1-2024

Using Modern Motivational Theories as a Design Framework for VR Gamification

Christopher Thomas Boardman

Abstract

Virtual reality has shown promise as a tool for organizational training. Implementing gamification, the use of game elements, has been shown to further improve learning outcomes. However, gamification has shown inconsistent results when applied to VR. To investigate this, we used motivational theory as a framework to select, justify, and implement game elements based on each theory to create different versions of a virtual grocery store training. Our results did not indicate any differences on performance, persistence, or error rate between each training condition, and exploratory analyses investigated the link between perceived competence and VR program quality. We discuss potential explanations for gamification's inconsistency and the effects of presence.

Keywords: virtual reality, gamification, motivational theory, psychology

MONTCLAIR STATE UNIVERSITY

Using Modern Motivational Theories as a Design Framework for VR Gamification

By

Christopher Thomas Boardman

A Master's Thesis Submitted to the Faculty of

Montclair State University

In Partial Fulfilment of the Requirements

For the Degree of

Master of Arts

Dec 2023

College of Humanities and Social Sciences

Department of Psychology

Thesis Committee

[redacted]

Dr. Kevin Askew

Thesis Sponsor

[redacted]

Dr. Valerie Sessa

Committee Member

[redacted]

Dr. Michael Bixter

Committee Member

[redacted]

Dr. Mele Kramer

Committee Member

USING MODERN MOTIVATIONAL THEORIES AS A DESIGN FRAMEWORK FOR VR GAMIFICATION

A THESIS

Submitted in partial fulfillment of the requirements.

For the degree of Master of Arts

By

Chris Boardman

Montclair State University

Montclair, NJ

2023

Acknowledgements

The author of this paper would like to thank the following:

Gilded Raven Games for providing a secondary lab location and for keeping spirits high during the data collection process.

Joe Bavazzano and the rest of ADP Center for Learning Technologies staff for assistance in procuring a VR headset for lab use.

The 2021 MSU I/O Psychology cohort for their encouragement and for listening to my 20-minute rants on gamification.

My family for supporting me throughout graduate school. Special thanks to my younger sister for pilot testing the training program almost daily.

Special thanks to Dr. Kevin Askew for being an amazing teacher, mentor, and friend and for guiding me through the thesis process, as I would have never completed it without his help. Thank you to the rest of the Montclair's Technology at Work lab team, especially Mike DeFabiis and Jordyn Kennedy.

This thesis is dedicated to Kelly Elizabeth Benning, for whom without her support this thesis wouldn't have been possible.

Contents

List of Figures

List of Tables

Using Modern Motivational Theories as a Design Framework for VR Gamification

Abstract

Virtual reality has shown promise as a tool for organizational training. Implementing gamification, the use of game elements, has been shown to further improve learning outcomes. However, gamification has shown inconsistent results when applied to VR. To investigate this, we used motivational theory as a framework to select, justify, and implement game elements based on each theory to create different versions of a virtual grocery store training. Our results did not indicate any differences on performance, persistence, or error rate between each training condition, and exploratory analyses investigated the link between perceived competence and VR program quality. We discuss potential explanations for gamification's inconsistency and the effects of presence.

Introduction

Virtual reality (VR) is a type of immersive technology that recreates vivid and immersive environments through the use of visual displays and other hardware. Users of VR may interact within this environment to varying degrees dependent on the specific type of VR implementation and do things such as moving around or interacting with objects. Software development programs allow for environments to be created for desired experiences, allowing VR developers great flexibility. While VR was originally created for entertainment purposes, it has been shown to have practical value in the fields of medicine (Li et al., 2017), education (Kavanagh et al, 2017), and military applications (Lele, 2013). Recently, VR has seen increased use for non-military, non-medical occupational trainings. (Cheng, 2010). Virtual reality is continuing to grow in size, and it is predicted that over 23 million jobs will be using VR by 2030 (PvC, 2019).

As virtual reality becomes more commercially available, many researchers have been examining the advantages of using the medium. VR has been associated with increases to learning motivation when compared to traditional instruction (Sattar et al., 2019) and has been used as an affordable alternative to more costly simulator training (Bhagat et al., 2016). An additional advantage is that the environments created by VR are immersive, providing a high degree of fidelity to the user (Berkman, 2020). Immersiveness has been associated with increases in learning outcomes (Baldwin & Ford, 1988), such as retention of declarative knowledge and job skill acquisition (Cheng, 2010; Lee, 2017), and could explain why a recent meta-analysis found that VR trainings are more effective than their traditional counterparts (Howard et al., 2020).

While VR training has been shown to be generally effective, the modern form of the technology is still relatively new, and more research is needed to determine best practices and how to make VR most effective for occupational training. One promising way to increase the effectiveness of VR training even further is by implementing gamification - the intentional use of typical elements from video games into another area of activity (Deterding et al., 2011). The use of gamification in non-VR contexts has been associated with increases in task performance, quality of work, and user engagement when compared to non-gamified approaches (Hosseini et al., 2022; Prasad & Mangipudi, 2021). Proponents of gamification believe this may be because gamified training is more engaging and therefore requires users to expend less cognitive effort to learn. Researchers have started to examine the effects of gamification in a virtual reality context, but the findings have been inconclusive so far. While some studies that add gamification to VR trainings find non-significant effects on learning (e.g., Palmas et al., 2019), other studies have found that adding gamification to VR trainings can increase learning (e.g., Chittaro & Buttussi, 2015; Lu & Davis, 2016). For this reason, Howard et al. (2020) called for more research into exploring the addition of gamification to VR trainings.

A challenge in applying gamification to a training context, and potentially a reason for the conflicting results in the literature, is that gamification is not a single technique but a collection of techniques. The exact number of techniques varies depending on the taxonomy being used, but a recently developed taxonomy identified 21 different *game elements -- e.g., points* for different in-training action, *leaderboards* to promote competition, adding a *narrative* to the experience to increase motivation -- which means there are 21 different ways to gamify a training program *if one is only using a single gamification technique*. In practice, gamification is usually done using multiple game elements, and as one might expect, the number of ways to implement gamification grows exponentially as the number of elements to be added increases. There are 210 different ways to gamify a training program if two gamification elements are to be used, and an astounding 1,330 different ways to gamify a training program if only three elements are to be implemented. The multiplicity of strategies to implement gamification is a problem because we do not know what gamification elements work 'best' in isolation or in combination with other game elements.

One way to address the "numbers problem" of what elements work best in gamification is by using well-established psychological motivational theories as a framework. Theories such as goal-setting theory, self-determination theory, and expectancy theory may provide guidance on what gamification elements work best during design implementation. Comparing different perspectives against each other could also reveal differences in what theory is the most appropriate when designing gamification for a VR training environment.

Accordingly, the purpose of this study is twofold. First, we seek to determine if gamification principles can be effective when applied to a VR training program. The findings from our study could help contribute to understanding some of the inconsistent findings in the literature and address the question about whether gamification is generally helpful in a VR training context. Second, we are interested in determining if motivational theories may be applicable to designing VR training environments and if so, what theory is most applicable. Implementation of game elements based on the tenets of each psychological theory may provide guidance on which elements work best and help researchers and practitioners address the challenges of implementing gamification in VR training.

Literature Review

Virtual Reality

In 1991, Coates defined virtual reality as “electronic simulations of environments experienced via head mounted eye goggles and wired clothing enabling the end user to interact in realistic three-dimensional situations” (Coates,1992). While other definitions expand on this to include non-head mounted display systems (HMDs) most researchers agree that virtual reality references the use of technological hardware to immerse the user in a simulated environment. Additionally, virtual reality may allow for multiple individuals to interact within this simulated environment, allowing for each user to represent themselves as avatars and shape their individual, or shared, spaces (Girvan, 2018). To create a more modern technology-centric definition, this research will use Jerald’s “computer-generated digital environments that can be experienced and interacted with as if that environment was real.” (Jerald, 2015)

A combination of different technologies allow users to see themselves in virtual environments through the use of VR. The most common virtual environment is displayed through the use of a standard computer monitor alongside a mouse and keyboard. However, as consumer market availability has increased, HMDs and motion-sensor controllers are seeing increased usage. Other techniques that create virtual environments include surrounding the user with projections (CAVE systems) or by encapsulating them with multiple monitors. Regardless of the medium used to create a virtual environment, differences were not found in device effectiveness in Howard’s meta-analysis (Howard, 2021). Researchers suggest that the effectiveness of the medium used may be mediated by *task technology fit*, or “the degree to which technology assists an individual in performing their portfolio of tasks” (Goodhue &

Thompson 1995). This construct suggests that the interaction between task and technology characteristics create an effect that enhances performance outcomes.

Different characteristics of virtual reality may also affect the degree to which users feel disconnected from the physical world. These feelings of disconnect may be explained by the concepts of presence and immersion. *Immersion* has been defined as the extent to which computer displays can deliver illusions of reality across four dimensions: the degree to which that they shut out physical reality, provide a range of sensory processes, surround the user with visual stimuli, and are vivid in resolution and richness (Slater & Wilbur, 1997). Immersion has been found to have positive effects on spatial understanding, may reduce information clutter, and can lead to higher levels of task performance, particularly with stereoscopic HMDs (Bowman & McMahan, 2007). *Presence* is defined as the psychological sense of being in a virtual environment (Slater & Wilbur, 1997). Multiple studies have been performed on the effects of presence on task performance (Slater et al., 1996, Witmer & Singer, 1998) but researchers have been unable to find a conclusive link between the two (van Baren & IJsselsteihn, 2004). Research has indicated that cybersickness - symptoms of discomfort or malaise in virtual reality - is negatively associated with presence (Weech et al, 2009). Higher ratings of cybersickness in turn have been associated with lower ratings on task performance, most likely due to discomfort and symptoms of motion sickness.

Due to the immersive nature of the medium, virtual reality has been utilized for a variety of different purposes. Creators of virtual reality software are able to design and cater virtual environments for unique and individual purposes. For example, in the medical field, virtual reality has been used to train junior surgeons on minimally invasive surgeries using an immersive training device. This allows junior surgeons to learn the same skills they would in a quicker time instead of over an 80-hour work week (Alaker et al., 2016). Research has also indicated that virtual reality may be used as an alternative analgesic for patients in pain as they allocate attention to the simulation and away from a painful procedure, e.g., chemotherapy (Wismeijer & Vingerhoets, 2005). Virtual reality has also been used for a variety of military training purposes. Modern flight simulators have been created to address the growing need of highly realistic air warfare simulations to train U.S. Air Force pilots (Huang, 2004). These simulators have also extended to maritime use for diving operations, military engineering for battlefield visualization (Strickland 2010), and may even evolve current ground operations using real-time goggle displays for soldiers (Cameron, 2010). Virtual reality has also seen increased use in job training and professional development fields as the technology becomes increasingly available. VR has been used in safety training for construction work for identifying risk assessment (Sacks et al., 2013), manufacturing training using gamified assessment (Ulmer et al., 2020), and research has found that it may be an effective tool for teaching interpersonal skills alongside real life trainers (Mast et al., 2018).

Virtual reality has been examined as a tool to teach both declarative and factual knowledge. Research has indicated that the use of a virtual environment may lead to higher declarative knowledge retention when compared to traditional instruction (Webster, 2015). Research has also demonstrated that when users were able to directly manipulate elements of the VR training, they would be more likely to perform better on knowledge-based assessment (Jang et al., 2017) In addition to knowledge retention, a literature review has also shown virtual reality training to have increased task performance on psychomotor activities, such as manufacturing assembly, when compared to traditional training methods (Albich, 2021). It is emphasized that to make these comparisons, researchers should use error tracking when measuring procedural, or skill-based, knowledge instead of using time. Because virtual reality training takes longer to complete, error tracking may provide better clarity on whether procedural knowledge was retained (Patel et al., 2017). Ultimately, the effectiveness of virtual reality training has been shown to be more effective when compared to traditional methods, yielding outcomes that were half a standard deviation higher (Howard, 2021).

