Expertise and Engagement: Re-Designing Citizen Science Games With Players' Minds in Mind

Josh Aaron Miller
Northeastern University
Boston, Massachusetts
miller.josh@husky.neu.edu

Uttkarsh Narayan
Northeastern University
Boston, Massachusetts
narayan.u@husky.neu.edu

Matthew Hantsbarger
Northeastern University
Boston, Massachusetts
m.hantsbarger@northeastern.edu

Seth Cooper
Northeastern University
Boston, Massachusetts
se.cooper@northeastern.edu

Magy Seif El-Nasr
Northeastern University
Boston, Massachusetts
m.seifel-nasr@northeastern.edu

ABSTRACT

Many studies have already shown that games can be a useful tool to make boring or difficult tasks more engaging. However, with serious game design being a relatively nascent field, such experiences can still be hard to learn and not very motivating. In this paper, we explore the use of learning and motivation frameworks to improve player experience in the well-known citizen science game Foldit. Using Cognitive Load Theory (CLT) and Self Determination Theory (SDT), we developed six interface and mechanical changes to the tutorial levels in Foldit designed to increase engagement and retention. We tested these features with new players of Foldit and collected both behavioral data, using game metrics, and prior experience data, using self-report measures. This study offers three major contributions: (1) we document the process of operationalizing CLT and SDT as new game features, a unique methodology not used in game design previously; (2) the user interface, specifically the level selection screen, significantly impacts how players progress through the game; and (3) a player's expertise, whether from prior domain knowledge or prior gaming experience, increases their engagement. We discuss both implications of these findings as well as how these implementations can generalize to other designs.

CCS CONCEPTS

• **Human-centered computing** → **User centered design**; *HCI theory, concepts and models*; *Empirical studies in HCI*.

KEYWORDS

Game-based learning, serious games, citizen science, motivation, personalization

ACM Reference Format:

Josh Aaron Miller, Uttkarsh Narayan, Matthew Hantsbarger, Seth Cooper, and Magy Seif El-Nasr. 2019. Expertise and Engagement: Re-Designing

Citizen Science Games With Players' Minds in Mind. In *The Fourteenth International Conference on the Foundations of Digital Games (FDG '19)*, August 26–30, 2019, San Luis Obispo, CA, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3337722.3337735>

1 INTRODUCTION

Video games are rapidly becoming popular across mainstream media. The Entertainment Software Association reported in 2017 that at least one person in 65% of U.S. households play video games 3 or more hours each week, with an average of 1.7 gamers per game-playing household [57]. With such a massive audience seeking entertainment, games are an excellent tool for engagement.

This is why many researchers have started using games as a platform for practical applications. *Serious games*, which include transformational games [10] and human computation games [50], can take advantage of the countless hours already spent on gaming to train users in the workplace, supplement formal teaching methods, and even crowdsource empirical data through *citizen science*, an increasingly common practice in which the general public contributes to the collection, processing, or analyzing of scientific data [47]. However, many of these games require users to voluntarily play them, and simply existing as a playful app doesn't guarantee that the software will hold the players' attention for long periods of time.

For these games to have a meaningful impact, it is crucial to understand and improve player engagement. By having a large and engaged audience, these games can successfully accomplish their intended purpose, whether that is user education or citizen science. But, importantly, this must be done without sacrificing the critical elements of the game, such as learning components or data collection.

Previous research on improving engagement with games has involved the application of two cognitive design frameworks: **Self Determination Theory** (SDT) and **Cognitive Load Theory** (CLT). SDT identifies key factors that are intrinsically motivating to players, while CLT focuses on minimizing mental load, allowing better focus on the game. On their own, CLT is most often applied to formal learning settings, while SDT has seen a range of applications including games. Yet, these two theories combined have not been tested thoroughly in games, and especially not in citizen science games. Because these games deal with complex topics, such as protein folding [6] and DNA sequence alignment [31], they often

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

FDG '19, August 26–30, 2019, San Luis Obispo, CA, USA

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-7217-6/19/08...\$15.00

<https://doi.org/10.1145/3337722.3337735>

require a great amount of mental demand. Moreover, citizen science projects struggle with “drop-outs” [18], suggesting that some players can often feel a lack of intrinsic motivation. However, it remains unknown how well these theories extend beyond their standard usage to apply to the problem of engagement with citizen science games.

In this paper, we use these theories in a more critical environment than previous research, applying and operationalizing them to *Foldit*, a citizen science puzzle game about folding proteins. Understanding the game requires a great amount of mental effort, and most of the players’ motivations come from a love for science rather than intrinsically from the game itself [11, 29]. Our hypothesis is that the combination of SDT and CLT to a citizen science game will improve engagement, since citizen science games require both learning and motivation.

In order to test this hypothesis, we operationalized SDT and CLT to create six experimental features which we tested with new players of *Foldit*. These features are: (1) reducing perceptual clutter (*Minimalist UI*), (2) imposing a minimal-load order on tasks (*Ordered Levels*), (3) providing meaningful choice in task order (*Branching Levels*), (4) encouraging curiosity through detail (*Science Info*), (5) teaching instead of telling (*Strategic Instructions*), and (6) providing context-insensitive help (*Help Panel*).

After implementing and deploying these features to new players of *Foldit*, we collected both game metrics data and self-report survey responses on demographic and psychographic information. Analysis of the game metrics data revealed two findings: first, that *Branching Levels* reduced level re-completions, suggesting that the UI discouraged players from returning to old levels. Second, the survey results identified that player expertise had a significant impact on engagement, as suggested by an increase in level completions and number of sessions played. Regardless of whether this expertise came from prior domain knowledge or gaming experience, both kinds of experience correlated with a significant rise in engagement with the game; conversely, not having this expertise correlated with below average engagement. This finding agrees with some of the more recent research that personalization is key to properly operationalize theories of learning and gamification [24, 44, 63].

This study offers three major contributions. First, we document the process of operationalizing SDT and CLT as concrete design objectives and then creating and implementing features into a citizen science game based on these design objectives. This study uniquely applies these theories toward new game features instead of using them to measure an existing game design, and we describe both the generalizable process as well as the difficulties encountered and how this impacts future citizen science game design. Second, we found that user interfaces, particularly our level selection screen, can affect how players progress through the game. In our case, by guiding players toward the next incomplete level, our new selection screen reduced level re-completions. Lastly, we present strong evidence that expertise, either in gaming or domain knowledge, correlates with game engagement, perhaps because both expertise and engagement are driven by the same underlying motivation.

