Original Research Article
The effect of interaction fidelity on stress relief and enjoyment for university students
Daniel M. Shafer*, Corey P. Carbonara, Michael F. Korpi
Department of Film & Digital Media, Baylor University, Waco 76798, Texas, United States of America
* Corresponding author: Daniel M. Shafer, Daniel_M_Shafer@baylor.edu

ABSTRACT
This study investigated the possible effect of varying levels of interaction fidelity (IF) in virtual reality (VR) games on several beneficial outcomes, including stress relief, calories burned, spatial presence, and enjoyment. Players were asked to play one of three VR games with varying IF levels. Results indicated that interaction fidelity was not a significant factor in these outcomes except for enjoyment; the high IF game was enjoyed more than the low IF game. However, playing any game did have a significant, if small, impact overall on both stress relief and calories burned. The results are encouraging due to the indication that spending time in VR and metaverse applications may have small but inherent health benefits.

Keywords: virtual reality; stress relief; enjoyment; ANOVA; interaction fidelity

1. Introduction
Virtual reality (VR) as an industry has seen relatively slow but steady growth since the first modern headsets, the HTC Vive, the Oculus Rift, and PlayStation VR were released in 2016[1-2]. Consistent technological improvement has been made as each new generation of headset has been released[3,4]. These improvements have resulted in better image processing and quicker response times, reducing lag and judder and increasing realism[5]. With that growth and advancement has come a continued interest in how VR games (not to mention video games in general) impact the mental well-being of those who play them. Despite the stigma that often comes attached to video games, much research has suggested that not only are they not particularly harmful[6-10], but video games are quite beneficial in many cases[11]. One of those benefits that has received recent attention is that video games can be used as a means of stress reduction[12-15]. Of course, findings in this line of research are mixed, because scholars’ operational definitions of stress differ, the actual type of stress measured differs between studies, and, perhaps most importantly, game content varies widely. There are certainly some types of games that are markedly stress-inducing (yet are still enjoyable for some players), and the type of stress induced by certain types of games varies as well.

This study seeks to investigate the relationship between VR gameplay, stress, and fitness benefits. The current VR literature seems to indicate that both exercise and video games are used as a way to combat stress. A recent rise in the popularity of exergames (video games played as exercise activities) suggests a merger of these two stress-reduction mechanisms. We also know that some games employ natural user interfaces (NUIs) which generally involve more bodily movements than games that use more traditional control methods.
However, not all games that use NUIs are exergames. As such, this study considers the variable known as interaction fidelity (IF). Interaction fidelity is an indicator of how much real, authentic motion is incorporated into a game. A game high in IF requires actions that closely mirror the actions one would execute if doing a similar activity in real life. Little to no current research exists that investigates games that vary in IF and their impact on fitness indicators and stress in players. Furthermore, it is useful to also investigate the impact of varying levels of IF on common user experience outcomes such as sense of spatial presence and enjoyment. This study seeks to investigate the relationships between these variables and fill this gap in knowledge.

2. Literature review

2.1. Exercise, stress, and video games

General stress due to life circumstances is quite common, particularly among college students. The American College Health Association (AHCA) reported in 2021 that nearly 79% of college students experienced moderate to high levels of stress over the 30 days preceding their survey[16]. This is surely understandable; college is difficult. But that fact does not minimize the toll stress can take on mental health[17], and it highlights the importance of finding ways to ameliorate or relieve stress on a regular basis in order to prevent it reaching problematic levels[18]. Daily exercise is often recommended by health professionals and organizations as a way to combat psychological stress[19–21]. College students exercise frequently as a means to maintain fitness and relieve stress[22].

Video games are another method sometimes used to combat stress. Russoniello and colleagues found that casual video games are often used by many players for stress reduction and other positive outcomes[23]. A recent meta-analysis of 28 research studies found that playing video games can reduce stress in both children and adults[24]. The use of video games by college students for stress relief has been found to be effective as well; nearly as effective as more accepted methods like mindfulness meditation[12].

