The Underlying Reason for Mixed Results in Gamification Studies

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Eric Weng
Spring, 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor
MC Forelle, Department of Engineering and Society

Introduction

If you asked a student to recount their favorite lesson in last semester's calculus class, chances are they would be hard-pressed to provide an answer. Now, if you asked that same student to describe their hardest boss fight in a video game they played a decade ago, they could probably spend hours recounting every intricate detail. All across the United States, students are losing faith in their country's education system. While 88% percent of college students state their reason for pursuing higher education is to get a job, only 34% of students believed their degree prepared them for the job market (Strada Education Foundation, 2018). On the other hand, Americans cannot get enough of video games. In 2023, the US video game market made $114 billion in revenue (Clement, 2024), exceeding the GDPs of all but the 17 wealthiest countries in the world (World Bank, 2023). Why have games captivated millions of young learners' minds in a way that school is unable to replicate, and how can we make education more exciting for children?

Luckily for students, the idea of using games to improve everyday life is not new. Gamification is the practice of applying game design elements to non-game settings with the goal of promoting desirable behavior through fun (Lee & Hammer, 2011). For example, teachers may hand out stickers to students who repeatedly excel in class, mirroring how video games award virtual achievements to players who complete certain feats. While formal literature surrounding gamification stretches back to the 2000s, thanks to the rise of digital learning technology and computer games, gamification has now become a burgeoning field of research and industry (Deterding et al., 2011).

However, it is not all smooth sailing for gamification (to use a game metaphor, this is where the boss reveals its second health bar). For educators, integrating new techniques into

classrooms could further deplete already scant resources that could be better allocated elsewhere. Likewise, being forced to engage in play risks both demotivating students who do not want to participate and conditioning them to learn only when presented with external rewards (Lee & Hammer, 2011). Moreover, despite what marketers would advertise to the public, the mechanics behind gamification are still not well understood. Scholars frequently express the need for more rigor and empirical data, without which multiple studies testing the same set of game elements have oftentimes obtained conflicting results (Dicheva et al., 2015). It seems the enormous hype surrounding gamification has pushed researchers to jump the gun on proper scientific etiquette .

Therefore, I argue that the inconclusive results among gamification studies are mainly due to researchers following the hype surrounding gamification, causing a lack of controlled experimental design and limited consideration for teacher and student needs. First, I will provide a background into the rationale, benefits, and challenges of gamification and explain how to successfully adopt it into classrooms. Then, by analyzing scholars' methodologies for designing studies, I will demonstrate that researchers' attachment to the expected image of gamification prevents them from exercising sufficient scientific rigor. I will also show that such studies are difficult to integrate into existing pedagogical practices. By the end, I will illuminate how gamification research can benefit by listening to the practices and needs of teachers and students.

Background and Literature Review

To understand why researchers are interested in gamifying education, let us first understand the shortcomings of the modern education system. During the Industrial Revolution, schools trained students to recall memorized facts and pass standardized tests—a learning style often termed the "factory model" (Bashore, 2022). Over a century later, schools in the US still

largely prioritize students getting the right answer and earning good exam scores, while STEM subjects are given higher job value compared to the humanities and arts. However graduates are finding themselves supplanted by machines and AI in the roles they were trained for, thanks to advancements in automation (Krishnan, 2020). Schools' overreliance on memorization and standardization not only promotes conformity and compliance and stifles creativity in students, but severely restricts the image of a "successful" student (Fredericks, 2021). As a result, students who do not fit such a rigid mold and are not supported by their schools may begin to view themselves as unintelligent and eventually drop out (Long, 2017). To keep students competitive in modern workplaces, schools should cultivate students' creativity and problem-solving skills.