2 BACKGROUND

In this section, we review the relevant ways in which players engage with a game. First, the player must be intrinsically motivated to engage; then they must be taught the relevant skills needed to play; and finally the gameplay must match their needs and expectations, which covers a range of design considerations from difficulty and rewards to choice and purpose. If these steps to engagement are successful, the players may enter a *flow state*. Every step of this process affects how much the player engages with the game.

2.1 Self Determination Theory

Much of our understanding of player motivation has come from Self Determination Theory (SDT) and its sub-theory, Cognitive Evaluation Theory [12, 54]. Namely, video games provide intrinsic motivation, rather than extrinsic motivation, by giving the players **competence** in the game mechanics, **autonomy** in the execution of their actions, and (to a lesser extent [13]) providing a sense of **relatedness** to others [14, 17, 48, 53]. The theory suggests that it is this satisfaction of psychological needs which makes video games more intrinsically engaging than other activities. In adapting SDT to video games, researchers developed the Player Experience of Need Satisfaction (PENS) model [52]. This model focuses on executing the principles of SDT in games, such as making the controls as intuitive as possible so that the player can feel competently in control, or providing meaningful choices for the player to clearly demonstrate autonomy.

However, before a player can demonstrate competence, they must first learn the game mechanics enough to understand and master them. This can be difficult, since by their nature serious games often involve more mentally challenging tasks than other games. Because of this, the road to competence is barred by the player’s limited *cognitive load* capacity—the player can handle only so much information at once. Before competence is possible, the player must learn the game’s rules, but this process can be made easier if the game’s design minimizes cognitive load. That is, in this conceptual model of player engagement, we use Cognitive Load Theory to establish better competence, then use Self Determination Theory as the primary model of motivation.

2.2 Cognitive Load Theory

Cognitive Load Theory (CLT) is a robust framework for describing how learning and processing are occurring. This makes it a useful tool for game design since it describes what mental difficulties the player will experience and how to alleviate those difficulties. Cognitive load is pivotal to player engagement because overloading the player mentally will cause frustration and a general negative affect, which can cause some players to disengage or otherwise stop playing [43, 58].

Learning, according to CLT, is the process of chunking individual elements into a *cognitive schema*. Qiao et al. [49] define cognitive schemata as highly organized and domain-specific knowledge which reduce element interactivity. This is ultimately what differentiates experts from novices for that subject. A schema is stored in long-term memory and represents the learned concept wholly as one element. Schemata also have faster retrieval speeds than their component elements, require fewer retrieval cues, and can be

automatically processed without additional resources or cognitive load [34, 49]. Consequently, the goal of learning is to form cognitive schemata. Barriers to this process are what create the difficulties that we are trying to remove in order to reduce cognitive load and improve learning. Cognitive load, which is perhaps best imagined as the demand on the learner's resources, is itself split into three categories: extraneous, intrinsic, and germane.

Extraneous cognitive load represents the inefficiencies in the method of presenting material [35]. Specifically, if a presented element cannot contribute to the formation of cognitive schemata, its presentation adds extraneous load. For example, if this sentence does not help you understand CLT, then it adds extraneous cognitive load by forcing you to process non-essential information.

Extraneous load often results in decreased performance, as shown in an experiment by Muth [40]. In this experiment, students who were given extraneous information on mathematical word problems performed worse than students with either no extraneous information or students who were told before the experiment that extraneous information would be present. It follows, then, that most uses of CLT have historically emphasized the reduction of extraneous load, rather than trying to manipulate the other two categories [62]. Extraneous load can take many forms, ranging from jargon to poor arrangement of material (e.g. from complex to simple rather than reversed). However, in some cases the "jargon" is central to the material, and other times not. (Imagine, for example, the word 'importunate' used in the context of a newspaper article, where another word would suffice, as opposed to the context of a vocabulary lesson, where the word itself is the critical information.) In this way, identical elements may be intrinsic or extraneous based on the learner's needs [62].

There are two other parts to cognitive load in CLT: **intrinsic load**, which represents the inherent difficulty of the interactions between elements, and **germane load**, which represents the cost of learning (i.e., chunking the interacting elements into a cognitive schema) [35, 49, 62]. However, most applications of CLT focus on extraneous load since it is easier and more worthwhile to minimize and because intrinsic and germane loads are linked to the learning process itself [62]. We believe that this precedent applies to the current study as well because intrinsic and germane loads can be affected only by inherently changing the nature of the material to be learned (in this case, the game mechanics and biochemistry), which we do not want to manipulate.

2.3 Usage of SDT and CLT

Although SDT and CLT have been thoroughly studied as frameworks (e.g., [52, 62]), most applications of these frameworks to games either don't make changes directly informed by the theories or don't empirically validate the approach. For example, many gamification studies use SDT to inform their general principles [59] such as by justifying the existing design elements of points, badges, and leaderboards [42]. Some studies, such as Mekler et al. [38] and Sailer et al. [56], empirically test these design elements for efficacy based on the theory, though results have been mixed. Entertainment games sometimes apply SDT as a lens to guide design, such as in Shadow of Mordor's Nemesis system [22]. However,

SDT has not been thoroughly studied in directly informing domain-specific mechanics, with the exception of two studies. In the health game *Spa Play* [17], developers used SDT to influence goal setting (autonomy), social connectivity (relatedness), and feedback (competence) among other gamification strategies. In a study by Peng et al. [46], researchers tested features intended to support autonomy and competence in an exergame and successfully increased need satisfaction. In summary, some preliminary work applied SDT in the forms of (1) a lens to guide designing entertainment games; (2) a justification for generic gamification elements; and (3) a design principle to create domain-specific features in serious games for health. The current research extends these findings by creating domain-specific features in a citizen science game which involves more learning than previous applications of SDT.

With respect to CLT, only a few studies have considered the importance of managing cognitive load in the context of game-based learning [33], and most work in this field measures the impact of game features on extraneous cognitive load and learning outcomes (e.g. [26]).

For both SDT and CLT, previous work with these frameworks uses them either theoretically or as a lens to interpret measurements of learning and engagement, such as the NASA Task Load Index (NASA-TLX) [25], the Intrinsic Motivation Inventory (IMI) [55], or through coding interview transcripts [17].