Recently there has also been a rise in the use of VR games as effective fitness tools[25,26]. Games designed for the purpose of exercise are known as exergames[27]. Exergames bring these two effective stress-reduction methods together. Several research studies have investigated the use of exergames for health and other benefits[28–30]. Staiano and Calvert reported encouraging results that indicate that exergames used in physical education courses have positive health benefits for students[28]. Eng’s research team looked at several studies that considered the effect of exergames on executive functioning (EF) in children and found that playing exergames improves EF in both acute and chronic cases[29]. Rüth and colleagues reviewed 20 studies for evidence of benefits of off-the-shelf exergames used for rehabilitation. The results were generally positive, with some mixed findings[30].

2.2. How interaction fidelity may impact fitness and stress

The mechanism by which VR games create a situation for effective exercise and stress relief may be a construct known as interaction fidelity (IF). IF is defined as “the degree to which actions used to perform tasks in the game mirror reality”[31]. Another way to think about it is “action realism”[32]. To unpack the term a bit further, fidelity when applied to computers is the degree to which a device is able to recreate a real-world experience accurately[33]. Interaction fidelity, then, has to do with the actions taken to accomplish tasks in the virtually recreated world. As such, interaction fidelity has been described as “the objective degree with which the actions (characterized by movements, forces, or body parts in use) used for a task in the UI correspond to the actions used for that task in the real world”[34]. For example, a baseball batting game in VR should feel like taking batting practice in the real world. The closer the VR experience mimics real life actions required to hit a ball—timing, force, swing angle, etc., the higher the IF. Swinging a virtual sword in a VR game should feel
similar to swinging a sword in the real world. Dodging an incoming projectile in VR should require the same
move set required to dodge a rock or a ball in real life. The closer movements required to accomplish VR tasks
match movements required to accomplish similar tasks in the real world, the higher IF will be. By the same
token, the closer the VR system’s response time mimics what is possible in the real world, the more realism
can be leveraged by the technology, resulting in a more authentic exercise experience. The WINReality
baseball and softball training program relies on realism and the implementation of accurate response times to
give batters the real experience of hitting a 100-mph fastball. That application is only possible because of
the significant technological advances that have been made in recent VR technology. Games high in IF,
therefore, should be more effective for exercise purposes, resulting in more calories burned than their lower IF
counterparts even if those games are not specifically designed to be exergames. Therefore, we predict that:

H1: A high interaction fidelity game will be more effective as a fitness activity (as measured by the
participants’ fitness tracker or smart watch, if present) than a moderate or low IF game.

H2: A high interaction fidelity game will be more effective at stress relief than a moderate or low IF game.

Since data on the effects of gameplay on stress generally is sparse, this project will also interrogate the
following research question:

RQ: What overall effect of gameplay on stress will there be?

2.3. Interaction fidelity and user experience

Interaction Fidelity has been shown to be predictive of a sense of controller naturalness, and, somewhat
paradoxically, a detractor from a sense of interactivity. However, IF has also been found to facilitate the
quick completion of some tasks if the experience is high in IF versus a moderate level of IF. Lower levels of
IF have also been found to be more effective than moderate levels of IF for manipulation, navigation, and
search tasks. As mentioned, IF has also been closely linked to controller naturalness. Controller
naturalness is a predictor of spatial presence as well as a player’s sense of self-efficacy at the game, both of which have been found to have a positive impact on enjoyment. The more enjoyable a game is, the more a player is likely to stay engaged with the game for a long enough period of time (a minimum of 20 min) to provide adequate stress reduction and a worthwhile exercise experience. Furthermore, games that induce Csikszentmihalyi’s flow state can potentially be more engaging than games that do not, resulting in increased time spent exercising.

Based on these previous findings and indications, this study also investigates the relationships between
IF, sense of spatial presence and enjoyment. Past research seems to suggest that certain lines of influence exist,
but video games are complex as are individual reactions to them. Commercially available games and VR
experiences may or may not leverage variables consistently for all individuals. Nevertheless, it is predicted
that:

H3: (a) Interaction fidelity will positively predict spatial presence and enjoyment, and (b) spatial presence
will partially mediate the relationship between IF and enjoyment.

