VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Predicting proteus effect via the user avatar bond: a longitudinal study using machine learning

This is the Published version of the following publication

Latifi, Mohammed Qasim, Poulus, Dylan, Richards, Michaella, Yap, Yang and Stavropoulos, Vasileios (2024) Predicting proteus effect via the user avatar bond: a longitudinal study using machine learning. Behaviour and Information Technology. ISSN 0144-929X

The publisher's official version can be found at
https://www.tandfonline.com/doi/full/10.1080/0144929X.2024.2363974
Note that access to this version may require subscription.

Downloaded from VU Research Repository https://vuir.vu.edu.au/49110/

Behaviour & Information Technology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tbit20

Predicting proteus effect via the user avatar bond: a longitudinal study using machine learning

Mohammed Qasim Latifi, Dylan Poulus, Michaella Richards, Yang Yap & Vasileios Stavropoulos

To cite this article: Mohammed Qasim Latifi, Dylan Poulus, Michaella Richards, Yang Yap & Vasileios Stavropoulos (08 Jul 2024): Predicting proteus effect via the user avatar bond: a longitudinal study using machine learning, Behaviour & Information Technology, DOI: 10.1080/0144929X.2024.2363974

To link to this article: https://doi.org/10.1080/0144929X.2024.2363974

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 08 Jul 2024.

Submit your article to this journal

Article views: 631

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=tbit20

BEHAVIOUR & INFORMATION TECHNOLOGY
https://doi.org/10.1080/0144929X.2024.2363974

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Predicting proteus effect via the user avatar bond: a longitudinal study using machine learning

Mohammed Qasim Latifi[a], Dylan Poulus[b,c], Michaella Richards[d], Yang Yap[a] and Vasileios Stavropoulos[a,e]

[a]RMIT University, Melbourne, Australia; [b]Southern Cross University, Gold Coast, Australia; [c]Manna Institute, Southern Cross University Australia, Australia; [d]Mighty Serious Health Games Studio, Victoria, Australia; [e]National & Kapodistrian University of Athens, Athens, Greece

ABSTRACT

The impact of an avatar on real-world behaviors of users is known as the Proteus Effect. Different user avatar bond (UAB) aspects, including identifying, immersing, and compensating via the avatar, influence an individual's Proteus Effect propensity. This study aimed to use machine learning (ML) classifiers to automate the prediction of those likely to experience Proteus Effect, based on their reports of identifying, immersing, and compensating with their avatar. Participants were 565 gamers (Mage = 29.3 years; SD = 10.6), assessed twice, six months apart, using the User-Avatar-Bond Scale and the Proteus Effect Scale. Tuned and untuned ML classifiers showed ML models could accurately identify individuals with higher Proteus Effect propensity, informed by a gamer's reported UAB, age, and length of gaming involvement, both concurrently and longitudinally (i.e., six months later). Random forests performed better than other MLs, with avatar identification as the strongest predictor. This suggests higher Proteus Effect propensity for those with a stronger user-avatar bond, informing gamified health applications to introduce adaptive behavioral changes via the avatar. Prevention and practice implications are discussed.

ARTICLE HISTORY
Received 23 July 2023
Accepted 17 May 2024

KEYWORDS
Gaming; Avatar; user-avatar bond; machine learning; artificial intelligence; Internet

Introduction

The Internet is an essential part of daily life, while gaming is a popular internet application that has been experiencing continuous growth (Gomez et al. 2022; Király et al. 2023; Stavropoulos et al. 2021a).[1] Not surprisingly, there has been increasing academic interest in the potential behavioural effects of gaming (Akbari et al. 2023; Ash 2016; Bowman, Kowert, and Cohen 2015; Caroux et al. 2022; Elson and Quandt 2016; Ferguson and Colwell, 2018; Galanis et al. 2023; Mancini and Sibilla 2017; Nowak and Fox, 2018; O'Brien et al. 2022; Stavropoulos, Gomez, and Motti-Stefanidi 2019; Stavropulos et al. 2020a; Wu, Hu, and Li 2022). For example, scholars are exploring how serious games, which are gamified applications for health and education, may help acquire skills and improve one's well-being (Derks et al. 2022; Zhonggen 2019). Indeed, promising preliminary studies demonstrate serious games' efficacy in treating mental health issues like depression and anxiety (Abd-Alrazaq et al. 2022; Carlier et al. 2020).

Despite this, empirical evidence on the adaptive and/or maladaptive impact of general gaming appears conflicting (Lee, Kim, and Choi 2021; Quandt and Kowert 2020). Indicatively, some research supports adverse effects like sexist behaviours, physiological arousal, and mental health problems (Gabbiadini et al. 2016; Stermer and Burkley 2015; Barlett et al. 2008; Coyne et al. 2018). Other studies present contrasting positive findings (Anderson and Carnagey 2009; Ferguson, 2015; Ferguson and Donnellan 2017; Ferguson and Wang 2019; Von Salisch et al. 2011). Specifically, moderate/adaptive internet gaming engagement has shown potential beneficial effects, including improved visual and spatial skills, openness to foreign cultures, a sense of accomplishment, increased prosocial behaviours, a cathartic/relief experience, and a sense of belonging (Colder Carras et al. 2021; Elson and Ferguson 2014; Elson and Quandt 2016; Raith et al. 2021).[2]

User-avatar interplay

Extant literature suggests that several factors may contribute to video games' positive and/or negative influences on the behaviour of the user (Stavropoulos, Ratan, and Lee 2022b). Among such factors, the relationship between the user and their in-game

CONTACT Dylan Poulus Dylan.poulus@scu.edu.au Locked Mail Bag 4, Coolangatta, QLD, 4225, Australia
Supplemental data for this article can be accessed online at https://doi.org/10.1080/0144929X.2024.2363974.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

representation figure, known as an avatar, has been identified as pivotal (Abd-Alrazaq et al. 2022; Carlier et al. 2020; Korkeila and Hamari 2020; Ratan, 2013; Ratan et al. 2020, 2022; Stavropoulos, Ratan, and Lee 2022b). Indeed, avatars may offer a high degree of customisation in role-playing game genres, which allows players to choose features like race, gender, clothing, accessories, and abilities (McKenna et al. 2022). In these games, avatars can evolve by gaining experience, knowledge, skills, achievements, and in-game wealth/possessions/currency (Liew et al. 2018; Ratan et al. 2020). Such features could encourage users to personify themselves in desired ways, often forming a connection between their offline self and ideal self via their avatar (Stavropoulos, Ratan, and Lee 2022b; Šporčić and Glavak-Tkalić 2018). Consequently, a dynamic bi-directional psychological process may occur between the two, influencing engagement in the game and likely affecting the online and offline behaviours of some users (Liew et al. 2018; Ratan et al. 2020; Stavropoulos, Ratan, and Lee 2022b; Watanabe and Ho 2023). Aligning with this hypothesis, recent work models the User Avatar Bond (UAB) as a complex space of interplay between gamers and avatars composed by four interconnected agents (i.e. player 1-player 2; player 1-avatar 1; player 1-player 2's avatar; avatar 1 with avatar 2; Banks and Carr 2019).

Aiming to address these issues, researchers have explored the UAB through various theoretical frameworks, including psychoanalysis, cyberpsychology, and communication studies (Blinka 2008; Kaye, Pennington, and McCann 2018; Liew et al. 2018; Zendle, Cairns, and Kudenko 2018). For instance, media/communication scientists have conceptualised the UAB as serving body, emotional, and identity links between gamers and avatars (Ratan 2013). Others emphasise the connection between who the gamer is and the behaviour of their avatar (Bowman, Kowert, and Cohen 2015; Elson and Quandt 2016; Ferguson and Colwell 2018; Mancini and Sibilla 2017). In this context, Banks and Bowman (2016) suggested that the UAB exists on a continuum (i.e. weak to strong) and is defined by three factors: (a) the degree of self-identification and self-differentiation, indicating how much the avatar is perceived as separate and distinct from the player; (b) emotional intimacy, characterised by a profound emotional attachment, potentially expressed through caring language and a sense of shared experiences; and (c) shared/distinct agency, encompassing aspects like moral decision-making and assuming responsibility (Banks and Bowman 2014; Banks and Bowman 2016). Reinforcing these views, and stemming from the field of psychology, Mancini and Sibilla (2017) identified different UAB profiles: 'Idealized' (avatar preferred over gamer), 'Actualized' (avatar more socially desirable), 'Alter Ego' (avatar less socially desirable), and 'Negative Hero' (avatar closer to gamer's ideal but less socially desirable).

Expanding such work, concepts such as *priming, game transfer phenomena (GTP), self-perception,* and *self-relevance* have been also employed to help understand the ways in which gamers interact with their offline and online environments via their avatars (Ash 2016; Bailenson and Blascovich 2004; Kaye, Pennington, and McCann 2018; Ortiz de Gortari and Diseth 2022; Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015; Ratan and Dawson 2016). *Priming* in the context of video games involves exposure to game stimuli that can influence a player's subsequent thoughts, attitudes, or behaviours. Specifically, it advocates the activation of cognitions, feelings, and actions through exposure to game elements, including a gamer's avatar, which can carry over into offline situations (Ash 2016; Kaye, Pennington, and McCann 2018). Building upon the concept of priming, GTP also describes the carryover of gaming experiences into offline perceptions, feelings, and behaviours (Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015; Kaye, Pennington, and McCann 2018; Matthews 2019; Zendle, Cairns, and Kudenko 2018).

Such views and concepts align with the self-perception theory (Green, Delfabbro, and King 2021; Ros et al. 2020). This supports the idea that individuals learn about their attitudes and emotions through self-observation, particularly in situations with unclear motivations (Bem, 1967). The UAB has been suggested to intertwine with self-perception theory in four ways. Firstly, behavioural reflections, where players' actions and choices for their avatars could reflect their preferences, values, and personality traits offline (Green, Delfabbro, and King 2021; Ros et al. 2020). Secondly, users project aspects of their identity onto avatars, and the traits attributed to avatars influence how players perceive themselves. Thirdly, emotional experiences through avatars can be incorporated into the players' self-concept (e.g. pride from completing in-game challenges). Lastly, avatar customisation may reflect users' self-image and desired identity, offering insight into their self-perception and identity goals (Green, Delfabbro, and King 2021; Ros et al. 2020).

In this context, the *self-relevance* theory complements *self-perception theory* and *priming* UAB explanations (Ratan and Dawson 2016). Indeed, *avatar self-relevance* gauges the extent to which users perceive their avatar as personally relevant to their identity/self-hood (Bailenson and Blascovich 2004). It reflects the connection and appropriateness of the avatar in representing the user's essential being and individuality (Ratan and Dawson 2016). The UAB is supported to be stronger when users closely identify with their avatars, as seen in embodiment and self-presence measures (Ratan and Dawson 2016; Ratan and Sah 2015). For instance, research indicates that gender-consistent avatars foster a stronger emotional bond, and customised avatars increase susceptibility to avatar gender influences in subsequent tasks (Ratan and Dawson 2016; Ratan and Sah 2015).

