ruggiero et al. (2013)	Boeker et al. (2013): the attitude of the GbEL[b] group with character and sense of control was significantly more positive (mean difference = 1.33, $p < .001$).
Competition	Chen et al. (2018) Plass et al. (2013) Vos et al. (2011) Yang et al. (2020)	Chen et al. (2018): the competition groups showed higher learning goals than the control group (mean = 3.60, SD = .71; mean = 3.26, SD = .67, respectively).
Competence/self-efficacy	Hawkins et al. (2019) Hong et al. (2017) Touati and Baek (2018) Vos et al. (2011)	Touati and Baek (2018): enjoyment was found to be especially associated with perceived competence ($\beta = .652$, $p < .01$).
Reward	Chang et al. (2020) De-Marcos et al. (2014) Iten and Petko (2016) Plass et al. (2013)	De-Marcos et al. (2014): the game provided rewards (trophies) on completion of each level and when looking at the practical acquisition of the gamified approach, this is significantly beneficial for the performance (e.g., post-test scores on databases, $F(4, 345) = 20.59$, $p < .001$).

Storytelling	Boeker et al. (2013) Laine et al. (2016) Yang et al. (2020)	Boeker et al. (2013): GbEL group with adventurous story significantly more positive attitude (mean difference = 1.33, $p < .001$).
Stressing benefits/ meaning-making	Hieftje et al. (2019) Ruggiero et al. (2013)	Hieftje et al. (2019): 'One Night Stan' stressed the benefits of the game in the participants' personal lives. Still, the game was experienced as fun and enjoyable, and the satisfaction with the intervention was reported as high.
Relatability	Baek and Touati (2017) Touati and Baek (2018)	Baek and Touati (2017): significant association between intrinsic motivation and enjoyment, assuming the relatability is the cause of enjoyment (r = .692, $p < .01$).
Avatar/host	Bouzid et al. (2016) Savulich et al. (2017)	Bouzid et al. (2016): 77.8% of the participants was satisfied with the game that implemented avatar technology. Due to absence of control group and the small sample size, significance cannot be indicated.
Risk	Devonshire et al. (2014) Howard-Jones and Demetriou (2009)	Howard-Jones and Demetriou (2009): a significant increase in electrodermal activity (EDA) was found for the gaming condition with uncertainty ($F(1, 16) = 11.12$, $p = .004$).
AR	Hao and Lee (2019) Laine et al. (2016)	Hao and Lee (2019): attention was significantly higher in the AR group than the traditional learning group ($F(1, 145) = 109.19$, $p < .001$).
Discovery learning (learning by means of trial-and-error)	Ruggiero et al. (2013) Vos et al. (2011)	Ruggiero et al. (2013): The qualitative results indicated that for participating in a camp for game development, discovery learning plays an important role in the motivation, engagement, and meaning-making of a juvenile offender.

[a] AR = augmented reality

[b] GbEL = game-based E-learning

Discussion

The research question constructed for this thesis is: "*How can existing intrinsically motivating experiments used in a variety of domains inform us about intrinsic motivation techniques to be used in the future designs of online memory-based listening games?*". The results of the scoping review revealed that most importantly collaboration, competition, challenge, risk, rewards, storytelling, and feedback play a crucial role in the intrinsic motivation of the player. Many participants prefer risk over certainty, and elements of collaboration or competition motivate participants to engage as well. Also, in the case of a user interface that is simple and easy to understand, players are optimally motivated to engage in the game. The results also indicate a significant role of AR in motivation, but this seems irrelevant for the development of a memory-based listening game, since AR influences the visual experience of a game, but does not affect the auditive experience.

Roles of the techniques in intrinsic motivation

The role of collaboration or competition in intrinsic motivation might be explained by social factors, like wanting to be valuable when working together, but also stimulating each other to participate and a drive to win in a competition.

The challenge, for example by means of CAT, motivates participants because boredom due to too easy items is prevented. Difficulty levels of the items are tailored to the performance level of the participant and overall testing time is typically lower for CAT (Weiss, 1973). In addition, by keeping up the challenge for participants, they tend to be more motivated during the experiment (Van der Linden & Glas, 2000, as cited in Wauters et al., 2010).

Regarding the role of risk and rewards in intrinsic motivation, one might think of the association between risk, dopamine, and the reward system. In short, dopamine neurons respond to differences in expected rewards and actual rewards. When taking a risk, one expects a reward, but the outcome might differ from the expectation. If an actual reward is higher than the expected reward, dopamine neurons increase their activation, which positively affects experienced emotions (Schultz, 2010). These dopamine boosts could affect intrinsic motivation. This effect can also be explained in the context of music perception. Research revealed that dopamine is produced as an anticipation to specific moments in one's favourite music (Salimpoor et al., 2013). This specific moment in the piece of music is comparable to the expected reward.

The role of storytelling in intrinsic motivation is explicitly mentioned by participants in the study of Laine et al. (2016). For example, one participant stated: "*The contents were*

interesting. It made me to play this game more and more" (p. 524). By creating this realistic story and context, participants were motivated to participate.

Lastly, the role of feedback in intrinsic motivation can be explained by the feeling of competence that is elicited by positive feedback. Indeed, obtaining feedback was found to improve intrinsic motivation (Iten & Petko, 2016). Though, previous research indicated that children might respond more aversive to negative feedback than to positive feedback (Van Leijenhorst et al., 2006). Nevertheless, all studies that included feedback implemented both positive and negative feedback. In addition, all studies that reported the association between feedback and intrinsic motivation, found a significant association (e.g., Iten & Petko, 2016). Still, the way in which feedback is implemented should be considered with caution, since Van Leijenhorst et al. (2006), for example, has shown that reactions to positive or negative feedback can differ.

An interesting finding was the contrast between factors that are significantly related to intrinsic motivation versus learning outcomes. For example, risk plays an important role in intrinsic motivation, while sense of control is strongly related to learning outcomes. Evidently, when risk is higher, the sense of control is lower and vice versa. One might assume risk is chosen over certainty when learning outcomes are not the main goal, since a loss does not affect the evaluation of performance, and the uncertainty is experienced as fun according to participants in the study by Howard-Jones and Demetriou (2009), hence increases motivation. On the other hand, the SDT states that a sense of control (described as *autonomy*) plays an important role in development, thus is important for learning outcomes (Deci & Ryan, 1980).