Gamification can address the motivation and engagement issues currently plaguing education by granting students more control over their learning. According to self-determination theory, gamification provides sources of both intrinsic and extrinsic motivation by fulfilling students' primary psychological needs: competence (perceived aptitude), autonomy (control over oneself), and relatedness (social belonging) (Mekler et al., 2017). First and foremost, a gamified learning environment acts as a sandbox for students to safely develop their critical-thinking skills where players do not fear failure but learn from it. Since games typically allow players to restart a level from an earlier checkpoint, students possess the freedom to experiment with different strategies to win (Berkeley Center for Teaching & Learning, 2015). Games also let players learn more quickly from their mistakes by providing instant feedback, as opposed to students waiting weeks for exams and papers to be graded (McGonigal, 2011). Compared to traditional lecture-based approaches, students can organically develop their knowledge at their own pace. Additionally, because games are better at inducing flow states—periods of intense concentration and enjoyment—students experience higher engagement by redirecting energy towards

productive tasks for longer periods of time (McGonigal, 2011). Lastly, students also have more avenues to exercise social interaction and teamwork in new and interesting ways, helping them feel more in touch with classmates they do not know well. In short, gamification motivates and engages students with a personalized and interactive experience that naturally facilitates higher knowledge retention and test scores.

Despite being a relatively new field, gamification has seen significant growth in numerous industries in the last decade and a half. The global gamification market was valued at around $12 billion in 2021 and is projected to grow to $30 billion in 2025, around 20% of which is expected to be occupied by education (Boksamp, 2023). Outside of education, gamification has also been applied to sectors such as business, healthcare, and government, and military, suggesting it will remain popular for many years to come. Although increased attention attracts more research and commercial investment into gamification, it also spreads misinformation and misconceptions that hinder adoption by educators (Kabilan et al., 2023). Teachers will want to acquire a strong understanding of how to implement it in the classroom to avoid harming student learning.

One such framework for adopting new learning practices and technologies is Technological, Pedagogical, and Content Knowledge (TPACK, or formerly TPCK). To start, an educator's overall competency is composed of knowledge in three domains—Technological, Pedagogical, Content—and their intersections (Koehler & Mishra, 2005). First, Technological Knowledge (TK) is the ability to use technology in the classroom, from whiteboard projectors to electronic learning management systems. Second, Pedagogical Knowledge (PK), is the experience in teaching practices and learning styles, such as gamification. Lastly, Content Knowledge (CK), is expertise in the subject matter being taught or learned. TPACK treats these

three domains with equal importance to ensure that instructors do not prioritize any over the others. From here, the three basic domains intersect to form TCK, PCK, and TPK, and the latter then combine to form TPCK. TPACK's philosophy is that teachers cannot improve education by just picking up new skills and techniques, but must learn how to best apply them to their existing classrooms (Koehler & Mishra, 2005). In the context of gamification, an instructor might start by searching for software relevant to their particular subject (TCK), such as voice recorders for language courses and graphing tools for mathematics courses. Similarly, the instructor may want to learn what game elements are most conducive for learning in their subject areas (PCK). For instance, a discussion-based class may not be suitable for a highly competitive environment. Next, the instructor could practice creating their desired gamified methods using the available online learning softwares (TPK). Finally, the instructor will strive to create a system in which they understand the role of each domain and its relationship to the others (TPCK). Moving forward, following TPACK allows us to think like educators and assess gamified systems based on how much or how little they disrupt instructors' capabilities to effectively teach and students' capabilities to learn.

Methods

With all the exaggerated headlines and purported figures claiming gamification to be a panacea for education's problem, educators probably desire to know which game elements they should use in their classrooms and what outcomes will arise in students. Therefore, a research question arises: How do the selected game elements and result metrics reflect the perceived success of gamification studies?

To answer this question, I conducted a critical literature review on past gamification studies, paying particular attention to the game elements they used and result metrics they recorded. Most articles were accessed from the ACM Digital Library, ScienceDirect, IEEE Xplore, and other computer science-related databases by searching for articles with the keywords "gamification", "student", "learning/education", and the name of some element or metric. I found that the references in meta-analyses were an excellent starting point. My research consisted of reading the introduction, methods and conclusion sections of each paper to determine which elements and metrics the authors chose, their justification for choosing them, and how they interpreted their results. Moreover, I evaluated each study for how well the researchers adhered to TPACK, such as how well the gamified system suited the subject area or fit with the technology used. Originally, I also planned a comparative literature review, intending to compare and contrast how choosing "traditional" elements (points or badges) and metrics (grades or motivation) versus "non-traditional" ones affected how positive or conclusive the results are. Unfortunately, I not could not accomplish the second review due to the sheer imbalance and overlap of those game elements and result metrics in studies.