3. Materials and methods

This study is a three-group randomized experiment taking place in a controlled gaming research lab
setting.

3.1. Participant recruitment and scheduling

Participants were recruited via in-class announcement, email, and announcements via a learning
management system (LMS) such as Canvas. Upon arrival in the research lab, participants were greeted by a 
graduate or undergraduate research assistant, who directed them to a seat at a computer desk. Participants 
completed the pre-stimulus questionnaire on a lab PC, consisting of all demographic questions and a brief 
questionnaire on their current level of stress. They were also asked if they have a fitness tracker or smart watch 
that records their fitness data. Those that did were asked to report their current calories burned for the day as 
well as their current heart rate. Participants then proceeded to play a randomly assigned game on the Meta 
Quest 2 headset for up to 20 min. After stimulus exposure, participants were asked to report on their calories 
and heart rate (if they had a fitness tracker). They were then returned to the questionnaire on the lab PC to 
complete the post-stimulus questionnaire, which assessed their current level of stress, their sense of spatial 
presence, and their enjoyment. Upon completion of these tasks, they were thanked and dismissed. Total 
experimental visit time was approximately 45 min.

3.2. Sanitation of headsets

After each session was completed, the RA of that session wiped all contact surfaces of the headset and 
controllers with an alcohol-based, electronics-safe disinfectant wipe. They then loaded each headset and its 
controllers into a medical grade ultraviolet light sanitation cabinet. The equipment was sanitized via a 5-minute 
cycle before the next session began.

3.3. Stimulus material

Games were selected as stimulus material based on an extensive pretest of 30 games for the Meta 
(formerly Oculus) Quest 2 headset. Three undergraduate student research assistants (RAs) and the PI played 
and evaluated each game for IF. Due to various technical difficulties, some games were only evaluated by two 
RAs, yet this was deemed acceptable for our purposes since there were at least three IF scores rendered for 
each game.

The game selected for the High IF condition was Beat Saber\(^{[51]}\), which was rated highest in IF (IF = 8.62). The game selected for the Low IF condition was Tsuro\(^{[52]}\), which was lowest in IF of all the tested games (IF = 4.10). The game selected for the Moderate IF condition was A Fisherman’s Tale\(^{[53]}\) (IF = 6.29). The measure 
for determining IF was the framework for interaction fidelity analysis (FIFA)\(^{[54,55]}\). Each evaluator filled out 
the FIFA questionnaire after playing each game. Their scores were evaluated for similarity via the scale 
reliability procedure in SPSS. This test produces a Chronbach’s alpha score which indicates the interrater 
reliability for IF evaluations of each game. Interrater reliability scores for each game indicated good reliability 
(see Table 1).

<table>
<thead>
<tr>
<th>Game</th>
<th>Cronbach’s (\alpha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beat Saber (high IF)</td>
<td>0.861</td>
</tr>
<tr>
<td>A Fisherman’s Tale (moderate IF)</td>
<td>0.819</td>
</tr>
<tr>
<td>Tsuro (low IF)</td>
<td>0.947</td>
</tr>
</tbody>
</table>

3.4. Game descriptions

- **Beat Saber**

  Beat Saber is a musical rhythm game\(^{[51]}\). The controllers become two laser swords (similar to lightsabers 
  from Star Wars), one purple or blue, one red, in the hands of the player. The player stands on a virtual platform 
as red and purple/blue cubes fly toward them in time with a musical track. The object of the game is to slash 
and cut the cubes before they pass by, ideally in time with the music. Obstacles such as flying walls also come 
toward the player, and they must be dodged. The game has several levels of difficulty that increase the tempo
of the music, the rate at which the cubes and other objects arrive, and the number of cubes. It is one of the highest rated VR games in existence\cite{56–58}.

