David Beyea*
University of Wisconsin –
Whitewater
Heide Hall, Room 434
800 West Main Street
Whitewater, WI 53190-1790

Rabindra (Robby) Ratan
Yiming (Skylar) Lei
Hanjie Liu
Gabriel E. Hales
Chaeyun Lim
Michigan State University

A New Meta-Analysis of the Proteus Effect: Studies in VR Find Stronger Effect Sizes

Abstract

The present study examines why some studies of the Proteus effect—the phenomenon that people tend to conform behaviorally and attitudinally to their avatars' identity characteristics—facilitate the phenomenon more effectively than others. A previous meta-analysis of the Proteus effect (Ratan et al., 2020) failed to examine potentially notable moderating factors of the phenomenon, so we examine such factors through a meta-analysis of the 56 quantitative experimental Proteus effect studies published at the time of this analysis. Studies that utilized virtual reality technology (e.g., head-mounted displays) elicited stronger effect sizes than those that utilized flat screens, as hypothesized. No support was found for the hypothesis that effect sizes differ by software type utilized (commercial or custom-built). We offer suggestions for future research into the Proteus effect, and how to best examine possible variables of the phenomenon.

1 Introduction

The Proteus effect, first demonstrated and named by Yee and Bailenson (2007), refers to the phenomenon of people behaviorally and attitudinally conforming to the characteristics of their virtual avatars. For example, using avatars with healthier body sizes increases the users' physical exercise intensity (Li et al., 2014; Peña & Kim, 2014). Using avatars who look like inventors (compared to avatars in casual clothing) increases creativity while brainstorming (Buisine & Guegan, 2020; Guegan et al., 2016). Many studies of the Proteus effect have focused on theoretically significant variables in predicting and measuring the effect, yet few have examined the underlying facilitating factors. A meta-analysis of 46 studies of the Proteus effect found that the phenomenon is reliable, with effect sizes generally in the small-to-medium range (Ratan et al., 2020). However, this analysis did not sufficiently identify what study design facets explain the underlying reasons for differing effect sizes. Ratan et al. (2020) compared previous Proteus effect studies' approaches to measurement, such as whether the studies included behavioral or attitudinal measures, at what point in the study these measures were taken, and whether main or interaction effects were assessed. The present study instead compares previous Proteus effect studies' approaches to operationalizing the avatar-use context, focusing on modality and software quality. Hence, the present study aims to address the

Presence, *Vol. 31, Winter 2022*, 189–202
https://doi.org/10.1162/PRES_a_00392

*Correspondence to beyead@uww.edu.

overarching research question of why some studies facilitate the Proteus effect to a greater extent than others.

Studies of the psychological mechanisms behind the Proteus effect provide clues to potential moderators. Yee and Bailenson (2007) argued that the avatar changes the user's self-perception, inducing behavioral conformity. Other researchers have suggested that avatar characteristics prime participants' cognitive schema in ways that lead to the observed Proteus effect (Peña, 2011; Peña et al., 2009). However, Yee and Bailenson (2009) found that the effect was more substantial when participants controlled an avatar instead of simply watching an avatar behave, which illustrates the importance of the psychological connection between the user and avatar (e.g., sense of embodiment) in the Proteus effect.

Building on this inference, the present meta-analytical study examines factors of Proteus effect studies that may influence the phenomenon's strength. We focus on two levels of modality employed by previous research: hardware and software. Regarding hardware, we compared Proteus effect studies that used virtual reality (VR) headsets with studies that used flat-screen computers. Regarding software, we compared studies that used commercial products with those that used custom-made platforms. Results of these comparisons suggest that study design factors do indeed influence the strength of Proteus effect outcomes.