Enhancing such work, studies stemming from the area of clinical psychology propose important, distinct UAB subdimensions (Blinka 2008; Stavropoulos et al. 2020b; Burleigh et al. 2018; Liew et al. 2018). These subdimensions include identification (recognising oneself in the avatar), immersion (experiencing avatar needs as real-life needs), and compensation (attributing qualities to the avatar that the player lacks and wishes to have in their off-game life; Blinka 2008; Stavropoulos et al. 2020a; Burleigh et al. 2018; Liew et al. 2018). Studies have suggested that identification, immersion, and compensation within the UAB may establish a connection influencing an individual's thoughts, emotions, and behaviours outside the game (Stavropoulos et al. 2022a; Burleigh et al. 2018; Liew et al. 2018). This could be of pivotal significance for serious games/gamified treatments using avatars to positively change one's behaviour (Stavropoulos, Ratan, and Lee 2022b). Overall, serious games differ in effectiveness across different populations (i.e. some benefit more than others; Stavropoulos, Ratan, and Lee 2022b). Therefore, it is important to understand who could benefit the most in order to match game and user profile to optimise outcomes (Stavropoulos, Ratan, and Lee 2022b).

The Proteus Effect

To address this question, the present study will employ the concept of Proteus Effect (PE), as it is more avatar specific than priming and GTP, to examine avatar related behavioural effects on the user (Ash 2016; Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015; Yee, Bailenson, and Ducheneaut 2009). PE describes how a gamer's behaviour in both the game and the real world is influenced by their avatar's characteristics (Blinka 2008; Fox, Bailenson, and Tricase 2013; Van Looy et al. 2012). Coined from 'protean', denoting versatility, it references the shape-changing abilities of the Greek God Proteus (Yee, Bailenson, and Ducheneaut 2009). The avatar is considered the player's entire self-representation and primary identity cue in the game environment (Yee, Bailenson, and Ducheneaut 2009). Thus, the avatar is anticipated to substantially influence gamers' behaviour in both online and offline realms (McKenna and Bargh 2000). PE is rooted in self-perception theory, suggesting individuals deduce internal states by observing their outward behaviour, including the appearance of avatars and the surrounding environment in the context where PE may occur (Bem 1972; Liu 2023).

PE is supported to be grounded in three key mechanisms: Behavioural confirmation, self-perception, and deindividuation (Yee and Bailenson 2007; Stavropoulos et al. 2020b). In brief, behavioural confirmation involves how an individual's expectations lead others to behave according to the perceiver's expectations. Self-perception involves the individual determining the reasons behind their behaviour by identifying the attitudes that give significance to their preferences (e.g. gamers using taller avatars exhibited increased confidence; Yee and Bailsenson 2007). Deindividuation suggests that individuals tend to distance themselves when part of a large group, often feeling more connected to their in-game avatar than their real selves (Yee and Bailsenson 2007). Concerning PE, studies hypothesise that as identification, compensation and idealisation with the avatar grow, a stronger influence on the offline behaviours of gamers may be exerted (De Gortari and Diseth 2022; Stavropoulos, Ratan, and Lee 2022b). For example, one may become more prosocial or confident due to their connection with their more prosocial/confident/high-achieving avatar representation (Greitemeyer and Osswald 2011).

The present study

In this regard, Artificial Intelligence (AI), specifically machine learning (ML), holds promise for exploring the relationship between the UAB and Proteus Effect propensity in gamers. Machine learning, a key AI methodology, excels in identifying complex patterns in data and using them to make predictions (Horton and Kelinman 2015; Kuhn and Wickham 2020; Lin, Lin, and Lane 2020). More specifically, ML procedures, particularly supervised learning, involve training models on a dataset with known predictors and outcomes, where they establish relationships between the two. The trained ML model later assesses predictive validity and accuracy by making predictions on a separate unseen dataset. Despite the increasing use of ML in human–computer interaction, no study has explored the relationship between UAB and PE propensity in gamers using ML to date (Horton and Kelinman 2015; Kuhn and Silge 2022; Kuhn and Wickham 2020; Gabrieli et al. 2023; Ibrahim, Clinch, and Harper 2022).

To contribute to this area of knowledge, the present study aims to examine a longitudinal dataset, including two time points 6 months apart, using ML classifiers to determine if/whether the PE propensity of a user can be predicted by their UAB identification, immersion, and compensation/idealisation. It does so while considering the gamers' age and years of gaming involvement, as past literature supports that younger gamers with lengthier gaming exposure might be more receptive to immersive gaming aspects, such as the UAB (Stavropoulos et al. 2021a). This is crucial, given the literature assuming the concurrent and prospective influence of the PE on users' behaviours and cognitions (Stavropoulos, Ratan, and Lee 2022b). Thus, identifying those more receptive to the PE is significant for informing clinical practices related to gamified treatments, and especially in determining users who might benefit the most when using avatars crafted for therapy (Stavropoulos, Ratan, and Lee 2022b).

Therefore, the following two research questions were explored:

(1) How can AI/ML supervised training be used/employed to identify the Proteus Effect (PE) propensity of a gamer concurrently, based on their reported user-avatar bond (UAB) aspects of identification, immersion, and compensation?
(2) How can AI/ML supervised training be used/employed to identify the Proteus Effect (PE) propensity of a gamer six months into the future, based on their reported user-avatar bond (UAB) aspects of identification, immersion, and compensation?

Considering the two research questions, and based on past literature, it was expected that the UAB components of identification, immersion, and compensation would predict those more likely to experience PE both at present and six months into the future (Ash 2016; Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015; Yee and Bailenson 2007). Specifically, those reporting higher UAB aspects were expected to be more likely to report PE experiences (Ratan and Dawson 2016). Furthermore, it was envisaged that using ML classifiers would leverage the accuracy/precision of prediction beyond traditional regression analysis methods, in line with past ML findings (Brown et al. 2024).

Methods

Participants

Participants were sampled from the community (e.g. RMIT, Victoria, Melbourne, and Deakin Universities), Victorian public and catholic schools, Australian gamers' groups (e.g. Aus Gaymers Network), venues (e.g. Fortress Melbourne), and online forums (e.g. https://www.ausgamers.com), as well as via the distribution of advertising YouTube videos (e.g. https://www.youtube.com/watch?v = LC1z-7LCArY). Adolescents and adults older than 12 years were eligible to voluntarily/anonymously participate, provided addressing the plain language information statement describing the study aims, risks and their participation rights (e.g. withdrawal without any penalties and/or repercussions at any point) and providing informed consent. For adolescents (i.e. 12–18 years), these were firstly addressed by their responsible parent/guardian and secondly by the adolescents themselves (i.e. assent). A sample of 627 gamers were initially recruited. Of those, seven were excluded as preview-only responses, 19 as spam, one as a bot, 12 due to lack of consent, eight for failing validity questions (e.g. claimed they played non-existing games; e.g. *Risk of Phantom*), and 15 for insufficient responses. Therefore, the final sample comprised 565 role-playing-gamers. ($M_{age} = 29.3$ years $SD = 10.6$, $Min_{age} = 12$ $Max_{age} = 68$; $Males_{cisgender} = 283$, 50.1%), who were longitudinally assessed in the community six months apart (two time-points, T1 and T2). Tables 1 and 2 provide a detailed description of the sample at T1.

Measures

The following two scales were utilised besides demographic and internet use general information.

Proteus Effect Scale (PES; as amended by Stavropoulos et al., 2020)

The PES was used to assess Proteus Effect tendencies/behaviours in real life. It consists of six questions under one factor mirroring the influence the virtual surrounding/character exert on the individual when offline (e.g. 'I see things differently when I play with

Table 1. Participant's age, game playing/social media usage years and daily week and weekend consumed time at time point 1.

	Age	Gaming Years	Mean Daily Gaming Time in the week	Mean Daily Gaming Time in The weekend	Social Media Years	Mean Daily Social Media Usage Time in the week	Mean Daily Social Media Usage Time on the Weekend
N	562	556	557	555	558	545	543
Mean	29.3	5.62	2.23	3.39	7.06	2.55	3.01
SD	10.6	4.49	1.82	2.40	4.41	2.16	2.48
Min	12.0	0.00	0.00	0.00	0.00	0.00	0.00
Max	68.0	30.0	15.0	18.0	17.0	15.0	16.0

Table 2. Participants' sociodemographic, gaming and social media usage information at Time point 1.

		N	Total N	Proportion	p
Gender	Man (cisgender)	283	565	0.501	1.000
	Woman (cisgender)	259	565	0.458	0.053
	Man (transgender)	4	565	0.007	< .001
	Woman (transgender)	1	565	0.002	< .001
	Nonbinary	12	565	0.021	< .001
	Not Listed	3	565	0.005	< .001
	Prefer not to say	3	565	0.005	< .001
Sexual Orientation	Heterosexual-Straight	359	488	0.736	< .001
	Homosexual	36	488	0.074	< .001
	Bisexual	75	488	0.154	< .001
	Asexual	5	488	0.010	< .001
	Other	13	488	0.027	< .001
Ancestry	Aus./Engl.	412	565	0.552	0.015
	Chinese	20	565	0.035	< .001
	German	7	565	0.012	< .001
	Indian	10	565	0.018	< .001
	Other	118	565	0.209	< .001
Occupational Status	Full-Time Employed	271	490	0.553	0.021
	Part-Time Employed	77	490	0.157	< .001
	Student	64	490	0.131	< .001
	Trainee	2	490	0.004	< .001
	Not Currently Working	32	490	0.065	< .001
	On Temporary Leave (Education Leave, Public Service Leave, Training, Maternity Leave)	5	490	0.010	< .001
	Other	39	490	0.080	< .001
Educational Status	Professional Degree (i.e. MD, JD etc completed)	10	489	0.020	< .001
	PhD Degree (Completed)	17	489	0.035	< .001
	Postgraduate Studies (MSc Completed)	67	489	0.137	< .001
	Undergraduate University Course (completed)	176	489	0.360	< .001
	Intermediate between secondary level and university (e.g. Technical training)	97	489	0.198	< .001
	Senior secondary school (Years 11–12)	101	489	0.207	< .001
	Secondary school (Years 7–10)	9	489	0.018	< .001
	Other	12	489	0.025	< .001
Livingwith_w1	Family of origin (two parents/partners, only child)	34	564	0.060	< .001
	Family of origin (two parents/partners and siblings)	108	564	0.191	< .001
	Mother (only child, parent divorced-separated-widowed)	19	564	0.034	< .001
	Mother and sibling(s) (parent divorced-separated-widowed)	17	564	0.030	< .001
	Father (only child, parent divorced-separated-widowed)	6	564	0.011	< .001
	Father and sibling(s) (parent divorced-separated-widowed)	5	564	0.009	< .001
	With Partner	149	564	0.264	< .001
	Alone	61	564	0.108	< .001
	With Friend(s)	28	564	0.050	< .001
	Temporary accommodation	4	564	0.007	< .001
	Other	18	564	0.032	< .001
	With Partner and Children	115	564	0.204	< .001
Relationship Status	Single	148	490	0.302	< .001
	In a romantic relationship (A romantic relationship is defined as a romantic commitment of particular intensity between two individuals of the same or the opposite sex (When you like a guy [girl] and he [she] likes you back).	157	490	0.320	< .001
	Engaged	24	490	0.049	< .001
	Married	145	490	0.296	< .001
	Defacto	16	490	0.033	< .001
Partner Games Together	Yes	99	344	0.288	< .001
	No	245	344	0.712	< .001
Partner Uses Social-Media Together	Yes	227	340	0.677	< .001
	No	113	340	0.333	< .001
Social Media Users	Yes	550	565	0.973	< .001
	No	15	565	0.027	< .001
FB users	No	168	565	0.297	< .001
	Facebook	397	565	0.703	< .001
Twitter users	No	320	565	0.566	0.002
	Twitter	245	565	0.434	0.002
Instagram users	No	195	565	0.345	< .001
	Instagram	370	565	0.655	< .001
Pinterest users	No	469	565	0.830	< .001
	Pinterest	96	565	0.170	< .001
TikTok users	No	368	565	0.651	< .001
	Tik Tok	197	565	0.349	< .001
Most preferred social media	Facebook	145	557	0.260	< .001
	Twitter	66	557	0.118	< .001

(*Continued*)

Table 2. Continued.