Gamification

The process of defining gamification has been difficult as there has been disagreement between researchers in creating a more operational definition. Using the early work from games philosopher Johan Huizinga, the most commonly cited definition of gamification suggests that gamification is not just a game, but instead the "use of game design elements in non-game contexts" (Deterding et al., 2014**)**. They emphasize that gamification involves a sense of rules or boundaries, that "gameful interactions" are the artifacts that determine the quality of the game experience, and that 'gameful designs' are those that design in mind with game elements. Huotari and Hamari also suggest that gamification is "the process of enhancing experiences with affordances for gameful experiences to support the user's overall value creation." This definition emphasizes that individual user differences matter, and that gamification is layered over a system designed with gamification in mind. For the purposes of this paper, we will be using a definition coined by Seaborn and Fels (2015), who synthesized these and other existing theories that define gamification as "the intentional use of game elements for a gameful experience of non-game tasks and contexts." This definition emphasizes the careful intention of implementing game elements like Huotari and Hamari while expanding the definition to include non-game contexts like Deterding.

There are a variety of approaches in taxonomizing game elements due to their subjective nature. To address this, Toda et al. (2019) developed a taxonomy of game elements for instructional environments. This taxonomy places game elements across five dimensions to appropriately categorize them. The performance dimension consists of elements that provide direct feedback based on user action such as *points* or *progression*. The ecological dimension affects the gamified environment's interaction with the user, using elements such as *chance* or implementing a user economy. Elements within the social dimension encourage user interactions within the virtual environment, such as *cooperation* or *competition*. The personal dimension provides direct meaning to the user by keeping users engaged with *puzzles* or *item rarity*. Finally, the fictional dimension overlays the system to create context or immerse the player further with a *narrative* or *storytelling device*. Ultimately, Toda et al. suggest that using a combination of these elements across dimensions will result in more motivated users by means of player engagement and motivation.

Unfortunately, there is no predetermined combination of these elements that have been found to be the most effective when applied. With 21 separate elements there are an overwhelming number of combinations. For example, trying to select 3 elements to implement means trying to select out of 1,330 possible combinations, 5,985 for 4 elements, and so on. While a taxonomy provides a means to clarify the subjective criteria of defining game elements, it does not provide guidance on which elements to select. To improve gamification effectiveness, a method that narrows down game elements to an effective few is needed. Usage of psychological motivational theories may aid in this 'numbers problem' by providing meaningful context for these elements.

Motivational Theories

In order to make better predictions about which theories could best inform gamification design, we conducted a literature review of relevant motivational theories. First, we determined whether each theory could predict motivation during the training process. For example, the theory of planned behavior explains intention to use training, rather than motivation during the training, so it was not used. Next, we selected from the remaining theories whether they could be used to narrow down potential game elements. For example, according to reinforcement theory, all game elements could be either punishers or reinforcers and therefore could not help to deliberate which elements would be best. After consideration, we arrived at three theories: self-determination theory, goal-setting theory, and expectancy theory. A list of theories that were considered for use is available in Table A.

To test these theories, we decided to apply them to a training program created for grocery store cashiers. Cashiers are responsible for receiving and disbursing money by using electronic scanners or cash registers. (O*NET) While the grocery cashier job is moving away from operating single-user registers to maintaining multiple self-checkout kiosks, cashiers are still responsible for the knowledge of operating these kiosks. The job essential tasks required of a grocery cashier are relatively simple: scanning and manually entering items, removing items if needed, and accepting payment from the customer. We believe that this job is simple enough that all three theories can be applied to this position in a purposeful way. Additionally, elements of the cashier role can be easily manipulated for an experimental setting. For example, buttons or icons on the cashier screen can be changed as desired, and the number of grocery items can remain the same between orders. Although we were able to fit these theories to this particular job, other job training programs may be applicable to these theories as well.

Self Determination Theory

Developed in 2000, self-determination theory (SDT) postulates that human motivation is based around self-regulation and human needs for development. Using an inductive approach, Deci & Ryan identified three primary needs: competence (feelings of mastery and effectiveness), relatedness (having a sense of belonging and connection with others), and autonomy (feeling in control of their life) that must be met for optimal social development and well-being (Deci & Ryan, 2000). SDT suggests a continuum of motivation ranging from behaviors that feel extrinsically motivated to those that are intrinsically motivated. Extrinsic motivation is driven from external sources such as promises of reward or compliance. Intrinsic motivation is driven from self-control and personal desire, whether it be from enjoyment or personal satisfaction of completion. Deci & Ryan suggest that intrinsically driven behaviors lead to better outcomes for personal and social growth so long as the theory's mediators are met.

We applied SDT to our cashier training program to determine what game elements the theory predicts are the most effective at increasing motivation and by extension learning. In order to reduce subjectivity, we took a three-step approach in applying the theory. First, we made judgments on each element about whether or not the theory predicts it would increase learning or motivation in general. We focused on the three primary needs of SDT: autonomy, competence, and relatedness, which SDT posits are necessary for increased intrinsic motivation. Next, we made judgments on whether or not these elements could be applied to a grocery training program. While some elements were predicted to increase motivation, they could not be applied in the grocery context such as *puzzles* or *levels*. Finally, once we established a list of elements that were predicted and applicable, we narrowed down the list to the three most strongly connected game elements. Three elements were selected for each theory to keep the number of elements per theory constant. These final three elements were then implemented into a gamified condition of the grocery store training to represent SDT. Elements that were considered but not selected for the SDT condition are listed in Table B.

The first element that was predicted by SDT was *renovation*, defined as any game element that allows the trainee to repeat or restart an action (Tondello et al., 2019). Renovation elements give trainees the freedom to fail at a task without consequence, allowing them to repeat the task as often as they would like (Kam & Umar, 2018), thus increasing perceptions of autonomy. SDT predicts that as autonomy increases, so does intrinsic motivation. We have implemented renovation by giving trainees the option to repeat each block of training upon completion as often as they chose. The second element that was predicted by SDT were *points*, defined as units of measurement that represent trainee performance on a task. Points provide players with feedback which may increase perceptions of competency on a task. (Kam & Umar, 2018). SDT predicts that as these perceptions of competency increase, so does intrinsic motivation. We implemented points by providing trainees with a screen above the register on which they could view their points total at any time. The final element that was predicted by SDT was *acknowledgement*, or all kinds of feedback that praise the trainee's specific actions. Acknowledgments may increase trainee competency as they receive them and increase intrinsic motivation according to SDT. We implemented acknowledgements by using trophies that appeared above the register as trainees completed orders. Trainees could view their progress towards earning trophies on the screen above the register next to their point total.

Expectancy Theory

Vroom's expectancy theory (ET) postulates that the intensity of work effort depends on the perception that an individual's effort will result in a desired outcome (Vroom, 1964). According to ET, individuals will choose to take an action that has the highest perceived motivational force. Vroom suggests that this motivational force is the product of three components: expectancy, instrumentality, and valence. Expectancy is the belief that an individual's efforts will lead to a high level of performance. Instrumentality is the degree to which the individual perceives their efforts will result in a specific outcome. Valence refers to the degree to which the individual desires the rewards or outcomes from performing the task. ET suggests that motivation occurs when these three factors are present, but motivation will disappear when one factor is equal to zero (Vroom, 1964).

We applied ET in a similar way to SDT to our cashier training program to determine which elements were the most effective at predicting motivation and, by extension, learning. We repeated the same three-step approach to reduce subjectivity. First, we made judgments on whether ET predicted each element and whether it would increase learning or motivation. We focused on the three main components of ET: instrumentality, valence, and expectancy, which ET posits are necessary for the highest motivational force. Next, we deliberated on whether these elements could be applied to the grocery store task in a meaningful context. While ET predicted some elements could increase motivation or learning, they could not be applied to the grocery store task. Finally, we narrowed down the list to the three most strongly connected elements. These elements were then implemented into a gamified version of the grocery store training to represent ET. Elements that were considered but not selected for the ET condition are listed in Table B.

The first element that was predicted by ET was *social pressure*, defined as any element that causes pressure from social interactions between other trainees. (Toda et al., 2019). One form of social pressure exists through gift-giving between trainees. Gift-giving may allow trainees to develop relationships with past and future trainees by receiving or giving a gift. After receiving a gift, trainees may greatly value the ability to reciprocate this gift to the next trainee. (Hamzah et al, 2014). ET predicts that as trainees place more value on the reward (or in this case, earning a reward for someone else), motivational force will increase, causing trainees to be more motivated to perform the task. We implemented social pressure via gift-giving in the form of a donut: trainees were informed at the start of the training that the previous trainee performed well enough to earn themselves a donut. They were also informed that if they perform well enough on the task that they will earn the next trainee a donut as well.

The second element predicted by ET was *acknowledgement*. Acknowledgements may act as a form of feedback to trainees to enhance the perception that they are responsible for task success (expectancy). ET predicts that as these perceptions of expectancy increase, so does motivational force. We implemented acknowledgement in the same way as the SDT condition: using trophies above the register that appeared as trainees complete grocery orders and providing a screen on which trainees could track their progress towards earning them. The final element predicted by ET was *competition*, defined as when two or more players compete against each other towards a common goal (Toda et al., 2019). When trainees have the opportunity to compete, they may feel more confident at a task if they outperform others, subsequently increasing expectancy perceptions. (Hamzeh et al., 2014). Just like the acknowledgement element, ET predicts that as expectancy perceptions increase, so does motivational force. We implemented competition by using a trainee leaderboard that appeared alongside the acknowledgement screen. Trainees could see their leaderboard position rise as they completed actions within the training, and the positions were easily attainable to prevent frustration (and to possibly prevent decreases in expectancy).

Goal Setting Theory

Goal-setting theory (GST) states that specific, challenging goals lead to higher levels of task performance than easy or vague goals (Locke & Latham, 1990). Goals affect performance through four mechanisms: *direction, effort, persistence, and knowledge*. Direction refers to the frequency at which an individual performs behaviors that align with the goal while stopping behaviors that are goal-irrelevant. Effort is the proportion of the goal's difficulty compared to the effort that is committed to completing the goal. Persistence refers to the degree to which the individual continues to work towards the goal across an extended period. Knowledge (or task strategy) refers to the specific knowledge and abilities that are either used or self-taught to reach a goal's completion. In addition to these mechanisms, the goal and performance relationship is moderated by ability (individual skill or talent), complexity (simplicity of the goal), goal commitment (determination to finish the goal) and feedback on goal progress. While goals can be set for others, Locke and Latham emphasize that trainees should participate in the goal-setting process (2006). Ultimately, goals increase performance through the above mechanisms and moderators.

We applied GST in the same manner as the other two theories to our cashier training program to determine which elements best predicted increases to motivation and learning according to the theory. We used the same three-step process in narrowing down gamification elements. First, we made judgments on what game elements GST would predict to increase motivation and learning. We focused on what game elements would address the mediators or moderators of goal setting, such as goal clarity, or would allow for implicit goal setting. Next, we determined if these game elements were applicable to the grocery store training. Finally, we narrowed down the list of remaining elements to the three most strongly connected elements. These elements were then implemented into a gamified version of the grocery store training to represent GST. Elements that were considered but not selected for the GST condition are listed in Table B

The first element that was predicted by GST was *progression*, defined as any element that provides the trainee with information on their progress and allows them to locate themselves within the training (Toda et al., 2019). Progression elements allow for specific goals to be set for trainees and provide feedback towards completing them (Tondello et al., 2018). GST posits that specific goals with accompanying feedback are effective at increasing motivation (Locke & Latham, 1990). We implemented progression using two progress bars; one indicated progress towards completing a full grocery order, while the other indicated progress towards completing the training program. The second element that was predicted by GST was *objectives*, or elements that guide the player's actions through goals and subgoals (Toda et al., 2019). Objectives may help to establish explicit goals for the trainee while also clarifying steps to accomplish them. GST predicts that simpler goals are more motivating to complete than complex goals. We implemented objectives using a screen that displayed the next steps for a trainee to complete a full grocery order. As trainees completed actions, the objective screen updated to display the next steps. For example, after a trainee selected the checkout icon, the objective element instructed them to accept payment. The final element that was predicted by GST was competition. Competitive elements may allow for implicit goal setting as trainees may desire to outperform other trainees and therefore be more committed to completing them. GST predicts that as goal commitment increases, subsequently so does the motivation to complete them. We implemented competition in the same way as ET: a leaderboard screen that displayed the trainees' current progress and used easily attainable leaderboard ranks to encourage trainees to set their own goals to reach them.