- **A Fisherman’s Tale**

  *A Fisherman’s Tale*\cite{53} is a highly interactive narrative puzzle game. The player starts as a puppet fisherman who lives in a lighthouse. The goal of the game is to solve puzzles throughout the lighthouse, eventually opening every room, getting to the top, and lighting the lighthouse’s lamp. It is an immersive, storybook style experience\cite{59,60}.

- **Tsuro**

  *Tsuro – The Game of the Path*\cite{53} is a VR port of a popular board game. The game takes place in a “tranquil… meditation garden”\cite{61}. Each player controls a stone that they move along an ever-developing path on a square board. The object of the game is to create a path for your stone that allows it to stay on the board, while trying to maneuver the other players’ stones off of the board. Following a path off of the board means that player loses. One can also lose by running into another player’s stone, eliminating both players. The last stone standing is the winner\cite{61,62}.

3.5. Power analysis and participants

A prospective power analysis using G*Power 3.1\cite{63,64} indicated that for an ANOVA test (H1 & H2) that includes 3 distinct groups, and the necessary reliability to find a medium effect size of 0.25 with a power of 0.85, \( n = 180 \) would be needed. G*Power analysis also indicated that for a multiple regression test with up to three predictor variables (H3) and the reliability to find a small effect size of 0.10 with a power of 0.95, the correct target sample size would be \( n = 176 \). Given the results of the power analyses for the intended tests, the target sample was \( n = 180 \). Due to moderate oversampling to ensure the target sample was achieved, data were collected from 228 participants. All were students at a mid-sized research university located in the south-central region of the United States. The sample was comprised of 112 males and 114 females (2 did not specify). Racial breakdown was 76% white, 10.5% Asian, 5% black or African-American, and the remainder other ethnicities. About 28% had more than average knowledge about how VR systems work, and approximately 36% had experience using a VR system in the past. About 30% reported that they play computer games quite often.

3.6. Measures

As noted, this study measured several variables that were hypothesized to vary based on the level of interaction fidelity of the game played. Interaction fidelity was previously determined by an extensive pretest using the FIFA scale developed by McMahan\cite{54,55} and used to determine IF of games in previous work\cite{31}.

3.6.1. Stress

A modified version of the stress measure, the 9th iteration of the Psychological Stress Measure (PSM-9)\cite{65} was used, as described by Desai and colleagues\cite{12}. The modification employed changes the instructions for the measure from “Check the number that best indicates the degree to which each statement has applied to you recently, that is, in the last 4 to 5 days”, to “Check the number that best indicates the degree to which each statement applies to you right now.” The PSM-9 was originally measured on an 8-point Likert-type scale ranging from (1) Not at all to (8) Extremely. For analysis purposes, this scale was changed from 1–8 to 0–7: (0) Not at all to (7) Extremely. This was done to give each scale variable a true zero point to indicate absence of the factor. Sample items include: “I feel rushed; I do not seem to have enough time”, “I feel preoccupied, tormented, or worried” and “I feel full of energy and keen (reverse scored)”. Scale reliability was good for both the pretest measure (\( \alpha = 0.854 \)) and the posttest measure (\( \alpha = 0.835 \)).
3.6.2. Spatial presence

Spatial presence was measured using the ITC-SOPI Spatial Presence Subscale. It consists of 19 items measured for this study on a Likert-type scale ranging from (0) Not at all to (7) Extremely. Sample items include: “I felt I could interact with the displayed environment”, “I felt that the characters and/or objects could almost touch me”, and “I felt I was visiting the places in the displayed environment”. Scale reliability was good, Cronbach’s α = 0.895.

3.6.3. Enjoyment

Enjoyment was measured using Klimmt and colleagues’ game enjoyment scale. It is an eight-item scale measured for this study on a Likert-type scale ranging from (0) Not at all to (7) Extremely. Sample items include “I found the game entertaining”, “It was great fun to take over the control in the game”, and “The game was not at all interesting to me (reverse scored)”. Scale reliability was good, Cronbach’s α = 0.876.