		N	Total N	Proportion	p
	Instagram	135	557	0.242	< .001
	Pinterest	5	557	0.009	< .001
	Tik Tok	99	557	0.178	< .001
	Other, please define which	107	557	0.192	< .001
Gaming with best friend	No	336	565	0.595	< .001
	Yes	229	565	0.405	< .001
Using social media with best friend	No	189	565	0.335	< .001
	Yes	376	565	0.665	< .001
Gaming with other friends	No	312	565	0.552	0.015
	Yes	253	565	0.448	0.015
Using social media with offline friends	No	154	565	0.273	< .001
	Yes	411	565	0.727	< .001
Gaming with family members	No	406	565	0.719	< .001
	Yes	159	565	0.281	< .001
Using social media with family members	Yes	472	564	0.837	< .001
	No	92	564	0.163	< .001

Note: H_a is proportion ≠ 0.5.

another character in my real-life'). Items are answered on a 5-point Likert scale (e.g. 1 = strongly disagree to 5 = strongly agree), with their addition resulting in an overall $PE_{\text{real-life}}$ score between 6 and 30 and higher numbers indicating higher experiences. Participants with higher PE propensity were classified as those who scored more than 3/5 in 4/6 items. The instrument demonstrated sufficient reliability across both waves in the current sample (Cronbach's $\alpha_{\text{PES wave 1}}$ = 0. 0.903, McDonald's $\omega_{\text{PES wave 1}}$ = 0.904, Cronbach's $\alpha_{\text{PES wave 2}}$ = 0. 0.914, McDonald's $\omega_{\text{PES wave 2}}$ = 0.915; see Appendix 1, Tables 8a-d).

User-Avatar Bond Questionnaire (UAQ-Q; Blinka 2008)

The UAB-Q was employed to measure different gamer-avatar bond dimensions. The 12 UAQ items are answered on a 5-point Likert scale (1 = strongly disagree-5 = strongly agree). The instrument examines three factors: identification (4-items; 'Both me and my character are the same'), immersion (5-items: 'Sometimes I think just about my character while not gaming'), and compensation (3-items: 'I would rather be like my character'). The overall score, as well as the subscale scores are extracted via the addition(s) of the respective items' points, with higher scores indicating stronger UAB experiences. The internal consistency rates of the scale were sufficient across both study waves (Cronbach's $\alpha_{\text{UAB-Q wave 1}}$ = 0.804; McDonald's $\omega_{\text{UAB-Q wave 1}}$ = 0.813, Cronbach's $\alpha_{\text{UAB-Q wave 2}}$ = 0.849; McDonald's $\omega_{\text{UAB-Q wave 2}}$ = 0.867, Cronbach's $\alpha_{\text{Ident. wave 1}}$ = 0.701; McDonald's $\omega_{\text{Ident. wave 1}}$ = 0.729, Cronbach's $\alpha_{\text{Ident. wave 2}}$ = 0.770; McDonald's $\omega_{\text{Ident. wave 2}}$ = 0,789 Cronbach's $\alpha_{\text{Immers. wave 1}}$ = 0.717; McDonald's $\omega_{\text{Immers. wave 1}}$ = 0.727, Cronbach's $\alpha_{\text{Immers. wave 2}}$ = 0.764; McDonald's $\omega_{\text{Immers. wave 2}}$ = 0.775, Cronbach's $\alpha_{\text{Comp. wave 1}}$ = 0.604; McDonald's $\omega_{\text{Comp. wave 1}}$ = 0.656, Cronbach's $\alpha_{\text{Comp. wave 2}}$ = 0.660; McDonald's $\omega_{\text{Comp. wave 2}}$ = 0.709 see Appendix 1, Tables 9–12, a–d).

Procedure

Approvals were granted by: (a) the Victorian University Human Research Ethics Committee [HRE21-044], the Department of Education and Training of The Victorian State Government, Australia [2022_004542] and the Melbourne Archdiocese of Catholic Schools [1179]. Data collection involved three data-streams, paired via a non-identifiable code, unique for each participant: (a) a battery of demographic, internet/gaming/social media use questions and psychometric questionnaires/scales available via an online Qualtrics link, that initially directed to the plain language information statement and then requested the provision of informed consent by ticking a box, for one to commence the survey; (b) wearing an actigraphy tracker (Fitbit) for seven days to monitor physical activity/sleep (e.g. daily steps and sleep duration), that was electronically paired with the other data-streams via a unique code (i.e. records were automatically collected via the Fitbit portal based on the participant's code and those not owning a Fitbit were provided with a device during a mutually arranged/agreed meeting with the research team) and; (c) carrying a mobile monitoring application, called Aware Light (Van Berkel et al. 2022) recording screen on/off time, number and length of calls (i.e. duration) and texts (i.e. length in characters) for 7 days (i.e. aware data was also matched with the other data-streams through the unique participants' code). The procedure was/is to be repeated four times, once every six months, with the current study being based on the first two completed collection waves (for detailed information, see Appendix 2).

Analyses

To address research question 1 (i.e. using UAB to predict PE propensity), ML procedures as per the Tidymodels package were conducted in R-Studio (Horton and Kelinman 2015; Kuhn and Wickham 2020). Firstly, data was balanced considering Yes/No Proteus Effect propensity cases to improve learning/ML prediction using the synthetic minority oversampling technique (SMOTE; DMwR package; Torgo et al. 2013). This algorithm introduces additional cases of the minority group by taking into consideration a potential number (k) of their nearest neighbours based on Euclidean distance (Chawla et al. 2002).[3] Secondly, data was split into 4/5 training and 1/5 testing, stratifying Yes/No Proteus Effect propensity proportions to be equal across the splits whilst adopting a conservative bell-shaped Bayesian prior distribution.[4] Finalised training and testing datasets were identical regarding Yes/No Proteus Effect propensity proportions ($X^2 = 0$, df = 1, $p = 1$). For cross-validation and ML hyperparameters' tuning, training data was additionally divided 10 times (i.e. folds) and training data bootstrapped versions were also created. Thirdly, the ML predictive equation (i.e. predictors, outcomes and organisation of training and testing data), called recipe in the context of tidymodels (Kuhn and Silge 2022), was introduced, such that: (a) the binary Yes/No Proteus Effect propensity at T1 was the outcome and UAB was the independent predictor; (b) a minimum ratio of 50% Proteus Effect propensity cases was maintained across all samples tested including the cross-validation and bootstrapped training data versions; (c) zero variance, strongly sparse/skewed and potentially highly intercorrelated predictors were excluded, to solidify findings.[5] Predictors were also scaled and centred prior to the analysis to accommodate classification (i.e. 0 = mean & 1 = Standard Deviation [SD]; Kuhn and Wickham 2020). Fourthly, seven supervised (i.e. models where the outcome is known in the training step/stage, namely: LASSO, SVM-Kernel, Random Forests, Naïve Bayes, Logistic Regression, XGB, and k-NN) ML models recommended for binary classification (see Table 3 for detailed descriptions) were introduced, alongside the null model (i.e. no ML prediction) in both their tuned and their untuned versions (Kuhn and Wickham 2020). At this point it should be noted that supervised classifiers were preferred as for the current analysis the correct predicted values (e.g. who is and/or not more likely to exhibit the Proteus Effect) were known in both the training and the testing data. Thus, the aim of the analysis was to test the capacity of ML models to actually learn. This is not the case with unsupervised algorithms, where there is no initial knowledge of the predicted values (Alloghani et al. 2020). Fifthly, model and algorithms were combined to create different workflows, which were: (a) trained in the default versions on the training data; (b) tuned considering their hyperparameters[6] via the bootstrapped versions of the training data for cross validation (i.e. a procedure for repetitive training of ML models on sub-segments of the training data and testing them on the remaining part of the data, such that distributional features of the sample do not interfere with the ML performance; Kuhn and Silge 2022), and; (c) tested across both their default and tuned versions on the testing data. To address research question 2 (i.e. UAB predicting future Proteus Effect propensity of users; six months later) the same procedure was repeated with Proteus Effect propensity, as reported in time point 2, utilised as the outcome/dependent variable. Model findings were compared based on their confusion matrices and several fit indices including model accuracy, precision, the area under the curve, recall and f-measures (see yardstick r package; Kuhn, Vaughan, and Vaughan 2020).

To prevent spurious findings, estimation of the necessary sample size was conducted a priori considering model overfitting. In machine learning, overfitting refers to models fitting training data too closely, indicating insufficient data and poor generalisation (Chawla et al. 2002). Overfitting was addressed by balancing the skewed dataset using Synthetic Minority Over-Sampling Technique (SMOTE; Chawla et al. 2002; Torgo et al. 2013), applying early stopping and LASSO regularisation with Random Forests, and performing cross-validation and hyperparameter tuning (see Table 3). These efforts help develop models that generalise beyond the training data and avoid overfitting. Overall, important steps were taken to preclude overfitting and erroneous conclusions through a priori sample size estimation and modelling techniques.

Results

Participant characteristics

With regards to gaming patterns at T1, participants reported having been a gamer for on average for 5.62 years (Min = <1 year, Max = 30 years; SD = 4.49), playing for an average of 2.23 h daily during weekdays (Min = <1 h, Max = 15 h; SD = 1.82) and 3.39 h during the weekend (Min = <1 h, Max = 18; SD = 2.40). The

Table 3. ML models trained, tuned and tested.