One notable weakness of my research method is that none of the aforementioned resources contain any direct student accounts of the gamification, since all student experiences were merely summarized in the results section. Additionally, any faults in experimental design present in these studies may limit how accurate or representative my conclusions are. Nonetheless, the evidence I gathered can still be considered primary sources, since it comprises the writers' own testimonies and reasonings.

Analysis

Much of gamification research is rooted in the popular but poorly substantiated claim that gamification improves engagement, motivation, and grades, allowing misinformation to propagate. Interviews with participating students reveal that lack of motivation and engagement due to environmental factors or challenging courses is a root cause of poor academic performance (Cao et al., 2023). Clearly, authors have a good reason for commonly selecting grades, engagement, and motivation—which I will refer to collectively as GEM—as metrics. On the other hand, research into validating said presumption has been scarce, compared to the amount of work done in other computing or educational fields (Attali & Arieli-Attali, 2015). Additionally, existing studies have been poorly-designed, with many mixing different elements, lacking control groups, or using limited sample sizes. (Hanus & Fox, 2015). Unsurprisingly, these studies frequently lead to inconclusive or contradictory results. While it is good that researchers continue to verify these beliefs, other scholars who are already operating on false assumptions may end up conducting biased or unreproducible research, harming the accuracy of future studies. By the time such beliefs are proven or rejected, other misinformation may already have become common practice. Likewise, educators who bring misconceptions to their classrooms will find themselves building their skills on incorrect pedagogical knowledge. As a result, students may experience adverse learning outcomes, and both parties may be dissuaded from further experiences with gamification.

Due to the high emphasis of gamification research on GEM metrics, scholars have inadequately explored the impact on learner and social factors. Out of 51 studies analyzed in a literature review, the majority (31) were classified as behavior and cognitive, or relating to GEM outcomes. In contrast, far fewer (10) studies were learner-centric, or concerned with how

students personally felt about the game elements and whether groups of students reacted differently to them (Dichev & Dicheva, 2017). A few papers explicitly mentioned studying social interaction, but these were not given a distinct category in the results. Despite the wealth of research focusing on GEM metrics, researchers have not explored less-documented factors such as student satisfaction, perception of learning, and social interaction. Nevertheless, detractors may argue investigation into those factors is not worth the proportional reward given how conditioned the current educational system is toward grades and test scores. While practically speaking, teachers and school administrators would more readily adopt gamification if they were confident that it would improve grades, such short-term thinking is what got us into trouble in the first place. If we do not break the mindset that grades are the most important goal of education, then we will inevitably end up in a similar—or worse—dilemma as we are in now. Furthermore, self-determination theory states that relatedness contributes towards students' intrinsic motivation (Krath et al., 2021). For example, a system in which students were not allowed to communicate with other students or voice how they felt about their experience would be less like a school, and more like a prison. In the worst case, this blind devotion to holding GEM as the only outcome runs the risk of discouraging new researchers from exploring other learning outcomes that are all equally as essential to a successful classroom. Combined with the existing misconceptions, this unbalanced research body means teachers looking to build their gamified classrooms must do so on unreliable and incomplete pedagogical knowledge.