By considering CLT during the development cycle, games can be designed to minimize extraneous load, thus limiting the learner's frustration and enabling engagement. Minimizing cognitive load in this way synergizes with the research on Self Determination Theory (SDT), since reducing cognitive load inherently makes the game easier to use, which in turn promotes competence, one of the key components of SDT. In fact, these two frameworks have been combined previously to great success (see, for example, these theories applied to an online medical course [23]). Both CLT and SDT together inform a holistic view of game design for improving player engagement. By focusing on reducing the player's mental demand and catering to their psychological needs, serious games can more easily and reliably put players in a flow state that they enjoy and engage with, satisfying the purpose of the game.

Yet, no research has observed the effects of operationalizing both theories in a game as this study does. Rather, the concepts are typically used for measuring the impact of one change instead of tailoring the game experience to follow these guidelines of learning and motivation. Our study uses the unique methodology of re-designing features to minimize extraneous load and maximize engagement, rather than using SDT and CLT as measurements of an existing design. We chose this method of intervention to explore the potential effectiveness of this approach, hypothesizing that designing with these frameworks in mind would produce more theoretically-driven (therefore more reliable) results than only measuring the design's effectiveness after production.

2.4 Personalized Game Design

Rather than applying these concepts in broad strokes, recent research (e.g., [45, 47, 58]) has discovered that different players, and different game genres, engage differently with certain game elements.

Notably, players seek different goals at different levels of game expertise. Park et al. [45] mined player data of a popular massively multiplayer online role-playing game (MMORPG) and discovered that players early in the game focus both on achievement and social interaction. That is, for a player to succeed, it is essential that they progress in the game by learning the game mechanics, demonstrating their understanding, and developing relationships with other players. Next, when a player is more advanced in the game, their focus turns entirely toward achievement. In this stage, social interaction is less important, and players seek to master the game. Finally, when a player has reached the maximum level, this finding is reversed; achievement is no longer relevant, and a player's engagement in the game becomes almost entirely social.

In a more genre-agnostic report, Schoenau-Fog [58] conducted a survey in 2011 to determine what categories were most important to players for their subjective engagement. The highest rated categories included achievement (both progression and completion), socializing, audiovisual sensation or stimulation, and experiencing the story. This report validates previous evidence that achievement and socialization are important for player engagement, and adds that players are motivated also by the surface ludic elements, such as art, music, and narrative. In applying these findings to the current study, we surveyed players on their expertise and found that, even in citizen science games, player expertise affects the type and magnitude of their engagement.

2.5 Flow

Engagement with games has been studied thoroughly and goes by many names, including flow, GameFlow, presence, immersion, motivation, enjoyment, and fun [3, 7, 9, 58, 61]. Dependent on the type of game, even more factors are involved, such as suspense in the case of match-based games [36] or body image in the case of health games [16]. While a complete review of these terms would be superfluous, understanding the main theme is crucial for the current research. Because Foldit is a game focused on creative solutions to complex problems, it relies very heavily on the flow state of interaction, since the solution to a puzzle (that “aha!” moment) requires focus and concentration.

A flow state is achieved by matching the challenge of the game to the skill of the player, providing a sense of control, removing distractions and fear of failure, and providing direct and immediate feedback and clear goals [4, 8, 9]. This leads to concentration on the task at hand, a loss of self-consciousness, and an altered sense of time. Once this change to consciousness occurs, players are immersed, or feeling present in the game despite only controlling it through technology [60].

A game's design significantly affects how easily the player can achieve this flow state. Notably, the sense of control and removal of distractions are partially dependent on the game's user interface (UI). Research suggests that a game's perceived ease-of-use is a key factor in both choosing to play a game and achieving a flow state within it [27]. Moreover, it should be noted that good game design and ease-of-use are not entirely independent from each other. For example, a game with clear goals and immediate feedback will likely be perceived as easier to use than a game without these elements. Conversely, a game with unresponsive controls and a cluttered UI

will challenge the player's sense of control and their attempts to remove distractions, further separating them from their flow and immersion. Games that can be adjusted to skill level and provide incremental levels of difficulty lead to increased engagement in players [2]. In this way, smart game design develops flow, and flow in turn develops player engagement, which finally allows serious games to accomplish their intended purpose.

There are many similarities between good game design and engaged learning. Dickey [15] notes that they share focused goals, challenging tasks, clear and compelling standards, protection from adverse consequences for initial failures, affirmation of performance, affiliation with others, novelty and variety, choice, and authenticity. By these similarities, flow is optimal for both games and learning, making it an ideal state of play for a serious game.

2.6 Citizen Science

Yet, it is unknown whether SDT and CLT, which have been examined in commercial games, can be similarly mapped to the design of serious games. In this study, we narrow that question to the scope of citizen science games, which we find have *separate* motivations than traditional games. To summarize, several surveys of citizen science [11, 29, 30] found that players play for the science, not the game. Their initial motivations do not depend on the game itself, but rather on their previous interest in science (and the specific research topic), curiosity, and desire to contribute to research [29, 30]. Furthermore, the players' continued engagement with the game is dependent on recognition and feedback for their contributions, both from the game and its developers. That is, the players want to feel rewarded for their contributions and know that they are making a difference. Continued interest then relies on enjoying the task itself and proper pacing [30]. In addition, teamwork in citizen science games motivates the players, not by competing against one another, but by creating subgroups that each have a clearly defined goal which contributes to the overall project of the citizen science game [29]. These findings confirm that socialization remains important for engagement, but rather than achievement being individualistic, players of citizen science games seek achievement as a community. In her 2015 survey of Foldit players, Curtis added that citizen science players are also motivated by interaction with others in the community and an intellectual challenge [11]. Jennett added to this a year later that players also hope to share the same goals and values as the project and enjoy helping others, learning new information, and feeling like part of a team [30].

Some recent research has also explored how players gain expertise in Foldit. In 2017, researchers showed that beginners rely more on trial and error, learning only from the game tutorial and being unsure why consequences were happening, whereas experts relied more on manual tuning, external information such as paratexts, and had a deeper understanding of why consequences were happening [47].

3 METHODOLOGY

3.1 Theory Adaptation

Foldit is a citizen science puzzle game in which the player attempts to discover the best fold of a protein [6]. Because Foldit is a complex game, it is greatly improved by its tutorial; in fact, previous work

on its tutorial increased play time by as much as 29% and player progress by as much as 75% [1]. However, because the design of the tutorial does not take into account the principles of SDT or CLT, this game is an excellent candidate for testing these frameworks.