Type	Operation	Hyperparameters Tuned	R-Package/ Engine Employed
Least Absolute Shrinkage Selection Operator (LASSO)	LASSO constitutes a regression analysis based, supervised ML classifier, that applies variable selection and regularisation to increase prediction accuracy. It achieves that via reducing noises and selecting certain features to regularise the model. From a calculation perspective lasso considers the magnitude rate of the coefficient, as a penalty to the loss function. Thus, the loss function is amended to reduce model complexity via restraining the sum of predictors' coefficients [Loss function = OLS + A (penalty) X summation (addition of s size[s] of coefficients)].	*penalty* = To perform regularisation (i. e. L1), LASSO considers/adds a penalty to the size of regression coefficients (i. e. predictor effects), aiming to minimise them. The optimum penalty value is obtained via the tuning process.	*glmnet*
K Nearest Neighbours (k-NN)	Th k-NN algorithm entails a supervised, non-parametric classification/prediction, that relies on estimating proximity/relevance/distance of one case with '*k*' others, as per their Euclidean distance. Alternatively, k-NN classifies/categorises a case taking into consideration its neighbouring cases (i.e. similarity of a case with previously identified cases).	*neighbours* = The number (k) of neighbouring points to be considered in order to optimise the learning/ prediction performance of the algorithm, as defined via the tuning process.	*Knn*
Support Vector Machine Kernel (SVM-K)	Kernel ML is based on pattern examination/analysis and is mostly known via its popular support-vector machine (SVM) version. The kernel function refers to a mathematic procedure, which enables SVM to pursue deep learning via conducting bidimensional classifications of uni-dimensional data through the projection of a lower-dimension to a higher one. Subsequently, a kernelised SVM employs a linear computation to address non-linear/classification problems.	*cost* = In SVM cost resembles/postulates the logistic function via a piecewise linear. In practice, the cost hyperparameter programs/guides the algorithm's optimisation regarding the rate/size of misclassification allowed in the training sample. Higher cost values indicate tighter margins and the opposite *degree* = The degree hyperparameter dictates the flexibility/boundaries of prediction(s), such that higher values allow higher flexibility *scale_factor* = The scaling hyper-parameter of categorical/classification kernel(s) reflects the optimum normalisation patterns/process (i. e. kernel width) required to avoid any data modification.	*Kernlab*
X Gradient Boosting (XGB)	XGBoost is an ML classifier recommended for structured/tabular data. It implements gradient boosted decision trees to optimise prediction. XGBoost does so via providing a parallel tree boosting those integrates/considers weak prediction/learner models/decision trees. However, and in contrast to random forest bagging of generated trees, XGBoosting operates in a sequential manner, with any subsequent tree being influenced by the previous/last tree outcome.	*mtry* = The number of independent variables to be randomly assessed at each decision tree split. *min_n* = An integer/value/number for the least data points in a node (i. e. tree branch) that enables further split. *tree_depth* = The value defining the highest tree depth (i. e. subsequent splits) suggested to optimise prediction. *Learn rate (i.e. shrinkage)* = The value/rate required for the boosting adaptation to occur over successive iterations. *loss_reduction* = The reduction rate of the loss function suggested to progress with tree splits. *sample_size* = The amount/proportion of data required to be utilised in the algorithm's fitting process over each iteration.	*xgboost*
Random Forests	Random forest is a flexible and broadly employed supervised, ensemble (i.e. composite) ML model, that integrates/considers the results of numerous decision trees (i.e. bagging), whilst being trained/ learning to address a prediction/classification task. Practically, random forests conduct a meta-estimation that averages/considers the outcomes of multiple decision tree classifiers, implemented on different data sub-samples, to improve accuracy and deter over-fitting.	*mtry* = The number of independent variables to be randomly assessed at each decision tree split. *min_n* = An integer/value/number for the least data points in a node (i.e. tree branch) that enables further split.	*Ranger*
Naïve Bayes	Naïve Bayes operates as a probabilistic, supervised, ML classifier, which functions generatively. This suggests that it aims to model the data class distribution, whilst assuming conditional independence probability (i.e. data characteristics/ measures are independent) to predict the way a specific class would generate input data.	*smoothness* = This refers to the Kernel component Smoothness, which defines the density value required for the algorithm to converge quicker to the real density of random numeric predictors. *Laplace* = Laplace transformation/smoothing refers to a technique/strategy/method that addresses the problem/risk of zero probability in the algorithm.	*naivebayes*

(*Continued*)

Table 3. Continued.

Type	Operation	Hyperparameters Tuned	R-Package/ Engine Employed
Logistic Regression	Logistic Regression is also considered a supervised ML classifier that employs a logistic function to predict/ model binary/dichotomous dependent outcomes.	*penalty* = In logistic regression, as with LASSO, the regularisation penalty hyperparameter aims to address generalisation error and thus reduce overfitting risks. As such, it enhances the probability of simpler concluded models. *mixture* = A regularisation parameter value ranging between 0–1 to enhance model accuracy [mixture 1 corresponds with LASSO; 0 with ridge regression and in the interim with elastic modelling in between LASSO and ridge].	*glm*

Note: Glmnet is derived from 'Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., & Qian, J. (2021). Package 'glmnet'. *CRAN R Repositary*'; Ranger is derived from 'Wright, M. N., Wager, S., Probst, *P*., & Wright, M. M. N. (2019). Package 'ranger'. *Version 0.11*, *2*'. Kernlab is derived from 'Karatzoglou, A., Smola, A., Hornik, K., & Karatzoglou, M. A. (2019). Package 'kernlab'. *CRAN R Project'*. Xgboost is derived from ' Chen, T., He, T., Benesty, M., & Khotilovich, V. (2019). Package 'xgboost'. *R version*, *90*, 1-66'. All other engines ae derived from ' Kuhn, M., & Silge, J. (2022)'. *Tidy Modelling with R*. ' O'Reilly Media, Inc.'.

maximum random sampling error for a sample of 565 at the 95% confidence interval (z = 1.96) equalled ±4.12% satisfying Hill's (1998) recommendations. Missing values of the analysed variables at T1 ranged between 3 (0.5% not stating their age) to 16 (2.83% not answering Item 9 on the User-Avatar Bond Scale) and were missing completely at random in the broader dataset ($MCAR_{test} = 37.9$, $p = 0.183$ $_{[9\ \text{missing patterns}]}$; Little, 1988). Attrition between waves was 276 participants (48.8%) and therefore, it was studied with low to moderate effect-sizes[7] regarding number of years spent gaming ($t_{Welch's} = 3.509$, df = 526, $p < 0.001$, Cohen's d = 0.296) and age ($t_{Student} = 4.967$, df = 560, $p < 0.001$, Cohen's $d = 0.4192$; see Appendix 1, Tables 1–7).

Analysis results

Before addressing research questions 1 & 2, Yes/No Proteus Effect propensity participants were identified with $N_{no\_PE\_propensity} = 420$ (78.95%) and $N_{Yes\_PE\_propensity} = 112$ (21.05%). Considering research question 1, to accommodate ML learning, oversampling of the minority class was conducted via k-NN SMOTE (Chawla et al. 2002; Torgo et al. 2013) resulting to a balanced dataset (i.e. $N_{Yes\_PE\_propensity} = 560$; 50%). Data was then split into 80% training and 20% testing and the proportions of Yes/No Proteus Effect propensity were compared across the two parts showing insignificant differences (*$X^2 = 0$, $df = 1$, $p = 1$, Cramer's $V = 0.00$; 50% Yes Proteus Effect propensity across both Training and Testing*). The prediction algorithm was introduced, scaling of predictors was conducted, descriptives of the training, testing and whole dataset were estimated (see ML models and algorithms; Appendix 3), while 10 sub-divisions and bootstrapped versions of the training data were produced for cross-validation and hyperparameter tuning (see folds & train_boot, Appendix 3). Models and workflows of the Null, LASSO, SVM-Kernel, Random Forests, Naïve Bayes and Logistic Regression (see Table 3) in their default hyperparameter versions (i.e. untuned) were then introduced, trained on the training data and tested on the testing data. Table 4 summarises their performance suggesting that, while all classifiers performed/learned acceptably and better than the null model, with Logistic Regression learning

Table 4. Null model and untuned algorithms performance on testing data (PE Wave 1).

	Null Model	Random Forests	Logistic Regression	LASSO	Naïve Bayes	SVM Kernel
ROC_AUC	0.5	0.81	0.998	0.812	0.861	0.82
PPV	0.5	0.755	0.98	0.75	0.826	0.793
F_meas	0.667	0.734	0.974	0.722	0.745	0.716
Recall	1	0.714	0.968	0.696	0.679	0.652
Accuracy	0.5	0.741	0.974	0.732	0.768	0.741

Notes: Accuracy reflects the ratio of correctly predicted cases, across the total number of cases. It is produced via the accumulation of the true positive and the true negative cases divided to the sum of all true positive, true negative, false positive and false negative cases. Accuracy values closer to 1 are considered desirable. Accuracy > .90 = Excellent; 70%<Accuracy < 90% = Very good; 60%<Accuracy < 70% = Good; Accuracy < 60% is poor. Area under the curve (AUC) refers to the area under the receiver operating characteristic (ROC) curve, as the latter is visualised in an orthogonal axis system/graph, where the horizontal line captures the false positive rate (FPR; 1 − specificity) and the vertical axis the sensitivity (True positive rate; values closer to 1 are considered better/ improved). AUC < .5 = No discrimination; 0.5 < AUC < .7 = Poor discrimination; .7 < AUC < .8 = Acceptable discrimination; .8 < AUC < .9 = Excellent discrimination; AUC > .9 = Outstanding discrimination. Positive Predictive Value [PPV] or Precision is irrespective of the prevalence of a condition, and reflects the proportion/ratio of all the true positive classified cases divided by the addition of the true positive and the false positive cases (i.e. how many of those classified as positive were actually positive? values closer to 1 are considered better/improved). Recall or sensitivity is associated to the prevalence of a condition and reflects the proportion/ratio of all the true positive classified cases divided by the sum of all the true positive and the false negative classified cases (i. e., how many of the true positive cases have been recalled? values closer to 1 are considered better/improved). F-Measure or F1-score/ F-Score reflects the ratio of the multiplication of recall and precision multiplied by two and divided by the accumulation of recall and precision such that the balance between precision and recall achieved by the model is captured. Higher values are considered better/improved (Jiao and Du 2016).

Table 5. Hyperparameter tuning summary across classifiers (PE Wave 1).

Type	Hyperparameters Tuned	Tuning Results
Least Absolute Shrinkage Selection Operator (LASSO)	*penalty*	0.00139
K Nearest Neighbours (k-NN)	*neighbours*	10
Support Vector Machine Kernel (SVM-K)	*cost*	10.1
	scale_factor	1
X Gradient Boosting (XGB)	*mtry*	6
	min_n	15
	tree_depth	11
	Learn rate (i.e. shrinkage)	0.0425
	loss_reduction	0.171
	sample_size	0.455
Random Forests	*mtry*	1
	min_n	6
Naïve Bayes	*smoothness*	0.763
	Laplace	0
Logistic Regression	*penalty*	0.0000000001
	mixture	0.05

See Table 3 for detailed information regarding the classifiers applied.

outperformed (comparatively) with excellent indicators across all criteria followed by the Naïve Bayes. Identification was the most significant predictor for Logistic Regression (i.e. > 25 points), while all other predictors exceeded 10 points (see VIP section, Appendix 3). The two lowest performing models besides the null model were Random Forest and SVM-Kernel.

To optimise learning and modelling capacity, the versions of LASSO, SVM-Kernel, Random Forests, Naïve Bayes and Logistic Regression, as well as XGB and k-NN were later tuned (see Table 3 regarding their respective hyperparameters' functions), trained on the training data and tested on the testing data. Table 5 summarises the tuned hyperparameters' values per classifier, and Table 6 their performance suggesting that, while all classifiers performed/learned acceptably and better than the null model, Random Forests outperformed comparatively with excellent indicators across

Table 6. Tuned algorithms performance on testing data (PE Wave 1).

	Null Model	Random Forests	Logistic Regression	LASSO	Naïve Bayes	SVM Kernel	XGB	k-NN
ROC_AUC	0.5	0.985	0.811	0.811	0.864	0.976	0.926	0.947
PPV	0.5	0.955	0.755	0.755	0.826	0.99	0.867	0.944
F_meas	0.667	0.946	0.734	0.734	0.745	0.939	0.839	0.842
Recall	1	0.938	0.714	0.714	0.679	0.893	0.812	0.759
Accuracy	0.5	0.946	0.741	0.741	0.768	0.942	0.844	0.857

Notes: Accuracy reflects the ratio of correctly predicted cases, across the total number of cases. It is produced via the accumulation of the true positive and the true negative cases divided to the sum of all true positive, true negative, false positive and false negative cases. Accuracy values closer to 1 are considered desirable. Accuracy > .90 = Excellent; 70%<Accuracy < 90% = Very good; 60%<Accuracy < 70% = Good; Accuracy < 60% is poor. Area under the curve (AUC) refers to the area under the receiver operating characteristic (ROC) curve, as the latter is visualised in an orthogonal axis system/graph, where the horizontal line captures the false positive rate (FPR; 1 − specificity) and the vertical axis the sensitivity (True positive rate; values closer to 1 are considered better/improved). AUC < .5 = No discrimination; 0.5 < AUC < .7 = Poor discrimination; .7 < AUC < .8 = Acceptable discrimination; .8 < AUC < .9 = Excellent discrimination; AUC > .9 = Outstanding discrimination. Positive Predictive Value [PPV] or Precision is irrespective of the prevalence of a condition, and reflects the proportion/ratio of all the true positive classified cases divided by the addition of the true positive and the false positive cases (i.e. how many of those classified as positive were actually positive? values closer to 1 are considered better/improved). Recall or sensitivity is associated to the prevalence of a condition and reflects the proportion/ratio of all the true positive classified cases divided by the sum of all the true positive and the false negative classified cases (i. e. how many of the true positive cases have been recalled? values closer to 1 are considered better/improved). F-Measure or F1-score/ F-Score reflects the ratio of the multiplication of recall and precision multiplied by two and divided by the accumulation of recall and precision such that the balance between precision and recall achieved by the model is captured. Higher values are considered better/improved (Jiao and Du 2016).