Consequently, lax experimental design has prevented researchers studying gamification from understanding how individual game elements affect students. Researchers often use particular game elements just because they are popular, to the extent that the point, badge, and leaderboard elements are grouped together as PBL (Dichev & Dicheva, 2017). In one study, the

authors even purposefully included a leaderboard, despite knowing it would blur the distinction between their treatment groups, writing that leaderboards were “an essential element in almost every gamification approach” (Krause et al., 2015). This conformity to expectation reinforces the preconceived image of what gamification should look like, ironically mirroring the predicament many students find themselves in. As with metrics, relying too heavily on what is popular precludes scholars from exploring alternate designs using mechanics such as narrative and randomness. Moreover, because many studies opt to test multiple elements simultaneously, researchers have difficulty isolating which element caused which result (Dicheva et al., 2015). Without the knowledge of single elements, authors cannot definitively tell if combining elements leads to different outcomes than those of the elements alone, or if a confounding variable arose. Although one could argue that single elements are difficult to test and that elements need to combine to have any effect, merely including an element because it is “standard” is insufficient justification for a lack of scientific rigor. Since gameplay arises from the interplay of all the mechanics and the people interacting with them, researchers should not assume a gamified system is the sum of its elements. Good experimental design does not require all scholars to become fluent in game design, but rather steadily and systematically build up knowledge using controlled experiments, like has been repeated time and time again in other fields. If educators are to be well-informed on how to use gamification to help students learn, researchers cannot resort to haphazardly throwing together game elements and expecting accurate results.

When researchers select game elements based on popularity instead of good game design or pedagogical practices, positive motivators may produce adverse effects on students. Competitive game elements like those in the PBL group are often believed to drive extrinsic motivation while potentially lowering intrinsic motivation (Ratinho & Martins, 2023). After

enough exposure to external rewards, students may begin to learn only when promised a badge, which is no better than students seeing grades as the only reason for learning (Toda et al., 2018). In one study, heatmaps used to visualize projected student performance (akin to leaderboards) motivated high-performing students to work harder and earn better grades. On the flip side, those heatmaps drove already low-performing students to stop interacting with the game elements or seeking out challenges (Auvinen et al., 2015). When students are already motivated, external rewards such as badges can actually make learning less satisfying and students feel controlled by the game, which defeats the principal purpose behind introducing gamification to classrooms. (Hanus & Fox, 2015). These results reveal how gamification is not a substitute for proper motivation, and a forced increase in external rewards and competition serves only to distract students from learning or widen the achievement gap between motivated and unmotivated students. Thus, teachers should not envision gamification as a replacement for existing teaching methods, but as a supplementary resource that can be molded to fit them.

In the same vein, gamification only benefits student learning when designers and teachers place proper measures to enforce desirable actions. On a gamified social platform where students received badges for sharing discoveries, very few students genuinely and meaningfully interacted with others under their own volition. Rather, many others only participated for the sake of collecting badges or when prompted by teachers. (Boticki et al., 2015). Without proper intervention by instructors, students in gamified environments will fail to properly utilize the game elements and provide low-quality contributions, quickly falling back to their existing habits. In another study, students who were placed into teams and were given a storyline to follow remained focused, staying after class to finish additional challenge tasks and asking questions to teachers (Cao et al., 2023). Despite the use of point and leaderboard elements, it

appeared that narrative and teamwork elements helped students belong to a group of supportive peers and gave them an extra sense of progression beyond winning first place. Together, studies like these highlight how teachers or other students are virtual for mediating interaction with the gamified system.

However, if researchers do not collaborate with teachers and students to design gamified learning environments, the resulting systems will poorly integrate into classrooms. Many methodologies follow a wide-range of design frameworks that share little in common (Mora et al., 2017). Out of 40 publications reviewed, 24 employed frameworks that were described as user-centered, focusing on human-computer interaction or user experience, while the rest were game-centered and technology-centered. Strangely though, none of the frameworks involved the relevant stakeholders (teachers and students) at any stage of design, with only half the studies mentioning stakeholder interaction (Mora et al., 2017). This blatant neglect of stakeholders constitutes a gross violation of engineering practice. Delivering a product that does not follow the client's requirements—no matter how amazingly constructed—means failing to honor the original contract, as well as wasting countless weeks of work. If researchers do not ask teachers or students about the type of learning environment they want, the gamified system risks clashing with the teaching style. In turn, instructors will see no incentive to collaborate with students, flocking to whatever digital gamification tools appear most trendy and attractive instead of understanding their benefits to pedagogy (Kabilan et al., 2023). In the original paper proposing TPACK, the authors described how teachers and students can work together to better adapt educational technology to the classroom through "learning by design" (Koehler & Mishra, 2005). By the same token, in order to successfully design and implement gamification, researchers, instructors, and students should all work together to discover each other's needs. Additionally,

putting equal effort game- and technology- centered design alongside user-centered design ensures teachers can leverage their existing pedagogical and technical knowledge to exert more control over how the new system interacts with those knowledge domains.

