This article is a part of a symposium titled: *On the Architecture of Game Science*

Simulation & Gaming
1–23

DOI: 10.1177/1046878118774385
journals.sagepub.com/home/sag
SAGE

Gamification Science, Its History and Future: Definitions and a Research Agenda

Richard N. Landers[1], Elena M. Auer[2], Andrew B. Collmus[2], and Michael B. Armstrong[2]

Abstract

Background. Definitions of **gamification** tend to vary by person, both in industry and within academia. One particularly popular lay interpretation, introduced and popularized by Ian Bogost, and reiterated by Jan Klabbers, is that gamification is “bullshit” and “exploitationware.” They describe gamification as a marketing term or business practice invented to sell products rather than to represent a real and unique phenomenon relevant to a nascent **game science**. However, this view is an oversimplification, one which ignores a growing body of theory development and empirical research on gamification within a **post-positivist epistemology**. In fact, because gamification is so much more outcome-focused than general game design, current gamification research in many ways has a stronger footing in modern social science than much games research does.

Aim. In this article, to address common misunderstandings like these, we describe the **philosophical underpinnings** of modern gamification research, define the relationship between games and gamification, define and situate **gamification science** as a subdiscipline of game science, and explicate a six-element framework of major concerns within gamification science: **predictor** constructs, **criterion** constructs, **mediator** constructs, **moderator** constructs, **design processes**, and **research methods**. This framework is also presented diagrammatically as a causal path model.

[1]University of Minnesota, Twin Cities
[2]Old Dominion University

Corresponding Author:
Richard N. Landers, 75 E River Road, University of Minnesota, Minneapolis, MN 55416
Email: rnlanders@tntlab.org

Conclusion. Gamification science refers to the development of **theories of gamification design** and their **empirical evaluation** within a post-positivist epistemology. The goal of gamification **scientist-practitioners** should be to understand how to best meet organizational goals through the design of gamification interventions, drawing upon insights derived from both gamification science and games research more broadly.

Keywords
constructs, criteria, criterion, design, game, game science, gamification, gamification science, mediator, moderator, post-positivism, predictor, research methods, science, theory

There are many different definitions of *gamification* to be found in both the research literature and among lay people. One particularly prominent interpretation, popularized by Ian Bogost in *The Atlantic* and the games industry blog *Gamasutra*, is that "gamification is bullshit" (Bogost, 2011a) and "exploitationware" (Bogost, 2011b). This view has made its way into the rhetoric of some games researchers, including that of Klabbers (2018), as a technique to discredit the entire field of gamification as legitimate scholarly enterprise. As Klabbers describes it, "Gamification is a business practice… a management method… [it] does not aim at making real-time business processes a playful game" (p. TBD). Klabbers goes so far as to presume the purpose of gamification despite neither himself being a practitioner nor citing any empirical evidence to support the point: "'gamifiers' apply a behaviorist approach to managing the workplace, to improve performance" (p. TBD). This unsupported proscription assumes a great deal about those practicing gamification that appears inaccurate given available data. In short, Klabbers (2018) is not describing gamification as practiced but rather as he imagines it to be. Condemning an entire field of study as illegitimate based upon nothing but assertion is dangerous and damaging to this new field's credibility.

Klabbers' (2018) stance here is particularly confusing given his claim that game science is rooted in modern physics, which adopts a hard logical positivist approach to the world: there is an underlying truth that must be discovered. In contrast, modern social scientists generally rely upon post-positivism, which is neither logical positivist nor postmodern in nature. Simply put, logical positivists generally believe in an objective and objectively measurable reality, postmodernists believe that all knowledge is constructed socially and therefore subjective, and post-positivists believe in an objective reality that is viewed through the lens of subjective interpretation. Post-positivism, typically credited to Karl Popper (1963), is the philosophical framework underlying essentially all modern social science. Adopting that epistemology, our mission as social science researchers is to minimize our bias and subjectivity as much as possible while investigating various truths, realizing that our interpretation of those truths is always through the lens of our experiences and that those truths may change while or after we attempt to observe them.

As can be easily seen by perusing even a small part of its sizable research literature (currently over 30,000 papers on Google Scholar; over 700 works indexed by Web of Science), gamification is frequently studied within this philosophical framework. This is evidenced whenever a stance is taken suggesting: 1) that properly designed gamification interventions have the potential to affect people in some specific desirable way, 2) that collection of the data is the best way to determine this effect or its boundary conditions, and 3) that various interpretive biases should be minimized while doing so, such as via experimental design (see Hamari, Koivisto, & Sarsa, 2014 for an early listing of such studies). Within this framework, the core of gamification is a design process (see Deterding, 2015; Deterding, Dixon, Khaled, & Nacke, 2011) intended to augment or alter an existing real-world process using lessons (initially) from the game design research literature to create a revised version of that process that users will experience as game-like. Gamification is not itself a product; one does not create a gamification as one creates a game. Instead, one adds game elements to change a process that already exists to change how that process influences people. This sort of augmentation has been tested in many contexts, including education (e.g., Landers & Landers, 2014), participation in government (e.g., Bista, Nepal, Paris, & Colineau, 2014), health (e.g., Pyky et al., 2017), marketing (Hamari, 2017), and management (e.g., Mekler, Bruhlmann, Tuch, & Opwis, 2017; Stanculescu, Bozzon, Sips, & Houben, 2016). This diversity of applying design principles inferred by successful game design is the core of gamification, not simply "applying a behavioral approach" (Klabbers, 2018, p. TBD).

This results in a more complex and nuanced definition of gamification than Klabbers' (2018) "behavioral management technique" (p. TBD). Instead, gamification research focuses upon the designer's intentions and implementation choices, which vary widely but are unified by the degree of attention paid to inspiration and lessons drawn from game design. When gamifying, the designer's goal is to create a specific change in a distal target outcome, such as increased learning (e.g., Landers & Landers, 2014), increased health (Pyky et al., 2017), increased job performance (Stanculescu et al., 2016), or increased civic engagement (e.g., Bista et al., 2014). The causal pathway from gamification to distal outcome is buffered by one or more intermediary causal changes, called mediators, in the target person's characteristics. Such mediators include attitudinal, motivational, and/or behavioral effects. Therefore, gamification is most conceptually similar to *game design*, not to *games*. This insight is key to understanding the gamification literature and its role in relations to game science. For gamification researchers and practitioners, creating a playful game is not necessarily a design goal. Gamified applications may not even be intended to be fun; for example, Armstrong and Landers (2017) demonstrated how the addition of a narrative alone to an existing employee learning activity improved learner reactions. In situations like these, play is often irrelevant.

Given this background and in pursuit of a unified game science, we present here a view of gamification science as a distinct subdiscipline of game science consisting of researchers adopting a social scientific epistemological footing. This subdiscipline shares many goals and perspectives as broader game science yet occupies a unique niche, similar to the positioning of constructs and concepts described by Deterding and colleagues (2011). This view integrates findings from the study of games and other fields to develop recommendations for how to create and evaluate gamification interventions to ensure they influence outcomes as their designers intended. Therefore, gamification science focuses upon the development of theories of gamification design and their empirical evaluation within a post-positivist epistemological framework. This distinguishes gamification science not only from game science but also from other types of gamification research that are not scientific in nature. The purpose of this distinction is not to condemn non-scientific approaches to studying gamification but rather to set specific boundaries upon what the subdiscipline of gamification science entails. Within this view, the goal of gamification scientists is to understand how to best influence human behavior, attitudes, and other states with designed interventions derived from games. To elucidate this view, we also present a research agenda and framework for the study of gamification science.