Table 7. Null model and untuned algorithms performance on testing data (PE Wave 2).

	Null Model	Random Forests	Logistic Regression	LASSO	Naïve Bayes	SVM Kernel
ROC_AUC	0.5	0.937	0.704	0.686	0.813	0.698
PPV	0.5	0.9	0.7	0.676	0.862	0.686
F_meas	0.667	0.867	0.675	0.597	0.694	0.615
Recall	1	0.837	0.651	0.535	0.581	0.558
Accuracy	0.5	0.872	0.686	0.64	0.744	0.651

Notes: Accuracy reflects the ratio of correctly predicted cases, across the total number of cases. It is produced via the accumulation of the true positive and the true negative cases divided to the sum of all true positive, true negative, false positive and false negative cases. Accuracy values closer to 1 are considered desirable. Accuracy > .90 = Excellent; 70%<Accuracy < 90% = Very good; 60%<Accuracy < 70% = Good; Accuracy < 60% is poor. Area under the curve (AUC) refers to the area under the receiver operating characteristic (ROC) curve, as the latter is visualised in an orthogonal axis system/graph, where the horizontal line captures the false positive rate (FPR; 1 − specificity) and the vertical axis the sensitivity (True positive rate; values closer to 1 are considered better/improved). AUC < .5 = No discrimination; 0.5 < AUC < .7 = Poor discrimination; .7 < AUC < .8 = Acceptable discrimination; .8 < AUC < .9 = Excellent discrimination; AUC > .9 = Outstanding discrimination. Positive Predictive Value [PPV] or Precision is irrespective of the prevalence of a condition, and reflects the proportion/ratio of all the true positive classified cases divided by the addition of the true positive and the false positive cases (i.e. how many of those classified as positive were actually positive? values closer to 1 are considered better/improved). Recall or sensitivity is associated to the prevalence of a condition and reflects the proportion/ratio of all the true positive classified cases divided by the sum of all the true positive and the false negative classified cases (i. e. how many of the true positive cases have been recalled? values closer to 1 are considered better/improved). F-Measure or F1-score/ F-Score reflects the ratio of the multiplication of recall and precision multiplied by two and divided by the accumulation of recall and precision such that the balance between precision and recall achieved by the model is captured. Higher values are considered better/improved (Jiao and Du 2016).

Table 8. Hyperparameter tuning summary across classifiers (PE Wave 2).

Type	Hyperparameters Tuned	Tuning Results
Least Absolute Shrinkage Selection Operator (LASSO)	*penalty*	0.00910
K Nearest Neighbours (k-NN)	*neighbours*	10
	cost	10.1
Support Vector Machine Kernel (SVM-K)	*scale_factor*	1
X Gradient Boosting (XGB)	*mtry*	1
	min_n	3
	tree_depth	11
	Learn rate (i.e. shrinkage)	0.00268
	loss_reduction	0.495
	sample_size	0.336
Random Forests	*mtry*	2
	min_n	3
Naïve Bayes	*smoothness*	0.658
	Laplace	0
Logistic Regression	*penalty*	0.00785
	mixture	0.75

Note: See Table 3 for detailed information regarding the classifiers applied.

all criteria followed by SVM-Kernel. The two lowest performing were the Logistic Regression and Lasso models.

The same process was repeated for research question 2, with Random Forests again outperforming relatively other classifiers in both their tuned and untuned versions followed by the LASSO and SVM-Kernal models respectively. The two lowest performing models in the were the LASSO and SVM-Kernal models in the untuned versions, and Logistic Regression and LASSO in the tuned versions. Table 7 and Tables 8 and 9 summarise the performance of the untuned versions, the tuned hyperparameters' values, and the performance of the tuned classifiers, respectively (for details, see Appendix 4). The tuned version of Random forests is the highest learning model with excellent predictive capacity.

Discussion

The current study employed a longitudinal design consisting of a normative sample of gamers, to be the first to train a series of tuned and untuned AI/ML automated procedures, to identify an individual's concurrent and prospective (i.e. 6 months later) PE propensity. To address these aims, ML models employed as predictors the user's reported avatar identification, immersion, and compensation, chronological age, and years of gaming involvement (Blinka 2008). All utilised ML classifiers were comparatively examined twice in relation to an individual's concurrent and prospective Proteus Effect propensity respectively. Accordingly, data was split into training and testing parts for the AIs to be trained and assessed, while a prediction ML model was introduced. The models underwent training, tuning, and testing to confirm their ability to determine if an individual exhibits or possesses greater susceptibility to the Proteus Effect. Findings demonstrated that while all AI classifiers tested in the present study could learn and perform better than the null model (i.e. random prediction), Random Forests had the strongest learning potential with identification being the most significant training predictor. Thus, as hypothesised, the UAB components were able to predict PE propensity of an individual, with higher UAB scores corresponding to higher PE scores.

Proteus effect and the user-avatar-bond

The study's findings were overall congruent with previous studies suggesting that a stronger bond with one's avatar (UAB) is more likely to be associated with Proteus Effect propensity (De Gortari and Diseth 2022; Šporčić and Glavak-Tkalić 2018; Stavropoulos et al. 2020a). Interestingly, at wave 1, while all three psychological aspects of the UAB, entailing

Table 9. Tuned algorithms performance on testing data (PE Wave 2).

	Null Model	Random Forests	Logistic Regression	LASSO	Naïve Bayes	SVM Kernel	XGB	k-NN
ROC_AUC	0.5	0.947	0.705	0.705	0.827	0.941	0.851	0.862
PPV	0.5	0.923	0.683	0.683	0.806	0.923	0.833	0.867
F_meas	0.667	0.878	0.667	0.667	0.734	0.878	0.759	0.712
Recall	1	0.837	0.651	0.651	0.674	0.837	0.698	0.605
Accuracy	0.5	0.884	0.674	0.674	0.756	0.884	0.779	0.756

Notes: Accuracy reflects the ratio of correctly predicted cases, across the total number of cases. It is produced via the accumulation of the true positive and the true negative cases divided to the sum of all true positive, true negative, false positive and false negative cases. Accuracy values closer to 1 are considered desirable. Accuracy > .90 = Excellent; 70%<Accuracy < 90% = Very good; 60%<Accuracy < 70% = Good; Accuracy < 60% is poor. Area under the curve (AUC) refers to the area under the receiver operating characteristic (ROC) curve, as the latter is visualised in an orthogonal axis system/graph, where the horizontal line captures the false positive rate (FPR; 1 − specificity) and the vertical axis the sensitivity (True positive rate; values closer to 1 are considered better/improved). AUC < .5 = No discrimination; 0.5 < AUC < .7 = Poor discrimination; .7 < AUC < .8 = Acceptable discrimination; .8 < AUC < .9 = Excellent discrimination; AUC > .9 = Outstanding discrimination. Positive Predictive Value [PPV] or Precision is irrespective of the prevalence of a condition, and reflects the proportion/ratio of all the true positive classified cases divided by the addition of the true positive and the false positive cases (i.e. how many of those classified as positive were actually positive? values closer to 1 are considered better/improved). Recall or sensitivity is associated to the prevalence of a condition and reflects the proportion/ratio of all the true positive classified cases divided by the sum of all the true positive and the false negative classified cases (i. e. how many of the true positive cases have been recalled? values closer to 1 are considered better/improved). F-Measure or F1-score/ F-Score reflects the ratio of the multiplication of recall and precision multiplied by two and divided by the accumulation of recall and precision such that the balance between precision and recall achieved by the model is captured. Higher values are considered better/improved (Jiao and Du 2016).

identification, immersion, and compensation, were significant predictors of an individual's PE propensity, the strongest was shown to be Identification. This appears to be consistent with previous literature supporting that a player tends to first identify with their respective avatar for immersion to occur (i.e. 'I am my avatar' precedes to feeling 'my avatar's needs are my needs'; Blinka 2008; Stavropoulos et al. 2021a). Such interpretations are reinforced by self-perception and self-relevance theorised effects (Bailenson and Blascovich 2004; Ratan and Dawson 2016). The avatar may operate essentially as a personification and representation of the gamer's identity within the video game's virtual world, fused to an extent with their desires regarding how they would prefer to be, and likely priming in turn (at least to an extent) their feelings, thoughts and behaviours out of the game (Hsu, Gross, and Hayne 2023; Kaye, Pennington, and McCann 2018; Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015; Matthews 2019; McKenna et al. 2022; Zendle, Cairns, and Kudenko 2018). Indeed, in line with the self-relevance theory, the use of the avatar can lead one to associate its characteristics with their self-perception, facilitating to a closer connection between the two and greater avatar effects on how one behaves out of the game (Chandler, Konrath, and Schwarz 2009; Klimmt et al. 2010). The closer the avatar is to the desired/congruent expressions of the gamer, the greater the identification experienced – *feeling like the avatar and the player are alike/similar* – resulting in a stronger UAB and thus allowing behaviour transference from one entity to the other (i.e. self and avatar; Burleigh et al. 2018; Green, Delfabbro, and King 2021; Liew et al. 2018; Šporčić and Glavak-Tkalić 2018; Stavropoulos et al. 2020b). The notion of alignment between player and avatar characteristics, as highlighted in the previous research above, underscores the significance of congruence in strengthening the UAB, aligning with the work of Banks and Bowman (2014, 2016). They highlighted the dynamic interplay between gamers and the gaming environment, suggesting that it's the nuanced interactions between the player and the game, including one's avatar, that end up shaping gaming effects (Elson et al. 2015; Elson and Quandt 2016; Ferguson and Colwell 2018).

Consequently, if a player does not develop a strong identification with their avatar, it is reasonably unlikely for them to align their offline cognitions, emotions, and behaviours with it, and thus to present with higher PE propensity (Ratan et al. 2020). This could be attributed, in part, to lower game immersion, which in turn reduces influences to a player's thoughts and behaviours, as also implied by the priming and GTP concepts (Ash 2016; Kaye, Pennington, and McCann 2018; Ortiz de Gortari, Oldfield, and Griffiths 2016, 2015). Thus, as the process of identification unfolds, the Proteus Effect may take place, and the player could start fusing their 'self' with their avatar being more likely to think, feel, and behave in accordance with it.