This study demonstrates a unique methodology: whereas most research citing SDT and CLT use them to measure outcomes of an unrelated variable, here the frameworks are the variables themselves. We define one possible operationalization of these frameworks and then implement features based on our operationalized definitions of these theories. In this way, we adapt the abstract theories into a concrete representation within Foldit and measure the behavioral outcomes of this representation.

This section describes not only how we operationalized SDT and CLT within Foldit, but also how to generally adapt SDT and CLT into the design of your own game, regardless of its genre, purpose, or educational value.

There were six operationalized changes: *Minimalist UI*, *Ordered Levels*, *Branching Levels*, *Science Info*, *Strategic Instructions*, and *Help Panel*. The first two, *Minimalist UI* and *Ordered Levels*, were designed to minimize cognitive load. *Branching Levels* and *Science Info* were designed to increase autonomy and motivation respectively, and *Strategic Instructions* and *Help Panel* were designed to improve competence. We confirmed that these features implemented the intended design changes through semi-structured interviews with two experts: an expert in human-centered design and Self Determination Theory, and an expert player of Foldit. These experts reported after two iterations that the designed features meet our operationalized definitions and impact player experience as intended.

Below, we review how each concept of CLT and SDT were operationally defined. Then we expand on the creation of each feature and how our operational definition could generalize for implementation into other games.

For this study, we operationally define extraneous load as perceptual clutter, irrelevant information, and introducing complex ideas before simpler ones. Autonomy is operationally defined as perceived choice by the player, specifically the choices they can make within the game. Purpose is operationally defined as any game element which emphasizes the player's interests, namely science. Lastly, competence is defined as the player's understanding of the tools available to them. Note that we do not examine relatedness in this study due to scope constraints, however, we believe that including both autonomy and competence is sufficient for representing SDT in this study, especially in light of the evidence of Rigby and Ryan [52], who show that relatedness is less correlated overall with desired player outcomes, such as fun/enjoyment, than competence and autonomy for popular genres such as first-person shooters and strategy games.

3.1.1 Minimalist UI. *Reducing Extraneous Load in the User Interface.* In its default version, the first few tutorial levels of Foldit display several user interface (UI) elements that both are not introduced explicitly to the player and also not necessary for completing the level. These elements include a control panel, a mode indicator, and a score history visualization. To reduce the extraneous load on the player, these elements were withheld from the early levels

Figure 1: Experimental Features. From the top left and continuing clockwise: part of the original UI, that same UI element reduced (*Minimalist UI*), on-demand basic game instructions (*Help Panel*), a pop-up text bubble asking if the player would like to know more about the science of Foldit and the following text if the user clicks “Sure!” (*Science Info*).

of the game and then introduced when the player is more familiar with the rest of the game (see Fig. 1 for example).

The generalized goal of this feature is to reduce cognitive load within the game elements, introduce only one mechanic at a time and explicitly so, thus segmenting the learning needs [37]. Many games understand to add in one element at a time, but the UI overwhelms the user from the beginning. Cognitive Load Theory would imply the need to remove most of the UI until the user has demonstrated a need for more control and information.

3.1.2 Ordered Levels. *Reducing Extraneous Load in the Level Ordering.* The Foldit tutorial previously succeeded in comprehensively covering many of the game elements, but was not structured in a way that minimizes cognitive load. The implemented approach to satisfy Cognitive Load Theory is perhaps best described by Lepink and van den Heuvel [35]: the task should begin as simple as possible, minimizing complexity and interacting elements, while giving a high level of support and instruction. After an element has been introduced, the task should increase in complexity but not interactivity. Particularly, no other elements should be introduced during this process. This method ensures well-ordered problems [21].

To generalize this feature, once the learner has mastered one element, a new one can be introduced with minimal increase in the cognitive load due to the interaction between elements. The concept of element interactivity applies to both actions and elements, and only one should be introduced at a time. During the learning process, support should be gradually reduced in proportion to the learner's proficiency, minding the expertise reversal effect (i.e. redundant support hinders and frustrates experts; see [51]). This process is repeated until the learner reaches a high-interactivity and high-complexity level of play, at which point they can be considered an expert on the subject. For this experiment,

Figure 2: Branching Levels. On the right, the original level selection screen is organized into sets and unlocks linearly. On the left, the new level selection screen branches at several points throughout the tutorial.

levels were simply re-ordered in a way which attempts to meet the criteria described above. In order to use this feature, designers must understand their game’s skill dependency tree [5] and use this knowledge to enforce well-ordered problems. Moreover, when a skill is introduced, it should be practiced in isolation against increasingly difficult problems before being mixed with other game mechanics.

3.1.3 Branching Levels. Autonomy through Meaningful Choice. Autonomy is afforded by game designs which increase choice, such as choice of goals [55]. To increase the autonomy players felt while playing Foldit, those in the *Branching Levels* condition saw a different level selection screen than other players. In this selection screen, rather than choosing levels linearly, the level options branched out, providing several choices for the order in which to complete levels (see Fig. 2).

The goal of this feature, more generally, is not only to increase perceived choice, but also to reduce the chance that players will not be able to proceed in the game, since they will have alternate paths of progress available to them.

3.1.4 Science Info. Motivation through Purpose and Relevance. When players of citizen science games were surveyed on their experiences, players requested more detail on the science behind the games [29]. By providing more in-game detail to players seeking this knowledge, it is predicted that this will give them more purpose. They will thus become more immersed in the game world, since their newly found purpose within the game appeals to the players’ willingness to concentrate [3]. Thus, at several opportunities throughout the tutorial, players in a *Science Info* condition may opt-in to know more relevant information on the science of the game element being introduced in that level (see Fig. 1 for example).

Some researchers have argued that Purpose belongs as a fourth dimension to the traditional three-factor model of Self Determination Theory [19], especially when applied to game-based learning. This internal drive is therefore hypothesized to increase player engagement. The generalization of this feature includes any appeal to the player’s personal interests.

3.1.5 Strategic Instructions. Context-Sensitive Help. Certain tutorial levels instruct the player on what actions to take, but not why those actions should be taken. For example, in the tutorial level 3-4, “Lock and Lower,” the instructions say “Pull this sheet so it lines

up with the others. Use [the] Shake and Wiggle [tools] to finish up.” However, to increase the player’s competence, the instructions should instead describe why sheets should line up with each other, what the Shake tool does, and what the Wiggle tool does. By giving the players more instruction on how to use the tools available to them and less instruction on what actions they should take, the player will become more competent in transferring these skills to new puzzle scenarios, instead of routinely and blindly following directions. In general, for the player to develop a sense of mastery, they must be able to understand the mechanics holistically, rather than as case-by-case usage.