2 Hardware

Given that avatar embodiment is a significant factor within the Proteus effect (Yee & Bailenson, 2009), the effect's magnitude may depend on factors that influence embodiment, such as media modality. Specifically, the effects of controlling an avatar on a two-dimensional (flat) screen may differ from controlling an avatar in VR. Controlling an avatar on a flat screen has been found to induce the Proteus effect (Ash, 2016; Li et al., 2014), sometimes with a relatively large effect size (e.g., $r = .69$, Guegan et al., 2016); however, Proteus effect studies that use VR (e.g., Aymerich-Franch et al., 2014; Banakou et al., 2016; Christou & Michael, 2014) might find stronger effects given that VR tends to induce a greater sense of embodiment than flat-screen media (Peck & Gonzalez-Franco, 2021; Shin, 2018; Yee & Bailenson, 2009). VR allows the user to temporarily extend their sense of self into the avatar (Slater et al., 2014), or even expand their own body schema (e.g., with a third arm, Won et al., 2015). In VR, a participant can move their bodies at will (e.g., tilt head, lift arm) and see the avatar's corresponding movements (e.g., in a virtual mirror), which is a common mechanism of increasing embodiment in Proteus effect research (Fox et al., 2013; Yee & Bailenson, 2007, 2009). In contrast, users often control avatars on a flat screen with a symbolic controller (e.g., keyboard and mouse) that requires minimal body movement, lacks a one-to-one parallel between the user's and avatar's body, and manifests only a limited number of preset animations. Additionally, VR environments are usually presented through head-mounted goggles that sufficiently block users' peripheral vision, leading to a higher level of immersion compared to flat screens. In turn, a higher level of spatial presence can be achieved in VR than on flat screens (Cummings & Bailenson, 2016), which potentially contributes to the Proteus effect. Thus, we pose the following:

Hypothesis 1: The hardware of avatar use moderates the effect sizes found in studies of the Proteus effect such that studies using VR show stronger effects compared to those using flat screens.

3 Software

The production type and software quality used in Proteus effect studies may also influence the strength of the effect. Experimental stimuli are often developed specifically for the study context and are operationally distinct from real-world media (Reeves et al., 2016). Some Proteus effect research has utilized custom-made virtual worlds and avatars created by the researcher(s) for testing purposes (e.g., Aymerich-Franch et al., 2014; Guegan et al., 2016; Yee & Bailenson, 2007), which are potentially experientially distinct from commercially available "off-the-shelf" software used in other Proteus effect research (e.g., Christou & Michael, 2014; Ratan

& Dawson, 2016; Via, 2016). Levels of production expertise, team characteristics, and time and resource investment tend to differ between these types of software. Commercial software is often developed by large, diverse teams that include professionals (e.g., writers) in the industry (Koster, 2018). In contrast, custom software created for Proteus effect research is more likely to result from the labor of a small team of students or faculty with relevant skills but limited professional experience. Hence, we might expect commercial software to have fewer technical "bugs," stronger narratives, and avatar-related mechanisms (e.g., appearance detail) that facilitate stronger Proteus effects. However, as a counterargument, commercial software is less likely to facilitate the researchers' intended experimental manipulations than software custom-built for a specific research project. Commercial software used in Proteus effect research might feel more forced or contrived toward the study goals, while custom software can be designed directly for these goals.

In sum, there are two competing arguments. Commercial software may lead to stronger Proteus effects because of its higher production quality. However, custom software may lead to stronger Proteus effects by addressing a study's research questions more directly. Hence, we posit two competing hypotheses.

Hypothesis 2a: Proteus effect studies that use commercial software tend to find *greater* effect sizes than studies that use custom-built software.

Hypothesis 2b: Proteus effect studies that use custom-built software tend to find *greater* effect sizes than studies that use commercial software.

4 Methods

This research responds to a previous meta-analysis of the Proteus effect (Ratan et al., 2020), which included 46 studies that utilized an experimental design to examine how avatar attributes assigned to participants influenced their behaviors or attitudes related to those attributes. The present study included 45 of these 46 studies—removing one due to published concerns regarding its data analysis (Hilgard, 2019). We followed the same inclusion criteria as in Ratan et al. (2020) to identify 14 additional experimental studies (e.g., published after 2020) to add to the present analysis. Three studies were dropped (Ogawa et al., 2020; Oyanagi & Ohmura, 2019; Stavropoulos et al., 2021) because they did not report statistics that could be converted to an r value, leading to a total K of 56 experimental studies.

4.1 Modality Coding

Three research assistants were trained by the principal researchers to code for Hardware Modality (VR vs. Flat Screen) and Software Type (Commercial vs. Custom) in each experiment. First, they denoted the hardware used as any VR head-mounted device (e.g., HTC Vive, Oculus Rift) or a flat screen (e.g., computers, video game consoles). Then, they denoted whether the software type was commercially available (e.g., *The Sims* or *Grand Theft Auto V*) or custom-built for the study (e.g., virtual worlds developed in Unity 3D). The full research team reviewed and reached complete agreement on these codings.