This interpretation is reinforced by wave 2 findings, where the prospective PE propensity was determined six months into the future, with identification being again the strongest indicator, followed by compensation. Several studies have suggested that there is a tendency for some individuals to use gaming as a form of escapism, due to identity-related issues/discomforts, including poor self-concept and self-esteem, psychological vulnerability (Green, Delfabbro, and King 2021; Lemenager et al. 2020; Šporčić and Glavak-Tkalić 2018; Stavropoulos et al. 2020a; Van Looy 2015). In that line, scholars suggest that an avatar may enable a gamer to compensate for negative self-perceptions in a virtual environment, presenting the way they would like to have been (Blinka 2008; Stavropoulos, Ratan, and Lee 2022b). As such, increased identification, and compensation with their avatar at an earlier stage may lead players to align more their offline behaviour with that of their persona six months later, enabling PE to take place (Ratan et al. 2020; Šporčić and Glavak-Tkalić 2018). According to Yee and Bailenson (2007), the avatar of the player is not simply a uniform that is worn within the game, but rather the players' entire self/identity represented within the game environment. In fact, the customisation of the avatar allows consciously and subconsciously the players wishes and characteristics to be carried into the avatar (Stavropoulos, Ratan, and Lee 2022b). Kiesler, Siegel, and McGuire (1984) and McKenna and Bargh (2000) have supported that it is exactly this connection that enables the avatar to significantly impact the players' behaviour, both within the virtual environment of the game, and in the real world.

Implications & limitations

Securing such knowledge (i.e. who is more likely to experience PE) may help to boost the effectiveness of serious games employing avatars by targeting individuals with higher UAB identification and compensation tendencies. This presents as a significant health opportunity, as several studies have linked video games with positive effects on mental health, suggesting that playing video games can help reduce stress levels and promote relaxation by helping individuals recover and regulate their mood, especially after a stressful task (Markey, Markey, and French 2019; Russoniello, O'Brien, and Parks 2009). Furthermore, other scholars have advocated that cognitive enhancements, including attention, memory, problem-solving, and decision-making skills, have been associated with video games, which could be maximised by higher PE (Green and Bavelier 2003; Powers et al. 2013). Finally, as multiplayer online role-playing video games also provide opportunities for social interaction, higher PE experiences are likely to enhance social connectivity and social support skills, while promoting a sense of belonging, fostering prosocial behaviours, and improving general social skills for those who mostly needed, if receptive to PE (Cole and Griffiths 2007; Greitemeyer and Osswald 2010).

However, caution should be exercised when interpreting the study's findings due to its use of a sample gathered from the community and its reliance on self-reported data, which could introduce potential biases and confounding influences (i.e. social compliance effects). Furthermore, the restricted adult sample may limit generalisability of findings to underaged gamers. Future research should aim to address such weaknesses to expand the available knowledge in this rapidly developing field.

Conclusion and further research

Overall, the present study supports that guided by the UAB identification, immersion, and compensation tendencies, while considering a player's age and years of gaming engagement, AI/ML Random Forests procedures can be trained to automatically determine gamers with higher Proteus Effect propensity. Such knowledge can be used to maximise the benefits of serious games employing avatars, primarily by emphasising user identification to cultivate pro-wellbeing Proteus Effect behaviours. For example, with full user consent and transparency, these ML tools could provide personalised insights to users regarding their PE susceptibility. This may assist gamers in making more informed, empowered choices about their avatars and gaming experiences. Participating gamers could in the future have the autonomy to utilise these prediction algorithms on an opt-in basis, equipping them to potentially self-regulate and optimise their avatar selections toward prosocial ends. In that context, this research establishes foundations for scalable procedures to evaluate individual PE propensity, enabled by ethical oversight of ML/AI technology applications. Additionally, gamers could voluntarily leverage these analytical techniques to gain greater self-awareness around other mental health factors encoded within gaming behaviours and choices. At this point it should be emphasised that, any deployment of AI/ML for the analysis of human psychology warrants thoughtful governance to align with user rights and interests.

Notes

1. In Australia, where the present study is conducted, over 17 million people engage in digital gaming, 76% of parents play games with their children, while 75% of players socialize through gaming (Digital Australia Report [DAR]; Brand and Jervis 2021). These findings align with global research highlighting the widespread appeal of gaming (Will 2019; Statista 2020).
2. Methodological elements, such as the utilization of partial effect sizes in meta-analyses and controls for publication bias (Ferguson and Kilburn 2009), along with the inherently complex nature of human-computer interaction and its associated confounding factors (e.g. user, application, and surrounding features influencing gaming experience), have been implicated in the discrepancies fuelling the, so called, 'Gaming Debate' (Ferguson and Colwell, 2018; Quandt and Kowert 2020). Experimental studies focusing on game activities and encompassing various sources of influence, such as game content and pace, reinforce the view that it is the gamer-game interplay that defines the nature of gaming effects on players (Elson and Ferguson 2014; Elson and Quandt 2016; Ferguson and Colwell 2018; Kowert and Ferguson 2021; Nowak and Fox 2018; Stavropoulos, Ratan, and Lee 2022b).
3. k-NN operates by identifying the distance between a suggested case and all other data cases identified. Firstly, it chooses a number (k) of cases nearest to the point of interest. Secondly, it attaches the most frequent class to that point (i.e. Yes/No GD risk; Chawla et al. 2002).
4. Adopting a Bayesian perspective, a potential distribution/variability is required for every model parameter before proceeding to data analysis. The range of these values was carefully/modestly/conservatively suggested here to follow a Cauchy shape (i.e. t-shape with seven degrees of freedom; Muth, Oravecz and Gabry 2018).
5. Steps c, d and e, were included as a precaution in the applied algorithm and did not effectively exclude any predictor.
6. A hyper-parameter constitutes an ML parameter, the value of which needs to have been specified prior to the learning ML being trained, in contrast to simple parameters which are 'learned' during the training of the model. Therefore, hyperparameters pose external model configurations (i.e. not based on the data) employed for the estimation of model parameters. Fine-tuned hyperparameters increase the capacity of a learning model to perform with higher accuracy, and are achieved through a 'grid' process in tidymodels (Kuhn and Wickham 2020).
7. Cohen's d, very small~0.01, small~0.20, medium~0.50, large, 0.80, very large~1.20 (Sawilowsky n.d.); Cramer's V, > 0.25 = very strong, > 0.15 = strong, >0.10 = moderate, > 0.05 = weak, > 0 no or very weak (Akoglu, 2018).

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Abd-Alrazaq, A., E. Al-Jafar, M. Alajlani, C. Toro, D. Alhuwail, A. Ahmed, … M. Househ. 2022. "The Effectiveness of Serious Games for Alleviating Depression: Systematic Review and Meta-Analysis." *JMIR Serious Games* 10 (1): e32331. https://doi.org/10.2196/32331.

Akbari, M., M. H. Bahadori, S. Khanbabaei, B. B. Milan, Z. Horvath, M. D. Griffiths, and Z. Demetrovics. 2023. "Psychological Predictors of the Co-Occurrence of Problematic Gaming, Gambling, and Social Media use among Adolescents." *Computers in Human Behavior* 140: 107589. https://doi.org/10.1016/j.chb.2022.107589.

Akoglu, H. (2018). "User's Guide to Correlation Coefficients." *Turkish Journal of Emergency Medicine/TüRkiye Acil Tıp Dergisi* 18 (3): 91–93. https://doi.org/10.1016/j.tjem.2018.08.001.

Alloghani, M., D. Al-Jumeily, J. Mustafina, A. Hussain, and A. J. Aljaaf. 2020. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. In *Supervised and Unsupervised Learning for Data Science: Unsupervised and Semi-Supervised Learning*, edited by M. Berry, A. Mohamed, and B. Yap. Cham: Springer. https://doi.org/10.1007/978-3-030-22475-2_1.

Anderson, C. A., and N. L. Carnagey. 2009. "Causal Effects of Violent Sports Video Games on Aggression: Is It Competitiveness or Violent Content?" *Journal of Experimental Social Psychology*, 45 (4): 731–739. https://doi.org/10.1016/j.jesp.2009.04.019.

Ash, E. 2016. "Priming or Proteus Effect? Examining the Effects of Avatar Race on in-Game Behavior and Post-Play Aggressive Cognition and Affect in Video Games." *Games and Culture* 11 (4): 422–440. https://doi.org/10.1177/1555412014568870.

Bailenson, J. N., and J. Blascovich. 2004. "Avatars." In *Berkshire Encyclopedia of Human-Computer Interaction*, edited by W. S. Bainbridge, 64–68. Great Barrington, MA: Berkshire Publishing Group.

Banks, J., and N. D. Bowman. 2014. "The win, the Worth, and the Work of Play: Exploring Phenomenal Entertainment Values in Online Gaming Experiences." *Proceedings of Meaningful Play*.

Banks, J., and N. D. Bowman. 2016. "Emotion, Anthropomorphism, Realism, Control: Validation of a Merged Metric for Player–Avatar Interaction (PAX)." *Computers in Human Behavior* 54: 215–223. https://doi.org/10.1016/j.chb.2015.07.030.

Banks, J., and C. T. Carr. 2019. "Toward a Relational Matrix Model of Avatar-Mediated Interactions." *Psychology of Popular Media Culture* 8 (3): 287–295. https://doi.org/10.1037/ppm0000180.

Barlett, C. P., C. L. Vowels, and D. A. Saucier. 2008. "Meta-Analyses of the Effects of Media Images on Men's Body-Image Concerns." *Journal of Social and Clinical Psychology* 27(3): 279–310. https://doi.org/10.1521/jscp.2008.27.3.279.

Bem, D. J. 1967. "Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena." *Psychological Review* 74 (3): 183–200. https://doi.org/10.1037/h0024835.

Bem, D. J. 1972. "Self-Perception Theory." In *Advances in Experimental Social Psychology*, edited by L. Berkowitz, Vol. 6, 1–62. Academic Press. https://doi.org/10.1016/S0065-2601(08)60024-6

Blinka, L. 2008. "The Relationship of Players to Their Avatars in MMORPGs: Differences Between Adolescents, Emerging Adults and Adults." *Cyberpsychology* 2 (1): 5.

Bowman, N. D., R. Kowert, and E. Cohen. 2015. "When the Ball Stops, the Fun Stops Too: The Impact of Social Inclusion on Video Game Enjoyment." *Computers in Human Behavior* 53: 131–139. https://doi.org/10.1016/B978-0-12.

Brand, J. E., and J. Jervis. 2021. *Digital Australia 2022*. Eveleigh: IGEA. Retrieved July 12, 2023, from: https://igea.net/wp-content/uploads/2021/10/DA22-Report-SUMMARY-FINAL-17-10-21.pdf.

Brown, T., T. L. Burleigh, B. Schivinski, S. Bennett, A. Gorman-Alesi, L. Blinka, and V. Stavropoulos. 2024. "Translating the User-Avatar Bond Into Depression Risk: A Preliminary Machine Learning Study." *Journal of Psychiatric Research* 170: 328–339. https://doi.org/10.1016/j.jpsychires.2023.12.038.

Burleigh, T. L., V. Stavropoulos, L. W. L. Liew, B. L. M. Adams, and M. D. Griffiths. 2018. "Depression, Internet Gaming Disorder, and the Moderating Effect of the Gamer-Avatar Relationship: An Exploratory Longitudinal Study." *International Journal of Mental Health and Addiction* 16: 102–124. https://doi.org/10.1007/s11469-017-9806-3.

Carlier, S., S. Van der Paelt, F. Ongenae, F. De Backere, and F. De Turck. 2020. "Empowering Children with ASD and Their Parents: Design of a Serious Game for Anxiety and Stress Reduction." *Sensors* 20 (4): 966. https://doi.org/10.3390/s20040966.

Caroux, L., M. Delmas, M. Cahuzac, M. Ader, B. Gazagne, and A. Ravassa. 2022. "Head-up Displays in Action Video Games: The Effects of Physical and Semantic Characteristics on Player Performance and Experience." *Behaviour & Information Technology* 42 (10): 1–21. https://doi.org/10.1080/0144929X.2022.2081609.