3.1.6 Help Panel. Context-Insensitive Help. In 2012, Andersen et al. [1] tested the option of providing a help button in Foldit and other games. They found that this yielded a 12% increase in play duration for Foldit but caused negative effects in the other games. They concluded that providing assistance gave mixed results which would require further study. However, these results agree with Cognitive Load Theory. Because Foldit is a more complex game, it demands more cognitive load than the other games tested. Moreover, players working through the early levels of Foldit can be considered novices at the game. Thus, these players would benefit from more instruction and guidance [28]. Indeed, in the other simpler games, the assistance is redundant, which causes the players to experience the expertise reversal effect [51]. For this reason, a help button was implemented which provides context-insensitive assistance, such as information on game controls and relevant vocabulary (Fig. 1). We predicted that, despite the mixed findings of Andersen et al., a help panel would have an overall positive effect on learning.

More generally, Gee [20] recommends that on-demand help be available for game-based learning, and this is hypothesized to become increasingly useful as the player develops autonomy and mastery over a wide set of precise controls, such as the complex combos in a fighting game or the detailed economies of 4X and city-building games.

4 EXPERIMENTAL EVALUATION

4.1 Experiment Design

New players of Foldit ($n=1957$) were randomly assigned to one of seven conditions, either one of the six experimental features listed above or the control condition. To avoid effects of player

interaction, social chat features were disabled for all individuals in the experiment. The data were collected for approximately 6 weeks beginning in November 2018 from the version of Foldit available on its website (<https://fold.it/>).

In order to observe the effects of our feature sets, we used the data collection model of Game Metrics Behavior Assessment [41]. In this way, we infer true effects from behavioral differences, such as number of play sessions and level completions.

In addition to game metrics, players in the first level were asked to fill out an optional^{*} survey of demographic information, including their level of education, how much prior knowledge they have in biochemistry, and how frequently they play games. Of the participants, 287 users (14.7%) responded to this survey. Note that before playing Foldit, all players, whether taking the survey or not, must first consent to a user agreement of data collection based on Institutional Review Board guidelines.

4.2 Analysis Methods

First, player expertise was determined along three axes: education, prior knowledge, and gaming, using their responses to the demographic survey (n=287). These were coded as binary variables with high education (n=122) referring to users with “some college” or more, high prior knowledge (n=136) referring to users with “approximately one undergraduate course in biology” or more, and high gaming frequency (n=140) referring to users who play games at least twice per week.

Next, several ANOVA and Kruskal-Wallis omnibus tests were performed to identify the effects of player expertise on our experimental conditions. We considered these game metrics as measures: total time played, number of sessions, unique level completions (i.e., completing a level for the first time), and level re-completions (i.e., completing an old level again). We performed a Kruskal-Wallis rank sum test (n=1957) to identify any main effects of the experimental conditions, and then examined three two-way ANOVAs (n=287) to identify the interactions between experimental condition and a dimension of the player’s expertise (education, prior knowledge, or game frequency). We used an Aligned Rank Transform [64] to adjust for non-normality and performed a Holm correction after the tests to account for multiple comparisons. Finally, pairwise comparisons using the Wilcoxon rank-sum test were performed as post-hoc analyses to understand the nature of significant main effects.

4.3 Results

For all ART-adjusted ANOVAs (which captured the interactions between player expertise and experimental condition but included only players who took the survey), the main effect of condition was ignored since this information was captured in the Kruskal-Wallis tests (which included all players). The Kruskal-Wallis rank sum test for condition was trending significance after Holm correction for unique level completions ($p = .058$). A post-hoc Wilcoxon rank-sum test indicated this was driven by the *Strategic Instructions* condition,

^{*}Because Foldit is an online game available to anyone and used for more than the purposes of this study (e.g., classroom instruction), we chose an opt-in method for surveying the players. To account for the potential bias in self-selection (e.g., players who answer the survey may have a higher baseline engagement), we restrict analyses that use survey data to only the sample of users who answered the survey.

All Players (n = 1957)			
Factor	Game Metric	χ^2	p
Condition	Level Completions	15.11	.058
Condition	Level Re-completions	46.41	< .001
Survey Respondents (n = 287)			
Factor	Game Metric	F-value	p
Gameplay Frequency	Level Completions	8.03	.020
Prior Knowledge	Sessions	7.45	.027

Table 1: Summary of Main Results

which had the lowest mean for unique level completions ($M=8.84$, $SD=7.56$). The highest mean for unique level completions was *Minimalist UI* ($M=10.15$, $SD=9.65$). For level re-completions, the Kruskal-Wallis rank sum test was significant ($p < .001$). A post-hoc Wilcoxon rank-sum test indicated that players in the *Branching Levels* condition had significantly fewer level re-completions (Fig. 3); while the control group had an average of 0.84 ($SD=2.08$) re-completions, players in the *Branching Levels* condition had an average of 0.29 ($SD=0.73$) re-completions. No significant effects were found for total time played.

Interestingly, two different kinds of player expertise showed a significant increase in engagement (Fig. 4). Game playing frequency was found to be significantly correlated with an increase unique level completions ($F=8.03$, $p < .05$) from a median of 6 levels (low gaming expertise) to a median of 8 levels (high expertise). Prior domain knowledge was significantly correlated with an increase in number of sessions ($F=7.45$, $p < .05$) from a median of 1 session to 2 sessions. No other significant main effects were found. These findings are summarized in Table 1; note that “Level Completions” refers specifically to unique levels completed by the player.

Although no significant effects were found for the *Help Panel*, it was used regularly by players in this condition (n=234). Players clicked it on average 2.92 times ($SD=7.45$). Similarly, players in the *Science Info* condition (n=215) requested to know more about the science of Foldit an average of 3.84 times ($SD=4.00$).