Game Science and Gamification, Defined

Game science, broadly defined, refers to the study of games using various tools and assumptions of many natural science, social science, and engineering fields (Klabbers, 2018). It is interdisciplinary in the sense that commonly used tools, relied-upon assumptions, and even epistemological foundations vary across clusters of researchers. Within this broad landscape, those researchers adopting a scientific epistemology, which is to say one adopting a logical positivist and/or post-positivist philosophical orientation, can reasonably call their research science. Within the values of mainstream science, effects exist in the real world and the goal of a scientist is to discover, measure, and predict these effects. For example, given the current capabilities of computers, there are theoretically completely optimized methods for procedural content generation, and research in this area coming from computer science (e.g., Shaker, Yannakakis, & Togelius, 2010) is intended to get closer and closer over time to uncovering this truth. As an example on the social sciences side of the spectrum, researchers are trying to understand the effects of violent video games on children (Ferguson, 2007), with the typically unstated assumption that various true effects of violent video games on children exist and that researchers must try to measure these effects in order to better understand them.

Gamification science can be defined as a social scientific, post-positivist subdiscipline of game science that explores the various design techniques, and related concerns, that can be used to add game elements to existing real-world processes. To draw meaningful conclusions, this science must be post-positivist, rather than logical positivist, because of its focus on human behavior, which is reactive to science and interventions. Unlike the natural sciences, where phenomena continue to exist regardless of our measurement of them, scientists should expect the effectiveness of gamification interventions to change. For example, once a gamification intervention has been implemented in a particular context, its removal does not result in an immediate return to the “ungamified” state (Thom, Millen, & DiMicco, 2012). Instead, the users who previously experienced gamification have themselves been changed through that experience, and this changed state becomes the new truth for this population. Over time, population effects within groups, organizations, and even entire cultures can shift similarly. For gamification science to draw meaningful conclusions about gamification, we must understand this broader context and attempt to integrate it into our thinking when designing studies, yet this sort of thinking is very uncommon in the broader study of game science that Klabbers (2018) describes.

Gamification science thus shares conceptual foundations and a vocabulary with game science but promotes research objectives that are unique to its particular application area, the creation of gamification interventions intended to influence human behavior. As practiced currently, in terms of citations, gamification science is most strongly influenced by the field of human-computer interaction (e.g., Deterding et al., 2011), which is itself at the interdisciplinary crossing of computer science and social science. Researchers within gamification science often play two roles: as observers/evaluators/theorists of gamification in the wild but also themselves as implementers of novel gamified approaches. This exposure to both sides suggests that the most effective gamification scientists are in fact scientist-practitioners (Kanfer, 1990), with experience and understanding of both perspectives, allowing their research to inform their practice and vice-versa.

Building a Science of Gamification

In the remainder of this article, we describe a framework for the science of gamification, dividing our treatment into two sections. The first section describes the four core construct classes in gamification science; specifically, we define and explain the theoretical concepts that are the focus of empirical gamification research. Broadly, theory development in the social sciences involves answering a series of four essential questions about a phenomenon of interest: *what, how/why, who*, and *where/when* (Whetten, 1989). In many cases, methods to address and answers to these questions overlap the methods and answers of the broader game science literature, but for the sake of simplicity, we will focus here on application to gamification science. Within gamification science, these questions are addressed by clearly defining constructs and studying causal direct effects, indirect effects, and boundary conditions amongst those constructs. Thus, when building a framework for the science of gamification, the initial step is to identify and define constructs of interest.

The four types of person-focused constructs studied within gamification science are game elements (predictors), targeted organizational outcomes (criteria), intermediary individual changes (mediators), and personal and situational contexts (moderators). As depicted in Figure 1, game elements (e.g., storylines, action languages, points) are the initial causal force in created change in distally desired outcomes, which in gamification science so far include improved product quality (Goomas, Smith, & Ludwig, 2011), decreased costs of healthcare (Pereira, Duarte, Rebelo, & Noriega, 2014),

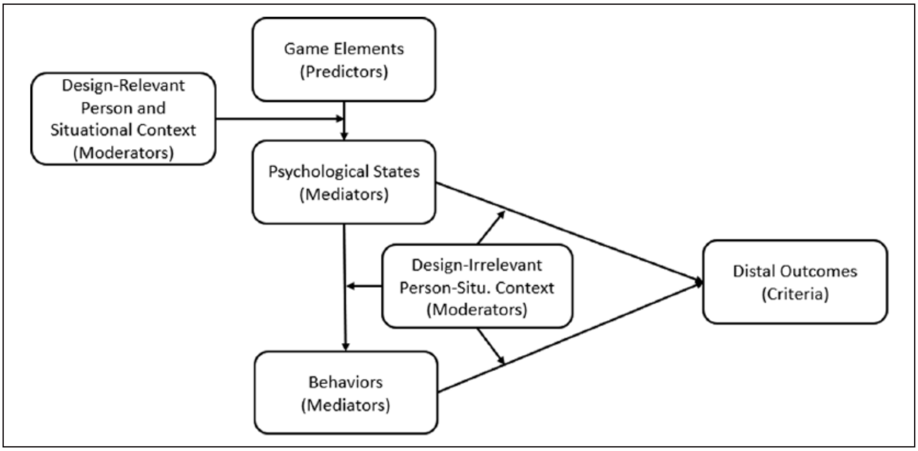

Figure 1. Theoretical causal relationships between constructs in gamification science.

reduced energy consumption (Gustafsson, Katzeff, & Bang, 2009), increased numbers of job applicants (Chow & Chapman, 2013) and many other domain-contingent outcomes. This relationship between the game elements and target outcomes is mediated by (i.e., causally contingent on) psychological and behavioral changes within an individual. For example, improved learning (e.g., Domínguez et al., 2013; Wilson et al., 2009) resulting from a school-wide gamification intervention would be an intermediate psychological goal in pursuit of a distal outcome, such as improved graduation rates. Importantly, psychological mediators are causally related to behavioral mediators, as described by Hamari et al. (2014), and behavioral mediators lead to changes in larger-scale valued criteria. For example, a leaderboard (predictor) might improve employee motivation to attend to training materials (psychological mediator), which in turn improves effort (behavioral mediator), which in turn improves learning (criterion). Each of these causal pathways may be moderated by other variables, in turn. Further, each of these relationships may be moderated such that the context in which gamification takes place, both in terms of the people experiencing it and the broader situation, may affect the direction and/or strength of the relationships between game elements, state changes, and target outcomes.

The second section of our treatment below describes meta-issues related to the successful use of these constructs in gamification science; specifically, implementation and broader methodological concerns must each be considered explicitly. First, the effectiveness of different design strategies must be understood in relation to the constructs involved. For example, once a target outcome has been selected and classes of game elements chosen in pursuit of that outcome, there are many different ways to engineer a particular system but theoretically one ideal way given the context in which that system will exist. Second, gamification science must concern itself with central aspects of research methodology, including psychometric measurement, experimental design, and generalizability, in order to maximize its trustworthiness and real-world value.