Chandler, J., S. Konrath, and N. Schwarz. 2009. "Online and on my Mind: Temporary and Chronic Accessibility Moderate the Influence of Media Figures." *Media Psychology* 12: 210–226. https://doi.org/10.1080/15213260902849935.

Chawla, N. V., K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. 2002. "*Smote: Synthetic Minority Over-Sampling* Technique." *Journal of Artificial Intelligence Research* 16: 321–357. https://doi.org/10.1613/jair.953.

Colder Carras, M., V. Stavropoulos, F. Motti-Stefanidi, A. Labrique, and M. D. Griffiths. 2021. "Draconian Policy Measures are Unlikely to Prevent Disordered Gaming." *Journal of Behavioral Addictions* 10 (4): 849–853. https://doi.org/10.1556/2006.2021.00075.

Cole, H., and M. D. Griffiths. 2007. "Social Interactions in Massively Multiplayer Online Role-Playing Gamers." *CyberPsychology & Behavior* 10 (4): 575–583. https://doi.org/10.1089/cpb.2007.9988.

Coyne, S. M., W. A. Warburton, L. W. Essig, and L. A. Stockdale. 2018. "Violent Video Games, Externalizing Behavior, and Prosocial Behavior: A Five-year Longitudinal Study During Adolescence." *Developmental*

Psychology 54 (10): 1868–1880. https://doi.org/10.1037/dev0000574.

De Gortari, A. B. O., and Å. Diseth. 2022. "Multidimensional assessment of Game Transfer Phenomena: Intrusive cognitions, Perceptual Distortions, Hallucinations and Dissociations." *Frontiers in Psychology* 13. https://doi.org/10.3389/fpsyg.2022.896238.

Derks, S., A. M. Willemen, M. Wouda, M. Meekel, and P. S. Sterkenburg. 2022. "The Co-Creation Design Process of 'You & I': A Serious Game to Support Mentalizing and Stress-Regulating Abilities in Adults with Mild to Borderline Intellectual Disabilities." *Behaviour & Information Technology* 41 (14): 2988–3000. https://doi.org/10.1080/0144929X.2021.1968034.

Elson, M., and C. J. Ferguson. 2014. "Twenty-five Years of Research on Violence in Digital Games and Aggression." *European Psychologist* 19 (1): 33–46. https://doi.org/10.1027/1016-9040/a000147.

Elson, M., and T. Quandt. 2016. "Digital Games in Laboratory Experiments: Controlling a Complex Stimulus Through Modding." *Psychology of Popular Media Culture* 5 (1): 52–65. https://doi.org/10.1037/ppm0000033.

Elson, M., J. Breuer, J. Van Looy, J. Kneer, and T. Quandt. 2015. "Comparing Apples and Oranges? Evidence for Pace of Action As a Confound in Research on Digital Games and Aggression." *Psychology of Popular Media Culture* 4 (2): 112–125. https://doi.org/10.1037/ppm0000010.

Ferguson, C. J. 2015. "Do Angry Birds Make for Angry Children? A Meta-Analysis of Video Game Influences on Children's and Adolescents' Aggression, Mental Health, Prosocial Behavior, and Academic Performance." *Perspectives on Psychological Science* 10 (5): 646–666. https://doi.org/10.1177/1745691615592234.

Ferguson, C. J., and J. Colwell. 2018. "A Meaner, More Callous Digital World for Youth? The Relationship Between Violent Digital Games, Motivation, Bullying, and Civic Behavior among Children." *Psychology of Popular Media Culture* 7 (3): 202–215. https://doi.org/10.1037/ppm0000128.

Ferguson, C. J., and M. B. Donnellan. 2017. "Are Associations Between "Sexist" Video Games and Decreased Empathy Toward Women Robust? A Reanalysis of Gabbiadini et al. 2016." *Journal of Youth and Adolescence* 46 (12): 2446–2459. https://doi.org/10.1007/s10964-017-0700-x.

Ferguson, C. J., & J. C. K. Wang. 2019. "Aggressive Video Games are Not a Risk Factor for Future Aggression in Youth: A Longitudinal Study." *Journal of Youth and Adolescence* 48 (8): 1439–1451. https://doi.org/10.1007/s10964-019-01069-0.

Ferguson, C. J., and J. Kilburn. 2009. "The Public Health Risks of Media Violence: A Meta-Analytic Review." *The Journal of Pediatrics/ the Journal of Pediatrics* 154 (5): 759–763. https://doi.org/10.1016/j.jpeds.2008.11.033.

Fox, J., J. N. Bailenson, and L. Tricase. 2013. "The Embodiment of Sexualized Virtual Selves: The Proteus Effect and Experiences of Self-Objectification via Avatars." *Computers in Human Behavior* 29 (3): 930–938. https://doi.org/10.1016/j.chb.2012.12.027.

Gabbiadini, A., P. Riva, L. Andrighetto, C. Volpato, and B. J. Bushman. 2016. "Acting like a Tough Guy: Violent-Sexist Video Games, Identification with Game Characters, Masculine Beliefs, & Empathy for Female Violence Victims." *PloS One* 11 (4): e0152121. https://doi.org/10.1371/journal.pone.0152121.

Gabrieli, G., M. H. Bornstein, P. Setoh, and G. Esposito. 2023. "Machine Learning Estimation of Users' Implicit and Explicit Aesthetic Judgments of web-Pages." *Behaviour & Information Technology* 42 (4): 392–402. https://doi.org/10.1080/0144929X.2021.2023635.

Galanis, C., N. Weber, P. H. Delfabbro, J. Billieux, and D. L. King. 2023. "Gaming Disorder and Stigma-Related Judgements of Gaming Individuals: An Online Randomized Controlled Trial." *Addiction* 118 (9): 1687–1698.

Gomez, R., V. Stavropoulos, D. Tullett-Prado, B. Schivinski, and W. Chen. 2022. "Network Analyses of Internet Gaming Disorder Symptoms and Their Links with Different Types of Motivation." *BMC Psychiatry* 22 (1): 76. https://doi.org/10.1186/s12888-022-03708-6.

Green, C. S., and D. Bavelier. 2003. "Action Video Games Modify Visual Selective Attention." *Nature* 423 (6939): 534–537. https://doi.org/10.1038/nature01647.

Green, R., P. H. Delfabbro, and D. L. King. 2021. "Avatar Identification and Problematic Gaming: The Role of Self-Concept Clarity." *Addictive Behaviors* 113: 106694. https://doi.org/10.1016/j.addbeh.2020.106694.

Greitemeyer, T., and S. Osswald. 2010. "Effects of Prosocial Video Games on Prosocial Behavior." *Journal of Personality and Social Psychology* 98 (2): 211–221. https://doi.org/10.1037/a0016997.

Greitemeyer, T., and S. Osswald. 2011. "Playing Prosocial Video Games Increases the Accessibility of Prosocial Thoughts." *The Journal of Social Psychology* 151: 121–128. https://doi.org/10.1080/00224540903365588.

Hill, R. (1998). "WHAT SAMPLE SIZE is 'ENOUGH' in INTERNET SURVEY RESEARCH?" In *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, Interpersonal Computing and Technology: An Electronic Journal for the 21st Century* (pp. 1064–4326).

Horton, N. J., and K. Kleinman. 2015. *Using R and RStudio for Data Management, Statistical Analysis, and Graphics*. CRC Press.

Hsu, C. W., J. Gross, and H. Hayne. 2023. "The Avatar Face-off: A Face (Less) Avatar Facilitates Adults' Reports of Personal Events." *Behaviour & Information Technology* 43 (4): 800–810. https://doi.org/10.1080/0144929X.2023.2187242.

Ibrahim, A., S. Clinch, and S. Harper. 2022. "Extracting Behavioural Features from Smartphone Notifications." *Behaviour & Information Technology* 42 (16): 1–19. https://doi.org/10.1080/0144929X.2022.2145996.

Jiao, Y., and P. Du. 2016. "Performance Measures in Evaluating Machine Learning Based Bioinformatics Predictors for Classifications." *Quantitative Biology* 4 (4): 320–330. https://doi.org/10.1007/s40484-016-0081-2.

Kaye, L. K., C. R. Pennington, and J. J. McCann. 2018. "Do Casual Gaming Environments Evoke Stereotype Threat? Examining the Effects of Explicit Priming and Avatar Gender." *Computers in Human Behavior* 78: 142–150. https://doi.org/10.1016/j.chb.2017.09.031.

Kiesler, S., J. Siegel, and T. W. McGuire. 1984. "Social Psychological Aspects of Computer-Mediated Communication." *American Psychologist* 39: 1123–1134. https://doi.org/10.1037/0003-066X.39.10.1123.

Király, O., P. Koncz, M. D. Griffiths, and Z. Demetrovics. 2023. "Gaming Disorder: A Summary of its Characteristics and Aetiology." *Comprehensive Psychiatry* 122: 152376. https://doi.org/10.1016/j.comppsych.2023.152376.

Klimmt, C., D. Hefner, P. Vorderer, C. Roth, and C. Blake. 2010. "Identification with Video Game Characters as Automatic Shift of Self-Perceptions." *Media Psychology* 13: 323–338. https://doi.org/10.1080/15213269.2010.524911.

Korkeila, H., and J. Hamari. 2020. "Avatar Capital: The Relationships Between Player Orientation and Their Avatar's Social, Symbolic, Economic and Cultural Capital." *Computers in Human Behavior* 102: 14–21. https://doi.org/10.1016/j.chb.2019.07.036.

Kowert, R., and C. Ferguson. 2021. "The Psychology of Digital Games." In *In Handbook of Esports Medicine: Clinical Aspects of Competitive Video Gaming*, 187–199. Cham: Springer International Publishing.

Kuhn, M., and J. Silge. 2022. *Tidy Modeling with R*. O'Reilly Media, Inc.

Kuhn, M., D. Vaughan, and M. D. Vaughan. 2020. Package 'Yardstick'.

Kuhn, M., and H. Wickham. 2020. "Tidymodels: A Collection of Packages for Modeling and Machine Learning Using Tidyverse Principles." https://www.tidymodels.org.

Lee, E. J., H. S. Kim, and S. Choi. 2021. "Violent Video Games and Aggression: Stimulation or Catharsis or Both?" *Cyberpsychology, Behavior, and Social Networking* 24 (1): 41–47. https://doi.org/10.1089/cyber.2020.0033.

Lemenager, T., M. Neissner, T. Sabo, K. Mann, and F. Kiefer. 2020. "Who am i" and "how Should i be": A Systematic Review on Self-Concept and Avatar Identification in Gaming Disorder." *Current Addiction Reports* 7: 166–193. https://doi.org/10.1007/s40429-020-00307-x.

Liew, L. W. L., V. Stavropoulos, B. L. M. Adams, T. L. Burleigh, and M. D. Griffiths. 2018. "Internet Gaming Disorder: The Interplay Between Physical Activity and User–Avatar Relationship." *Behaviour & Information Technology* 37 (6): 558–574. https://doi.org/10.1080/0144929X.2018.1464599.

Lin, E., C.-H. Lin, and H.-Y. Lane. 2020. "Precision Psychiatry Applications with Pharmacogenomics: Artificial Intelligence and Machine Learning Approaches." *International Journal of Molecular Sciences* 21 (3): 969. https://doi.org/10.3390/ijms21030969.

Little, R. J. A. (1988). "A Test of Missing Completely at Random for Multivariate Data with Missing Values." *Journal of the American Statistical Association* 83 (404): 1198–1202. https://doi.org/10.1080/01621459.1988.10478722.