5 Preliminary Results

Aligning with previously outlined best practices and procedures of meta-analyses (Rosenthal & DiMatteo, 2001), the sample size of each of the analyzed studies was calculated and used to establish the corresponding r value. When there were multiple conditions that represented a Proteus effect within a given study, we selected the most conservative r-value, even if other conditions had a higher rate of significance. With these r values, we calculated the corresponding weighted averages, followed by 95% confidence intervals, variance, and expected variance from the potential sampling errors of the studies. Lastly, we calculated the unexplained variance of sets and subsets within the gathered studies and their results.

We started our analysis by examining all studies included in our sample as one grouping (see Table 1). For these 56 selected studies, weighted r was .23

Table 1. *Overview of Studies Included in the Present Analysis*

Study	Effect Size (r)	N	Hardware	Software
Ash (2016)	0.16	80	Non-VR	OTS
Aviles (2017)	0.15	236	Non-VR	OTS
Aymerich-Franch et al. (2014)	0.30	54	VR	Custom
Banakou et al. (2013)	0.54	30	VR	Custom
Banakou et al. (2016)	0.37	59	VR	Custom
Banakou et al. (2018)	0.30	30	VR	Custom
Beaudoin et al. (2020)	0.11	52	VR	Custom
Beyea et al. (2021)	0.19	406	Non-VR	*
Bian et al. (2015) [Study 1]	0.36	90	Non-VR	Custom
Bian et al. (2015) [Study 2]	0.18	90	Non-VR	Custom
Buisine & Guegan (2020)	0.30	72	Non-VR	Custom
Buisine et al. (2016)	0.76	12	Non-VR	OTS
Chang et al. (2019)	0.39	76	VR	Custom
Chen et al. (2012)	0.45	30	Non-VR	OTS
Christou & Michael (2014)	0.32	48	VR	Custom
De Rooij et al. (2017)	0.32	61	VR	Custom
Fox & Bailenson (2009) [Study 1]	0.39	63	VR	Custom
Fox & Bailenson (2009) [Study 2]	0.22	73	VR	Custom
Fox et al. (2009)	0.29	69	VR	Custom
Fox et al. (2013)	0.28	86	VR	Custom
Gomes (2012)	0.10	145	Non-VR	Custom
Guegan et al. (2016)	0.41	54	Non-VR	Custom
Hershfield et al. (2011) [Study 1]	0.26	50	VR	Custom
Hershfield et al. (2011) [Study 2]	0.42	21	VR	Custom
Kaye et al. (2018) [Study 1]	0.02	120	Non-VR	Custom
Kilteni et al. (2013)	0.39	36	VR	Custom
Kocur et al. (2019)	0.33	40	Non-VR	OTS
Kocur et al. (2020)	0.42	32	VR	Custom
Koda & Oguri (2019)	0.43	22	Non-VR	Custom
Lee et al. (2014)	0.20	120	Non-VR	Custom
Lee-Won et al. (2017)	0.20	238	Non-VR	OTS
Leung et al. (2021)	0.67	10	VR	Custom
Li et al. (2014)	0.22	140	Non-VR	OTS
McCain et al. (2018)	0.21	131	VR	Custom
Navarro et al. (2020)	0.13	305	Non-VR	Custom
Palomares & Lee (2010)	0.16	151	Non-VR	Custom
Peck et al. (2013)	0.39	30	VR	Custom
Peña & Kim (2014)	0.21	94	Non-VR	OTS
Peña et al. (2009) [Study 1]	0.51	51	Non-VR	OTS
Peña et al. (2009) [Study 2]	0.25	78	Non-VR	OTS

Table 1. *Continued.*

Study	Effect Size (r)	N	Hardware	Software
Peña et al. (2016)	0.30	96	Non-VR	OTS
Ratan & Sah (2015)	0.42	64	Non-VR	OTS
Ratan et al. (2016)	0.11	229	Non-VR	OTS
Ratan & Dawson (2016)	0.44	76	Non-VR	OTS
Reinhard et al. (2020)	0.03	67	VR	Custom
Sah et al. (2017)	0.21	44	Non-VR	OTS
Sherrick et al. (2014)	0.15	142	Non-VR	*
Sylvia et al. (2014)	0.02	47	Non-VR	OTS
Van Der Heide et al. (2013)	0.35	48	Non-VR	*
Via (2016)	0.13	90	Non-VR	OTS
Yee (2007)	0.23	57	VR	Custom
Yee & Bailenson (2007) [Study 1]	0.40	32	VR	Custom
Yee & Bailenson (2007) [Study 2]	0.33	48	VR	Custom
Yee & Bailenson (2009)	0.27	70	VR	Custom
Yee et al. (2009) [Study 2]	0.33	40	VR	Custom

*These text-based studies were removed from the software portion of the analysis.

and variance was .017. The expected variance due to sampling error was .01, resulting in an unexplained variance of .007 (expected variance subtracted from observed variance).