In summary, these four classes of constructs and two meta-issues are critical to successfully researching and implementing gamification and therefore to understanding how gamification is situated as a social science. Thus, we will describe each of these six major concerns in turn.

Predictor Constructs: The Science of Game Elements

Game elements are artifactual or social elements that are characteristic to games (Deterding et al., 2011); the interpretation of *characteristic*, however, is disputed amongst many subdisciplines of game science. Deterding et al. (2011), from a human-computer interaction perspective, broadly defined characteristic as "elements that are found in most (but not necessarily all) games, readily associated with games, and found to play a significant role in gameplay" (p. 12). Calvillo-Gámez, Cairns, and Cox (2015) further classified elements into the Core Elements of Gaming Experience (CEGE) framework, which consists of game play (i.e., the scenarios and rules of the game, and the environment), defined as how the game is presented to players. From the perspective of a game developer, game elements might be described using the Mechanics, Dynamics, and Aesthetics (MDA) framework in which mechanics are the basic rules or components of the game, dynamics are the behavior of the player with the mechanics, and aesthetics are the emotional responses of the player (Hunicke, LeBlanc, & Zubek, 2004). From a procedural content generation perspective (i.e., Hendrikx, Meijer, Van der Velden, & Iosup, 2013), game elements can be categorized into six classes, including game bits (e.g., textures, sounds buildings), game space (e.g., indoor maps, bodies of water), game systems (e.g., ecosystems, urban environment), game scenarios (e.g., puzzles, storyboards, story), game design (e.g., system design, world design), and derived content (e.g., news and broadcasts, leaderboards). Considering the differences in goals and focus for different fields, the variety in approaches and frameworks regarding game elements is appropriate.

In the context of gamification research, game elements are operationalized as causes of effects of interest in processes that have been gamified and can be borrowed from any of these frameworks. Before gamification, there are no (or fewer) game elements in some existing process, but after gamification, more game elements are present. Adding game elements is an incremental and versatile process, ranging from the simple addition of one element to the addition of a complex set of elements (Armstrong, Ferrell, Collmus, & Landers, 2016). The goal of gamification research regarding elements is thus to draw counterfactual causal conclusions (c.f., Pearl, 2000); specifically, scientific gamification researchers want to know what would happen differently if game elements were added to a situation where they were not already being used. To that end, game elements should generally be experimentally or quasi-experimentally manipulated, clearly defined, and used purposefully to induce outcomes of interest.

The specific list of game elements that can be manipulated this way is also an issue of dispute among gamification scholars. There are a few existing gamification-oriented game element taxonomies, most of which differ based on outcome of interest. For example, Bedwell, Pavlas, Heyne, Lazzara, and Salas (2012) developed a concise framework of game attributes for learning contexts that included elements such as adaptation, challenge, control, fantasy, and progress, and Landers (2014) incorporated these as the foundational constructs in his theory of gamified learning. Robinson and Bellotti (2013) put forth a preliminary taxonomy of gamification elements for varying anticipated commitment, a set of elements theorized to engage users with different levels of expected engagement with a computer-based service. In the context of enterprise-related gamification, game elements have been divided into core game play (e.g., survival, social, collection, trading) and key game mechanics (e.g., status, points, leaderboards; Raftopoulos, Walz, & Greuter, 2015). While none of these taxonomies are exhaustive beyond the context of their outcome of interest, they provide guidance for researchers to use when choosing which game elements to study or implement by focusing upon those elements most likely to prove valuable in achieving outcomes in the context for which they are designed. Thus, the development and testing of taxonomies like these, and perhaps eventually, a unified framework, is a central concern in gamification science. Importantly, such a framework may not ultimately categorize game elements by their implementation characteristics but by their anticipated effects; for example, if leaderboards and badges were ultimately to be found to affect identical mediators, there would be no reason to distinguish between these elements in terms of design.

Criterion Constructs: Outcomes of Interest From Gamified Processes

In the extant literature, the addition of game elements in non-game contexts has been related empirically and theoretically to a variety of distal outcomes ranging from commonly studied outcomes like student learning (e.g., Domínguez et al., 2013; Wilson et al., 2009) to more novel outcomes such as exercise and nutrition (Pereira et al., 2014) or energy consumption (Gustafsson et al., 2009). Ultimately, however, the final outcome of gamification is whatever change a gamification researcher-practitioner wishes to effect. Such changes typically meet an organizational need. For gamification designers in for-profit organizations, criteria are typically defined by the organization and are commonly tied to return on investment (Conley & Donaldson, 2015). In non-profit organizations, such as public research institutions, criteria vary more widely. For example, Yan, Conrad, Tourangeau, and Couper (2011) examined the effects of progress bars, a simple game element that provides the user with feedback about their progress, on survey completion. With the addition of a progress bar, Yan and colleagues found that respondents were less likely to quit the survey when they expected a short task rather than a longer task. In this scenario, progress bars were added by the researchers to increase intentions to complete the survey (psychological mediator), reduce respondent attrition (behavioral mediator) and thus ultimately increase the overall number of completed surveys (distal outcome), an important job-related goal for a professional researcher.

As exhibited, the ultimate intended outcomes of gamification are quite broadly defined, because they are contextualized to the application domains in which gamification is practiced. Gamification could theoretically be applied to an infinite number of domains, so application domain is the primary driver of a researcher's choice of criterion. Distal goals such as increased retention rates in MOOCs (e.g., Krause, Mogalle, Pohl & Williams, 2015), return on investment (e.g., Conley & Donaldson, 2015), civic engagement (e.g., Bista et al., 2014), and resilience against natural disasters (e.g., Horita et al., 2014), in the domains of education, business, governance, and healthcare, respectively, have all been empirically studied as outcomes of gamification interventions. Thus, a key aspect of gamification research involves clearly defining the goals of gamification interventions and empirically studying the success or failure of game elements to achieve those goals.

Mediator Constructs: How Gamification Results in Criterion Change

Gamification can only achieve its distal goals, such as increased organizational return on investment (Conley & Donaldson, 2015), by achieving proximal goals of change within individuals (Hamari et al., 2014). Fundamentally, gamification design succeeds based upon its effect on individuals, not organizations. These individual-level effects are aggregated to influence organization-relevant criteria via a variety of causal pathways; they cannot influence the organization as a distinct entity directly. Thus, understanding proximal changes in a target person's psychological states and the effect of those state changes on their behaviors is key to understanding when and why gamification creates distal change desired by organizational or other stakeholders.

Such changes can be understood through the lens of *mediation*. A *mediator* is defined as a variable that occurs causally between two others; specifically, the effect of a predictor on an outcome can be explained by the predictor's effect on the mediator and the mediator's subsequent effect on the outcome (Hayes & Preacher, 2010). Mediators can and often do occur in series. In Figure 1, such a series of mediators can be observed: game elements affect outcomes of interest only via indirect effects on other states. These direct and indirect relationships can be formally stated:

1. Game elements may have a causal, *direct effect* on psychological states.
2. Psychological states may have a causal, *direct effect* on behaviors.
3. Both psychological states and behaviors may have a causal, *direct effect* on outcomes.
4. Game elements may have a causal, *indirect effect* on target outcomes via the intermediary causal effect of psychological states.
5. Game elements may have a causal, *indirect effect* on target outcomes via the intermediary causal effects of psychological states on behaviors and behaviors on outcomes.