3.6.4. Cybersickness

Cybersickness was measured using the simulator sickness questionnaire (SSQ). It is a 16-item list of symptoms which asks participants to indicate their level of each symptom on a scale ranging from 0 (none) to 3 (severe). The instructions read: “Please indicate how much each symptom is affecting you right now”. Example items from the list include “headache”, “eyestrain”, “sweating”, “nausea”, and “vertigo”. Certain terms such as vertigo, stomach awareness and fullness of the head have associated definitions in the instructions: *“Fullness of the head is experienced as a feeling of head or ear pressure, but not an actual headache… **Vertigo is experienced as a loss of orientation with respect to vertical upright… ***Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea”. Scale reliability was good, Cronbach’s α = 0.874.

3.6.5. Physical activity

Physical activity was measure by first asking if participants had a smart watch or activity tracker, and then asking them to report their current heart rate and total calories burned both before and after engaging with the VR game.

4. Results

This project was preregistered prior to data analysis with Open Science Framework (OSF).

4.1. Hypothesis 1

Hypothesis 1 predicted that a high interaction fidelity game will be more effective at stress relief than a moderate or low IF game. Results indicated that H1 was not supported. A repeated measures analysis of variance (RMANOVA) indicated that stress was not significantly impacted by level of interaction fidelity of the game (see Figure 1).

![Figure 1. Pregame and postgame stress by interaction fidelity level.](image-url)
Results indicated no significant interaction between IF and stress within subjects: Wilks’ $\lambda = 0.99$, $F_{(2, 150)} = 0.71$, $p = 0.494$, $\eta^2_{\text{part}} = 0.009$. Reduction of stress did not significantly differ between conditions according to the between-subjects test: $F_{(2, 150)} = 1.42$, $p = 0.421$, $\eta^2_{\text{part}} = 0.011$.

4.2. Hypothesis 2

Hypothesis 2 predicted that a high interaction fidelity game would be more effective as a fitness activity than a moderate or low IF game. This was tested by collecting participants’ activity tracker or smart watch data before and after gameplay if they had such a device. The two metrics of interest are heart rate and calories. Results of the RMANOVA indicated that IF level did not significantly impact heart rate: $F_{(2, 59)} = 0.33$, $p = 0.715$, $\eta^2_{\text{part}} = 0.011$, see Figure 2. IF level also did not significantly impact calories burned: $F_{(2, 57)} = 2.14$, $p = 0.128$, $\eta^2_{\text{part}} = 0.070$, therefore H2 is not supported. Although interaction fidelity level did not significantly impact fitness metrics, gameplay generally did significantly impact calories burned: Wilks’ $\lambda = 0.55$, $F_{(2, 57)} = 47.22$, $p < 0.001$, $\eta^2_{\text{part}} = 0.453$. In fact, playing the game explained 45.3% of the total calories burned.

4.3. Research question

The research question asked what overall effect of gameplay on stress there might be. Evidence from the test performed in the investigation of H1 indicates that gameplay did indeed serve to reduce stress generally. Average stress levels went from $M = 2.60$ before gameplay to $M = 2.26$ after gameplay. Results of the RMANOVA analysis indicates that this difference, although practically small, is significant: Wilks’ $\lambda = 0.85$, $F_{(1, 152)} = 25.91$, $p < 0.001$, $\eta^2_{\text{part}} = 0.146$. In fact, playing a VR game explained 14.6% of the reported stress reduction for participants.