The Current Study

The aim of the study was to examine the effectiveness of gamification on virtual reality training with design informed by motivational theory. Gamification elements were selected using each motivational theory as a design framework. These elements were then applied to separate conditions of a VR-based grocery store cashier training program based on each theory. We theorize that training designed with gamification elements which are informed by psychological theories will have higher performance outcomes than non-gamified trainings. Thus, we hypothesize:

Hypothesis 1. Users taking gamified conditions of the grocery store training will have significantly increased performance outcomes when compared to the non-gamified control group.

One additional method of measuring performance is through error tracking. Error tracking has been shown to be an effective way to measure trainee procedural knowledge retention (Bailey et al, 2017). However, game elements could be distracting and may be detrimental to performance (Kim & Werbach,2016). We are unable to predict whether the inclusion of gamification may cause more errors in a virtual reality setting. Therefore, we present two competing hypotheses:

Hypothesis 2a. Users taking gamified conditions of the grocery store training will have significantly *more errors* when compared to the non-gamified control group.

Hypothesis 2b. Users taking gamified conditions of the grocery store training will have significantly *fewer errors* when compared to the non-gamified control group.

Gamification has also been hypothesized to increase the motivation of trainees. We expect that as motivation increases so will its associated factors. One factor that has been associated with motivation is persistence (Vollmeyer & Rheinberg, 2000). If trainees have sufficient training time and are highly motivated, they will spend more time or interact with learning tasks on their own volition (Atkinson & Raynor, 1974). We therefore hypothesize that:

Hypothesis 3. Users taking one gamified condition will have significantly increased persistence outcomes compared to the other gamified conditions.

In addition to analyzing the effects of informed gamification design in a VR context, we aimed to examine which of these theories may be most effective as a design framework. While each motivational theory has a different approach in ways to increase motivation, one theory may prove to be more effective than others to inform gamification design in VR-based training. We have stated this as three competing hypotheses.

Hypothesis 4a. GST is the superior gamification framework -- Users taking the GST condition will have significantly increased performance and persistence and fewer errors compared to the other two gamified conditions.

Hypothesis 4b. SDT is the superior gamification framework -- Users taking the SDT condition will have significantly increased performance outcomes compared to the other two gamified conditions.

Hypothesis 4c: EV is the superior gamification framework -- Users taking the EV condition will have significantly increased performance outcomes compared to the other two gamified conditions.

Method

Participants

Participants (*N* = 126) were recruited from a board game store located in New Jersey. Additional participants were recruited through Montclair State University's SONA participant system or through on-campus canvassing. Participants were offered entry into a raffle for a $20 Amazon gift card or were given SONA participation credits. Participants were screened before starting the study on whether they were able to use a virtual reality headset and if so, whether they experienced motion (simulator) sickness. Participants were at least 18 years of age and were primarily between the ages of 25-34 years old (62.5%). Out of the 130 participants, the sample identified as 63.2% male, 26.6% female, and 10.2% non-binary/other. Two participants were unable to complete the full length of the training, while two did not attempt to make any progress towards training goals, and thus all four were dropped.

Measures

Pre-test Scales

Personality Factors: Personality factors were measured using items from Goldberg's (1992) IPIP Big-Five Factor Inventory subscales for Conscientiousness, Openness, and Agreeableness. Participants rated 30 personality statements on a 5-point scale (1 = Very Inaccurate, 3 = Neither Accurate or Inaccurate, 5 = Very Accurate). One item from the Intellect scale ("I am full of ideas") was dropped due to low item validity. Coefficient alpha for conscientiousness was reported as $\alpha = .84$, agreeableness was reported as $\alpha = 0.82$, openness was reported as $\alpha = 0.55$ before Item #7 was removed, and as $\alpha = 0.70$ after it was removed.

Familiarity and VR Experience: Familiarity with VR and Video Games were each assessed using a single item "How familiar are you with (Virtual Reality/Video Games)" on a five-point scale (1 = Not at all, 5 = Extremely Familiar.) On average, participants had moderate familiarity with VR (M = 2.78) and were very familiar with video games (M = 4.31). Prior experience of using a virtual reality headset was assessed using a single item (1= No Experience, 3 = Moderately Experienced, 5 = Extremely Experienced). 26.2% of participants had never used VR, 42.1% had slight experience, 22.2% had moderate experience, and 9.5% of participants were very or extremely experienced. Additionally, VR Ownership was assessed using a yes or no question: 19% of participants stated that they currently owned a VR headset.

Prior Cashier or Supermarket Experience: Prior experience of working in a cashier role or in a supermarket was assessed using two yes or no questions that asked, “Have you ever worked in a (cashier position/supermarket) before?” Out of 126 participants, 68.2% stated they previously worked a cashier job, 30.2% of which were in a supermarket setting.

Post-test Scales

Cybersickness: Cybersickness was measured using Kim’s (2018) Virtual Reality Sickness Questionnaire (VRSQ). Participants rated the intensity of symptoms immediately after the training across nine items on a four-point scale (1 = None, 2 = Slight, 3 = Moderate, 4 = Severe). Examples of symptoms included ‘Eye strain’ and ‘Fullness of the Head’. Coefficient alpha was α = .85 in the present study. Out of 126 participants, 95 experienced little to no cybersickness, 16 reported slight symptoms, 5 reported moderate symptoms, and 4 reported severe symptoms, *M* = 6.45, SD = 8.95.

Intrinsic Motivation: Intrinsic motivation was measured using the Task Evaluation Questionnaire, a shortened version of Ryan et al’s Post-Intrinsic Motivation Questionnaire (1983) that measured intrinsic motivation using four subscales: interest/enjoyment, perceived choice, perceived competence, and pressure/tension. Participants rated statements across 22 items on a 7-point Likert scale (1= not at all true, 4 = somewhat true, 7 = very true). Cronbach’s alpha for the interest subscale was measured as α = 0.95, competence as α = 0.90, choice as α = 0.80, and pressure as α = 0.80. Example items from this measure included “I found the task very interesting” and “I felt pretty skilled at this task.” In addition to the four subscales, perceived effort was measured using a subscale of the full 45-item IMI measure. Validity for this scale was calculated as α = 0.89. An example item from the effort subscale was “I tried very hard on this activity.”

Presence: Presence was measured using the SUS (Slater-Usoh-Steed, 2000) presence questionnaire. Participants rated their experiences in the virtual environment across six items on a 7-point Likert scale. An example of an item measuring presence was: “To what extent were there times during the experience when the grocery store was the reality for you?” Reliability for presence was measured as α = 0.88.

Performance: Performance was measured as a total point value tallied through correct in-training actions during the five-minute grocery task and was recorded using the program itself. Participants earned a total of 820 points on average (*M* = 820.48), and a table of these correct actions and their associated point values is available in **Table D**

Errors: Errors were measured as incorrect actions taken during the five-minute grocery tasks and were recorded using the program itself. Participants made four errors on average (*M* = 4.27), and a table of these errors is available in **Table E**

Persistence: Persistence was measured by tracking the time users remain in the headset after the five-minute period. Time was captured in seconds using a researcher stopwatch that began at the start of the task and was stopped when the participant stated they were finished using the program. On average, participants stayed an extra 24 seconds in the training before asking to remove the headset (*M* = 24.57).

VR Errors: A measure of glitches, technical errors, or any extraneous variance caused by the headset or computer was captured by the researcher (“0” as no errors, “1” as bugs within the program, “2” requiring the headset to be remounted, and “3” required a complete reset of the training program.)

VR Training Program

Program Description

The virtual reality training program was created using Unity version 2020.3.44f1 and programmed using Microsoft Visual Studio version 17.7. The virtual environment mimicked the viewpoint from behind a checkout register at a grocery store. Elements of this environment included food aisles, a produce section, cash checkouts, an entrance/exit to the store, and virtual customers waiting in line (represented by blue wireframe virtual avatars). The controllers were transformed into two white gloves that acted as virtual hands, allowing users to pick up items and interact with them in the environment. Additionally, a half-foot long ray extended from the palm of each hand to allow users to interact with buttons on the register screen. Users were not able to move from the established one-foot boundary of the headset and had no other means of locomotion within the environment. Images of this environment are available in Figures A, B, and C

The grocery register included a belt on the right where grocery items were spawned, a bay to the left in which scanned items could be placed, and a scan window in the middle which recognized items as they were scanned and weighed. All register information was displayed and updated on the register screen as the user performed actions such as logging in, scanning items, entering produce, or completing transactions. The right half of the screen contained a 12-key number pad (with buttons to clear and enter user input respectively), while the left half displayed scanned items and buttons for certain interactions (e.g., voiding items). To avoid making the training too novel, extra buttons were added in this space that serve no function other than to display an error screen. Additionally, a 'cheat sheet' of all available produce codes (PLU codes) were available to the right of the cashier on a small table for reference.

The virtual training program required users to first complete three blocks of grocery cashier training and then to complete as many full, errorless orders as they can within a five-minute time span. Upon completion of each training block, users saw and grabbed a labeled orb that allowed them to continue to the next block. During the first training block, users were required to use the grab buttons on the controller to pick up a cloth and then to clean the register belt by wiping it across four locations on the belt. Upon completion, they were then asked to log into the register- using the triggers on the controller to activate the number pad on the touch screen. The second training block taught users how to scan items and enter price lookup codes (PLU) for produce items. Items were spawned on the register belt, scanned, and then placed to the left bay. Produce items were entered by first placing the item on the scale, selecting the 'PLU Code' button, and then entering the correct code based on the cheat sheet.

The third training block taught users how to complete an order and two common problem scenarios: being asked to remove an item and manually entering an item's barcode due to failure to scan. Users began this final block by scanning an item, which caused the customer to make a request (visualized in the form of a pop-up expression). Users selected the 'Void Item' button, scanned the item in question, and then moved the item aside. Non-scanning items were entered by selecting the 'UPC Code' button and entering the item code found on the cheat sheet. For the purposes of the training, only one item would not scan- cake mix. This was done to lower the amount of time during the five-minute block that employees would have to meticulously enter codes with the VR controllers. Finally, users selected the "Finalize Transaction" button, prompting them to either accept cash, credit, or EBT/Food Stamps. The customer NPC stated their form of payment through the pop-up expression, and the correct form of payment was selected.

Upon completion of the third training session users started the five-minute test. First, users cleaned their register and logged in to the system- starting a five-minute timer unseen to the trainee in which they were required to complete as many grocery orders as possible. Each grocery order began by spawning 9 pre-selected items: 7 standard grocery items, and two produce items that required PLU codes. After all items were entered, users finalized the transaction, accepted payment, and nine new items were spawned on the grocery belt. Users would also be occasionally prompted with item void requests and occasionally must manually enter a UPC code. This process was continued until the headset is removed and the program was terminated. After the five-minute period is finished, the program takes a snapshot of all available performance data: the number of scanned items, orders completed, and the user's score which was calculated through correct in-program actions.

Conditional Differences

Training content, grocery items, and training blocks were the same between all versions of the training, However, the control and motivational conditions differed within the virtual environment and training block navigation. All three motivation conditions contained a display board floating at eye level that contained contextual information dependent on the condition. Figures representing each respective condition are available in Figures D, E and F.