In gamification, researchers have examined a variety of mediating variables as explanatory links between game elements and outcomes of interest, although this research is still in initial stages in both gamification science and game science more broadly. For example, in a review of research on games and learning, Garris, Ahlers, and Driskell (2002) theorized that the relationship between game elements and learning outcomes was mediated by user judgments and user behavior. In support of this, Landers and Landers (2014) found in an empirical study that time spent on participating in a learning activity (behavior) mediated the relationship between the use of a leaderboard (predictor) and learning (outcome). Some commonly studied mediators in gamification studies are motivation/engagement (Denny, 2013; Downes-Le Guin, Baker, Mechling, & Ruylea, 2012), flow (Eickhoff, Harris, de Vries, & Srinivasan, 2012; Witt, Schneiner, & Robra-Bissantz, 2011), and enjoyment (Cheong, Cheong, & Filippou, 2013; Flatla, Gutwin, Nacke, Bateman, & Mandryk, 2011).

In most empirical gamification research studies, mediators are left unspecified and unmeasured. For example, although Landers and Landers (2014) identified the mediating effect of time-on-task between leaderboards and learning, there must also be some additional psychological change within those learners that brought about their increase in time-on-task. For example, the leaderboard may have helped them better recognize meaningful intermediate goals, increased their confidence in their work tasks, communicated to them the importance of completing goals, or any number of additional possibilities. Without measuring psychological mediators in addition to behavioral ones, this study supports the theory that learners behaved differently because of the game elements implemented but not why they behaved differently. Although measuring both types of mediators is not necessary for a study to draw meaningful conclusions, the most informative empirical gamification studies from gamification science will do so as relevant to the outcomes being targeted.

Moderator Constructs: Circumstances Under Which Gamification Is Successful

Although gamification may (or may not) be effective on average across typical intervention designs, there is incremental value in understanding more specifically when and for whom gamification is effective, a concept called *moderation*. A *moderator* is defined as a variable that affects the direction and/or strength of the relationship between a predictor and outcome (Baron & Kenny, 1986). Thus, the estimate of the effect of a given predictor on an outcome is conditional on the value of the moderator variable (Cohen, Cohen, West, & Aiken, 2003). A moderator variable may be either categorical (e.g., sex, race) or continuous (e.g., attitudes, personality) in nature, as well as either organismic (e.g., human perceptions, abilities) or situational (e.g., environmental conditions; Cohen et al., 2003). In Figure 1, moderators are captured as *person and situational context* to indicate how this class of variables can affect the strength all other relationships of interest in gamification. For example, Toril, Reales, and Ballesteros (2014) examined the effect of video game training on the cognitive performance of older adults, finding that the training effect was larger for those aged 71-80 than those aged 60-70. Thus, age was a moderator of the game training-performance relationship. Studying moderation effects like these allows scientists to investigate the generalizability and external validity of a given effect across different groups and contexts, which better informs practice (Fairchild & McQuillin, 2010).

Importantly, moderators can exert causal influences in two major ways within the causal diagram depicted in Figure 1, and the specific causal path involved influences how they should be interpreted. First, *design-relevant moderators* are moderators that influence the effectiveness of game elements on immediate, targeted psychological state changes. For example, if a gamification designer designs a leaderboard (predictor) to increase engagement (psychological state), attitudes towards leaderboards (design-relevant person context) may change the strength of that effect. If a person has negative attitudes towards leaderboards, the addition of leaderboards may in fact harm that person's engagement. Such moderators are labeled *design-relevant* because a gamification designer should try to anticipate such effects and design gamification interventions around them. If, for example, it is known that most users are already overexposed to leaderboards, leaderboards will probably not be an effective design choice.

In contrast, *design-irrelevant moderators* are those that alter the strength of other causal relationships further downstream of the immediate psychological state changes created by effective gamification. Inadequate consideration of such moderators can cause gamification to appear unsuccessful even if designed effectively to cause a targeted psychological state change. For example, consider a design context in which leaderboards (predictor) are intended to improve sales performance (criteria) by increasing motivation to sell (psychological state) and thus salesperson work intensity (behavior). The gamification designer can only affect the first relationship with gamification design. This relationship is the effect of leaderboards on motivation to sell, and the designer can influence it primarily by redesigning the leaderboards or implementing additional game elements. All other relationships are causally downstream. Given this, consider what would happen if a design-irrelevant situational moderator, such as a downturned economy, were to be introduced into this system. With an economic downturn, sales in general would drop; sales would become more difficult for all salespeople in this organization. In such a situation, gamification may still improve motivation to sell, but because of external forces (i.e., the economic moderator), motivation to sell may no longer lead to improved sales performance. A research study examining the effects of leaderboards among salespeople by measuring gamification's impact on sales alone might conclude that leaderboards were ineffective, whereas the gamification did in fact bring about the psychological change targeted. In short, the gamification was effective but appeared ineffective in terms of sales performance due to a design-irrelevant moderator. In such a situation, it would be incorrect to attribute this failure to the gamification itself.

Thus, the distinction between design-relevant and design-irrelevant moderators becomes critical when assessing the effectiveness of gamification interventions. If only causally downstream outcomes are used to assess the effectiveness of gamification interventions, researchers and practitioners could be misled. For example, if a study is conducted that fails to find a relationship between game elements (predictor) and learning (outcomes), it is unknown if this is because the game element failed to change a targeted psychological state, if that state failed to lead to learning, if that state failed to lead to behavioral change, if that behavioral change failed to lead to the outcome, or if any moderator of any of those pathways made the context of the study somehow unusual in comparison to the overall population of such effects. Because of this, greater causal distances between game elements and measured outcomes are associated with greater possibilities of confounds if the intermediary variables are not explicitly measured. Thus, measurement of all targeted mediators and careful consideration of both design-relevant and design-irrelevant moderators are necessary to effectively isolate the causal effects of interest to gamification science.

In terms of specific moderators of gamification intervention effectiveness, gender has been found in some studies to be design-relevant. Males and females tend to differ in their motivation to play and preferences for playing games (Greenberg, Sherry, Lachlan, Lucas, & Holstrom, 2010) such that females on average are more motivated by the social aspects of games whereas males on average are more motivated by elements of achievement (Williams, Consalvo, Caplan, & Yee, 2009), suggesting differential effectiveness of game elements on psychological state changes based upon the demographics of the group targeted by gamification. Because of this, achievement-oriented game elements, like competition, may more greatly impact outcomes of interest for males than for females (Shen, Liu, Santhanam, & Evans, 2016) whereas the converse may occur for social-oriented game elements (Koivisto & Hamari, 2014). Accordingly, gamification interventions designed to improve outcomes such as learning may need to account for the motivational and preferential differences in games and game elements across genders in order to maximize their effect rather than applying a one-size-fits-all approach. Importantly, such designs should consider gender norms as relevant to the locale in which the gamification will be used instead of relying on generalities like this one.

A variety of other person-level constructs may also moderate the effectiveness of particular gamification interventions, although it is important to distinguish between proxy variables and psychological constructs. Proxy variables, like gender, are not themselves causal but instead are correlated with constructs that are. For example, another proxy variable, age, has been found to moderate gamification effects such that older adults view game elements as less appropriate in serious contexts and also find them harder to use than younger adults (Koivisto & Hamari, 2014; Thiel, Reisinger, & Röderer, 2016). However, it is not age itself that likely causes this moderating effect but instead correlates of age, such as generational influences and physical well-being. In terms of psychological constructs, Landers and colleagues found that attitudes toward games and gamification and experience with games moderate the impacts of gamification interventions on learning (Armstrong & Landers, 2017; Landers & Armstrong, 2017; Landers & Callan, 2012). In general, once an effect of a proxy variable is found, research should be initiated to identify the true underlying cause of the observed moderation.