The study by De-Marcos et al. (2014) found significantly higher learning outcomes for the control group than for the gamified group. The difference between the two conditions was that the control group was instructed to just study the original learning material, while the participants in the gamified condition were given additional online reading material. These results suggest that participants might have been overloaded with work. Therefore, it is important to consider replacement of the original approach when implementing a gamified approach, instead of adding a gamified task to the original material.

Evidently, one should take into account the age of the participants when choosing the intrinsic motivation technique to implement, since previous research has shown that children respond differently to certain techniques than adults. For example, Van Leijenhorst et al. (2006) has shown that children respond more aversive to negative feedback than adults. It is therefore important to avoid negative feedback when the sample includes mostly children.

This is of less importance when the sample includes mostly adults. In addition, one might assume children to be more inclined to engage in adventurous stories that include a lot of fantasies, while adults might rather be driven by competition for example. The current scoping review describes studies that mostly included children that attend to primary or middle school (16 out of 22 studies). This suggests that the results should be considered with caution, since the results might be biased for primary or middle school children. Out of the 22 studies, 6 studies report the statistics on intrinsic motivation in (young) adults.

Ideas for future designs of online memory-based listening games

In short, the design of a new online memory-based listening game that is easy to understand, could include elements of risk and collaboration and/or competition. Based on the findings, I have thought of the following examples of future online memory-based listening games. Firstly, a game with four rounds of games and a reward system. During all rounds, points can be won, but also lost (the risk element). The player starts off with 10 points and firstly plays the round with a game that is similar to the Matching Pairs Game. The player can turn cards around, but instead of seeing a picture, the player hears a sound. The sounds are fragments from music of different cultures/regions. The cards with the same sounds should be paired. When more than half of the cards are still present, the player is asked to double the bet. If the two cards that are turned are not a match, one point is lost if the bet was not doubled, and two points are lost if the bet was doubled. If the two cards that are turned are a match, one point is won if the bet was not doubled, and two points are won if the bet was doubled. When less than half of the cards are left, only one point can be won or lost each turn. The first round is completed when all sound pairs are found. In the second round, the player will hear all sounds again and sees a map of the world. The player is instructed to drag each sound to the correct region of origin. Again, for each turn the bet can be doubled. For each correct answer, either one or two points are awarded. For each incorrect answer, either one or two points are lost. Lastly, the player is asked to rank all fragments from catchy to not catchy. The player can hear all sounds again and can divide all the points over all the fragments, awarding the most points to the catchiest fragment, according to the player. In the end, a ranking indicates the players with the most correct responses in the game.

Another example of a design of an online memory-based listening game, is realized by means of a character that the player can personalise and that walks around at a festival terrain. The character can visit several stages and at every stage, the character hears two fragments of music. Afterwards, the character has to report which fragment was catchier. This is reported by walking either to the left side or the right side of the stage, to vote for the

first or second fragment, respectively. Other random characters are around, walking to either the left or right side of the stage as well, creating a competition like experience. For every correct response (the catchiest fragment, based on the mean answer of other players), the player receives a coin. The player is able to type in an argument that supports their opinion, to earn five extra coins. In addition, at every stage (except the first one), the player is asked whether the person has heard the fragment before. Another way to earn two extra coins, is by approaching random characters at the terrain and answer their questions. But, when the player approaches the wrong character, namely a pick pocketer, five coins are stolen. When the player has collected enough coins, the character can attend to the final show, at which the player gets a list of the fragments that were in the game. The player is allowed to hear them all again and is then instructed to list them from catchy to less catchy.

Limitations

This scoping review was guided by the PRISMA-ScR (Tricco et al., 2018) guidelines; however, some limitations were encountered during the systematic process. One of the limitations is the type of studies that is included in the review. Only peer-reviewed journal articles are included. This resulted in the exclusion of unpublished studies, that might have yielded contrary findings. Most techniques described in the current review, revealed significant associations with either intrinsic motivation, learning outcomes, or both. However, this number of significant outcomes might indicate a publication bias, suggesting that authors of studies with insignificant results had trouble publishing their results. These unpublished articles might have been relevant for the purposes of the current review, hence imply that a knowledge gap exists.

Another limitation of the process of the scoping review includes the definition of reward. An exclusion criterion was the reward for participating in the study. However, not all studies mention whether participants were rewarded for participation. However, this is rather a limitation of the included study than of the scoping review process, but studies that did not mention a reward were included. It is unknown if this is justified, therefore, the extent to which true intrinsic motivation is measured is not guaranteed.

The paradox regarding the in-game rewards and extrinsic motivation being caused by rewards, is also an incentive for the questionability whether true intrinsic motivation is measured in the included studies. Since rewards can be viewed as extrinsic motivation, some might wonder whether these studies truly intrinsically motivate participants to participate in the study. However, participants were not rewarded for merely participating in the study. Only in-game rewards were given, and participants were not aware of in-game rewards

before consenting to participate in the study. Therefore, it can be argued that these in-game rewards did not interfere with the intrinsic motivation of the participants to engage in the experiment. In fact, some studies revealed that these in-game rewards even enhance intrinsic motivation or participation (e.g., Iten and Petko, 2016).

Lastly, a limitation of the study is the comparability of the results of different studies. When considering the different techniques to implement, one should not base their decision on the differences in significance ($p < .001$ versus $p < .05$) of the results of the studies. One cannot state that an association with $p < .001$ is more robust than with $p < .05$, since one does not know to what extent the different studies are comparable. When not taking into account the sample size for example, conclusions on differences in robustness between studies cannot be drawn. For studies with a greater sample size, the p-value will automatically be smaller than for studies with a smaller sample size. Therefore, I would advise to base the choices of techniques on the purposes of the future study.

In conclusion, based on the results of the scoping review one should consider implementing elements of collaboration or competition, challenge, risk (related to rewards), storytelling, as well as positive and negative feedback when designing a new online memory-based listening game.