5 DISCUSSION

In this study, we applied Self Determination Theory and Cognitive Load Theory in order to increase engagement and retention in a citizen science game. Unlike previous studies, which used these theories to measure cognitive outcomes, we operationalized these principles in order to construct and implement new game features. This process was described in our methods in a generalizable format which allows other designers to implement similar features for their games. Yet, only a couple of these features had a significant impact on behavioral measures of engagement, i.e., through game metrics data. Although we discuss practical issues in the limitations section below, these data largely suggest that creating effective features based on these theories, at least in an existing game framework, is a difficult process. We believe that the efficacy (or lack thereof) of these features does not imply that the cognitive and motivational frameworks are wrong. Rather, these principles cannot be easily tacked on as after-thoughts to a game’s design. Instead, we

Figure 3: Level Re-completions by Condition. *Branching Levels* significantly reduced level re-completions, possibly because the UI more clearly guided the players toward new levels, such as through the color scheme and by highlighting current and recommended levels. Mean (SD) for each condition: Control 0.84 (2.08); Minimalist UI 0.81 (2.81); Ordered Levels 0.96 (2.47); Branching Levels 0.29 (0.73); Science Info 0.55 (0.97); Strategic Instructions 0.85 (2.22); Help Panel 0.68 (1.59).

expect that these frameworks of learning and motivation can have a significant impact if woven into the design from the start of the design process. Thus, our hypothesis that these features would positively impact engagement was partially supported, as shown by the following examples.

First, the *Branching Levels* condition showed reduced level re-completions (Fig. 3). This may be because the UI was more clear that a level was completed, since the *Branching Levels* condition also used color and other visual attention design to indicate completed (green), unlocked (yellow and highlighted), and locked (red and silhouetted) levels. In this way, the UI design encouraged players to progress and visualize their progress, rather than re-complete previous levels. Additionally, since the new level selection screen shows what the level looks like before the player enters the puzzle, this may have reduced confusion and re-completions resulting from players being unsure whether they had solved the puzzle before or not.

The experimental condition with the biggest impact on level completions was *Strategic Instructions*, which was counter-intuitively negative. The change was intended to increase competence by replacing exact commands (e.g. “first Shake, then Wiggle this protein”) with explanations of how to use the tools and why. However, this lack of direct guidance may have led to an increase in cognitive load, since the instructions don’t contribute toward forming a cognitive schema for players who are unfamiliar with Foldit, namely all new players in our experiment. Hawlitschek and Joeckel [26] found a similar result when testing whether a digital educational game should include explicit learning instructions. They concluded

that the framing added increased extraneous load which led to the observed decreased learning outcomes. In our case, players may have been so unfamiliar with the game mechanics, regardless of background, that more nuanced instructions on their usage were overwhelming.

No significant effects were found in *Minimalist UI*, *Ordered Levels*, *Science Info*, or *Help Panel*, yet these features largely trended toward more engagement through level completions. As future work, we will continue to examine these features, since they may have impacted perceived autonomy and cognitive load, despite showing no significant effects on game metrics data.

Perhaps most notably, players with more expertise engaged more with the game, regardless whether the expertise was a biochemistry background or experience playing games (Fig. 4). These axes of expertise increased engagement in distinct ways: frequent game players completed more levels, and players with more domain knowledge played more sessions. This result is unsurprising, since Self Determination Theory would suggest that the extra expertise gives these players increased competence. Moreover, game players seem to apply their expertise by making more game progress, and players with more domain knowledge seem to have more interest in returning to the game. Critically, this finding emphasizes that Foldit is both a citizen science project *and* a game, and invites players from both communities.

This finding also agrees with recent research suggesting that personalization is vital for effective learning and gamification [24, 44, 63]. Since the player experience (i.e., the way the player perceives the game) varies with their expertise, the game’s design ought to

Figure 4: Prior domain knowledge and game expertise affect engagement. (Top) Players who play games frequently (median 8) completed on average more levels than players without that expertise (median 6; $p < .05$). Note that the median number of level completions across all players ($n=1957$) was 7, between low and high gaming expertise. **(Bottom)** Players with high prior domain knowledge (median 2) played significantly more sessions than players without prior domain knowledge (median 1; $p < .05$). Both figures show that player expertise correlates with an increase in behavioral engagement. Note that medians are shown to more accurately represent the data which are not normally distributed.

consider the player’s skills and preferences in order to construct the best experience for that particular player. Future work in this field should investigate the efficacy of dynamically tailoring tutorials to the player, not just in difficulty, but in how the material is presented, how rewards are structured, or even how the goal is framed.

5.1 Limitations

Although significant effects were found in the applications of these frameworks, the features implemented were relatively small compared to Foldit’s larger game design. Since this work was built

on the existing game structure of Foldit (now 11 years old), large-scale implementation proved difficult and time consuming. But being mindful of these frameworks at the beginning of a game’s design process could address this issue and allow more in-depth implementation of the operationalized design goals. Moreover, the operationalization process itself is a limitation in that there is a layer of abstraction between the theory and the implementation. This adds uncertainty to whether our findings (e.g., in level completions) were directly influenced by the constructs of SDT and CLT, or whether confounds exist in the design. We believe this limitation is acceptable to study theory-driven game development, since the alternative is ignoring the theory in the design process entirely.

The second major limitation to this study is that it examined only one game, and the effects may not generalize to other games or designs. However, this work demonstrates that even small changes to a game’s design can have a significant impact on engagement, and this line of research will be worth continuing for the efficiency of serious games in the future.

Additionally, in representing SDT in this study, we do not create any features which capture an operationalized definition of relatedness. However, since autonomy and competence have been shown to have a large impact on similar games [13, 52], we believe that including autonomy and competence is sufficient to approximate the effectiveness of this approach.

6 CONCLUSION

This study operationalized the learning framework of Cognitive Load Theory (CLT) and the motivation framework of Self Determination Theory (SDT) to create concrete design objectives. These objectives inspired new features which were implemented into the citizen science game Foldit in order to understand the effects of learning and motivation on a complex serious game. Uniquely, this study applied SDT and CLT in order to change the design of a game, rather than using these frameworks to measure the efficacy of existing game features. Moreover, we describe in our methodology how these operationalized definitions can be generalized for feature implementation in other games. In addition to this documentation, we present two other major contributions. First, we found that an updated level selection screen significantly impacted how players progressed through the game. This finding suggests that user interface design guides the player’s choices. Secondly, the player’s background strongly influences engagement in multiple ways. Since citizen science games are both citizen science projects and games, the player experience is influenced by both previous game experience and prior domain knowledge in the game’s subject.

