Full length article

The higher the score, the higher the learning outcome? Heterogeneous impacts of leaderboards and choice within educational videogames

Steve Nebel[*], Maik Beege, Sascha Schneider, Günter Daniel Rey

E-Learning and New Media, Faculty of Humanities, Technische Universität Chemnitz, Germany

ABSTRACT

Interactive multimedia environments such as educational videogames offer great potential for learning in groups with multiple players. Multiplayer games might lead to competition among the learners which is frequently used to motivate them to play again. Additionally, competitive outcomes as discrepancies to a desired standard might differ between players and this type of feedback might influence learning. Therefore, the experiment seeks to investigate learning effects of different amounts of standard discrepancy and the choice to repeat levels. Standard discrepancy was operationalized by either showing a high learner score (low standard discrepancy) or a medium learner score (large standard discrepancy) at a leaderboard. Choice to repeat a level (possibility to repeat vs. no possibility to repeat) was manipulated by presenting or hiding a repeat button. An experiment was conducted with 85 students who played a jump-and-run game in order to learn facts about three allegorical paintings. Results revealed an effect of standard discrepancy on retention performance with higher scores for the high standard discrepancy condition. Choice did not influence learning outcomes, but improved motivational and emotional measures. Findings underpin the new role of leaderboards as feedback mechanisms.

© 2016 Elsevier Ltd. All rights reserved.

Leaderboards as displays of ranks for comparison (Seaborn & Fels, 2015) are popular among different applications in digital media. They are frequently used within internet-based games as they can be persistent, accumulated long-term and represent skill levels better than single play sessions (Wang & Sun, 2011). Since the early days of arcade cabinets, leaderboards are used within commercial games to increase replayability. For example, ARPG's (Action Role Playing Games) like Diablo III (2012) or Path of Exile (2013) utilize season based leaderboards. As a consequence of their popularity, leaderboards are interesting for gamification. This concept is defined as "the intentional use of game elements for a gameful experience of non-game tasks and contexts" (Seaborn & Fels, 2015, p. 17). For this, leaderboards are often included in quizzes like a usability-quiz (Hemke, Meyer, Hühne, Schneider, & Wohlgemuth, 2014) or a quiz for hearing-impaired learners (Glova, Asuncion, Martin, Manzan, & Pagtaconan, 2014). They can even be found in applications for project management (e.g., Kudos Badges, 2015) or within tools to foster user participation during software development (Halan, Rossen, Cendan, & Lok, 2010). Considering these examples, it is not surprising that leaderboards appear advantageous to creators of educational learning materials. However, sufficient research analyzing individual learner outcomes when interacting with leaderboards is missing. Even if leaderboards are included in empirical comparisons they are often part of a larger gamification strategy (Bajko, Hodson, Seaborn, Livingstone, & Fels, 2015) or included in combination with achievements and other competitive mechanics (Landers & Landers, 2015) which limits the interpretation tremendously. Empirical studies addressing gamification elements often lack of sufficient statistical analysis to generate comparable effects (Seaborn & Fels, 2015). Even if valuable data is collected, the researchers regularly target other factors than learning (e.g. fun, Butler, 2013). Therefore, creators of educational content still need sufficient empirical studies to base their decision on whether to include leaderboards or not. Additionally, studies do not provide sufficient information on how leaderboards influence the individual learner. With this experiment, we seek to provide further empirical evidence to guide the

* Corresponding author. E-Learning and New Media, Faculty of Humanities, Technische Universität Chemnitz, Straße der Nationen 12, 09111 Chemnitz, Germany.
E-mail address: steve.nebel@phil.tu-chemnitz.de (S. Nebel).

implementation of leaderboards and to close the scientific gap of individual impacts.