5.1 Modality Results

We analyzed subsets of the Proteus effect studies in our sample based on hardware (VR or flat-screen technology). In the subset of studies that used VR technology ($k = 25$), the weighted r was .29 with all variance explained, a variance of .012, and an expected variance of .015. In studies performed on a flat screen ($k = 31$), the weighted r was .20 with an unexplained variance of .01. Non-overlapping confidence intervals (.05 and .03) indicated a significant difference between the two subgroups. We then ran a meta-regression, which found significant results (see Table 2), indicating that studies using the VR modality had higher effect sizes than those using the flat-screen modality. Hypothesis 1 (hardware moderates the effect sizes found in studies of the Proteus effect such that studies using VR show stronger effects than those using flat screens) was supported.

Table 2. *Meta-Regression Results for Modality*

	Estimate (r)	Standard error	Z-value	p
Intercept	0.24	0.02	11.42	<.001
Modality (VR)	0.07	0.34	2.13	0.033

5.2 Software Results

To examine H2 (commercial software leads to *greater* or *smaller* effect sizes compared to custom-built software), studies were divided into subsets that used commercial or custom-built software. Three studies were removed from this analysis as they examined the Proteus effect in a text-based digital space, such as a mock social media site or a text-based story (Beyea et al., 2021; Sherrick et al., 2014; Van Der Heide et al., 2013). These three studies represent a third environment where the Proteus effect occurs, separate from the high-fidelity digital environments of VR and flat-screen games. A third grouping was not performed in this step of the analysis, as a sample size of three studies would not present meaningful results.

In the subset of studies that used commercial software ($k = 17$), the weighted r was .22 with an unexplained variance of .01. In studies using custom-built environments ($k = 36$), the weighted r was .24 with an unexplained variance of .006. Overlapping confidence intervals (.04 and .03) did not indicate a significant difference between subgroups. Neither Hypothesis 2a nor 2b was supported.

Next, we examined software type as a moderator for the flat-screen subgroup (all VR-based studies utilized custom software). In the subset of flat-screen studies using commercial software ($k = 17$), the weighted r was .22 with an unexplained variance of .01. In flat-screen studies using custom-built software ($k = 11$), the weighted r was .18 with an unexplained variance of .006. Overlapping confidence intervals (.04 and .05) did not indicate a significant difference between subgroups.

5.3 Assessing Publication Bias

We calculated the correlation between the effect sizes and sample size of studies in order to assess the possibility of publication bias. The resulting negative correlation ($r = -.51$) indicates a possible publication bias in the literature (Levine et al., 2009). Thus, we performed multiple analyses to ascertain the possible influence of publication bias on our results. First, we calculated a fail-safe N (Borenstein, 2005), which indicated that the current meta-analysis results would become nonsignificant if there were a number of unpublished studies with a total of 8,400 participants with insignificant results, which seems unlikely. We then conducted a trim-and-fill analysis (Carpenter, 2012; Duval & Tweedie, 2000) which resulted in 19 imputed studies. Accounting for these 19 studies resulted in an estimated effect size of $r = .19$. Next, we performed a precision-effect test and precision-effect estimate with standard errors (PET-PEESE). The PET-PEESE analysis estimates the effect size of a meta-analysis by adjusting for any correlation between effect sizes and standard errors of component studies (Bartoš et al., 2022). Results indicated that the adjusted mean-effect-size estimate was nonsignificant, $\rho = .021$, $p = .530$. Finally, we ran a step-weight-function selection model, an adjustment method that accounts for publication bias by using weighted likelihood (Bartoš et al., 2022). Results indicated an estimated effect size of .20, adjusted for publication bias.

Of the multiple publication bias analyses, the PET-PEESE was the only test that suggests a negligible Proteus effect. When interpreting these results, note that the PET-PEESE examines the relationship between effect size and standard error, and that such a relationship may not occur due to publication bias (Lau et al., 2006). It is possible that studies with small sample sizes actually have a larger effect size due to differences in the protocols of the studies (Borenstein, 2005). While this possibility should not be assumed, it should be considered when interpreting these results given that many of the present studies ($k = 25$) were reliant on expensive VR equipment and complicated experimental setups, which may have resulted in larger effect sizes with small sample sizes. A correlation was observed between sample size and study modality ($r = .4$), with VR studies having fewer subjects (M = 53.32, SD = 24.7) than flat-screen studies (M = 112.07, SD = 88.57). Examined together, these analyses of publication bias suggest that there is a potential for bias that should be considered when interpreting our results.