References

Baek, Y., & Touati, A. (2017). Exploring how individual traits influence enjoyment in a mobile learning game. *Computers in Human Behavior*, *69*, 347-357. https://doi.org/10.1016/j.chb.2016.12.053

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2013). Improving participation and learning with gamification. *Proceedings of the First International Conference on Gameful Design, Research, and Applications*. https://doi.org/10.1145/2583008.2583010

Bateman, R. M., Sharpe, M. D., Jagger, J. E., Ellis, C. G., Solé-Violán, J., López-Rodríguez, M., Herrera-Ramos, E., Ruíz-Hernández, J., Borderías, L., Horcajada, J., González-Quevedo, N., Rajas, O., Briones, M., Rodríguez de Castro, F., Rodríguez Gallego, C., Esen, F., Orhun, G., Ergin Ozcan, P., Senturk, E., . . . Prandi, E. (2016). 36th International Symposium on Intensive Care and Emergency Medicine : Brussels, Belgium. 15-18 March 2016. *Crit Care*, *20*(Suppl 2), 94. https://doi.org/10.1186/s13054-016-1208-6

Boeker, M., Andel, P., Vach, W., & Frankenschmidt, A. (2013). Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. *PLoS One*, *8*(12), e82328. https://doi.org/10.1371/journal.pone.0082328

Bouzid, Y., Khenissi, M. A., Essalmi, F., & Jemni, M. (2016). Using educational games for sign language learning: A signwriting learning game: Case study. *Journal of Educational Technology & Society*, *19*(1), 129-141.

Bridges, D., Pitiot, A., MacAskill, M. R., & Peirce, J. W. (2020). The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ*, *8*, e9414. https://doi.org/10.7717/peerj.9414

Burgoyne, J. A., Bountouridis, D., Van Balen, J., & Honing, H. (2013). Hooked: A game for discovering what makes music catchy. *Proceedings of the 14th International Society for Music Information Retrieval Conference, ISMIR 2013*, 245-250.

Butler, L. P., & Walton, G. M. (2013). The opportunity to collaborate increases preschoolers' motivation for challenging tasks. *J Exp Child Psychol*, *116*(4), 953-961. https://doi.org/10.1016/j.jecp.2013.06.007

Chang, C.-S., Chung, C.-H., & Chang, J. A. (2020). Influence of problem-based learning games on effective computer programming learning in higher education. *Educational Technology Research and Development*, *68*(5), 2615-2634. https://doi.org/10.1007/s11423-020-09784-3

Chen, C.-H., Law, V., & Chen, W.-Y. (2018). The effects of peer competition-based science learning game on secondary students' performance, achievement goals, and perceived ability. *Interactive Learning Environments*, *26*(2), 235-244. https://doi.org/10.1080/10494820.2017.1300776

Chen, C.-H., Wang, K.-C., & Lin, Y.-H. (2015). The comparison of solitary and collaborative modes of game-based learning on students' science learning and motivation. *Journal of Educational Technology & Society*, *18*(2), 237-248.

De-Marcos, L., Domínguez, A., Saenz-de-Navarrete, J., & Pagés, C. (2014). An empirical study comparing gamification and social networking on e-learning. *Computers & Education*, *75*, 82-91. https://doi.org/10.1016/j.compedu.2014.01.012

Deci, E. L., & Ryan, R. M. (1980). Self-determination theory: When mind mediates behavior. *Journal of Mind and Behavior*, *1*(1), 33-43.

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. In *Cognitive evaluation theory* (pp. 43-85). Springer. https://doi.org/10.1007/978-1-4899-2271-7_3

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227-268.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie canadienne*, *49*(3), 182-185. https://doi.org/10.1037/a0012801

Denham, A. R. (2015). Supporting conceptual understanding of the associative and distributive properties through digital gameplay. *Journal of Computer Assisted Learning*, *31*(6), 706-721. https://doi.org/10.1111/jcal.12113

Devonshire, I. M., Davis, J., Fairweather, S., Highfield, L., Thaker, C., Walsh, A., Wilson, R., & Hathway, G. J. (2014). Risk-based learning games improve long-term retention of information among school pupils. *PLoS One*, *9*(7). https://doi.org/10.1371/journal.pone.0103640

Ferrara, J. (2013). Games for Persuasion. *Games and Culture*, *8*(4), 289-304. https://doi.org/10.1177/1555412013496891

Hao, K.-C., & Lee, L.-C. (2019). The development and evaluation of an educational game integrating augmented reality, arcs model, and types of games for english experiment learning: An analysis of learning. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2019.1619590

Hawkins, I., Ratan, R., Blair, D., & Fordham, J. (2019). The effects of gender role stereotypes in digital learning games on motivation for STEM achievement. *Journal of Science Education and Technology*, *28*(6), 628-637. https://doi.org/10.1007/s10956-019-09792-w

Honing, H., & Reips, U.D. (2008). Web-based versus Lab-based Studies: A Response to Kendall (2008). *Empirical Musicology Review*, *3*(2), 73-77.

Hieftje, K., Duncan, L., Florsheim, O., Sawyer, B., & Fiellin, L. E. (2019). One Night Stan: Feasibility study of an HIV prevention and sexual risk reduction social card game for young Black women. *Games for Health*, *8*(2), 112-120. https://doi.org/10.1089/g4h.2017.0106

Hong, J.-C., Hwang, M.-Y., Tai, K.-H., & Lin, P.-H. (2017). Intrinsic motivation of Chinese learning in predicting online learning self-efficacy and flow experience relevant to students' learning progress. *Computer Assisted Language Learning*, *30*(6), 552-574. https://doi.org/10.1080/09588221.2017.1329215

Honing, H. (2021). Lured Into Listening. *Zeitschrift für Psychologie*, *229*(4), 266-268. https://doi.org/10.1027/2151-2604/a000474

Howard-Jones, P. A., & Demetriou, S. (2009). Uncertainty and engagement with learning games. *Instructional Science*, *37*(6), 519-536. https://doi.org/10.1007/s11251-008-9073-6

Iten, N., & Petko, D. (2016). Learning with serious games: Is fun playing the game a predictor of learning success? *British Journal of Educational Technology*, *47*(1), 151-163. https://doi.org/10.1111/bjet.12226

Korsmit, I. R., Burgoyne, J. A., & Honing, H. (2017). If You Wanna Be My Lover … A Hook Discovery Game to Uncover Individual Differences in Long-term Musical Memory. *Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music*, 106-111.