Self-Determination Theory

Points, trophies, and renovation were the three gamification elements selected for the SDT condition of the virtual training. To implement renovation, an additional orb was displayed at the end of each training block labeled “Repeat”. At the end of each training block users could repeat the training segment to practice their skills without consequence. Once satisfied, users could continue to the next section of the training. Points were implemented with a point tracker on the display board that increased as the user performed correct in-training actions. Users could view this total at any time to track and monitor their progress. Trophies were implemented in two ways: a tracker (similar to points) that displayed the necessary number of orders to obtain the next trophy (e.g., “Two orders to reach Silver”) and a physical trophy that appeared over the register screen as they are earned and remained for the rest of the training.

Goal-Setting Theory

Progression, objectives, and competition were the three gamification elements selected for the GST condition of training. The progression element was implemented using two progress bars visible on both sides of the display board. The left progress bar indicated progress across the entire training and filled as users completed individual training blocks. The right bar indicated progress on completing a full grocery order and filled as users scanned items, entered PLU codes, finalized orders, and collected payment. Once the right bar was completed, it reset to track the next order. Both the objective and competition elements were visible on the display board. To represent competition, a leaderboard was used to show the progress of six competitors. Each competitor has an associated rank, name, and point value associated with them with the ‘Player’ rank (representing the user) at the bottom. As the user completed correct in-training actions, the Player rank rose and overtook the rest. Finally, the objective element was represented using a “Current Objective” section on the visual board that guided the user with sub-goals to complete a full grocery order. For example, the section displayed “Enter PLU Code” after the user selected the PLU Code button from the register and changed to “Scan Items” after the code was entered.

Expectancy Theory

Social Pressure, acknowledgement, and competition were the three gamification elements selected for the ET condition of training. Both the acknowledgement and competition elements were implemented in the same way as the SDT and GST conditions respectively. In addition to these two elements, social pressure was implemented in the form of a "Points to Donut" section on the display board. Before the training, users were informed that the previous participant earned them a donut and that their performance could earn the next user one as well. Users could track how many points were required to earn the next trainee a donut on the visual board. Upon earning enough points, a virtual donut was displayed above the register screen alongside the trophies earned. The donut remained until the training was finished.

Procedure

Participants were first given an informed consent form and were briefed on the risks of cybersickness from using the virtual reality headset. Participants were then informed how to use the headset and controllers before being started into their training program condition. Participants were then guided through each training block before arriving at the five-minute task. The researcher facilitating the training informed the participant to complete as many complete and errorless grocery orders as they could and began a stopwatch that started the moment the participant logged into the register. The participant was informed when the five-minute period was up, but were allowed to remain in the training until they wished to stop. Once the participants stated they were finished, the stopwatch was stopped, and the time was recorded. Finally, participants completed a post-test survey and completed a debrief. This procedure is demonstrated in Figure G.

Results

Primary Analyses

All 126 participants in each of the VR training conditions were successful at using the training program: Participants completed 5 full grocery orders on average and were able to enter PLU items, void items, and accept payment all of which were the learning objectives of the training program. Standard deviations indicated that there was good variability on all the dependent variables, participant scores (SD = 185.5), errors (SD = 3.38), and persistence (SD = 30.58). There were few errors within the program itself (*M* = 0.39, SD = 0.55) and only two errors that required that the headset either be removed temporarily or the program to be restarted fully. In short, we believe the training program created for this study showed acceptable validity based on the face validity of the training program and these simple descriptive statistics. Table C contains descriptive statistics for all scales and the bivariate correlations among study variables.

Hypothesis 1 stated that users taking a gamified version of the grocery store training would have significantly increased performance outcomes when compared to the non-gamified control group. To test this prediction, we first performed a one-way ANOVA with condition as the factor and performance as the dependent variable with the intention to conduct post hoc tests comparing each of the gamification conditions to the control condition if the F-test was significant. Contrary to Hypothesis 1, the omnibus F-test was not significant, $F(3,122) = 2.09$, $p = .106$. To test Hypothesis 1 further, we ran an additional ANCOVA with conscientiousness as a covariate since meta-analytic research has shown conscientiousness to be a predictor of training performance (Blume, 2010). However, this test too was not significant, $F(4,121) = 2.09$, $p = 0.106$. Thus, Hypothesis 1 was not supported.

Hypotheses 2a and 2b stated that users taking a gamified version of the grocery store training would make significantly more or fewer errors when compared to the non-gamified control group. Similar to Hypothesis 1, we performed a one-way ANOVA with condition as the factor and errors as the dependent variable. The omnibus F-test was not significant, $F(3,122) = 1.76$, $p = .158$, and thus post-hoc tests were not performed. We performed an additional ANCOVA using conscientiousness as covariate but this test was also not insignificant, $F(4,121) = 1.76$, $p = 0.158$. Thus, Hypotheses 2a and 2b were not supported.

Hypothesis 3 stated that users taking a gamified version of the grocery store training would persist in the training significantly longer when compared to the non-gamified control group. We performed a one-way ANOVA similar to both Hypothesis 1 and 2, this time using persistence as the DV. The omnibus F-test was not significant, $F(3,122) = 2.38$, $p = .073$, and thus post-hocs were not performed. An additional ANCOVA using conscientiousness as a covariate was also not significant, $F(4,121) = 2.43$, $p = 0.068$. Thus, Hypothesis 3 was not supported.

Hypotheses 4a, 4b, and 4c stated that users in one gamified version of the grocery store training would be superior to the other gamified versions of the training across the three dependent variables. To test this hypothesis, we planned to examine the post hoc tests of differences between different gamification groups for the three dependent variables after observing significant omnibus F-tests. One gamification condition would be deemed superior if participants in that condition scored higher, made fewer errors, or persisted more than the other conditions. Since no omnibus F-tests were significant, no post hoc tests were examined, and thus, hypotheses 4a, 4b, and 4c were also not supported.

Exploratory Analyses

In the interest of further understanding the connection between virtual reality and learning, we performed a series of exploratory analyses in addition to our main analyses. The first exploratory analysis aimed to create a model to best predict performance in the training program. To do this, we performed a series of multiple regressions by selecting factors that were most highly associated with performance. We found three variables that had significant bivariate relationships with performance, were likely antecedents of performance in the program, and were significantly related with performance in a multiple regression model using other predictors; VR Familiarity, Presence, and VR Errors (see Table F for the multiple regression results). Participants in our study who were familiar with VR might have performed better because they had higher mechanical skills from their prior VR experience. Presence, which interestingly was *negatively* associated with performance, might have had a "distraction effect" in that users paid more attention to the virtual environment instead of the training task. Finally, the number of technical errors that occurred during the training program had the strongest association with performance, with more errors being associated with lower performance. These three predictors explained about 12% of the variance in trainee performance in the present study. Interestingly, conscientiousness, an established predictor of trainee performance, was not found to be significant in the present study.

The second set of exploratory analyses sought to understand the factors that contribute to the intrinsic motivation of the participants in this training program. Three variables had significant bivariate relationships with intrinsic motivation, made theoretical sense as antecedents to intrinsic motivation in this study, and were significantly related to intrinsic motivation when included in a multiple regression model with the other predictors: agreeableness, VR familiarity, and presence (Results available in Table G). Participants who were familiar with using VR did not have to learn how to use it, and may felt the most engaged and intrinsically motivated. Agreeableness is a trait that is generally associated with intrinsic motivation (Ariani, 2013) and the data suggest this was true for our training program as well. Thus our final model of intrinsic motivation consisted of two predictors specific to virtual reality and one general predictor of intrinsic motivation, predicting 24% of the variance in intrinsic motivation

Our final exploratory analysis examined potential factors that influenced user perceptions of competence after the training. Perceived competence has been identified as a predictor of intrinsic motivation (Li, 2005) and as a potential mediator in performance (McEnrue, 1984). Therefore, creating a training environment that increases perceptions of competency is important for trainee success. We hypothesized that in our trainee program, program errors or glitches may have caused users to feel responsible for causing them and thereby reducing their perceptions of competency and subsequently performance. To test this, we performed a mediation analysis using bootstrapped confidence intervals (95%) using the *lavaan v0.6.16* package in R, predicting perceived competence by using VR program errors as the sole predictor and performance as the mediator. The results of the mediation analysis are available in Figure H. There was a significant indirect effect of VR program errors on perceived competence through performance, b = -0.080, 95% BCa CI [-0.162, -0.028], supporting our hypothesis

Discussion

Our study tested the effectiveness of gamification that was informed and designed by motivational theories on VR-based learning. To accomplish this, we first began by selecting a task (a grocery store cashier), assessing which motivational theories were appropriate for the training's design, and then narrowing down a list of gamification elements that could be applicable per each theory. We then created different versions of the training to represent each theory and compared each against a control. In addition to our main analyses, we performed exploratory analyses on the factors that affected performance, motivation, and perceived competence during the training program. The results of our main analyses did not support our predictions; there were no significant differences among conditions on performance, error rate, or persistence, a behavioral proxy for motivation. While previous studies on the effects of gamification on performance have been inconsistent (e.g., Makransky et al, 2021; Hicks, 2019; Chodan et al, 2017), researchers suggest that giving consideration to how game elements are selected and implemented may help to clarify this inconsistency (Howard et al., 2017). We used motivational theory as the lens to select and implement elements, but, interestingly, were unable to find an effect.

A potential explanation for this finding is that gamification operates differently in a virtual environment, and thus game elements which target intrinsic motivation alone are not sufficient for a significant difference in performance. Gamification researchers such as Kapp (2012) have advocated that a holistic approach should be used when implementing gamification, and that in addition to motivation, researchers should implement gamification in ways that can “engage people, promote learning, and solve problems.” (Kapp, 2012). While motivational theories may help us to select and implement game elements to increase intrinsic motivation, elements that address different antecedents of performance may also be necessary in order to facilitate the success of gamification in a VR context. For example, a study using a gamified VR program that employed a framework for experiential learning was unable to find effects for performance when compared to a non-gamified group but found significant increases to engagement and enjoyment in the VR-gamified group (Alrehaili & Osman, 2022). In addition to targeting performance-related attitudes like engagement, game elements which provide stimulating feedback may lead to higher performance as well. A study using SDT examining the effects of game elements with ‘juiciness’ (those with highly stimulating visual and audio feedback) were unable to find effects for gamification on performance but discovered that elements with more ‘juiciness’ (such as the sensation game element) were associated with increases to perceived competence- a known predictor of performance (Hicks, 2019). In short, it might be the case that gamification elements need to be engaging and distinctive to stand out in their virtual environment in order to be effective. Ultimately, while motivational theory may provide a solid foundation, it might not fully encompass the requirements for gamification’s success in a VR context. We suggest that a follow-up study examine this possibility, by implementing elements that not only target intrinsic motivation but also additional performance antecedents, and simultaneously examine the impact of salience of the elements on performance.

Our exploratory analyses identified factors that best predicted performance and other associated factors in our training program. One factor, presence, was identified as a negative predictor for task performance. Contrary to our findings, other research has discovered null associations between presence and task performance (van Baren and Ijsselsejn, 2004.) While this relationship continues to be examined, it has been hypothesized that while presence does not facilitate performance, “having some sense of presence in an environment is a necessary condition for performance to occur.” (Bystrom et al., 1999). One potential explanation for this negative association is Adams et al.’s “distraction hypothesis” (2012), which suggests that game elements may distract users from their environment or training goals if implemented suboptimally. The game elements we implemented may have caused users to feel *too* present, and thus not pay attention to the grocery task. This may suggest an optimal range for presence, in which too much may cause the user to feel distracted or overwhelmed and diminish performance while too little may prevent the user from feeling engaged and therefore prevent performance from occurring. However, contrary to this line of reasoning, we did not find a difference between our control group and experimental groups on presence, indicating that the game elements we selected did not appear to have any influence. Therefore, the mystery of the effects of presence are still unknown, suggesting that more research is needed on how presence operates within a virtual environment.