Aside from person-level moderators, different situations and broader contextual variables may also moderate the effectiveness of gamification interventions. For example, in a workplace training setting, shared perceptions within the organization of policies, practices, and procedures (Ostroff, Kinicki, & Muhammad, 2013) may moderate the effectiveness of gamified learning. If a trainee's coworkers and supervisor all perceive that gamified learning was useful, enjoyable, and effective, what could be labelled a positive climate for gamified learning, the trainee might be more likely to put effort into learning that had evidently been gamified, enhancing the effect of the game elements employed. Thus, gamified learning climate would be a design-relevant moderator in this example. In an empirical study, Mollick and Rothbard (2014) found that consent to the addition of games in the workplace was a vital moderator such that the addition of games increased positive affect except when consent was lacking, in which case the addition of game elements decreased positive affect. Klabbers (IN PRESS) described situational moderators like these in terms of "organizational configurations," but the concept is much broader than that, especially in the context of gamification.

Meta-Issue: Gamification as a Family of Design Methodologies

The specific techniques by which game elements are added to an existing process to make it seem game-like are a major focus of design researchers. One prominent research area within gamification science concerned with this is called gameful design. Gameful design can be defined as the process of "designing for gamefulness, typically by using game design elements" (Deterding et al., 2011, p. 3). Gameful design is a technique of gamification but is differentiated from the broader concept of gamification in that it involves pursuit of a specific design goal, that of designing a *gameful experience*, presumably a psychological mediator in terms of the model in Figure 1, whereas gamification more broadly can have any goal. For example, the addition of progress bars to non-game processes (c.f., Yan et al., 2011) is a gamification intervention *without gameful design* in that it is inspired by games but is not intended to make the survey to which the progress bars are added gameful. Instead, the only goal of such an intervention is to capitalize on the psychological effects of progress bars to change survey completion behavior. In some cases, gamification without gameful design can be legitimately criticized as manipulative or unethical (Rehn, 2016). For example, in managerial gamification, the addition of points, badges and leaderboards to increase control over currently discretionary employee behavior likely does not involve gameful design, although it may still be effective at directing employees to change their behavior.

Deterding et al. (2011) summarized the extant gameful design research literature into a framework which consists of underlying levels of gameful design elements. These levels are distinct given the difference in goals of each element, which range in level of abstraction. The first level of game design elements is *game interface design patterns*, which involves the addition of interaction-related design components to fix a known problem (Crumlish & Malone, 2009). Examples include designing badges, leaderboards, and levels. The second level, *game design patterns and mechanics*, consists of common parts of a game that reoccur such as time constraints, turns, or limited resources (Björk & Holopainen, 2005; Taylor, 2009). The third level, *game design principles and heuristics*, describes the approach to design using evaluative guidelines such as in the case of enduring play and clear goals (Isbister & Schaffer, 2008). The fourth level is *conceptual game models* and describes the use of design methodologies, such as MDA, CEGE, challenge, fantasy, curiosity, and game design atoms (Brathwaite & Schreiber, 2008; Calvillo-Gámez, Cairns, & Cox, 2010; Fullerton, 2008; Hunicke et al., 2004). Finally, the most abstract game design level is *game design method*, which refers to practices and processes specific to game design. This level includes, for example, playtesting and play centric design (Belman & Flanagan, 2010; Fullerton, 2008). These levels demonstrate the variety of approaches and goals of the gameful design process while highlighting the lack of consensus as to which are most effective and in what situations.

Any type of gamification design requires the choice of a specific approach, and these approaches vary widely in complexity and appropriateness for a given context. Although there is no agreed-upon method for such design (Bernhaupt, Isbister, & De Freitas, 2015), a few have been proposed. One method for gameful design, for example, proposes using skill atoms, which are the smallest differentiation in user's skills that is identifiable, as a design lens (Deterding, 2015). In this method, each skill atom is intended to create a feedback loop with the user until that user masters the goal of the skill atom, which in Deterding's (2015) case includes goals actions, tokens, feedback, a rule system, and challenge. Thus, the goal of each step in the design process is to facilitate changes in and progress toward achieving the designated goal. In contrast, Armstrong and Landers (2017) designed a narrative intervention for a training program using Thorndyke's (1977) approach to simple story development, integrating a developed setting, theme, and plot to create episodes, subgoals, and events in existing training content to make it seem more like a game. Although this process required a thorough development of narrative, it did not involve looking at skill atoms as described in Deterding's (2015) approach nor did it implement any game design methodologies described here. Nevertheless, Armstrong and Landers found that trainees were significantly more satisfied with the game fiction training than the original training with essentially equivalent declarative knowledge gains, demonstrating a successful change in outcome using a very simple design approach.

As the number of game elements in an existing system increases, the distinction between *gamified system* and *game* becomes increasingly muddied. Games are a structured type of play (Makedon, 1984), but play is only one potential design goal of gamification. Gamified systems may not involve play whatsoever whereas in the games literature, play and game tend to be used interchangeably (Klabbers, 2009). Makedon (1984) argued that games are a special form of play within specified rules and that play is tied to the player whereas a game is tied to its rules or elements, yet, as Klabbers (*in press*) observed, a game must be played to be a game, so they are inherently linked. This linkage does not necessarily exist in the context of gamification. Rather, users of gamified products may not have the opportunity to play even when interacting with a complete gamified system, and there may be no game rules to follow. Although a user may engage in play, a gamified product does not necessarily rely on play to achieve its intended outcome unless the gamification designer explicitly sets play as a goal.

Meta-Issue: Research Methods in Gamification Science

Across all gamification research, just as in the broader social sciences, high quality research methods are key to interpretable and generalizable research findings. Because of the substantial variation in both the extensiveness and quality of training in statistics and research methods across programs where any game science is taught, methodological rigor is a particularly salient challenge for gamification science. Fortunately, these techniques have a long history in social science, so the path ahead is straightforward. Specifically, the conclusions drawn in gamification science will be of much higher quality if researchers focus on three key goals of research design: psychometric measurement (i.e., test validity), experimental design (i.e., internal validity), and generalizability (i.e., external validity). These issues are a concern across game science broadly, but we will focus here on gamification science.

The first goal to increase scientific rigor in gamification science is to increase the use psychometrically reliable and valid measurement tools (i.e., test validity). Reliability refers to the degree to which a variable is measured consistently (Kimberlin & Winterstein, 2008) and can be expressed as a proportion. For example, a reliability estimate of .8 suggests that 80% of the observed variance in scores was caused by at least one underlying shared source; the remainder is essentially noise (i.e., *unsystematic measurement error*). The process of measuring and accounting for reliability can be conceptualized as the process of removing systematic measurement error from a given measurement tool, such as a survey. The most commonly reported estimate of reliability, coefficient alpha, refers to internal consistency reliability, the degree that items of a given scale are measuring the same construct (Cortina, 1993). In this regard, reliable survey measures tend to have at least three items[1] that measure the same construct and are correlated with each other at roughly the same magnitude. If survey measures cannot be trusted, neither can any conclusions drawn from them; for example, if a measure does not reliably assess task engagement yet task engagement is being tested as an outcome of a gamified system, that finding is not trustworthy.