Laine, T. H., Nygren, E., Dirin, A., & Suk, H.-J. (2016). Science Spots AR: A platform for science learning games with augmented reality. *Educational Technology Research and Development*, *64*(3), 507-531. https://doi.org/10.1007/s11423-015-9419-0

Plass, J. L., O'Keefe, P. A., Homer, B. D., Case, J., Hayward, E. O., Stein, M., & Perlin, K. (2013). The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of Educational Psychology*, *105*(4), 1050-1066. https://doi.org/10.1037/a0032688 (Advanced Learning Technologies)

Ruggiero, D., de Hurtado, B. G., & Watson, W. R. (2013). Juvenile offenders: Developing motivation, engagement, and meaning-making through video game creation. *International Journal of Game-Based Learning*, *3*(2), 112-129. https://doi.org/10.4018/ijgbl.2013040107

Salimpoor, V. N., Van den Bosch, I., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions Between the Nucleus Accumbens and Auditory Cortices Predict Music Reward Value. *Science*, *340*(6129), 216-219.

Sansone, C., & Tang, Y. (2021). Intrinsic and Extrinsic Motivation and the Self-Determination Theory. *Motivation Science*, *7*(2), 113-114. https://doi.org/10.1037/mot0000234

Savulich, G., Piercy, T., Fox, C., Suckling, J., Rowe, J. B., O'Brien, J. T., & Sahakian, B. J. (2017). Cognitive training using a novel memory game on an iPad in patients with amnestic mild cognitive impairment (aMCI). *International Journal of Neuropsychopharmacology*, *20*(8), 624-633. https://doi.org/10.1093/ijnp/pyx040

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behavioral and Brain Functions*, *6*(1), 24. https://doi.org/10.1186/1744-9081-6-24

Touati, A., & Baek, Y. (2018). What leads to player's enjoyment and achievement in a mobile learning game? *Journal of Educational Computing Research*, *56*(3), 344-368. https://doi.org/10.1177/0735633117713022

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., . . . Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.*, *167*(7), 467-473. https://doi.org/doi:10.7326/M18-0850

Van Leijenhorst, L., Crone, E. A., & Bunge, S. A. (2006). Neural correlates of developmental differences in risk estimation and feedback processing. *Neuropsychologia*, *44*(11), 2158-2170. https://doi.org/10.1016/j.neuropsychologia.2006.02.002

Vos, N., van der Meijden, H., & Denessen, E. (2011). Effects of constructing versus playing an educational game on student motivation and deep learning strategy use. *Computers & Education*, *56*(1), 127-137. https://doi.org/10.1016/j.compedu.2010.08.013

Wauters, K., Desmet, P., & Van den Noortgate, W. (2010). Adaptive item-based learning environments based on the item response theory: possibilities and challenges. *Journal of Computer Assisted Learning*, *26*(6), 549-562. https://doi.org/10.1111/j.1365-2729.2010.00368.x

Weiss, D. J. (1973). *The Stratified Adaptive Computerized Ability Test*. https://files.eric.ed.gov/fulltext/ED084301.pdf

Wiersma, U. J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology*, *65*(2), 101-114. https://doi.org/10.1111/j.2044-8325.1992.tb00488.x

Yang, Q.-F., Chang, S.-C., Hwang, G.-J., & Zou, D. (2020). Balancing cognitive complexity and gaming level: Effects of a cognitive complexity-based competition game on EFL students' English vocabulary learning performance, anxiety and behaviors. *Computers & Education*, *148*. https://doi.org/10.1016/j.compedu.2020.103808

Appendix

Table A

Summary of Studies Included in the Scoping Review on the Implementation of Intrinsic Motivation in Memory-Based Experiments.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Baek and Touati (2017)	Can individual traits (e.g., learning styles, collaboration skills, intrinsic motivation, and computer game attitude) predict the levels of enjoyment? What is the relationship between enjoyment and achievement?	164 participants from elementary schools in South Korea were instructed to play the mobile version of Minecraft for 1.5 hours every day for 3 weeks and improve their skills by engaging in discussions every day. Two tasks were given: (1) replicate a maze of their choice in Minecraft, being able to choose from 3 different difficulty levels, (2) replicate one's school building with a playground. For pre-test scores, the Computer Game Attitude Scale (CGAS) was administered on day one. After 3 weeks of playing, the students engaged in the Enjoyment Test, the Learning Style Inventory of Felder-Silverman's, the CGAS, the	Relatability (through replicating personal real-life situations). Collaboration.	The results of the different tests were correlated with learning styles. These correlations revealed that there is no significant correlation between learning style and collaboration skills (r = -.036, p = .645), which was also the case for the association between enjoyment and collaboration skills (r = .141, p = .072). The correlation between learning style and game enjoyment is significant (r = .210, p < .01), as was the case for the correlation between intrinsic motivation and enjoyment (r = .692, p < .01), for the correlation

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		Intrinsic Motivation Inventory, and the Collaboration Attitude Test.		between computer game attitude and enjoyment (r = .335, p < .01), and for enjoyment and achievement (r = .441, p <.01).
Boeker et al. (2013)	The aim of the study was to show the superiority of Game-based E-learning to a conventional instructional method for medical students learning phase contrast microscopy of urine specimens.	145 third-year medical students at the urology department of the UMC Freiburg in Germany were divided in two groups: one group was trained by means of an educational adventure-game (GbEL group) and the other group was trained by means of a written script-based approach (script group). The script group was provided of an 8-page script containing all the information, while the GbEL group played a game on the PC. In the game, the player must master a series of quest, while navigating through a landscape with different scenarios. The player controls the encounters of the character.	Adventurous and educational gamification (storytelling). Sense of control over the character and its encounters.	The script group had a mean score of 26.0 out of 34 points, compared to 28.6 points by the GbEL group, which is a significant difference (t = not mentioned, p < .001) with a Cohen's d effect size of 0.71. Regarding the attitude towards the learning method, the attitude of the GbEL group was significantly more positive than the attitude of the script group (mean difference = 1.33, p < .001).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		After the training, learning outcomes were measured with a 34-item single choice test (either true or false), in addition to the collection of student's attitudes towards the training.		
Bouzid et al. (2016)	What are deaf learners' interests in using the educational game MemoSign for learning sign writing notations and new vocabularies? In particular, do the learners enjoy the game, is it useful, and enables the game them to learn new vocabularies?	9 deaf Tunisian children of ages between 9 and 16 years played the MemoSign game, which is based on a learning version of the memory match game (LMMG) and contains 3 pair of cards (word/visual/calculates paired with sign writing). When a sign writing card is turned, a virtual avatar presents the visual-gestural modality. By means of this game, the researcher's goal is to promote a fun way of learning sign writing notations.	Gamification through Memory Match Game. Avatar technology.	Since there was no control group, no conclusions can be drawn from the results regarding significant improvements in learning. The results of this study show that 77.77% of the participants was satisfied with the game and 22.22% was not. Moreover, 88.88% thought the game was useful and effective for learning their vocabulary.
Chang et al. (2020)	What are learning outcomes of game-based learning (GBL) in a	69 Taiwanese students who enrolled for a programming course at a university were part of this study. All played the	Personal feedback through the login process.	The pre-test and post-test scores regarding 'understanding' and 'application' differed