Our analyses also highlighted the importance of VR familiarity on both task performance and intrinsic motivation. This makes logical sense as users who are familiar with VR can allocate more cognitive resources to the training task instead of learning how to use the medium. This highlights the importance of ensuring that trainees understand how to use VR before training in order to reduce the potential variance from learning it. For example, a medical study used a fifteen-minute acclimation period to ensure users were comfortable in the environment before proceeding with the surgical exercise, and found significant increases to engagement, performance, and training length when compared to a non-VR group. A “VR training period” gives newer VR users time to familiarize themselves with the medium and may increase performance. Our study provides additional support for the importance of ensuring VR familiarity when conducting VR training.

Another interesting finding was that conscientiousness, one of the best predictors of training success (e.g., Schmidt & Hunter, 1998; Blume et al, 2010), was conspicuously absent as a significant predictor of performance in our study. While sampling error is a possible explanation, this finding is consistent with a result found by Thorp et al., (2023) who also found null results despite having adequate power in their study. Thorp et al., suggested that users high in conscientiousness tend to prefer organization and planning, and that the chaotic nature of some virtual environments (such as a fast-paced grocery store) may cause users to attribute more cognitive resources in keeping things organized instead of focusing on the training task. Our findings are consistent with Thorp et al.’s explanation, although more research is needed to determine if conscientiousness is generally not a predictor of trainee performance in VR environments as it is in non-VR ones.

Finally, we found evidence that errors in a VR training program might impact a trainee’s perceived competence through task performance. This suggests that, in our study, users may have attributed deficiencies in their performance because to their own abilities rather than errors caused by the program. If this is found to be generally true, it highlights the importance of pilot testing VR training programs so that users have a good first experience with them, feel competent, and are more likely to continue to use those VR trainings in the future.

Contributions

Despite the null results of the main analyses, the current investigation makes multiple contributions to the understanding of VR-based learning. First, we examined the effectiveness of applying motivational theory to gamification design in VR training. A meta-analysis performed by Howard et al. tested the effectiveness of gamification by using a non-gamified condition of a training against a gamified one in a VR context. In the discussion section, Howard et al. (2021) pointed out that without a more granular approach to gamification, it would be difficult to understand why it works in some contexts but not others (Howard, 2021). To address this, our study used three motivational theories as a framework to justify and select individual game elements but was unable to find an effect for increasing performance. We suggest that gamification operates differently in a virtual environment, and that researchers should consider implementing additional elements to address other performance-related factors or ones that are highly stimulating in order for them stand out. Researchers may also consider using additional frameworks to select these elements, such as Robson's gamification principles (2015) or Zaric et al.'s gamified learning theory (2021).

Second, we have established a methodology for selecting motivational theories and subsequently game elements that align with the theories. In practice, game elements are often selected for use based on popularity (e.g. the 'PBL' triad of points, badges, and leaderboards is a commonly used combination) or seemingly at random. Additionally, multiple game elements are available and subset are used, and so choosing which game elements are effective leads to the "numbers problem", i.e., the number of game element combinations is too large. To address this, we used a methodical approach by narrowing down potential theories while discarding ones that were incongruous within a gamification context. Game elements that were selected addressed the needs of each motivational theory and were feasible to implement based on the training task. Thus, we have provided a roadmap for future researchers to use in connecting psychological theory to gamification elements if desired. Should future studies clarify the use cases of gamification in VR, the work described in Tables A and B provides an example of how to go from psychological theories to specific gamification elements.

Our final contribution is examining the effects of VR program errors and how they affect performance and perceived competence. VR is a relatively new medium for organizational training and thus training programs may act unpredictably due to programming errors or technical bugs. These technical issues may arise when trainees are taking the training and could affect performance or related factors. We identified these program errors as a negative predictor of performance in our training program. Additionally, results of our mediation analysis indicated that program errors do not directly affect perceived competence, but rather through an indirect effect mediated by performance. We suggest that future research should explore these program errors to better understand their effects while the medium continues to be developed.

Limitations

The current study had a number of limitations that should be acknowledged. First, in order to obtain a large sample to have sufficient power for our analyses, it was necessary to expand our data collection to a second location, a board game store, in addition to our original location of a university campus. Ideally, data collection would be done in a single location with the same number of participants specified by our power analysis. While we balanced conditions across location, this extraneous variable could theoretically explain the null findings of our main analysis. To address this, we performed the main analyses again separated by location and found the same pattern of results as our analyses run at both locations, thereby suggesting that running our data at two locations did not impact the conclusions. Additionally, we did not observe any significant difference between locations on performance, error rate, or persistence. Nonetheless, we cannot rule out the possibility that running the experiment in two locations impacted the results.

Our second limitation was the placement of gamification elements within the virtual environment. Gamification elements were placed behind the register screen near the next grocery checkout (see Figure D). While administering the training, we often found that participants would not assess their progress using the training board and instead pay attention only to the grocery task. Users may have missed these elements during the training and the potential benefits of gamification may have been missed entirely. For example, points served as a feedback indicator and failing to see them may have reduced the potential benefits of intrinsic motivation according to SDT. To address this, and consistent with the discussion above about the potential importance of element salience in VR, we recommend that future studies place game elements in a way that is always presented to the user, without being too distracting. Designers may consider a HUD, or “heads-up display”, game element that is visible at all times to the user for the purpose of relaying information to the user.

Our third limitation is that the training program only targeted a single occupation, and therefore the findings might not generalize to other occupations. Additionally, selecting a different occupation may not only result in different findings, but may also change which motivational theories are applicable. For example, a position that requires the employee to interact with multiple coworkers may find Bandura’s Social Cognitive Theory (1986) to be more applicable when selecting gamification elements. While we acknowledge other jobs could have been used instead, we believed that the grocery cashier role was a good fit for the purposes of our study. For one, the cashier task may be easily manipulated to match the needs of an experimental study on gamification. Furthermore, the position of a grocery store cashier is common across the world and will remain so in the future (albeit at reduced numbers) because even with self-checkout registers, it is still necessary to train cashiers. Thus the study has good ecological validity. While we believe that the grocery task was an appropriate starting task, we encourage other researchers to examine the generalizability of our findings.

Fourth, there was subjectivity in terms of connecting gamification elements to each motivational theory. We first selected elements from Toda et al.’s taxonomy (2019) based on how well they best addressed the requirements of each theory (e.g., points served as feedback to address SDT’s need for competency.) Then, we considered which of these elements would best be implemented for the purposes of the grocery store training task. The results of our selection (available in Table B) show each game element and our justification for why we selected the elements used in our final program. While matching elements was a subjective process, we believe that other researchers would arrive at similar conclusions using the methodological approach in matching elements that we did. However, This belief should be examined in a future study.

Our final limitation was the use of a non-standard training program for the grocery cashier task. We created our own grocery store training program as we were unable to use an established one. When designing our training program, we proceeded in a systematic fashion to improve its validity: We built our training environment to best match current grocery stores to the best fidelity possible - for example, we matched product codes and register screens to their real-world counterparts. We also a subject matter expert (i.e., someone who worked as a cashier) involved in creating the program and referenced current grocery store registers. Additionally, performance indicators from the training program indicated that our program was successful-users were able to perform each part of the grocery task and completed the target number of orders, on average. Ultimately, while we believe that our grocery training was sufficient enough for our study, we recommend that future researchers use existing validated training programs if possible.

Conclusion

In this study, we answered the call by Howard et al. (2021) to test gamification in a VR environment by using motivational theories to select game elements. We were unable to find an effect when using game elements to increase intrinsic motivation by addressing the needs of each theory. While our research did not find a significant effect of gamification on performance, we believe that further research is needed before abandoning the idea that gamification is useful in a VR environment. We suggest, based on the current study's results and a re-examining of the extant studies on gamification in VR, that gamification's inconsistent efficacy in VR across studies may be caused by attributes of virtual environments (i.e., if everything stands out, nothing stands out), how elements are implemented (e.g., elements may only be effective if they are visually stimulating), or properties of our study (e.g., where elements were placed in VR may have affected their success.) -- in short salience might be key to successfully implementing gamification in VR. Research should continue in continuing Howard's call for a careful examination of VR, and that examining gamification by using elements to target salience and antecedents of motivation may help us to understand what is generally true about its role in the virtual world.

References

Abich IV, J., Parker, J., Murphy, J. S., & Eudy, M. (2021). A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality*, *25*(4), 919–933.

Adams, D. M., Mayer, R. E., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology*, *104*(1), 235.

Adams, J. S., & Freedman, S. (1976). Equity theory revisited: Comments and annotated bibliography. *Advances in Experimental Social Psychology*, *9*, 43–90.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211.

Alaker, M., Wynn, G. R., & Arulampalam, T. (2016). Virtual reality training in laparoscopic surgery: a systematic review & meta-analysis. *International Journal of Surgery*, *29*, 85–94.

Albert, B. (1986). Social foundations of thought and action: A social cognitive theory. *NY: Prentice-Hall*.

Alrehaili, E. A., & Al Osman, H. (2022). A virtual reality role-playing serious game for experiential learning. *Interactive Learning Environments*, *30*(5), 922–935.

Ariani, D. W. (2013). Personality and learning motivation. *European Journal of Business and Management*, *5*(10).

Atkinson, J. W., & Raynor, J. O. (1974). Motivation and achievement. *(No Title)*.

Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, *41*(1), 63–105.

Barnes, S. (2016). Understanding virtual reality in marketing: Nature, implications and potential (November 3, 2016). *URL: Https://Ssrn. Com/Abstract*, *2909100*.

BERKMAN, M., ÇATAK, G., & EREMEKTAR, M. (2020). Comparison of vr and desktop game user experience in a puzzle game:“keep talking and nobody explodes.” *AJIT-e: Academic Journal of Information Technology*, *11*(42), 180–204.

Bhagat, K. K., Liou, W.-K., & Chang, C.-Y. (2016). A cost-effective interactive 3D virtual reality system applied to military live firing training. *Virtual Reality*, *20*, 127–140.

Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, *36*(4), 1065–1105.

Bowman, D. A., & McMahan, R. P. (2007). Virtual reality: how much immersion is enough? *Computer*, *40*(7), 36–43.

Bystrom, K.-E., Barfield, W., & Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *8*(2), 241–244.

Cameron, C. (2010). Military-grade augmented reality could redefine modern warfare. *New York Times*, *12*.

Cheng, Y., & Wang, S.-H. (2011). Applying a 3D virtual learning environment to facilitate student's application ability–The case of marketing. *Computers in Human Behavior*, *27*(1), 576–584.

Chittaro, L., & Buttussi, F. (2015). Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety. *IEEE Transactions on Visualization and Computer Graphics*, *21*(4), 529–538.

Deterding, S. (2014). *Eudaimonic design, or: Six invitations to rethink gamification*.

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

Fortes Tondello, G., Premsukh, H., & Nacke, L. (2018). *A theory of gamification principles through goal-setting theory*.

Girvan, C. (2018). What is a virtual world? Definition and classification. *Educational Technology Research and Development*, *66*(5), 1087–1100.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly*, 213–236.

Hamzah, W. A. F. W., Ali, N. H., Saman, M. Y. M., Yusoff, M. H., & Yacob, A. (2014). Enhancement of the ARCS model for gamification of learning. *2014 3rd International Conference on User Science and Engineering (i-USEr)*, 287–291.

Hicks, K., Gerling, K., Richardson, G., Pike, T., Burman, O., & Dickinson, P. (2019). Understanding the effects of gamification and juiciness on players. *2019 IEEE Conference on Games (CoG)*, 1–8.