In contrast, validity refers to the degree to which measurement measures the construct it is intended to. Reliability is a prerequisite for validity; by definition, unreliable measures are not valid. Whereas reliability is concerned with how consistently items are inter-related, validity is concerned with how well they measured a specified construct (i.e., the one that the scale is supposed to measure). Validity is a matter of degree, and no measure can ever be labeled valid. Instead, modern social science seeks to provide incremental validity evidence over time to develop an increasingly strong theoretical case that a measure is trustworthy. For example, criterion-related validity evidence is provided when a measured construct predicts variables on par with those established in prior theoretical and empirical studies. For example, if a cognitive ability test was gamified to make it more engaging or less intimidating, that test should still correlate with other traditional cognitive ability tests and outcomes of cognitive ability. It should also have an internal statistical structure consistent with its underlying measurement theory. If any of these tests failed, this could be interpreted as evidence that gamification harmed the test's ability to measure cognitive ability. The use of low quality, untested, measures is a common issue throughout game science broadly, and the impact of this cannot be overstated. As Korman (1974) wrote, "The point is not that adequate measurement is 'nice.' It is necessary, crucial, etc. Without it, we have nothing" (p. 194).

The second goal to increase rigor is to use experimental controls to manipulate study variables (i.e., internal validity). Internal validity describes the strength of inferences that can be drawn from a study given the way that its focal constructs were measured or manipulated. The most straightforward way to ensure a high degree of confidence is via random assignment in a between-subjects experimental design. In this design, there is high internal validity because it can be more easily concluded that the differing levels of the experimental variable caused any observed differences in the outcome of interest because nothing else differed between the experimental groups. In the context of gamification, it is recommended that researchers experimentally assign game elements, using experimental design to compare the effects of interventions with and without individual, targeted game elements in a purposeful and systematic fashion. Commonly, in gamification research there is a tendency to implement multiple game elements as a bundle. Unless the study is explicitly designed to examine the interactive effect of elements, this introduces a confound, or more formally, a threat to internal validity. It becomes impossible to conclude which game elements, combination of elements, or interaction between elements caused any observed change in outcomes. Practical takeaways from this type of study are limited. It is therefore recommended that gamification researchers isolate the effects of independent variables or explicitly model any expected interactions (Sailer, Hense, Mayr, & Mandl, 2017; Seaborn & Fels, 2014). Broadly, studies comparing gamification to no gamification without carefully isolating elements or meaningful element clusters are of limited theoretical value and should not be conducted.

The third goal to increase rigor is to design for replicability and to sample with cross-validation in mind (i.e., external validity). External validity describes the strength of the inference that results of a given study will generalize to other settings and samples. A key issue in external validity is sampling (i.e., asking if it can be expected that the effects observed in the study sample would be observed in the more general population). In many common games and gamification research scenarios, sampling is by convenience, which can impact the generalizability of results. In particular, snowball sampling on listservs is essentially never an acceptable sampling strategy (Landers & Behrend, 2015). In this approach, a researcher posts a general invitation to complete a study and ostensibly hopes that the sample they end up with represents a population they might be interested in. Unfortunately, if the people who received or responded to the email about gaming are different in some way relevant to the outcome being studied, then the results may not generalize beyond those respondents. In gamification research, if individuals who enjoy games are selected for or volunteer to be study participants, it is likely that they differ from the general population in some important variables associated with the outcome (e.g., interest in games, enjoyment from games, ability to learn new games or controls). The best way to strengthen external validity is to ensure that the study sample is heterogeneous and to replicate findings in other samples. Gamification researchers should make effort to use diverse samples when possible and to seek replication with other samples. One way to quickly replicate a new finding is with online panels such as Amazon Mechanical Turk, which is appropriate for many, but not all, research questions (Landers & Behrend, 2015).

Conclusion

In summary, we presented here three perspectives from which to refute Klabbers' (2018) dismissive claims about "what gamification is all about" (p. XXX). First, we described the philosophical foundations of gamification science. Specifically, gamification science operates within a post-positive epistemology, one in which scientists admit that it is only through the imperfect lens of human interpretation that underlying truths about gamification can be identified, but that this lens does not diminish the importance or validity of the questions being pursued. Ultimately, if we can help people or organizations reach their goals at a higher success rate despite uncertainty, we will have created a useful science. Second, we formally stated the purpose of gamification science and described how it can be identified as a subdiscipline within the broader field of game science. Specifically, it is similar in that it shares a research literature and game elements toolkit with many other subdisciplines, but it is unique in that it focuses on the design of gamification interventions for changing existing processes without necessarily creating a product that most would call a game. Third, we laid out a comprehensive research agenda for this field, highlighting the six major areas of concern for gamification science, unifying them all within one framework from which to ask research questions: predictor constructs, criterion constructs, mediator constructs, moderator constructs, design processes, and research methods. In doing so, we also provided numerous references to the sizable body of existing gamification science in a wide variety of contexts that are clearly distinct from the literature of other subdisciplines.

Although we expect significant overlap between these concerns and those of game science, gamification science's parent field, these arguments also reveal how gamification science may currently have a firmer footing in modern social science than this broader literature. Specifically, gamification interventions always have a change-related goal, whether that goal is explicitly stated or an implicit motivation of the designer. Gamification is intended to change specific outcomes in specific ways; designers want to use lessons from game science to *change human behavior*. This makes gamification more conceptually like other subdisciplines of game science in which games are used for a specific purpose, such as games for learning or persuasive games for instigating social change. Furthermore, gamification interventions are well-defined with known parameters, so specification of precise causal impacts of gamification interventions is much simpler than when hypothesizing effects caused by complete games. If a researcher wants to test if progress bars can effect change in a person, that researcher can randomly assign people to experience progress bars or the absence of progress bars and draw meaningful conclusion based upon statistical tests that should generalize to other contexts where people are considering adding progress bars. If a researcher believes that *violent video games* can effect change in a person, they have innumerable types and variations of such games to test, no single version of which is necessarily representative of violent video games, nor could it be. Beyond that, an immense variety of outcome constructs are potentially of interest, each with its own unique measurement concerns. Thus, construct specification and experimental design are much more easily managed in gamification science than in its parent field.

This distinction does not imply that other areas of game science are more important than gamification science or that gamification science is necessarily more scientific than those areas. Instead, it suggests only that scientific gamification research faces its own unique challenges related to but distinct from those other areas. In short, games science does not benefit from researcher in-fighting regarding which topics within which subdomains are worthy to be included. In an era when researcher credibility is openly questioned by the public, staking such claims only serves to further split an already fragmented field struggling to be heard. To make a difference in the world via the study of games, we must put such pedantry behind us.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Note

1. Increasing the number of items tends to increase reliability, but too many items can artificially inflate alpha as a function of inter-item correlation. See Cortina (1993).