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	computer programming course? What are students' perceptions of GBL in a computer programming course?	'Programmer Adventure Land' game and filled in a questionnaire on their attitudes towards problem-based game learning. A control group was absent. With role development and peer interaction, learning interest was promoted. During the game, students practice the course learning material through an adventurous game with a point system and personalized feedback to reflect achievement. Moreover, the game consists of two recursive levels.	Role development. Peer interaction. User interface design that is simple and easy to understand. Rewards (point system). Challenge.	significantly, indicating that these game-based learning tasks result in improved knowledge. Regarding the attitudes, the user interface seemed to play the most important role in satisfaction, followed by motivation and enjoyment.
Chen et al. (2015)	How do students in the individual mode differ from students in the collaborative mode with respect to acquiring knowledge?	50 seventh grade students played the game called 'Carrot Land' in which they learned about the effects, types and impact of force and the force equilibrium condition. First, the pre-test scores were measured. Then, the	Collaboration vs. individual working.	No significant between group differences were found. Within groups, both had a significant difference between pre-test and post-test scores (M = 43.80, SD = 11.20 for individual; M =

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	How do students in the individual mode differ from students in the collaborative mode with respect to motivation for learning? How is the game perceived/experienced by students in both the individual and collaborative modes?	students were divided in two conditions: individual or collaboration. The collaboration group played the game together and had to discuss a worksheet. After 20 minutes of playing, a post-test and motivational survey followed. Then, there were 15 to 20 more minutes of interviews.		45.40, SD = 14.54 for collaboration). Regarding the intrinsic motivation, this was higher for the collaboration group than for the individual group (M = 3.42, SD = 1.10 vs. M = 3.10, SD = .98). Though, the difference between groups is not significant.
Chen et al. (2018)	Do different modes of competition in science GBL have an effect on student's performance? Do different modes of competition in science	195 Taiwanese 14-15-year-olds were divided in three different conditions: playing anonymous competition games, playing non-anonymous competition games, the control group that played the game without competition. The participants completed three tasks that	Competition. Anonymous vs. non-anonymous (not significant). Feedback.	One-way analysis of covariance (ANCOVA) showed a significant main effect of competition on performance ($F(2, 194) = 6.633$, $p = .002$). The main effect of competition on learning goal was also significant for both

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	GBL have an effect on student's goal orientations?	increased in difficulty. Every task was firstly to locate the carrot and subsequently, based on the concept of force and balance that they learn about, to decide what is the best way to pull the carrot out of the ground. Scores were based on task time and their choices during the game. In addition, the Approaches to Learning scale was used to assess the learning and performance goals and students' perceived ability.		competition groups ($F(2, 194) = 5.66$, $p = .004$). For both main effects, there was no significant difference between anonymous and non-anonymous competition. Regarding the learning goals, the anonymous competition group reported significantly higher learning goals than the control group ($p < .05$). Again, there was no significant difference between the two competition groups.
De-Marcos et al. (2014)	Will gamification impact learning in large classroom environments? Will gamification impact participation rates?	265 first- and second-year students were divided over 3 conditions: the gamification plugin group, the social networking site group (irrelevant for this scoping review), and the control group. With the gamification plugin, students were able to complete course activities and compete/collaborate with	Rewards (trophies) on completion of each level.	The gamified group scored higher on the practical tests than the control group. A one-way ANOVA on the results of the final examination revealed that the control group scored higher than the gamified group ($F(4, 345) = 37.42$, $p < .001$). The

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	Will students have a positive attitude towards these tools?	other students. Activities were posed as challenges and divided into three or four levels, with rewards (trophies) for completion of each level. The control group learned through three hours of lectures combined with additional online reading material and activities. Next to pre-test and post-test scores, an attitudinal survey was used.		researchers expect that an overload of learning materials have influenced the results. However, when looking at the practical acquisition of the gamified approach, this is significantly beneficial for the performance (e.g., post-test scores on databases, $F(4, 345) = 20.59$, $p < .001$).
Denham (2015)	What effect does the application of endogenous game design principles have on participants' conceptual understanding of the associative and distributive properties? What are participants' attitudes towards the game-	111 fourth and fifth graders from the US were spread across three conditions: endogenous, exogenous, and control. Each spent 50 minutes playing their version of the game 'Shipping Express'. The goal is to load a certain number of trucks with a certain number of boxes within a set time. For each correct answer, bonus time is added. The levels increase in difficulty after	Endogenous game design principles (fluent and continuing relationship between factual information or instructions and the fantasy context of	The application of endogenous game design principles turned out not to result in any significant improvements in the post-test scores compared to the control group. On average, the post-test scores of the endogenous game group were even poorer than the scores of