Howard, M. C. (2017). A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Computers in Human Behavior*, *70*, 317–327.

Howard, M. C., & Gutworth, M. B. (2020). A meta-analysis of virtual reality training programs for social skill development. *Computers & Education*, *144*, 103707.

Howard, M. C., Gutworth, M. B., & Jacobs, R. R. (2021). A meta-analysis of virtual reality training programs. *Computers in Human Behavior*, *121*, 106808.

Huang, C.-M., Wang, S.-C., Chang, C.-F., Huang, C.-I., Cheng, Y.-S., & Fu, L.-C. (2004). An air combat simulator in the virtual reality with the visual tracking system and force-feedback components. *Proceedings of the 2004 IEEE International Conference on Control Applications, 2004.*, *1*, 515–520.

Jang, S., Vitale, J. M., Jyung, R. W., & Black, J. B. (2017). Direct manipulation is better than passive viewing for learning anatomy in a three-dimensional virtual reality environment. *Computers & Education*, *106*, 150–165.

Jerald, J. (2015). *The VR book: Human-centered design for virtual reality*. Morgan & Claypool.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Kavanagh, S., Luxton-Reilly, A., Wuensche, B., & Plimmer, B. (2017). A systematic review of virtual reality in education. *Themes in Science and Technology Education*, *10*(2), 85–119.

Kim, H. K., Park, J., Choi, Y., & Choe, M. (2018). Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, *69*, 66–73.

Kim, T. W., & Werbach, K. (2016). More than just a game: ethical issues in gamification. *Ethics and Information Technology*, *18*(2), 157–173.

Lee, S. H., Sergueeva, K., Catangui, M., & Kandaurova, M. (2017). Assessing Google Cardboard virtual reality as a content delivery system in business classrooms. *Journal of Education for Business*, *92*(4), 153–160.

Lele, A. (2013). Virtual reality and its military utility. *Journal of Ambient Intelligence and Humanized Computing*, *4*, 17–26.

Li, L., Yu, F., Shi, D., Shi, J., Tian, Z., Yang, J., Wang, X., & Jiang, Q. (2017). Application of virtual reality technology in clinical medicine. *American Journal of Translational Research*, *9*(9), 3867.

Li, W., Lee, A. M., & Solmon, M. A. (2005). Relationships among dispositional ability conceptions, intrinsic motivation, perceived competence, experience, and performance. *Journal of Teaching in Physical Education*, *24*(1), 51–65.

Locke, E. A., & Latham, G. P. (2006). New directions in goal-setting theory. *Current Directions in Psychological Science*, *15*(5), 265–268.

Locke, E. A., & Latham, G. P. (2013). *Goal setting theory, 1990.*

Lu, X., & Davis, S. (2016). How sounds influence user safety decisions in a virtual construction simulator. *Safety Science*, *86*, 184–194.

Luthans, F., & Stajkovic, A. D. (1999). Reinforce for performance: The need to go beyond pay and even rewards. *Academy of Management Perspectives*, *13*(2), 49–57.

Makransky, G., Andreasen, N. K., Baceviciute, S., & Mayer, R. E. (2021). Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality. *Journal of Educational Psychology*, *113*(4), 719.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, *60*(1), 48–58.

McEnrue, M. (1984). Perceived competence as a moderator of the relationship between role clarity and job performance: A test of two hypotheses. *Organizational Behavior and Human Performance*, *34*(3), 379–386.

Microsoft (2023) Microsoft Visual Studio v17.7. Microsoft

Mirvis, P. H. (1991). *Flow: The psychology of optimal experience*. JSTOR.

O*Net. (2023). *O*NET*. Retrieved from O*NET Online: https://www.onetonline.org/link/summary/41-2011.00

Palmas, F., Labode, D., Plecher, D. A., & Klinker, G. (2019). Comparison of a gamified and non-gamified virtual reality training assembly task. *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, 1–8.

Patel, A. D., Gallagher, A. G., Nicholson, W. J., & Cates, C. U. (2006). Learning curves and reliability measures for virtual reality simulation in the performance assessment of carotid angiography. *Journal of the American College of Cardiology*, *47*(9), 1796–1802.

Prasad, K., & Mangipudi, M. R. (2021). Gamification Mechanics: A Framework for Business Perspective. *Turkish Online Journal of Qualitative Inquiry*, *12*(4).

PwC. (2019). *Seeing is believing.* PwC.

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons*, *58*(4), 411–420.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68.

Sacks, R., Perlman, A., & Barak, R. (2013). Construction safety training using immersive virtual reality. *Construction Management and Economics*, *31*(9), 1005–1017.

Sattar, M. U., Palaniappan, S., Lokman, A., Hassan, A., Shah, N., & Riaz, Z. (2019). Effects of virtual reality training on medical students' learning motivation and competency. *Pakistan Journal of Medical Sciences*, *35*(3), 852.

Schmid Mast, M., Kleinlogel, E. P., Tur, B., & Bachmann, M. (2018). The future of interpersonal skills development: Immersive virtual reality training with virtual humans. *Human Resource Development Quarterly*, *29*(2), 125–141.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Slater, M., Linakis, V., Usoh, M., & Kooper, R. (1996). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 163–172.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, *6*(6), 603–616.

Strickland, J. (2007). How virtual reality military applications work. *Phương Thức Áp Dụng Công Nghệ Áp Dụng Vào Quốc Phòng. HowStuffWorks. Com.(Https://Science. Howstuffworks. Com/Virtual-Military. Htm). Truy Cập Tháng*, *10*.

Thorp, S. O., Rimol, L. M., & Grassini, S. (2023). Association of the Big Five Personality Traits with Training Effectiveness, Sense of Presence, and Cybersickness in Virtual Reality. *Multimodal Technologies and Interaction*, *7*(2), 11.

Toda, A. M., Oliveira, W., Klock, A. C., Palomino, P. T., Pimenta, M., Gasparini, I., Shi, L., Bittencourt, I., Isotani, S., & Cristea, A. I. (2019). A taxonomy of game elements for gamification in educational contexts: Proposal and evaluation. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, *2161*, 84–88.

Ulmer, J., Braun, S., Cheng, C.-T., Dowey, S., & Wollert, J. (2020). Gamified virtual reality training environment for the manufacturing industry. *2020 19th International Conference on Mechatronics-Mechatronika (ME)*, 1–6.

Unity (2020) Unity v.2020.3.44f1. Unity Technologies

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using presence questionnaires in reality. *Presence*, *9*(5), 497–503.

Vallerand, R. J. (2000). Deci and Ryan’s self-determination theory: A view from the hierarchical model of intrinsic and extrinsic motivation. *Psychological Inquiry*, *11*(4), 312–318.

Van Baren, J. (2004). Measuring presence: A guide to current measurement approaches. *Deliverable of the OmniPres Project IST-2001-39237*.

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, *10*(4), 293–309.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational effects on self-regulated learning with different tasks. *Educational Psychology Review*, *18*(3), 239–253.

Vroom, V. H. (1964). *Work and motivation.*

Webster, R. (2016). Declarative knowledge acquisition in immersive virtual learning environments. *Interactive Learning Environments*, *24*(6), 1319–1333.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in Psychology*, *10*, 158.

Wismeijer, A. A., & Vingerhoets, A. J. (2005). The use of virtual reality and audiovisual eyeglass systems as adjunct analgesic techniques: a review of the literature. *Annals of Behavioral Medicine*, *30*(3), 268–278.

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence*, *7*(3), 225–240.

Zaric, N., Roepke, R., Lukarov, V., & Schroeder, U. (2021). Gamified Learning Theory: The Moderating role of learners' learning tendencies. *International Journal of Serious Games*, *8*(3), 71–91.

Figure A)

Views of grocery store within VR Environment

Figure B)

Views of Register Screen, PLU cheat sheet and entering PLU codes

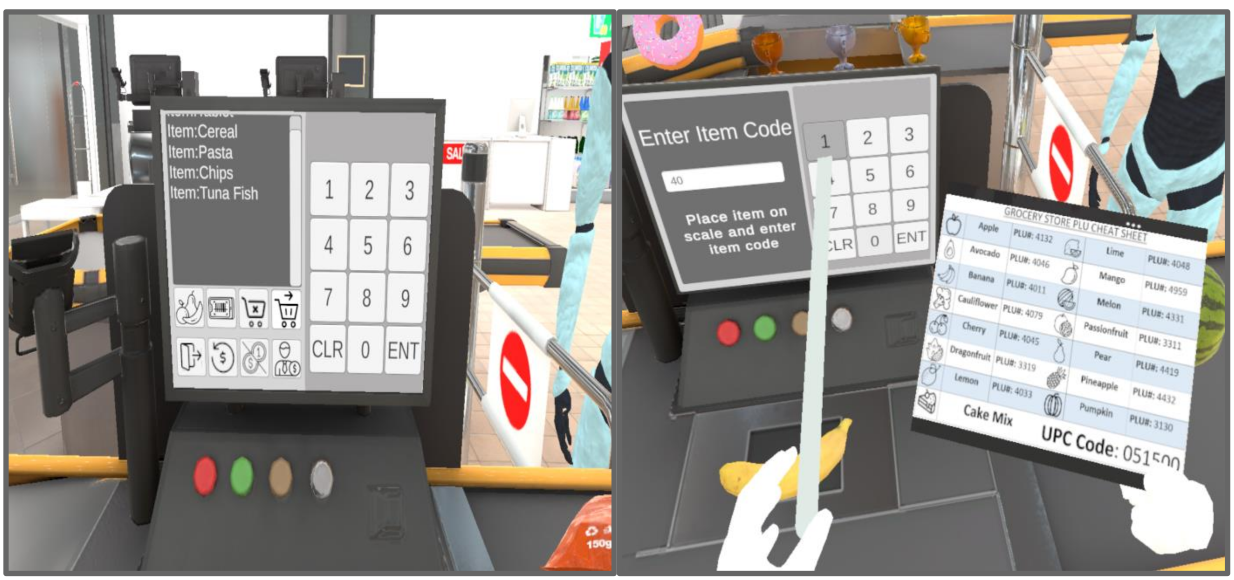

Figure C)

View of the register buttons and NPC void request

Figure D)

View of self-determination theory condition training

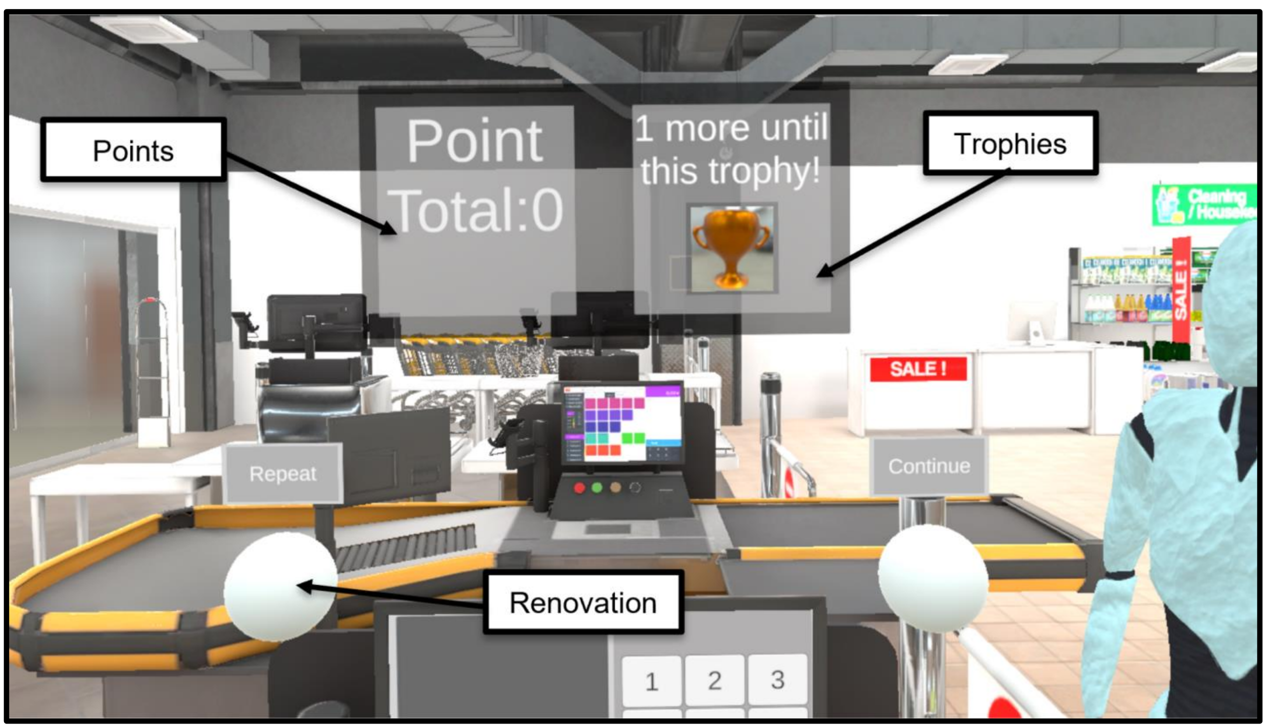

Figure E)