This study is an exploration of theory-driven game design. As opposed to applying traditional game design principles, by operationalizing SDT and CLT for the creation of features, we are ensuring that the design satisfies the learning and motivational needs of the players, whereas traditional game design principles do not capture these needs as directly. Notably, these design implementations were performed on an existing game. This is both a limitation to the methodology, since the existing structure of the game necessarily restricts the design, and a strength, since this study captures the reality that many practitioners have to work within the constraints of existing designs. Our findings would undoubtedly have been

different had we used a different operationalized definition of the frameworks or different frameworks entirely, such as flow [8], Information Processing Theory [39], or Keller's ARCS model [32]. Because autonomy, competence, and cognitive load are strongly validated in the literature (see Background) and applicable to design goals of citizen science games, we chose these frameworks and operationalizations as the first step in exploring this methodology. Future work may explore how different models can differentially affect the game's design. Moreover, while this work explored the behavioral effects of these implementations, future work should investigate the cognitive and motivational effects of these designs.

ACKNOWLEDGEMENTS

The authors would like to thank all of the Foldit players as well as the two consulted experts, Elín Carstensdóttir and Patrick 'S0ckrates' Camarador. This work was supported by the National Institutes of Health grant 1UH2CA203780. This material is based upon work supported by the National Science Foundation under grant no. 1629879.

REFERENCES

- [1] Erik Andersen, Eleanor O'Rourke, Yun-En Liu, Rich Snider, Jeff Lowdermilk, David Truong, Seth Cooper, and Zoran Popovic. 2012. The impact of tutorials on games of varying complexity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 59–68.
- [2] Patricia Belchior, Michael Marsiske, Shannon Sisco, Anna Yam, and William Mann. 2012. Older adults' engagement with a video game training program. *Activities, adaptation & aging* 36, 4 (2012), 269–279.
- [3] Emily Brown and Paul Cairns. 2004. A grounded investigation of game immersion. In *CHI'04 extended abstracts on Human factors in computing systems*. ACM, 1297–1300.
- [4] Jenova Chen. 2007. Flow in games (and everything else). *Commun. ACM* 50, 4 (2007), 31–34.
- [5] Daniel Cook. 2007. The chemistry of game design. *Gamasutra*. http://www.gamasutra.com/view/feature/1524/the_chemistry_of_game_design.php.
- [6] Seth Cooper, Firas Khatib, Adrien Treuille, Janos Barbero, Jeehyung Lee, Michael Beenen, Andrew Leaver-Fay, David Baker, Zoran Popović, et al. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466, 7307 (2010), 756.
- [7] Ben Cowley, Darryl Charles, Michaela Black, and Ray Hickey. 2008. Toward an understanding of flow in video games. *Computers in Entertainment (CIE)* 6, 2 (2008), 20.
- [8] Mihaly Csikszentmihalyi. 1997. Flow and the psychology of discovery and invention. *HarperPerennial, New York* 39 (1997).
- [9] Mihaly Csikszentmihalyi and Isabella Csikszentmihalyi. 1975. *Beyond boredom and anxiety*. Vol. 721. Jossey-Bass San Francisco.
- [10] Sabrina Culyba. 2018. *The Transformational Framework: A Process Tool for the Development of Transformational Games*. ETC Press.
- [11] Vickie Curtis. 2015. Motivation to participate in an online citizen science game: A study of Foldit. *Science Communication* 37, 6 (2015), 723–746.
- [12] Edward L Deci and Richard M Ryan. 1985. Cognitive evaluation theory. In *Intrinsic motivation and self-determination in human behavior*. Springer, 43–85.
- [13] Edward L Deci and Richard M Ryan. 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry* 11, 4 (2000), 227–268.
- [14] Edward L Deci and Richard M Ryan. 2008. Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne* 49, 3 (2008), 182.
- [15] Michele D Dickey. 2005. Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. *Educational Technology Research and Development* 53, 2 (2005), 67–83.
- [16] Shree Durga, Magy Seif El-Nasr, Mariya Shiyko, Carmen Sceppa, Pamela Naab, and Lisa Andres. 2013. Leveraging Play in Health-Based Games to Promote Sustained Behavior Change in Healthy Eating and Exercise. In *Proceedings of the DiGRA Conference*.
- [17] Magy Seif El-Nasr, Shree Durga, Mariya Shiyko, and Carmen Sceppa. 2015. Unpacking Adherence and Engagement in Pervasive Health Games. (2015).
- [18] Alexandra Eveleigh, Charlene Jennett, Ann Blandford, Philip Brohan, and Anna L Cox. 2014. Designing for dabblers and deterring drop-outs in citizen science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2985–2994.
- [19] David Farrell and David Moffat. 2014. Applying the Self Determination Theory of Motivation in Games Based Learning. In *European Conference on Games Based Learning*, Vol. 1. Academic Conferences International Limited, 118.
- [20] James Paul Gee. 2004. Learning by design: Games as learning machines. *Interactive educational multimedia: IEM* 8 (2004), 15–23.
- [21] James Paul Gee. 2005. Good video games and good learning. In *Phi Kappa Phi Forum*, Vol. 85. The Honor Society of Phi Kappa Phi, 33.
- [22] Kris Graft. 2015. Designing Shadow of Mordor's Nemesis system. *Gamasutra*. http://www.gamasutra.com/view/news/235777/Designing_Shadow_of_Mordors_Nemesis_system.php.
- [23] Tiago de Araujo Guerra Grangeia, Bruno de Jorge, Daniel Franci, Thiago Martins Santos, Maria Silvia Vellutini Setubal, Marcelo Schweller, and Marco Antonio de Carvalho-Filho. 2016. Cognitive load and self-determination theories applied to e-learning: impact on students' participation and academic performance. *PloS one* 11, 3 (2016), e0152462.
- [24] Richard Halverson. 2005. What can K-12 school leaders learn from video games and gaming? *Innovate: journal of online education* 1, 6 (2005), 3.
- [25] Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage Publications Sage CA: Los Angeles, CA, 904–908.
- [26] Anja Hawlitschek and Sven Joekel. 2017. Increasing the effectiveness of digital educational games: The effects of a learning instruction on students' learning, motivation and cognitive load. *Computers in Human Behavior* 72 (2017), 79–86.
- [27] Chin-Lung Hsu and Hsi-Peng Lu. 2004. Why do people play on-line games? An extended TAM with social influences and flow experience. *Information & management* 41, 7 (2004), 853–868.
- [28] Yuling Hsu, Yuan Gao, Tzu-Chien Liu, and John Sweller. 2015. Interactions Between Levels of Instructional Detail and Expertise When Learning with Computer Simulations. *Educational Technology & Society* 18, 4 (2015), 113–127.
- [29] Ioanna Iacovides, Charlene Jennett, Cassandra Cornish-Trestrail, and Anna L Cox. 2013. Do games attract or sustain engagement in citizen science?: a study of volunteer motivations. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1101–1106.
- [30] Charlene Jennett, Laure Kloetzer, Daniel Schneider, Ioanna Iacovides, Anna Cox, Margaret Gold, Brian Fuchs, Alexandra Eveleigh, Kathleen Methieu, Zoya Ajani, et al. 2016. Motivations, learning and creativity in online citizen science. *Journal of Science Communication* 15, 3 (2016).
- [31] Alexander Kawrykow, Gary Roumanis, Alfred Kam, Daniel Kwak, Clarence Leung, Chu Wu, Eleyine Zarour, Luis Sarmenta, Mathieu Blanchette, Jérôme Waldispühl, et al. 2012. Phylo: a citizen science approach for improving multiple sequence alignment. *PloS one* 7, 3 (2012), e31362.
- [32] John M Keller. 1987. Development and use of the ARCS model of instructional design. *Journal of instructional development* 10, 3 (1987), 2.
- [33] Kristian Kiili. 2005. Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education* 8, 1 (2005), 13–24.
- [34] Jimmie Leppink, Fred Paas, Tamara Van Gog, Cees PM van Der Vleuten, and Jeroen JG Van Merriënboer. 2014. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction* 30 (2014), 32–42.
- [35] Jimmie Leppink and Angelique van den Heuvel. 2015. The evolution of cognitive load theory and its application to medical education. *Perspectives on medical education* 4, 3 (2015), 119–127.
- [36] J Derek Lomas, Kenneth Koedinger, Nirmal Patel, Sharan Shodhan, Nikhil Poonwala, and Jodi L Forlizzi. 2017. Is difficulty overrated?: The effects of choice, novelty and suspense on intrinsic motivation in educational games. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1028–1039.
- [37] Richard E Mayer and Roxana Moreno. 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist* 38, 1 (2003), 43–52.
- [38] Elisa D Mekler, Florian Brühlmann, Alexandre N. Tuch, and Klaus Opwis. 2017. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior* 71 (2017), 525–534. <https://doi.org/10.1016/j.chb.2015.08.048>
- [39] George A Miller. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review* 63, 2 (1956), 81.
- [40] K Denise Muth. 1991. Effects of cuing on middle-school students' performance on arithmetic word problems containing extraneous information. *Journal of Educational Psychology* 83, 1 (1991), 173.
- [41] LE Nacke, Anders Drachen, and Stefan Göbel. 2010. Methods for evaluating game-play experience in a serious gaming context. *International Journal of Computer Science in Sport* 9, 2 (2010), 1–12.
- [42] Lennart E Nacke and Christoph Sebastian Deterding. 2017. The maturing of gamification research. *Computers in Human Behaviour* (2017), 450–454.
- [43] Thomas E Nygren. 1991. Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload. *Human Factors* 33, 1 (1991), 17–33.