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	based learning environment?	completion. In the endogenous version, understanding associative and distributive properties and knowledge of multiplication facts is required. The player is provided of single numbers and needs to select one of those to be multiplied in order to fill a truck. For the exogenous version, participants get a short description of associative and distributive properties, which they must read. During gaming, participants were instructed on the application of these properties. Also, there are no single numbers to select and to multiply, but multiplication pairs are on the boxes, which are the number of boxes needed for a certain truck to leave the dock. The control version is similar to the exogenous version, except it is lacking the gamification elements.	the game). Not significant, though.	the exogenous and the control groups.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Devonshire et al. (2014)	Can risk-based learning games for school pupils improve later recall of information?	448 pupils of ages 9-10. The independent variable (IV) was workshop delivery style and had three conditions: no reinforcement (control group), five multiple-choice questions (MCQs) answered in small teams (no risk group), and the risk group that answered the same MCQs but were to compete with one another by betting tokens on the correct answer. By betting correctly, the team will receive the double number of tokens back. All groups received a workshop, from which they learned about neuroscience. The dependent variables (DVs) are scores from a written neuroscience quiz made on the day of the workshop and one week later, an evaluation of the quiz regarding the extent to which it	Risk.	The test scores on the day of the workshop did not differ significantly between the three conditions. However, the longer-term retention is improved by using risk-based learning games (mean difference(risk-no risk) = 46.84, $p < .01$). Also, the workshop was experienced as significantly more interesting and informing for the risk condition than for the control group ($F(2, 326) = 7.366$, $p \leq .05$).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		was engaging, appropriate, informing, and enjoyable.		
Hao and Lee (2019)	Objectives: Comparison of the learning motivation and effectiveness between students who received AR English game learning and students who taught by traditional learning. Which type of AR games is the students' favourite? The correlations between AR game design elements and the attention relevance confidence-satisfaction (ARCS) model.	147 fifth graders were spread over two conditions: an experimental group (AR learning), and a control group (traditional learning). The control group learned English through slideshows and textbooks, while every 2 participants in the experimental group got an iPad Mini in pairs, to play learn English with an AR game after they got the instructions from their English teacher. This continued for 2 hours every week for 4 weeks.	Peer discussions (collaboration). Challenge and adaptability. AR.	For all dimensions (i.e., attention, relevance, confidence, and satisfaction), significant differences were found between the AR games teaching and the traditional teaching condition in favour of the AR games teaching condition ($F(1, 145) = 109.19$; 92,64; 104.60; 71.19, $p < .001$). These results suggest that AR games teaching would be an effective way to improve students' way of learning. However, the learning outcomes did not differ significantly between the two conditions ($t(145) = .957$, $p = .34$).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
				With respect to motivation, the implementation of peer discussions, challenge and adaptability would significantly improve motivation.
Hawkins et al. (2019)	How does scientist-character gender expression impact science, technology, engineering, and mathematics (STEM)-based learning motivation?	4514 Michigan State University Museum visitors engaged in a digital learning game, in which they play as a flying bird that can eat black or white moths by touching them. Afterwards, a scientist character explains something about the moths, and this scientist has four conditions: low/high-masculine, or low/high-feminine. The players' score indicated the motivation to perform.	Self-affirmation and self-efficacy though identifying with the game or scientist character.	A significant main effect of scientist gender was found, regardless of the gender of the participant ($F(1, 4510) = 2.98$, $p = .085$). For children under the age of 18, motivations for STEM-based learning is significantly affected by sex and gender of the scientist ($F(1, 304) = 8.25$, $p = .004$). As people grow up, they become less affected by sex-role stereotypes, since these differences were not found for older age groups.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Hieftje et al. (2019)	A social card game intervention 'One Night Stan' was developed to increase partner HIV testing and condom use. The aim of this study is to assess player's satisfaction and gameplay experience, and the game's preliminary impact on psychosocial variables related to behaviours associated with HIV prevention.	21 young black women from Connecticut played the multiplayer social card game for up to 5 players. The player with the most empowerment points (self-efficacy) wins, meaning that the person has the most positive health outcomes. Empowerment points can be received through different cards that are related to healthy sexual behaviours, like getting potential partners tested (test cards), being protected (condom cards), or refuse risky encounters (response cards). For the outcomes, gameplay satisfaction and experience, and impact on psychosocial variables of behaviour change around HIV prevention were measured pre-test, after completion (2 weeks) and at 6-week follow-up.	Identification through role-play situations. Stressing benefits by creating learning opportunities.	The game was experienced as fun and enjoyable, and the satisfaction with the intervention was reported as high. Results revealed a significant change in self-efficacy and intentions of the participants, with higher scores at post-test than pre-test. No significant effects were found for knowledge, attitudes, perceived susceptibility, and perceived norms.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Hong et al. (2017)	Objectives: Develop conceptual framework for how individual differences are relevant to learners' intrinsic motivation to learn Chinese radicals by examining online learning self-efficacy (OLSE) and flow experience through degree of learning progress (DLP). To verify the research model in relation to Students of Southeast Asian Heritage Learning Chinese (SSAHLC) and flow experience as reflected in their DLP.	78 SSAHLC fourth graders took the pre- and post-test to assess OLSE and IMCL. As intervention, once a week for 5 weeks, the participants engaged in a game that asks players to identify the semantic radical of a character. Players could interact with the animation of the game through gesture-based interaction (GBI). Some questions are posed in the game, and the participant is required to wave their hands to answer the question.	Feeling of competence (self-efficacy).	For Chinese learning intrinsic motivation, a significant relationship was found with OLSE ($t = 4.35$, $p < .001$) and with flow experience ($t = 4.43$, $p < .001$), suggesting that intrinsic motivation might serve as a predictor of OLSE and flow experience when learning Chinese radicals. Moreover, OLSE and flow experience can predict outcomes of learning progress.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Howard-Jones and Demetriou (2009)	Can the emotional response to a learning task be influenced by presenting it in a context that includes gaming uncertainty?	**Study 1** included 50 pupils from a primary school in Cyprus. The participants were instructed to report if a mathematical statement was true or false. They could choose whether the question was posed by Mr Certain or Mr Uncertain. Mr Certain would give one point for each correct answer, Mr Uncertain would either give zero or two points after tossing a coin after the correct answer. Afterwards, 10 students were randomly selected to participate in an interview to discuss their choices. **Study 2** included 20 pupils from a Year 9 science class in the UK. They played 'Wipe Out' in pairs, a game in which a question must be answered, and the participants roll two dices beforehand of which the sum would be the points awarded for a correct answer. After	Risk/uncertainty. Sense of control. Collaboration. Feedback.	For **Study 1** a significant preference was found for Mr Uncertain over Mr Certain ($\chi^2(1) = 77.98$, $p < .001$). Generally, participants preferred Mr Uncertain because they liked the risk and thought it was exciting. For **Study 2** a significant learning effect was found ($t = 8.87$, $p < .001$). For **Study 3** a significant increase in EDA was found for the gaming condition with uncertainty ($F(1, 16) = 11.12$, $p = .004$), suggesting enhanced engagement with learning through uncertainty, rather than just gaming.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		answering the question, the participants had the option to pass or to roll the dices again to earn extra points, but if one 1 was rolled, all points for that turn are lost. If two 1s were rolled, all points of the game are lost. The participants played against a computer and the one to first achieve a score of 100, is the winner. **Study 3** included 16 post-graduate students in two conditions: the gaming condition (played 'Wipe Out') and the non-gaming condition (played 'Wipe Out' without uncertainty). During the experiment EDA was measured.		
Iten and Petko (2016)	What kind of interrelations exist between general attitudes towards serious games, especially the expectation that this kind of	74 children from 5 Swiss primary schools played 'AWWWARE' for 30 minutes. This is an online serious game that promotes media competency. Participants are instructed to locate web	Reward system. Feedback.	The motivation to engage in a serious game in the future is significantly influenced by one's expectations of the usefulness and anticipated simplicity of the