View of goal-setting theory condition training

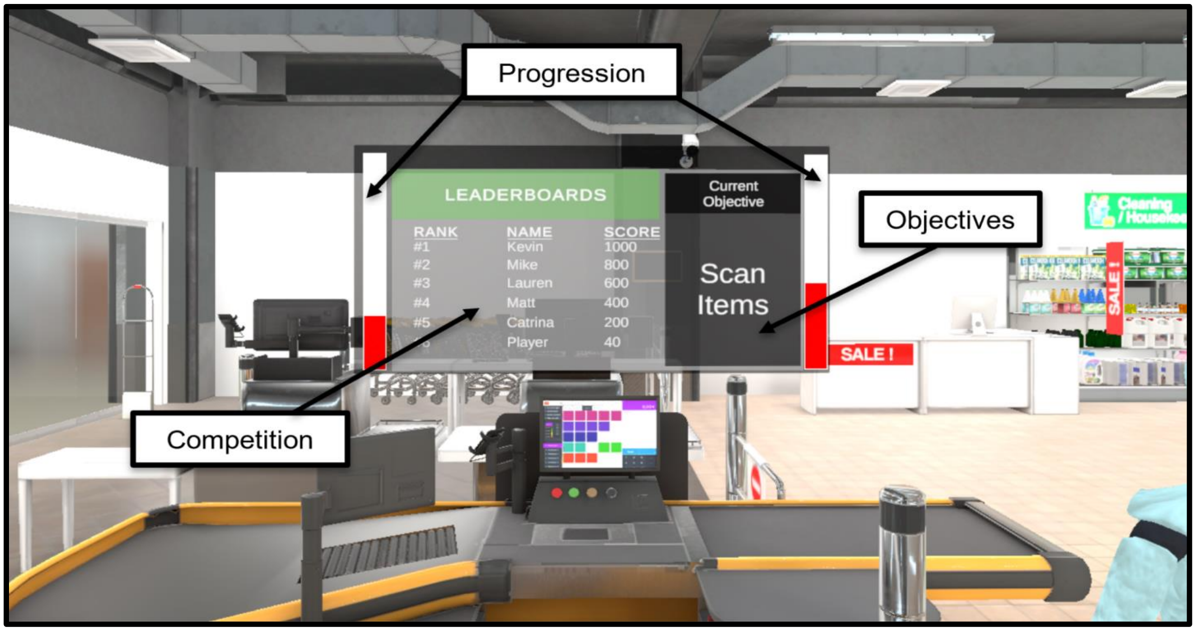

Figure F)

View of expectancy condition training

Figure G

Measures captured during each step of the training process.

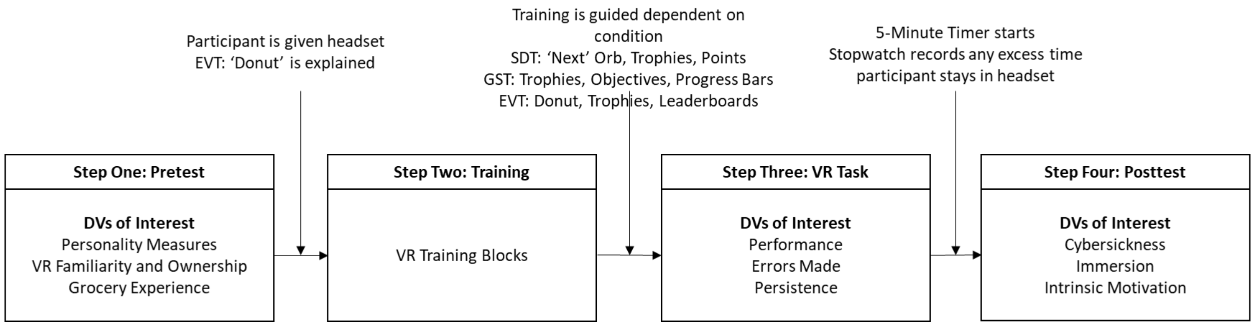

Figure H

Model of VR program errors as a predictor of perceived competence, mediated by performance.

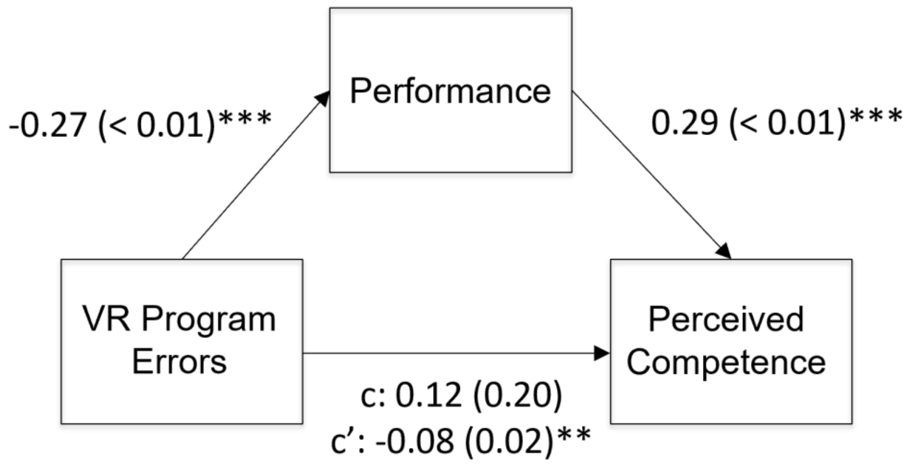

Table A)

Motivational Theories considered for use

Name of Theory	Choosing Theory
Self-Determination Theory (Deci & Ryan, 2000)	Gamification utilizes intrinsic motivation which is one of the core features of SDT. Additionally, each of the three needs of SDT can be addressed via different elements in a meaningful way
Goal-Setting Theory (Locke & Latham, 1990)	Performance goals can be set or suggested via gamification elements. Game elements can be selected to provide goal progress indicators and can allow users to course correct if needed
Expectancy Theory (Vroom, 1964)	Gamification can directly affect the valence and expectancy components of the theory. Elements can increase user confidence to increase task involvement and can increase satisfaction to develop intrinsic satisfaction of taking the training
Equity Theory (Adams 1976)	Equity theory states that individuals are motivated by a desire for fairness in their social relationships. We were unable to make predictions on why any game element is considered more fair than the rest.
Social Cognitive Theory (Bandura, 1986)	Social Cognitive Theory suggests that learning occurs through observation in a social context. We could not find a way to implement social game elements (outside of leaderboards) in a meaningful way based on the single-user design of the grocery simulation
Reinforcement Theory (Luthans, & Stajkovic,1999)	Reinforcement theory is a theory that suggests that reinforcers can be used to increase the likelihood an individual will exhibit a behavior. While aspects of the theory give guidance on design principles (eg: variable reinforcers are more effective than fixed reinforcers), they cannot make predictions on which elements are better reinforcers than others.
Theory of Planned Behavior (Ajzen, 1991)	Theory of Planned Behavior states that subjective norms and attitudes and behavioral control can predict an individual's intent to exhibit a behavior which in turn influences their behavior. Because users already volunteered to take part in the training, the theory cannot provide guidance on how to motivate *during* the training. Instead, we suggest this theory be used to promote intention to use the training or make predictions on gauge interest in applying to a cashier role.
Flow Theory (Csikszentmihalyi, 1990)	Flow theory suggests tasks that users enter a ‘flow state’ when they have optimally difficult tasks that allows them to optimally demonstrate their ability, Users in flow states feel fully immersed and are more engaged while participating in their task. While this theory provides guidelines on design principles (ie: how best to design for flow), reaching flow state depends on individual characteristics and we were unable to use this theory to predict which element promoted flow best.

Table B

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Acknowledgement	All kind of feedback that praises the players' specific actions	SDT	Acknowledgements could increase a trainee's sense of competence. SDT predicts that competence increases intrinsic motivation and performance.	Acknowledgements can be implemented in a supermarket cashier training by using trophies.	Self-Determination Theory (Trophies)
		GST	Acknowledgement may increase individual performance if goals are set to attain that acknowledgement	While GST predicts acknowledgement as an effective element, they may not be effective unless goals are set to attain them. Instead, we chose three different elements that were more aligned with explicity set goals (e.g., progress bars)	No
		EVT	Providing players with a feedback system like acheivements or trophies will increase trainee expectancy, which EVT predicts will increase motivation	Acknowledgements can be implemented in a supermarket cashier training by using trophies.	Expectancy Theory (Trophies)
Chance	Randomness and probability characteristics to increase or decrease the odds of certain actions or outcomes	SDT	No clear prediction	May negatively influence instrumentality according to EVT. Was not predicted by other theories	No
		GST	No clear prediction	Was not predicted by theory	No
		EVT	EVT predicts chance may decrease motivation because instrumentality (belief that performance will lead to desired outcomes) is lowered by the addition of randomness	Was not predicted by theory	No
Competition	When two or more players compete against each other towards a common goal	SDT	Competitive elements allow users to connect with fellow teammates against others, addressing the need for relatedness	Competition in teams allows for users to relate to fellow teammates against other teams. While team competition may be beneficial in group training (i.e., trainees are placed in groups and individual performance adds to a team score), team competition is not possible in a single use cashier training.	Goal Setting Theory (Leaderboards)
		GST	Competitive game elements could provide the basis for and encourage the setting of specific goals (e.g., reach 3rd place on a leaderboard)	Leaderboards allow for participants to set implicit goals to achieve a certain leaderboard rank.	No
		EVT	Competion may increase trainee expectancy as they achieve ranks on a leaderboard. EVT predicts that as expectancy increases, so does motivation	Leaderboards with easily attainable ranks on the cashier task may increase user expectancy as they are reached, subsequently increasing trainee motivation	Expectancy Theory (Leaderboards)
Cooperation	When two or more players collaborate to achieve a common goal.	SDT	Element that allow users to feel connected with each other, meeting the need for relatedness	Not viable for a single-user grocery store training. May be better suited to teach grocery/bagger dyads instead	No
		GST	No clear prediction	Was not predicted by theory	No
		EVT	No clear prediction	Was not predicted by theory	No

Table B (cont.)