Conclusion

Ultimately, the answer to the question, "How do the selected game elements and result metrics reflect the perceived success of gamification studies?", turns out to be that they do not—not completely. In a rush to capitalize on gamification's hype, scholars have grown too focused on what has become standard or expected, while neglecting to question whether their most basic assumptions are fact or myth. It is true that certain game elements do tend to create mixed or negative outcomes in students when implemented poorly. Still, more often than not the underlying variable is flawed experimental design, whether that be picking too many game elements or believing grades are the be-all and end-all to gamification. Some of this behavior is understandable, since it is easier to achieve concrete findings on more obvious areas such as grades. However, this penchant for overspecialization and trend-chasing endangers educators and students by allowing misinformation to go unchallenged.

So what can we as researchers do to improve future research? First and foremost, we must exercise stronger scientific rigor in our designs in regards to game elements and result metrics. I have not even begun to mention other demographics factors such as age, gender, and subject that we will need to consider. Above all, though, I believe we need to involve teachers and students more actively in research—not just as subjects, but as consultants. Afterall, they are the ones most directly impacted from gamification. Ignoring input from teachers and students has led to techniques that only work in ideal conditions and do not adapt easily to existing

classrooms. It is my hope that readers will share this paper with colleagues, instructors, and students to spread awareness of the limitations of the current body of research.

As noted before, I would have liked to read more student testimony. Further studies should strive to include a few interviews with participating students (with identifying information removed) describing how easy they adjusted to the gamified system, and if they preferred it to their current one. In addition, teachers could also weigh in on how costly implementing the tested system into their classrooms full time would be. Perhaps gamification research will always remain difficult due to how subjective games can be to their players. Though our quest for greater knowledge is filled with deadly dungeons and treacherous boss fights, I believe we will complete it by avoiding the pitfalls of hype and trends, listening to all voices in the party, and remaining methodological and objective in our game plan.

References

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: Do points affect test performance? *Computers & Education*, *83*, 57–63. https://doi.org/10.1016/j.compedu.2014.12.012

Auvinen, T., Hakulinen, L., & Malmi, L. (2015). Increasing Students' Awareness of Their Behavior in Online Learning Environments with Visualizations and Achievement Badges. *IEEE Transactions on Learning Technologies*, *8*(3), 261–273. IEEE Transactions on Learning Technologies. https://doi.org/10.1109/TLT.2015.2441718

Bashore, J. (2022, May 5). *The Factory Model: Does It Work? Did It Ever?* One-Room Education. https://one-roomeducation.com/the-factory-model-does-it-work-did-it-ever/

Berkeley Center for Teaching & Learning. (2015). *Implementing Gamification*. Berkeley Center for Teaching & Learning. https://teaching.berkeley.edu/teaching-guides/running-your-course/implementing-gamification

Boksamp, E. (2023, February 23). 25 Gamification Statistics [2023]: Facts + Trends You Need To Know. *Zippia*. https://www.zippia.com/advice/gamification-statistics/

Boticki, I., Baksa, J., Seow, P., & Looi, C.-K. (2015). Usage of a mobile social learning platform with virtual badges in a primary school. *Computers & Education*, *86*, 120–136. https://doi.org/10.1016/j.compedu.2015.02.015

Cao, C., Zhao, X., & Hopfgartner, F. (2023). Enhancing Chinese International Students' Motivation and Engagement in a STEM Subject in UK Higher Education: A Case of Gamification in Programming. *Proceedings of the 14th International Conference on Education Technology and Computers*, 6–11. https://doi.org/10.1145/3572549.3572551

Clement, J. (2024, October 21). *Video gaming market in the United States—Statistics & Facts*. Statista. https://www.statista.com/topics/8696/video-games-market-in-the-united-states/

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education*, *14*(1), 9. https://doi.org/10.1186/s41239-017-0042-5

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society*, *18*, 75–88.