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	learning will be fun, the actual perception of the fun of a specific serious game, perceived learning motivation and learning gains as well as improvements in post-test measures? Is expected enjoyment a predictor of actual enjoyment when learning with serious games? And is experienced enjoyment a predictor of learning motivation and learning gains?	pages that answer specific questions, like "what are the main reasons for traffic accidents?". While seeking for suitable information, participants need to avoid inappropriate content. This is made more difficult by wind effects that manipulate the cursor (raven that holds a kite). Avoiding inappropriate pages and selecting suitable ones is rewarded with points. Pre- and post-test scores and questionnaires were conducted to assess knowledge in the domain of critical internet literacy.		game ($F(6, 67) = 14.13$, $p < .001$). Also, the enjoyment of the game during playing, significantly influenced the gain in motivation to learn ($\beta = .22$, $p < .05$). The use of prior knowledge, experience of flow in the game, and obtaining feedback are also significantly related to a gain in motivation to learn ($\beta = .50$, $p < .001$; $\beta = .22$, $p < .01$; $\beta = .26$, $p < .05$). However, no significant influence of experience of fun on learning was found.
Laine et al. (2016)	Objectives: The current study contributed to the concept and architecture of Science Spots AR, the design and	61 Korean fifth grade elementary school children engaged in the Leometry game, which uses storytelling to optimize engagement. The task of the players is to help escaped leopards to	Challenge. Fantasy through storytelling (with characteristics of	The qualitative outcomes of the study indicated positive responses towards the combination of storytelling, challenges, and AR for the

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	implementation of the Leometry game prototype, and to the mixed-method formative evaluation implementation of the Leometry game prototype.	find their way back home, away from poachers who illegally captured them. By completing geometry challenges that they encounter along the way, they can bring the leopards back home. The challenges increase in difficulty level as the players get further.	humour and continuity). Augmented reality. Feedback.	learning effects of science games. Especially solving problems and playing with friends were features that were appreciated.
Plass et al. (2013)	How do the three modes of play (individual, competitive, and collaborative) affect learning, game performance, and motivation?	58 sixth to eighth graders joined the weekly visits of the study, during which they played the game 'FactorReactor' on Xbox either individually, collaboratively, or competitively. The goal is to transform the centre number into one of the surrounding goal numbers. This can be achieved through adding, subtracting, multiplying, or dividing it by one of the numbers from the inner ring. After each correct transformation, players were rewarded with rings. The more difficult the	Collaboration. Competition. Achievement goal orientations (mastery vs. performance). Feedback. Reward.	Regarding the learning effect, the post-test showed higher scores for the competitive condition (p = .02), but not for the collaborative condition, compared to the individual condition. This was analysed with a hierarchical linear modelling. For the mastery goal orientation, both the competitive and collaborative conditions scored higher than the individual condition (p = .01; p = .04), but

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		transformation, the more rings are awarded. These rings are needed to execute a transformation. Levels increased in difficulty.		there was no significant difference between the two experimental conditions. For the performance goal orientation, no significant difference was found between all conditions. Regarding the enjoyment, competitive and collaborative playing were significantly more enjoyable than playing alone (p = .03; p < .001).
Ruggiero et al. (2013)	In what ways does participating in Project Tech influence juvenile offender motivation, engagement, and meaning-making in a discovery learning game development camp?	10 adjudicated 13-to-19-year-olds participated in the two-week Project Tech camp. As a part of the project, the participants played several serious games, under more to gain simulated experience and to learn about social issues through programming. They reported their daily activities in a journal.	Immersion (identification). Discovery learning. Adjusted levels of academic challenge.	The qualitative results indicated that for participating in a game development camp discovery learning plays an important role in the motivation, engagement, and meaning-making of a juvenile offender.

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
	What insights emerge from learners' feedback, especially in relation to specific attributes of Project Tech?		Meaning-making. Positive interaction with the teacher.	The learners made little effort to make conversation with each other about any topic other than the camp itself. There was a competitive atmosphere between the individuals.
Savulich et al. (2017)	What are the effects of "Game Show", a novel learning and memory game, on cognition and motivation in patients with amnestic mild cognitive impairment (aMCI)?	42 patients were divided in two groups: a cognitive training group that received 8 hours of supervised gameplay of Game Show on an iPad (spread over 4 weeks), and a control group who attended clinic as usual. In Game Show, one can win gold coins by correctly associating different geometric patterns with different spatial locations.	Adjusted levels of difficulty. Implementation of a "host" that encourages to maintain and progress beyond their last played level.	The total number of errors was significantly reduced in the cognitive training group ($t(20) = 3.20$, $p = .005$), but not in the control group ($t(20) = -.17$, $p = .86$). Also, the post-test scores on the Mini-Mental State Examination (MMSE) were significantly higher for the cognitive training group than the control group ($t(40) = 2.15$, $p = .038$).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Touati and Baek (2018)	What factors influence enjoyment and achievement in a mobile learning game? How does one's perceived competence relate to their attitude toward the game and frequency of play? How does prior experience relate to one's perceived competence?	164 South-Korean students of ages 12-13 years played the Mobile version of Minecraft every day after school for 3 weeks. Two tasks were given: (1) replicate a maze of their choice in Minecraft, being able to choose from 3 different difficulty levels, (2) replicate one's school building with a playground. The participants were spread over 55 groups and groups discussed strategies before playing Minecraft individually. For pre-test scores, the CGAS was administered on day one. After 3 weeks of playing, the students were given the Enjoyment Test and Perceived Competence Test. Participant's final mazes from each task could get 100 points based on several criteria. The mean score would be the final score to reflect the achievement.	Relatability (through replicating own real-life situations). Collaboration. Perceived competence (self-efficacy)	Enjoyment was found to be especially associated with perceived competence ($\beta = .652$, $p < .01$) and game attitude ($\beta = .123$, $p < .05$). Moreover, enjoyment was the only factor that is significantly related to achievement ($\beta = .370$, $p < .01$). The strongest relationship was found between prior game experience and perceived competence ($\beta = .868$, $p < .01$), but perceived competence is also strongly related to intensity of use ($\beta = .830$, $p < .01$) and to game attitude ($\beta = .353$, $p < .01$).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
Vos et al. (2011)	How do different interactive tasks, in which a game was included, affect student intrinsic motivation and deep strategy use?	235 students of ages 10-12 years participated. The quasi-experimental study distinguished two groups: the construction group (constructed a game on Dutch proverbs), and the play group (played an existing game on Dutch proverbs). Both groups first were given a worksheet to introduce Dutch proverbs. The goal was to identify the meaning of eight proverbs of their selection from the sheet. Next, the construction group constructed a 'drag and drop game', so drag one textual picture and drop it next to the picture that relates to it. They could use the selected proverbs from the worksheet. The play group played an existing drag and drop game about proverbs. The goal was to drag all eight proverbs to their meanings as quickly as possible,	Discovery learning. Collaboration. Competition. Perceived competence. Feedback.	For the play group, the student intrinsic motivation and the deep strategy use scored lower during playing than during regular school lessons. ANCOVA revealed that the construction group scored significantly higher at deep strategy use than the play group ($F(1, 228) = 73.69$, $p < .001$). Regarding the interest, ANCOVA showed that the construction group scored significantly higher than the play group ($F(1, 228) = 119.33$, $p < .001$). So, scores on motivation and deep strategy use are higher for the construction group than for the play group.