References

Armstrong, M. B., Ferrell, J., Collmus, A. B., & Landers, R. N. (2016). Correcting misconceptions about gamification of assessment: More than SJTs and badges. *Industrial and Organizational Psychology*, *9*, 671-677. doi:10.1017/iop.2016.69

Armstrong, M. B., & Landers, R. N. (2017). An evaluation of gamified training: Using narrative to improve reactions and learning. *Simulation & Gaming*, *48*, 513-538. doi:1046878117703749

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173-1182. doi:10.1037/0022-3514.51.6.1173

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a taxonomy linking game attributes to learning: An empirical study. *Simulation & Gaming*, *43*(6), 729-760. doi:10.1177/1046878112439444

Belman, J., & Flanagan, M. (2010). Exploring the creative potential of values conscious design: Students' experiences with the values at play curriculum. *Eludamos. Journal for Computer Game Culture*, *4*, 57-67.

Bernhaupt, R., Isbister, K., & De Freitas, S. (2015). Introduction to this special issue on HCI and games. doi:10.1080/07370024.2015.1016573

Bista, S. K., Nepal, S., Paris, C., & Colineau, N. (2014). Gamification for online communities: A case study for delivering government services. *International Journal of Cooperative Information Systems, 23*(2), 1-25.

Björk, S., & Holopainen, J. (2005). *Patterns in game design*. Boston, MA: Charles River Media.

Bogost, I. (2011a, August 9). Gamification is bullshit. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2011/08/gamification-is-bullshit/243338/

Bogost, I. (2011b, May 3). Persuasive games: Exploitationware. *Gamasutra*. Retrieved from http://www.gamasutra.com/view/feature/6366/persuasive_games_exploitationware.php

Brathwaite, B., & Schreiber, I. (2008). *Challenges for game designers*. Rockland, MA: Charles River Media.

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2010). Assessing the core elements of the gaming experience. In *Evaluating user experience in games* (pp. 47-71). London, UK: Springer.

Calvillo-Gámez, E. H., Cairns, P., & Cox, A. L. (2015). Assessing the core elements of the gaming experience. In *Human–Computer interaction series game user experience evaluation* (pp. 37-62). doi:10.1007/978-3-319-15985-0_3

Cheong, C., Cheong, F., & Filippou, J. (2013, June). Quick Quiz: A Gamified Approach for Enhancing Learning. In *PACIS 2013 Proceedings* (pp. 206-219). Association for Information Systems.

Chow, S., & Chapman, D. (2013). Gamifying the employee recruitment process. In *Proceedings of the First International Conference on Gameful Design, Research, and Applications* (pp. 91-94). New York, NY: ACM.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). New York, NY: Routledge.

Conley, K., & Donaldson, C. (2014). Gamification: The measurement of benefits. In *Gamification in education and business* (pp. 673-688). doi:10.1007/978-3-319-10208-5_34.

Conley, K., & Donaldson, C. (2015). Gamification: The measurement of benefits. In *Gamification in education and business* (pp. 673-688). Cham, Switzerland: Springer.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98-104.

Crumlish, C., & Malone, E. (2009). *Designing social interfaces: Principles, patterns, and practices for improving the user experience*. Sebastopol, CA: O'Reilly Media.

Denny, P. (2013, April). The effect of virtual achievements on student engagement. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 763-772). New York, NY: ACM.

Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. *Human-Computer Interaction*, *30*, 294-335. Retrieved from http://dx.doi.org/10.1080/07370024.2014.993471

Deterding, S., Dixon, D., Khaled, R. & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In *MindTrek'11*. ACM. doi:10.1145/2181037.2181040

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380-392. doi:10.1016/j.compedu.2012.12.020

Downes-Le Guin, T., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research, 54*(5), 1-21.

Eickhoff, C., Harris, C. G., de Vries, A. P., & Srinivasan, P. (2012, August). Quality through flow and immersion: gamifying crowdsourced relevance assessments. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval* (pp. 871-880). New York, NY: ACM.

Fairchild, A. J., & McQuillin, S. D. (2010). Evaluating mediation and moderation effects in school psychology: A presentation of methods and review of current practice. *Journal of School Psychology*, *48*, 53-84. doi:10.1016/j.jsp.2009.09.001

Ferguson, C. J. (2007). The good, the bad, and the ugly: A meta-analytic review of positive and negative effects of violent video games. *Psychiatric Quarterly*, *78*, 309-316. doi:10.1007/s11126-007-9056-9

Flatla, D. R., Gutwin, C., Nacke, L. E., Bateman, S., & Mandryk, R. L. (2011). Calibration games. In *Proceedings of the 24th annual ACM symposium on user interface software and technology - UIST '11*. doi:10.1145/2047196.2047248

Fullerton, T. (2008). *Game design workshop: A playcentric approach to creating innovative games.* Amsterdam, NL: Morgan Kaufmann

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*, 441-467. doi:10.1177/1046878102238607

Goomas, D. T., Smith, S. M., & Ludwig, T. D. (2011). Business activity monitoring: Real-time group goals and feedback using an overhead scoreboard in a distribution center. *Journal of Organizational Behavior Management*, *21*, 196-209.

Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holstrom, A. (2010). Orientation to video games among gender and age groups. *Simulation & Gaming*, *41*, 238-259. doi:10.1177/1046878108319930

Gustafsson, A., Katzeff, C., & Bang, M. (2009). Evaluation of a pervasive game for domestic energy engagement among teenagers. *Computers in Entertainment (CIE)*, *7*, 54. doi:10.1145/1658866.1658873

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469-478. Retrieved from http://dx.doi.org/10.1016/j.chb.2015.03.036

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *2014 47th Hawaii International Conference on System Sciences (HICSS)* (pp. 3025-3034). IEEE. doi:10.1109/HICSS.2014.377

Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research*, *45*, 627-660. doi:10.1080/00273171.2010.498290

Hendrikx, M., Meijer, S., Van Der Velden, J., & Iosup, A. (2013). Procedural content generation for games: A survey. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *9*, 1. doi:10.1145/2422956.2422957

Horita, F. E., Assis, L. F. F. G., Castanhari, R. E., Isotani, S., Cruz, W. M., & de Albuquerque, J. P. (2014). *A gamification-based social collaborative architecture to increase resilience against natural disasters*. *Simpósio Brasileiro de Sistemas de Informação (SBSI)*, May 2014. Londrina, Brazil.

Hunicke, R., LeBlanc, M., & Zubek, R. (2004, July). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI* (Vol. 4, No. 1). Game Developers Conference, San Jose 2001-2004. Retrieved from https://www.researchgate.net/publication/228884866_MDA_A_Formal_Approach_to_Game_Design_and_Game_Research

Isbister, K., & Schaffer, N. (2008). Heuristic evaluation of games. In *Game usability* (pp. 79-89). doi:10.1016/b978-0-12-374447-0.00006-8

Kanfer, F. H. (1990). The scientist-practitioner connection: A bridge in need of constant attention. *Professional Psychology: Research and Practice*, *21*, 264-270.

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, *65*, 2276-2284. doi:10.2146/ajhp070364

Klabbers, J. H. (2009). *The magic circle: Principles of gaming & simulation*. Rotterdam, The Netherlands: Sense Publishers.

Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, *35*, 179-188. doi:10.1016/j.chb.2014.03.007

Korman, M. (1974). National conference on levels and patterns of professional training in psychology: The major themes. *American Psychologist, 29*(6), 441-449. doi:10.1037/h0036469

Krause, M., Mogalle, M., Pohl, H., & Williams, J. J. (2015). A playful game changer: Fostering student retention in online education with social gamification. In *Proceedings of the Second ACM Conference on Learning @ Scale* (pp. 95-102). New York, NY: ACM.