- [44] Rita Orji, Regan L. Mandryk, and Julita Vassileva. 2017. Improving the Efficacy of Games for Change Using Personalization Models. *ACM Trans. Comput.-Hum. Interact.* 24, 5, Article 32 (Oct. 2017), 32:1–32:22 pages. <https://doi.org/10.1145/3119929>
- [45] Kunwoo Park, Meeyoung Cha, Haewoon Kwak, and Kuan-Ta Chen. 2017. Achievement and friends: key factors of player retention vary across player levels in online multiplayer games. In *Proceedings of the 26th International Conference on World Wide Web Companion*. International World Wide Web Conferences Steering Committee, 445–453.
- [46] Wei Peng, Jih-Hsuan Lin, Karin A Pfeiffer, and Brian Winn. 2012. Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology* 15, 2 (2012), 175–196.
- [47] Marisa Ponti, Igor Stankovic, Wolmet Barendregt, Bruno Kestemont, and Lyn Bain. 2018. Chefs Know More than Just Recipes: Professional Vision in a Citizen Science Game. *Human Computation* (2018).
- [48] Andrew K Przybylski, C Scott Rigby, and Richard M Ryan. 2010. A motivational model of video game engagement. *Review of general psychology* 14, 2 (2010), 154.
- [49] Yu Qi Qiao, Jun Shen, Xiao Liang, Song Ding, Fang Yuan Chen, Li Shao, Qing Zheng, and Zhi Hua Ran. 2014. Using cognitive theory to facilitate medical education. *BMC medical education* 14, 1 (2014), 79.
- [50] Alexander J Quinn and Benjamin B Bederson. 2011. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1403–1412.
- [51] Günter Daniel Rey and Florian Buchwald. 2011. The expertise reversal effect: Cognitive load and motivational explanations. *Journal of Experimental Psychology: Applied* 17, 1 (2011), 33.
- [52] Scott Rigby and Richard Ryan. 2007. The player experience of need satisfaction (PENS) model. *Immersive Inc* (2007), 1–22.
- [53] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.
- [54] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.
- [55] Richard M Ryan, C Scott Rigby, and Andrew Przybylski. 2006. The motivational pull of video games: A self-determination theory approach. *Motivation and emotion* 30, 4 (2006), 344–360.
- [56] Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. 2017. How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior* 69 (2017), 371–380.
- [57] Demographic Sales and Usage Data. 2017. Essential Facts About the Computer and Video Game Industry, ESA Entertainment Software Association. 2017.
- [58] Henrik Schoenau-Fog. 2011. The Player Engagement Process-An Exploration of Continuation Desire in Digital Games.. In *Proceedings of the DiGRA Conference*.
- [59] Katie Seaborn and Deborah I Fels. 2015. Gamification in theory and action: A survey. *International Journal of human-computer studies* 74 (2015), 14–31.
- [60] Paul Skalski, Francis Dalisay, Matthew Kushin, and Yung-I Liu. 2012. Need for presence and other motivations for video game play across genres. *Proceedings of the Presence Live* (2012).
- [61] Penelope Sweetser and Peta Wyeth. 2005. GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment (CIE)* 3, 3 (2005), 3–3.
- [62] John Sweller. 2010. Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review* 22, 2 (2010), 123–138.
- [63] Gustavo F Tondello, Rita Orji, and Lennart E Nacke. 2017. Recommender systems for personalized gamification. In *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. ACM, 425–430.
- [64] Jacob O Wobbrock, Leah Findlater, Darren Gergle, and James J Higgins. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 143–146.