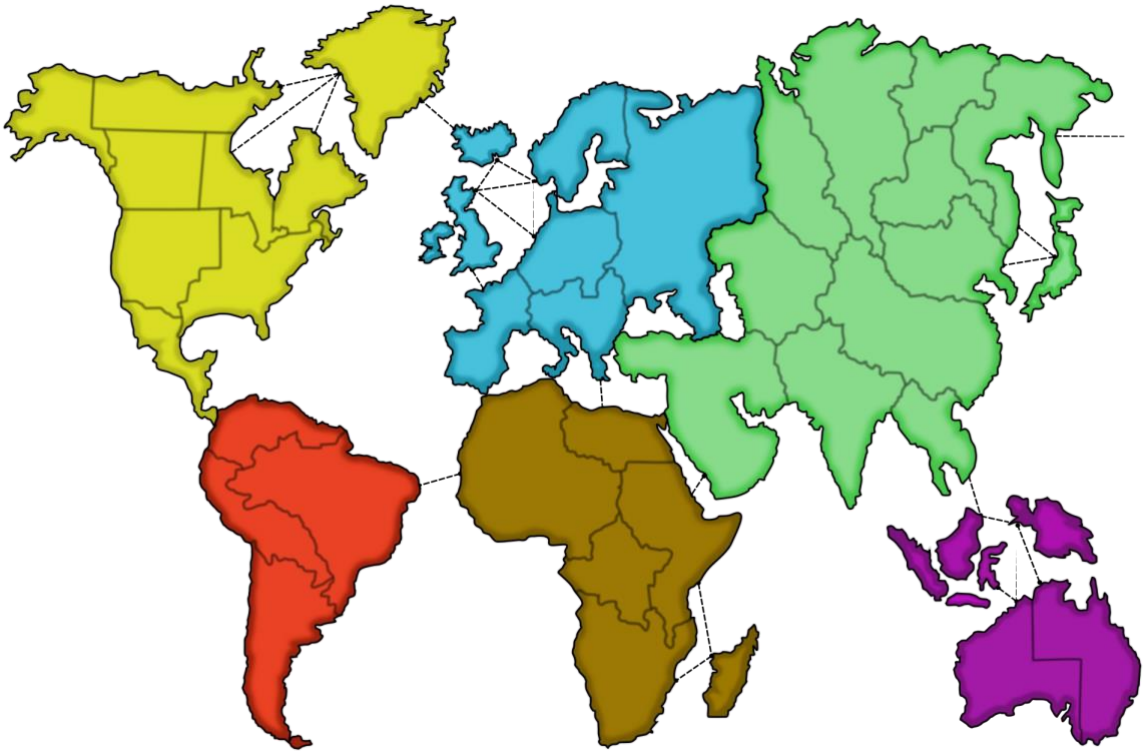
Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		with as little as possible errors. The students could either help each other, but also compete against each other. Intrinsic motivation and deep strategy use were measured through a questionnaire.		
Yang et al. (2020)	Could the cognitive complexity-based situation English vocabulary gaming (CLBG) approach enhance students' learning performance more than the conventional situational English vocabulary gaming (CG) approach? Could the cognitive CLBG approach improve students' learning motivation more than the CG approach?	51 Taiwanese students were assigned to either the CLBG group (learned with the game based on the cognitive complexity-based competition strategy) or the CG group (learned with the game based on the common competition strategy). The Chinese speaking students learned English vocabulary through an interactive game. During the game, a relevant storyline and character conversation enable the players to learn English vocabulary. For all groups there are 3 levels of difficulty. However, the CLBG group is able to	Competition. Implementation of relevant storyline. Challenge.	Results of ANCOVA showed that students from the CLBG group had significantly higher post-test scores than the CG group ($F(1, 50) = 5.79$, $p < .05$). Regarding the students' motivation, ANCOVA revealed no significant difference between the two groups ($F(1, 50) = 0.00$, $p = .997$).

Reference[a]	Research question	Participants and study design	Intrinsic motivation	Outcome
		upgrade or downgrade in level of difficulty based on learning performance during gaming. The CG group could only increase when they reached a certain score. Their system did not support shifting the task level and competition based on performance level.		

[a] Colours represent the origin of the sample. Green = Asia, blue = Europe, brown = Africa, pink = Oceania, yellow = North-America, and red = South-America.

Figure A

Demographical Distribution (Using the colour pattern used by Risk: Retrieved from https://godatadriven.com)



Note. Yellow = North America, red = South America, blue = Europe, brown = Africa, green = Asia, and pink/purple = Oceania.