Fredericks, A. D. (2021, August 13). How Education Quashed Your Creativity. *Psychology Today*. https://www.psychologytoday.com/us/blog/creative-insights/202108/how-education-quashed-your-creativity-0

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, *80*, 152–161. https://doi.org/10.1016/j.compedu.2014.08.019

Kabilan, M. K., Annamalai, N., & Chuah, K.-M. (2023). Practices, purposes and challenges in integrating gamification using technology: A mixed-methods study on university academics. *Education and Information Technologies*, *28*(11), 14249–14281. https://doi.org/10.1007/s10639-023-11723-7

Koehler, M. J., & Mishra, P. (2005). What Happens When Teachers Design Educational Technology? The Development of Technological Pedagogical Content Knowledge. *Journal of Educational Computing Research*, *32*(2), 131–152. https://doi.org/10.2190/0EW7-01WB-BKHL-QDYV

Krath, J., Schürmann, L., & von Korflesch, H. F. O. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, *125*, 106963. https://doi.org/10.1016/j.chb.2021.106963

Krause, M., Mogalle, M., Pohl, H., & Williams, J. J. (2015). A Playful Game Changer: Fostering Student Retention in Online Education with Social Gamification. *Proceedings of the Second (2015) ACM Conference on Learning @ Scale*, 95–102. https://doi.org/10.1145/2724660.2724665

Krishnan, K. (2020, April 13). *Our education system is losing relevance. Here's how to update it*. World Economic Forum. https://www.weforum.org/agenda/2020/04/our-education-system-is-losing-relevance-heres-how-to-update-it/

Lee, J. J., & Hammer, J. (2011). Gamification in Education: What, How, Why Bother? *Academic Exchange Quarterly*. https://api.semanticscholar.org/CorpusID:220866115

Long, C. (2017, December 19). *Some of the Surprising Reasons Why Students Drop Out of School*. National Education Association. https://www.nea.org/nea-today/all-news-articles/some-surprising-reasons-why-students-drop-out-school

McGonigal, J. (2011). *Reality is Broken: Why Games Make Us Better and How They Can*

Change the World. Penguin Books.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior*, *71*, 525–534. https://doi.org/10.1016/j.chb.2015.08.048

Mora, A., Riera, D., González, C., & Arnedo-Moreno, J. (2017). Gamification: A systematic review of design frameworks. *Journal of Computing in Higher Education*, *29*(3), 516–548. https://doi.org/10.1007/s12528-017-9150-4

Ratinho, E., & Martins, C. (2023). The role of gamified learning strategies in student's motivation in high school and higher education: A systematic review. *Heliyon*, *9*(8), e19033. https://doi.org/10.1016/j.heliyon.2023.e19033

Strada Education Foundation. (2018, February 18). *New Survey Reveals Crisis of Confidence in Workforce Readiness Among College Students*. https://web.archive.org/web/20240526222148/https://stradaeducation.org/press-release/new-survey-reveals-crisis-of-confidence-in-workforce-readiness-among-college-students/

Toda, A. M., Valle, P. H. D., & Isotani, S. (2018). The Dark Side of Gamification: An Overview of Negative Effects of Gamification in Education. In A. I. Cristea, I. I. Bittencourt, & F. Lima (Eds.), *Higher Education for All. From Challenges to Novel Technology-Enhanced Solutions* (pp. 143–156). Springer International Publishing. https://doi.org/10.1007/978-3-319-97934-2_9

World Bank. (2023). *GDP (current US$)*. World Bank Open Data. https://data.worldbank.org