4.4. Hypothesis 3

Hypothesis 3 predicted that interaction fidelity would positively predict spatial presence and enjoyment, and that spatial presence would partially mediate the relationship between IF and enjoyment. Results of a one-way ANOVA procedure indicated a significant (although small) main effect of IF: Wilks’ $\lambda = 0.96$, $F_{(2, 219)} = 2.49$, $p = 0.043$, $\eta^2_{\text{part}} = 0.022$. Further analysis of these results indicates that there was no significant effect of IF on spatial presence ($F_{(2, 220)} = 0.65$, $p = 0.525$, $\eta^2_{\text{part}} = 0.006$), but there was a significant effect of IF on enjoyment ($F_{(2, 220)} = 3.03$, $p = 0.05$, $\eta^2_{\text{part}} = 0.027$). Beat Saber (high IF; $M = 3.35$, $SD = 0.69$) was significantly more enjoyed than Tsuro (low IF; $M = 3.06$, $SD = 0.78$); $p = 0.05$, $\eta^2_{\text{part}} = 0.027$. Since there was no significant impact of IF on spatial presence, the prediction of mediation of spatial presence between IF and enjoyment is rendered moot; therefore, H3 is only partially supported.
5. Discussion

This investigation into the possible impact of variance in interaction fidelity in VR games offers food for thought. Interaction fidelity, it seems, does not have a significant impact on user reactions. Higher interaction fidelity does not equate to increased stress reduction and does not seem to offer greater health benefits than playing a moderate or low interaction fidelity game. Before we discuss these results more extensively, allow us to offer an explanation of the limitations of the study which may have impacted our findings.

5.1. Limitations

Participants only played each game for approximately 20 min. It is quite possible that this was not a long enough play session for significant differences to emerge regarding stress reduction. In the case of physiological markers for health benefits as well, it seems that little occurred in the 20 min of gameplay. A close look at Figure 3, however, does indicate that players of the high IF game, Beat Saber, burned nearly twice as many calories as their counterparts in the moderate and low IF conditions, which is probably what ultimately drove the overall significant results regarding calories burned. Added to this is the fact that not all participants were wearing activity trackers, making the available sample from which to glean physiological data much smaller than the entire sample. This means that our results may suffer from Type II error due to insufficient power. This seems likely given previous evidence that 20 min of gameplay might in fact be sufficient to produce effects on stress reduction and health indicators[44–49]. However, a longer exercise/game time may give participants time to slip into a flow state and reap more benefits than a short time period may have allowed. At any rate, the high IF game was the most enjoyed, which is a positive indicator that VR games that have more impactful potential health benefits may be more attractive to players who may choose them for fun but reap positive side effects.

5.2. Implications and suggestions for future research

Our results indicate that interaction fidelity, despite suggestions from past research, does not seem to be a driver of stress relief and significant health outcomes when it comes to games played for health purposes. Focusing on whether or not a game requires players to move in ways that correlate with real-world movements is less important than simply playing a game that gets one up and active. However, we would point to the small but notable difference in calories burned when playing Beat Saber versus A Fisherman’s Tale and Tsuro. These figures did not indicate statistical significance in the small portion of our sample who had fitness trackers, but a future study should ensure that all participants’ physiological indicators are measured by providing heart rate monitors and calorie tracking devices as an integral part of the study. In this study, such metrics were only gathered if the participant had such a device. This method was cost saving, but ultimately resulted in unconfirmed (but promising) results.
What we must acknowledge is that VR games do hold promise for both mental and physical health and well-being. They are enjoyable and increasingly accessible, meaning users are likely to find a beneficial, healthy, stress-reducing experience that keeps them engaged. As such, metaverse-based activities like games that require a certain level of activity may have positive side effects. We must continue to investigate and raise awareness of the potential positive health outcomes use of this technology can bring.

Author contributions

Conceptualization, DMS, CPC and MFK; methodology, DMS, CPC and MFK; formal analysis, DMS; writing—original draft preparation, DMS; writing—review and editing, CPC and MFK; All authors have read and agreed to the published version of the manuscript.

Acknowledgments

The authors gratefully acknowledge the support of the Baylor University Department of Film & Digital Media: Prof. Christopher J. Hansen, Chairman; and the support of the Baylor University College of Arts and Sciences, Dean Dr. Lee Nordt; Assoc. Dean Dr. Kim Kellison. Also appreciated is the technical support given by Mr. Ron Garrett and Mr. Bobby Frillou, and administrative support given by Mr. Kemper Beard; Dept. of Film & Digital Media, Baylor University.

Conflict of interest

The authors declare no conflict of interest.

References