Liu, Y. 2023. "The Proteus Effect: Overview, Reflection, and Recommendations." *Games and Culture*. https://doi.org/10.1177/15554120231202175.

Mancini, T., and F. Sibilla. 2017. "Offline Personality and Avatar Customisation. Discrepancy Profiles and Avatar Identification in a Sample of MMORPG Players." *Computers in Human Behavior* 69: 275–283. https://doi.org/10.1016/j.chb.2016.12.031.

Markey, P. M., C. N. Markey, and J. E. French. 2019. "The Effects of Playing Video Games on Mood, Anxiety, and Stress-Related Outcomes: A Meta-Analysis." *Psychology of Popular Media Culture* 8 (4): 403–416.

Matthews, J. S. 2019. "Issue Priming Revisited: Susceptible Voters and Detectable Effects." *British Journal of Political Science* 49 (2): 513–531. https://doi.org/10.1017/S0007123416000715.

McKenna, K., and J. Bargh. 2000. "Plan 9 from Cyberspace: The Implications of the Internet for Personality and Social Psychology." *Personality and Social Psychology Review* 4 (1): 57–75. https://doi.org/10.1207/S15327957PSPR0401_6.

McKenna, J. L., Y. C. Wang, C. R. Williams, K. McGregor, and E. R. Boskey. 2022. "You Can't be Deadnamed in a Video Game": Transgender and Gender Diverse Adolescents' use of Video Game Avatar Creation for Gender-Affirmation and Exploration." *Journal of LGBT Youth* 21 (1): 21–49. https://doi.org/10.1080/19361653.2022.2144583.

Muth, C., Z. Oravecz, and J. Gabry. 2018. "User-Friendly Bayesian Regression Modeling: A Tutorial with Rstanarm and Shinystan." *The Quantitative Methods for Psychology* 14 (2): 99–119. https://doi.org/10.20982/tqmp.14.2.p099.

Nowak, K. L., and J. Fox. 2018. "Avatars and Computer-Mediated Communication: A Review of the Definitions, Uses, and Effects of Digital Representations on Communication." *Review of Communication Research* 6: 30–53. https://doi.org/10.12840/issn.2255-4165.2018.06.01.015.

O'Brien, O., A. Sumich, T. Baguley, and D. J. Kuss. 2022. "A Partial Correlation Network Indicates Links Between Wellbeing, Loneliness, FOMO and Problematic Internet use in University Students." *Behaviour & Information Technology* 42 (16): 1–18. https://doi.org/10.1080/0144929X.2022.2142845.

Ortiz de Gortari, A. B., and Å Diseth. 2022. "Multidimensional Assessment of Game Transfer Phenomena: Intrusive Cognitions, Perceptual Distortions, Hallucinations and Dissociations." *Frontiers in Psychology* 13: 896238. https://doi.org/10.3389/fpsyg.2022.896238.

Ortiz de Gortari, A. B., B. Oldfield, and M. D. Griffiths. 2016. "An Empirical Examination of Factors Associated with Game Transfer Phenomena Severity." *Computers in Human Behavior* 64: 274–284. https://doi.org/10.1016/j.chb.2016.06.060.

Ortiz de Gortari, A. B., H. M. Pontes, and M. D. Griffiths. 2015. "The Game Transfer Phenomena Scale: An Instrument for Investigating the Nonvolitional Effects of Video Game Playing." *CyberPsychology, Behavior & Social Networking* 18 (10): 588–594. https://doi.org/10.1089/cyber.2015.0221.

Powers, K. L., P. J. Brooks, N. J. Aldrich, M. A. Palladino, and L. Alfieri. 2013. "The Effects of Video Games on Cognition: A Meta-Analysis." *Psychological Bulletin* 139 (1): 66–76.

Quandt, T., and R. Kowert. 2020. "The Video Game Debate: Where Do We Go from Here?" In *The Video Game Debate*, edited by T. Quandt and R. Kowert, *2*, 121–128. New York: Routledge.

Raith, L., J. Bignill, V. Stavropoulos, P. Millear, A. Allen, H. M. Stallman, … L. Kannis-Dymand. 2021. "Massively Multiplayer Online Games and Wellbeing: A Systematic Literature Review." *Frontiers in Psychology* 12: 698799. https://doi.org/10.3389/fpsyg.2021.698799.

Ratan, R. 2013. "Self-presence, Explicated: Body, Emotion, and Identity Extension Into the Virtual Self." In

Handbook of Research on Technoself, edited by R. Lippicini, 322–336. Hershey, USA: IGI Global.

Ratan, R., D. Beyea, B. J. Li, and L. Graciano. 2020. "Avatar Characteristics Induce Users' Behavioral Conformity with Small-to-Medium Effect Sizes: A Meta-Analysis of the Proteus Effect." *Media Psychology* 23 (5): 675. https://doi.org/10.1080/15213269.2019.1623698.

Ratan, R., and M. Dawson. 2016. "When Mii is me: A Psychophysiological Examination of Avatar Self-Relevance." *Communication Research* 43 (8): 1065–1093. https://doi.org/10.1177/0093650215570652.

Ratan, R., M. S. Klein, C. R. Ucha, and L. L. Cherchiglia. 2022. "Avatar Customization Orientation and Undergraduate-Course Outcomes: Actual-Self Avatars are Better Than Ideal-Self and Future-Self Avatars." *Computers & Education* 191: 104643. https://doi.org/10.1016/j.compedu.2022.104643.

Ratan, R., and Y. J. Sah. 2015. "Leveling up on Stereotype Threat: The Role of Avatar Customization and Avatar Embodiment." *Computers in Human Behavior* 50: 367–374. https://doi.org/10.1016/j.chb.2015.04.010.

Ros, S., S. González, A. Robles, L. Tobarra, A. Caminero, and J. Cano. 2020. "Analyzing Students' Self-Perception of Success and Learning Effectiveness Using Gamification in an Online Cybersecurity Course." *IEEE Access* 8: 97718–97728. https://doi.org/10.1109/ACCESS.2020.2996361.

Russoniello, C. V., K. O'Brien, and J. M. Parks. 2009. "The Effectiveness of Casual Video Games in Improving Mood and Decreasing Stress." *Journal of Cyber Therapy & Rehabilitation* 2 (1): 53–66.

Sawilowsky, S. S. n.d. New effect size rules of thumb. DigitalCommons@WayneState. https://digitalcommons.wayne.edu/coe_tbf/4/.

Šporčić, B., and R. Glavak-Tkalić. 2018. "The Relationship Between Online Gaming Motivation, Self-Concept Clarity and Tendency Toward Problematic Gaming." *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 12 (1): article 4.

Statista. 2020, April 16. Topic: Video game industry. https://www.statista.com/topics/868/video-games/

Stavropoulos, V., R. Gomez, and M. D. Griffiths. 2021a. "In Search of the Optimum Structural Model for Internet Gaming Disorder." *BMC Psychiatry* 21 (1): 1–12. https://doi.org/10.1186/s12888-020-02964-8.

Stavropoulos, V., R. Gomez, and F. Motti-Stefanidi. 2019. "Internet Gaming Disorder: A Pathway Towards Assessment Consensus." *Frontiers in Psychology* 10: 1822. https://doi.org/10.3389/fpsyg.2019.01822.

Stavropoulos, V., R. Gomez, A. Mueller, M. Yucel, and M. Griffiths. 2020a. "User-avatar Bond Profiles: How do They Associate with Disordered Gaming?" *Addictive Behaviors* 103: 106245. https://doi.org/10.1016/j.addbeh.2019.106245.

Stavropoulos, V., F. Motti-Stefanidi, and M. D. Griffiths. 2022a. Risks and Opportunities for Youth in the Digital Era: A Cyber-Developmental Approach to Mental Health. European Psychologist 27 (2): 86–101. https://doi.org/10.1027/1016-9040/a000451.

Stavropoulos, V., R. Ratan, and K. M. Lee. 2022b. "User-Avatar Bond: Risk and Opportunities in Gaming and Beyond." *Frontiers in Psychology* 13: 923146. https://doi.org/10.3389/fpsyg.2022.923146.

Stavropoulos, V., J. Rennie, M. Morcos, R. Gomez, and M. D. Griffiths. 2020b. "Understanding the Relationship Between the Proteus Effect, Immersion, and Gender among World of Warcraft Players: An Empirical Survey Study." *Behaviour & Information Technology* 40 (8): 821–836. https://doi.org/10.1080/0144929X.2020.1729240.

Stermer, S. P., and M. Burkley. 2015. "SeX-Box: Exposure to Sexist Video Games Predicts Benevolent Sexism." *Psychology of Popular Media Culture* 4 (1): 47–55. https://doi.org/10.1037/a0028397.

Torgo, L., R. P. Ribeiro, B. Pfahringer and P. Branco. 2013. "SMOTE for Regression." In *Lecture notes in computer science* (pp. 378–389). https://doi.org/10.1007/978-3-642-40669-0_33.

Van Berkel, N., S. D'Alfonso, R. K. Susanto, D. Ferreira and V. Kostakos. 2022. "AWARE-Light: A Smartphone Tool For Experience Sampling and Digital Phenotyping." *Personal and Ubiquitous Computing* 27 (2): 435–445. https://doi.org/10.1007/s00779-022-01697-7.

Van Looy, J. 2015. "Online Games Characters, Avatars, and Identity." In *The International Encyclopedia of Digital Communication and Society*, edited by R. Mansell and P. Hwa Ang, 1–11. Malden, MA: Wiley-Blackwell.

Van Looy, J., C. Courtois, M. De Vocht, and L. De Marez. 2012. "Player Identification in Online Games: Validation of a Scale for Measuring Identification in MMOGs." *Media Psychology* 15 (2): 197–221. https://doi.org/10.1080/15213269.2012.674917.

Von Salisch, M., J. Vogelgesang, A. Kristen, and C. Oppl. 2011. "Preference for Violent Electronic Games and Aggressive Behavior among Children: The Beginning of the Downward Spiral?" *Media Psychology* 14 (3): 233–258. https://doi.org/10.1080/15213269.2011.596468.

Watanabe, K., and B. Q. Ho. 2023. "Avatar-mediated Service Encounters: Impacts and Research Agenda." *The Service Industries Journal* 43 (3-4): 134–153. https://doi.org/10.1080/02642069.2023.2169277.

Wills, J. 2019. *Gamer Nation: Video Games and American Culture*. Johns Hopkins University Press ebooks. https://doi.org/10.1353/book.66180

Wu, Y., J. Hu, and W. Li. 2022. "The Link Between Online Gaming Behaviour and Unethical Decision-Making in Emerging Adults: The Mediating Roles of Game Cheating and Moral Disengagement." *Behaviour & Information Technology* 42 (10): 1534–1547. https://doi.org/10.1080/0144929X.2022.2087539.

Yee, N., and J. Bailenson. 2007. "The Proteus Effect: The Effect of Transformed Self-Representation on Behavior." *Human Communication Research* 33 (3): 271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x.

Yee, N., J. N. Bailenson, and N. Ducheneaut. 2009. "The Proteus Effect: Implications of Transformed Digital Self-Representation on Online and Offline Behavior." *Communication Research* 36 (2): 285–312. https://doi.org/10.1177/0093650208330254.

Zendle, D., P. Cairns, and D. Kudenko. 2018. "No Priming in Video Games." *Computers in Human Behavior* 78: 113–125. https://doi.org/10.1016/j.chb.2017.09.021.

Zhonggen, Y. 2019. "A Meta-Analysis of use of Serious Games in Education Over a Decade." *International Journal of Computer Games Technology* 2019: 1–8.