1. Literature review

1.1. Leaderboards in (educational) videogames

Leaderboards can contain and categorize various elements, such as the number of correct answers, achieved goals or time spent. They can also be used to evaluate and categorize solutions for given tasks (e.g., Foldit, Beta). The required data is usually easy to acquire as it is already part of the game. The resulting list is comparably easy to integrate into an educational concept as no interference with gameplay occurs. These technical advantages might partially explain the popularity of leaderboards. However, leaderboards influence several other aspects of educational gaming. For example, the increased difficulty by leaderboards also increases perceived value of the achieved results, thus increasing memorability (Wang & Sun, 2011). When presented at the end of the game, leaderboards are typically used to increase the motivation to play again and subsequently increase time-on-task (Landers & Landers, 2015). Leaderboards provide entertainment, a sense of accomplishment and memories linking play events to specific rewards (Wang & Sun, 2011). Leaderboards serve as source of motivation (Schubert, Paulsen, & Hager, 2014; Willems et al., 2014) as learners see that their work is recognized (Domínguez et al., 2013). Beyond these motivational impacts, leaderboards can influence goals and the perception of progress within educational videogames by mapping progress and inciting actions (Seaborn, Pennefather, & Fels, 2013). For example, if students lack of engagement in the early stages of a class or game they fall back in a leaderboard and start taking actions (Barata, Gama, Jorge, & Gonçalves, 2013). Therefore, leaderboards provide stimuli or goals (e.g., Glova et al., 2014; Hemke et al., 2014) which might help to complete learning tasks (Domínguez et al., 2013). Leaderboards might also be useful when players are not able to beat the complete game. Players can try to beat their own records (Wang & Sun, 2011) and, subsequently, perceive achievement even during the early stages of the learning process. In sum, leaderboards can provide long-term goals (e.g., be the best) and short-term goals (e.g., improve a self-determined number of places), which both support motivation and orientation during playtime. In addition, leaderboards, as the easiest form of social interaction (Domínguez et al., 2013; Wang & Sun, 2011), are perceived as more influencing than achievements or similar progression indicators. Wang and Sun (2011) claim that the reward of a satisfying position within the leaderboard might encourage social interaction as exchanging information might improve performance and, subsequently, closeness to a desired status. Leaderboards increase collaboration (Schubert et al., 2014) and competition (Sarangi & Shah, 2015) as well. This might be amplified even further, as providing information serves as a possibility to show off learned skills. Regarding these factors, it should be noted that the acceptance of (public) competitive factors might be culturally dependent (Barata et al., 2013; Cheong, Filippou, & Cheong, 2014; Schubert et al., 2014) and the impacts of leaderboards might vary within different cultural backgrounds. Additionally, the limited information value of some leaderboards might lead to learners feeling uncomfortable with their position (Domínguez et al., 2013). For example, learners do not know if others actually learned more or if they just played better. In this vein, it is shown that competition within leaderboards could lead to demotivation if the distance to other students gets too high. Therefore, leaderboards require enough users in order to ensure the existence of comparable competitors (Willems et al., 2014).

1.2. Heterogonous effects of competition

It is hypothesized that leaderboards affect motivation, post-test performance and behaviors through competition (Cagiltay, Ozcelik, & Ozcelik, 2015; Simões, Redondo, & Vilas, 2013). Competitive gameplay might increase interest (Plass et al., 2013), enjoyment (Vorderer, Hartmann, & Klimmt, 2003), or attention, excitement and involvement (Vandercruysse, Vandewaetere, Cornillie, & Clarebout, 2013). However, recent research on the effect of competition within various group constellations has shown heterogeneous effects (Nebel, Schneider & Rey, 2016). Among other things, it can be derived from this study that the impact of competition might vary from player to player relative to the personal gameplay experience and personal traits. Although competition as a form of social comparison should lead to an unidirectional drive upward (Festinger, 1954), this is not distributed homogenously among the players, and subsequently, within the leaderboard. Some learners might perceive stronger competition as others relative to their proximity to standards (Garcia & Tor, 2007, 2009; Garcia, Tor, & Gonzalez, 2006). Therefore, the effects of competition might vary among the players as well. For example, players close to a standard are more likely to replay parts of the game (Butler, 2013). In contrast, learners that are allowed to control the task themselves tend to practice a different task after good trials (Wu & Magill, 2011). Therefore, players in good positions might not choose to play again but rather play another level to learn something different. In addition, personal traits influence how enjoyable leaderboards actually are. For example, introverts might appreciate an offline-leaderboard more than extroverts (Codish & Ravid, 2014). Nonetheless, generally leaderboards are rather motivating than demotivating (Schubert et al., 2014).

1.3. Leaderboards as feedback mechanism

Leaderboards do not only induce competitive effects, they also provide information on how a player performed. The simplest element is the rank itself (e.g., 34rd). It can be regarded as a *praise for task performance*, although this kind of feedback contains very limited learning related information (Hattie & Timperley, 2007). Achievements can be classified as glory mechanisms (Wang & Sun, 2011) as well, although it is important to differ between achievements and leaderboards. The first are permanently granted and indicate an already reached standard and signal no need for further behavior. The latter also indicates a certain skill level but, additionally, a difference towards a potentially more desirable standard. However, the information leaderboards provide does not represent continuous feedback as it appears only punctual. Therefore, leaderboards can be compared to feedback interventions (FI), which are defined as "actions taken by external agents to provide information regarding some aspects of one's performance" (Kluger & DeNisi, 1996, p. 255), whereas the external agent is the game mechanic that forces the player to recognize the leaderboard. These FI impact pleasantness and arousal. Depending on its type (positive or negative) FI can result in positive or negative mood (Kluger & DeNisi, 1996). In line with some aspects of competition, the *Feedback Intervention Theory* (FIT) states that behavior is regulated by a comparison of feedback to standards (Kluger & DeNisi, 1996). Learners with a negative feedback (feedback that the performance is below a desired standard) either increase their effort, abandon or modify the standard, or reject the feedback message (Kluger & DeNisi, 1996). In line with the goal setting theory (Locke & Latham, 1990, 2002), Feedback Intervention Theory argues that arousal is elevated because the feedback-standard gap increases (Kluger & DeNisi, 1996). As a consequence, learners who perform already within the boundaries of a standard might not increase