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Economy	Transactions within the game, monetising game values and other elements.	SDT	No clear prediction	Was not predicted by theory	No
		GST	Specific goals may be set for trainees to earn an amount of in-game currency or item	This gamification element might be beneficial if a more comprehensive grocery store training that involves multiple days and multiple duties.	No
		EVT	Rewards or items from the economy element may be valued by the trainee. EVT predicts that as value increases, so does motivation	This gamification element might be beneficial if a more comprehensive grocery store training that involves multiple days and multiple duties.	No
Imposed Choice	Decisions that the player is obliged to make in order to advance the game.	SDT	Choice elements allow for the trainee to make their own decisions, meeting the need for autonomy	Grocery tasks were simple enough and did not require meaningful choices from the trainee	No
		GST	No clear prediction	Was not predicted by theory	No
		EVT	No clear prediction	Was not predicted by theory	No
Level	Hierarchical layers present in a game, which provide a gradual way for the player to obtain new advantages	SDT	Levels may increase a trainee's sense of competence. SDT predicts that when competence increases, so does intrinsic motivation	Design of the training did not feature any way to obtain advantages or 'upgrades' while performing basic grocery tasks on a single-use training. This element may be beneficial in more comprehensive trainings that involves multiple days and multiple duties	No
		GST	Levels may be set as a specific or difficult goal for players to reach	Was not predicted by theory	No
		EVT	Advantages from levels may cause trainees to perceive them as more valuable. EVT predicts that as this value increases, so does motivation	Was not predicted by theory	No
Narrative	Order of events where they happen in a game that are influenced by player actions	SDT	No clear prediction	We believe grocery training task is simple enough that narratives could not clarify them any further	No
		GST		Was not predicted by theory	
		EVT		Was not predicted by theory	

Table B (cont.)

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Novelty	Updates or changes that occur within the training environment (e.g., new information or new game elements)	SDT	No clear prediction	Was not predicted by theories	No
		GST			No
		EVT			No
Objectives	Guide the players' actions through goals and subgoals	SDT	No clear prediction	Was not predicted by theory	No
		GST	Objectives may increase motivation by clarifying goals. GST predicts that specific goals can increase motivation	Objectives can break the grocery store tasks into specfic goals for the trainee. GST predicts users will be more motivated to complete specific tasks	Goal Setting Theory (Objectives)
		EVT	No clear prediction	Was not predicted by theory	No
Point	Unit used to measure users' performance	SDT	Receiving points may increase feelings of competence, which SDT predicts increases intrinsic motivation	Trainees will earn high values of points by completing grocery tasks and will see their point total in front of the register.	Self-Determination Theory (Points)
		GST	Points may be motivating if reaching a specific or difficult point value is set as a trainee goal	Points are not inherently motivating unless a certain number of points is set as a goal. While we could have used this element by pairing it with an explicit point-value goal, we instead chose elements that either used an explicit goal (e.g., completing an objective), provided the ability for implicit goal-setting (e.g., filling a progress bar), or encouraged the user to set goals for themselves without a prior understanding of how the element works(e.g., reaching a rank on a leaderboard)	No
		EVT	Points provide feedback to the user to increase perceptions of expectancy on the task. EVT predicts as this perception increases, so does motivation	While points may increase perceptions of expectancy to the user, they have no instrinic meaning outside of a value. We have instead decided to use Acknlowedgement element as it has more inherent meaning (i.e., trainees know a trophy is for completing an amount of orders, whereas points are earned for completing any step of the task)	No
Progression	Alows players to locate themselves (and their progress) within a game.	SDT	Progression elements may increase feelings of competence as users can visuzalize their progress (e.g.., through progress bars). SDT predicts as the need of competence is met, so does intrinsic motivation	Progression elements may increase feelings of competence from trainees as bar is filled, but are so easily attainable that they are not meaningful representations of trainee progress	No
		GST	Progress bars may provide feedback on progress toward goal attainment. Trainees may set implicit goals to fill these bars	Progress bars can show progress on entire training and progress on individual orders. Trainees can set goals to complete training and set simpler goals to finish orders.	Goal Setting Theory (Progress Bars)
		EVT	Progress bars may provide feelings of trainee expectancy as they are filled. EVT predicts as expectancy increases, so does motivation	Progression elements may increase expectancy from trainees as bar is filled, but are so easily attainable that they are not meaningful representations of trainee progress	No

Table B (cont.)

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Puzzles	Challenges within the game that should make a player think	SDT	Meets the need for competence by allowing for optimal challenges to the user to feel difficult by not overwhelming	Grocery tasks were not complex enough to meaningful puzzles without making the tasks arbitrarily more difficult	No
		GST	Addresses the task strategy mediator by allowing user to practice skills through puzzle solving before attempting more difficult tasks.	Was not predicted by theory	No
		EVT	No clear prediction	Was not predicted by theory	No
Rarity	Limited resources and collectables	SDT	No clear prediction	Limited resources or collectables could not be implemented in a meaningful way in a short-term (5 minute) training. This element may have more value in repeated-use trainings or those that encourage trainee exploration	No
		GST	Specific or difficult goals may be set for users to attain specific items or resources within the training.	Was not predicted by theory	No
		EVT	Trainees may place more value on resources or collectables within the training, which EVT predicts increases motivation	Was not predicted by theory	No
Renovation	When players are allowed to redo/restart an action	SDT	Renovation gives trainees the 'freedom to fail', allowing them to act in the training without consequence. This may increase feelings of autonomy on the task, which subsequently increase motivation as predicted by SDT (Kam & Umar, 2018)	Trainees will have the opportunity to retry elements of the training after completing each portion. This may increase perceptions of autonomy during the training to increase motivation	Self-Determination Theory (Renovation)
		GST	No clear prediction	Was not predicted by theory	No
		EVT	No clear prediction	Was not predicted by theory	No
Reputation	Titles that the player accumulates within the game	SDT	No clear prediction	Reputation titles could not be implemented due to the single use training design. These elements may have more merit in multiple use trainings or those that allow titles to be shared outside of training (i.e., training pages)	No
		GST	Reputation titles may be motivating to the player if specific or difficult goals are set for the trainees to attain them	Was not predicted by theory	No
		EVT	Trainees may place more value on resources or collectables within the training, which EVT predicts increases motivation	Was not predicted by theory	No

Table B (cont.)

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Sensation	Use of players' senses to create new experiences	SDT	No clear prediction	Was not predicted by theories	No
		GST			No
		EVT			No
Social Pressure	Pressure through social interactions with other players	SDT	Personalized learning pages can be shared with other trainees to increase perceptions of relatedness. SDT predicts as these perceptions increase, so does intrinsic motivation	Short-term training would not benefit from personalized training pages or have a meaningful way to share training progress	No
		GST	No clear prediction	Was not predicted by theory	No
		EVT	Gift giving allows trainees to develop relationships with others who have taken the training in a satisfying way. Trainees may place a higher value on the training reward to reciprocate this. EVT predicts as value increases, so does motivation (Hamzah et al, 2014)	Trainees are informed the previous trainee 'earned' them a donut based on their past performance. Trainees may be more motivated to reciprocate and earn the next trainee a donut as well. As rainees value the donut reward more, they are more motivated to earn it	Expectancy Theory (Altruism)
Stats	Visible information used by the player related to their outcomes within the game	SDT	No clear prediction	Player stats on the training cannot be displayed until the end of the five-minute task. Motivation cannot occur due to temporal precedence of the effects of stats element. Other trainings may use a HUD element (Heads Up Display) to show player stats in real time in more comprehensive trainings.	No
		GST	Goals to attain specific statistic values may be set for or by the trainee, which GST predicts will increase motivation	Was not predicted by theory	No
		EVT	Player statistics may increase expectancy beliefs of task ability. EVT predicts as expectancy increases subsequently so does motivation	Was not predicted by theory	No
Storytelling	The way in which the story is told to the player.	SDT	No clear prediction	Due to the shorter length of the training, storytelling cannot be implementing in a meaningful way.	No
		GST	No clear prediction	Was not predicted by theory	No
		EVT	Storytelling may reframe rewards to be more valuable to the trainee. EVT predicts that as value perceptions increase, so does motivation	Was not predicted by theory	No

Table B (cont.)

List of gamification elements that were considered for the virtual grocery store training

Gamification Element	Description of Gamification Element (from Toda et al, 2019)	Theories	Does the theory predict increased motivatinal outcomes from element? (and why?)	Reason for implementing element (or not) on supermarket cashier training	Was the element selected for the supermarket cashier training?
Time Pressure	Pressure through time within the game.	SDT	No clear prediction	Was not predicted by theories	No
		GST			
		EVT			

Table C:

Bivariate correlations between variables of interest

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16
1. Score	820.5	185.5	--														
2. Errors Made	4.27	3.38	-0.17	--													
3. Persistence	324.6	30.58	0.02	-0	--												
4. VR Familiarity	2.78	1.19	.18*	0.01	0.1	--											
5. Conscientousness	3.35	0.71	-0.1	-0.1	0.17	-0	**0.84**										
6. Openness	3.69	0.42	0.03	-0.1	-0.1	0.12	.19*	**0.7**									
7. Agreeableness	3.96	0.62	0.05	0	0.11	0.02	0.15	.32**	**0.82**								
8. Cybersickness	6.45	8.95	0.05	-0.2	-0.1	-0.1	-0	.19*	0.15	**0.76**							
9. TEQ_Interest	5.58	1.81	0.12	-0.1	.32**	.30**	.18*	0.01	.26**	-.21*	**0.95**						
10. TEQ_Competence	6.07	1.4	.30**	-0.1	0.14	.24**	0	.23**	0.08	-0.1	.50**	**0.9**					
11. TEQ_Choice	6.32	1.44	.18*	-.19*	0.14	0.14	-0	0.1	0.14	-0.1	.35**	.23**	**0.8**				
12. TEQ_Pressure	2.86	1.51	-0.16	0.05	-0.1	-0.2	0.05	-0	0.03	0.1	-.18*	-.36**	-.38**	**0.8**			
14. Effort	5.68	1.67	0.01	-0.1	0.08	0.05	.25**	0.09	.24**	0.02	.37**	0.15	0.07	.34**	**0.89**		
15. Presence	4.19	1.36	-0.14	-0	.20*	0.16	0.15	0.01	0.15	0.02	.40**	.29**	-0.05	0.1	.55**	**0.88**	
16. VR Errors	0.39	0.55	-.27**	-0	-0.1	-0.1	0.08	0.13	-0.1	-0.1	-0.2	-0.12	0.09	-0	-0.1	-0.2	--

Note: Reliability information for scales is bolded on the diagonal. * p < .05, **p < .01

Table D:

Tracked Correct User Actions in VR Cashier Training

Action	Description	Points
Scan Grocery Item	User correctly scans a grocery item for the first time	10
Enter PLU Item	User correctly enters a produce item for the first time	20
Void Item	User voids an item correctly for the first time	10
Take Payment	User accepts the correct form of payment from customer	30
Perfect Order	User correctly scans all items	50

Table E:

Tracked User Errors in VR Cashier Training

Error	Description
Extra Scans	User scans an scanned item an additional time
Items Not Removed	User does not void a requested item
Wrong Items Voided	User voids the incorrect item
Wrong Payment Method	User accepts the incorrect form of payment
Incorrect Button	User selects a button that causes an error message
Incorrect PLU Code	User inputs an incorrect PLU code
Items Missed	User fails to scan one or more items

Table F

Results of multiple regression predicting performance

Predictor	*b*	**β**	**β** 95% CI	sr^2
(Intercept)	774.53**			
VR Familiarity	30.86*	0.20	[0.03, 0.37]	.04
Presence	-28.67*	-0.21	[-0.38, -0.04]	.04
VR Errors	-100.24**	-0.30	[-0.47, -0.13]	.09

Note: $p < .05$, * indicates $p < .05$, ** indicates $p < .01$

Table G

Results of multiple regression predicting intrinsic motivation

Predictor	*b*	**β**	***β*** 95% CI	sr^2
(Intercept)	0.37			
Agreeableness	0.59*	0.20	[0.04, 0.36]	.04
Presence	0.44**	0.33	[0.17, 0.49]	.11
VR Familiarity	0.37**	0.24	[0.09, 0.40]	.06

Note: $p < .05$, * indicates $p < .05$, ** indicates $p < .01$