Landers, R. N. & Callan, R. C. (2012). Training evaluation in virtual worlds: Development of a model. *Journal of Virtual Worlds Research, 5*(3), 1-20.

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming*, *45*, 752-768. doi:10.1177/1046878114563660

Landers, R. N., & Armstrong, M. B. (2017). Enhancing instructional outcomes with gamification: An empirical test of the Technology-Enhanced Training Effectiveness Model. *Computers in Human Behavior*, *71*, 499-507. doi:10.1016/j.chb.2015.07.031

Landers, R. N., & Behrend, T. S. (2015). An inconvenient truth: Arbitrary distinctions between organizations, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology*, *8*, 142-164. doi:10.1017/iop.2015.13

Landers, R. N., & Landers, A. K. (2014). An empirical test of the theory of gamified learning: The effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, *45*, 769-785. doi:10.1177/1046878114563662

Makedon, A. (1984). Playful gaming. *Simulation & Gaming*, *15*, 25-64. doi:10.1177/0037550084151003

Mekler, E. D., Bruhlmann, 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. doi:10.1016/j.chb.2015.08.048

Mollick, E. R., & Rothbard, N. (2014). Mandatory fun: Consent, gamification and the impact of games at work. *The Wharton School Research Paper Series*. doi:10.2139/ssrn.2277103

Ostroff, C., Kinicki, A. J., & Muhammad, R. S. (2013). Organizational culture and climate. In I. B. Weiner. (Ed.), *Handbook of psychology* (2nd ed., pp. 643-676). Hoboken, NJ: John Wiley & Sons.

Pearl, J. (2000). *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press. doi:10.1017/cbo9780511803161

Pereira, P., Duarte, E., Rebelo, F., & Noriega, P. (2014). A review of gamification for health-related contexts. In *Design, user experience, and usability. User experience design for diverse interaction platforms and environments lecture notes in computer science* (pp. 742-753). doi:10.1007/978-3-319-07626-3_70

Popper, K. (1963). *Conjectures and refutations: The growth of scientific knowledge*. London, UK: Routledge.

Pyky, R., Koivumaa-Honkanen, H., Leinonen, A.-M., Ahola, R. Hirvonen, N., . . .Ikaheimo, T. (2017). Effect of tailored, gamified, mobile physical activity intervention on life satisfaction and self-rated health in young adolescent men: A population-based, randomized controlled trial (MOPO study). *Computers in Human Behavior*, *72*, 13-22. doi:10.1016/j.chb.2017.02.032

Raftopoulos, M., Walz, S., & Greuter, S. (2015). How enterprises play: Towards a taxonomy for enterprise gamification. In Staffan Björk, Annika Waern, Mathias Fuchs (Eds.), *DiGRA 2015: Diversity of play* (pp. 1-17). Lüneburg, Germany.

Rehn, A. (2016). Gamification as ideological practice: On play, games and the modding of management. In M. Dymek & P. Zackariasson. (Eds.), *The business of gamification: A critical analysis*, pp. 62-78. New York, NY: Routledge.

Robinson, D., & Bellotti, V. (2013, April). A preliminary taxonomy of gamification elements for varying anticipated commitment. In *Procceedings of ACM CHI 2013 Workshop on Designing Gamification: Creating gameful and playful experiences*. New York, NY: ACM.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371-380. doi:10.1016/j.chb.2016.12.033

Seaborn, K., & Fels, D. I. (2014). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*. doi:10.1016/j.ijhcs.2014.09.006

Shaker, N., Yannakakis, G., & Togelius, J. (2010). Towards automatic personalized content generation for platform games. In G. M. Youngblood & V. Bulitko. (Eds.), *Proceedings of the sixth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (pp. 63-68). Menlo Park, CA: The AAAI Press.

Shen, W.-C. M., Liu, D., Santhanam, R., & Evans, D. A. (2016, June). *Gamified technology-mediated learning: The role of individual differences*. *Pacific Asia Conference on Information Systems (PACIS) 2016 Proceedings*, 47. Retrieved from http://aisel.aisnet.org/pacis2016/47

Stanculescu, L. C., Bozzon, A., Sips, R.-J., & Houben, G.-J. (2016). Work and play: An experiment in enterprise gamification. In *CSCW '16* (pp. 346-358). New York, NY: ACM. doi:10.1145/2818048.2820061

Taylor, T. L. (2009). The assemblage of play. *Games and Culture*, *4*, 331-339. doi:10.1177/1555412009343576

Thiel, S.-K., Reisinger, M. & Röderer, K. (2016, December). *"I'm too old for this!": Influence of age on perception of gamified public participation. Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia*, Rovaniemi, Finland. doi:10.1145/3012709.3016073

Thom, J., Millen, D., & DiMicco, J. (2012). Removing gamification from an enterprise SNS. In *Proceedings of the ACM 2012 Conference on Computer Supported Collaborative Work (CSCW '12)* (pp. 1067-1070). New York, NY: ACM. doi:10.1145/2145204.2145362

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*, 77-110. doi:10.1016/0010-0285(77)90005-6

Toril, P., Reales, J. M. & Ballesteros, S. (2014). Video game training enhances cognition of older adults: A meta-analytic study. *Psychology and Aging*, *29*, 706-716. doi:10.1037/a0037507

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, *14*(4), 490-495. doi:10.5465/amr.2011.55662499

Williams, D., Consalvo, M., Caplan, S., & Yee, N. (2009). Looking for gender: Gender roles and behaviors among online gamers. *Journal of Communication*, *59*, 700-725. doi:10.1111/j.1460-2466.2009.01453.x

Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J., …Conkey, C. (2009). Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & Gaming*, *40*, 217-266. doi:10.1177/1046878108321866

Witt, M., Scheiner, C., & Robra-Bissantz, S. (2011). Gamification of online idea competitions: Insights from an explorative case. *Informatik schafft Communities, 192*, 49-64.

Yan, T., Conrad, F. G., Tourangeau, R., & Couper, M. P. (2011). Should I stay or should I go: The effects of progress feedback, promised task duration, and length of questionnaire on completing web surveys. *International Journal of Public Opinion Research*, *23*, 131-147. doi:10.1093/ijpor/edq046

Author biographies

Richard N. Landers is an Associate Professor of Industrial-Organizational Psychology at the University of Minnesota and holder of the John P. Campbell Distinguished Professorship in Industrial & Organizational Psychology. His research program concerns the use of innovative technologies in assessment, employee selection, adult learning, and research methods.

Contact: rnlanders@tntlab.org

Elena M. Auer is an Industrial-Organizational psychology doctoral student. Her research interests include the use of innovative technology in employee selection and methodology including machine learning, big data, natural language processing, automatic interview/resume scoring, gamification, and unproctored internet testing.

Contact: eauer001@odu.edu

Michael B. Armstrong is a doctoral candidate in industrial-organizational psychology. His research focuses on technology in employee assessment and learning. He has published several papers and chapters on gamification in these contexts while also working as a gamification consultant.

Contact: marms018@odu.edu

Andrew B. Collmus is data scientist at Flex and doctoral student in industrial-organizational psychology at Old Dominion University. He has published several papers and chapters on gamification. His current work revolves around using machine learning to improve human resource processes.

Contact: acoll042@odu.edu