6 Discussion

The present study utilized a meta-analytical approach to examine two possible moderating factors that could influence the magnitude of the Proteus effect based on a sample of relevant literature. The initial results of the meta-analysis (see Table 3) showed the Proteus effect to have a weighted $r = .23$, with a CI = .03, which is not significantly different from the findings of a previous meta-analysis of $r = .24$ (Ratan et al., 2020) despite the additional studies examined in the present analysis. Potential moderating factors, located within varying subsets of previous studies on the Proteus effect ($k = 55$), included the hardware (i.e., VR or flat-screen technology) and software (i.e., commercial or custom-built) utilized. Our findings supported the expectation

Table 3. *Summarization of Meta-Analysis*

	K	Weighted *r*	Expected Variance	Unexplained Variance	95% CI	
					LL	UL
Overall	56	0.23	0.009	0.007	0.20	0.26
Modality						
VR Tech	25	0.29	0.015	0.012	0.24	0.34
Flat Screen	31	0.20	0.015	0.007	0.17	0.23
Software						
Commercial	17	0.22	0.008	0.009	0.17	0.27
Custom Built	36	0.24	0.012	0.006	0.20	0.28
Software Type ∗ Flat Screen						
FS∗Commercial	17	0.22	0.009	0.019	0.17	0.27
FS∗Custom	11	0.18	0.008	0.006	0.13	0.24

that studies that utilized VR technology found larger effect sizes ($r = .30$) than those that utilized flat screens ($r = .20$) (see Table 2). Conversely, no support was found for expectations regarding the type of software utilized (commercial or custom-built). Altogether, these results suggest that the size of the Proteus effect depends on facets of avatar use, with evidence that VR technology is more effective at inducing the phenomenon but without evidence for differences by software type. This latter lack of findings may have resulted from inherent limitations in the present study method, discussed below. Future research should use other approaches that afford more control than meta-analytical methods (e.g., experiments) to test how aspects of the software platform and other factors may moderate the Proteus effect.

VR likely induces larger Proteus effects because it facilitates greater immersion and presence in the virtual environment and a stronger sense of embodiment and identification with the avatar (Cummings & Bailenson, 2016; Peck & Gonzalez-Franco, 2021; Shin, 2018; Slater et al., 2014). This argument is consistent with previous findings that controlling an avatar in VR leads to stronger Proteus effects than simply watching an avatar behave (Yee & Bailenson, 2009), but the present research examined studies in which participants controlled their avatars in both VR and flat-screen settings. However, the experience of controlling an avatar is richer in VR, which may be associated with meaningful qualitative differences. For example, VR generally allows for a greater avatar-user perceptual link than flat screens because the avatar's movements correspond with the user's body movement. When you move your arm up, the avatar's arm moves up. Although this is more common in VR, though rare, this type of avatar-body tracking is currently available through commercial flat-screen platforms (e.g., *Animaze*). Hence, future research could examine whether such body tracking contributes to the Proteus effect independently of the visual modality or if some other facet of VR is responsible for the differences found in the present study (e.g., sense of presence in the virtual environment). Additionally, when examining modality-independent Proteus effect variables, researchers should, if possible, use VR technology. The higher effect sizes found in VR would provide the best chance of capturing evidence of changes to the Proteus effect due to non-modality-based variables, with fewer participants needed. A power analysis using the results of the present study suggests that identifying evidence of the Proteus effect with VR technology would require fewer participants ($n = 84$) than the use of a flat-screen modality ($n = 193$).

VR environments within this meta-analysis do not include environments created through the VR Cave Automatic Virtual Environment (CAVE) modality. The VR CAVE modality consists of a room-scale area where the VR environment is projected on the walls, floors, and ceilings. While the VR CAVE is a fully immersive virtual environment, that is, the user is in a room with the environment projected on the room surfaces, there is no avatar for the individual to embody. As the Proteus effect occurs when the avatar influences the individual, without an avatar, there can be no Proteus effect. A VR CAVE would provide an interesting third modality in which to examine the Proteus effect if "reflective" surfaces in the VR environment presented an avatar for the user. To the knowledge of the researchers of the present study, there is no current study that examines the Proteus effect in the VR CAVE modality.

This study did not find a difference between commercial and custom-built software in Proteus effect sizes. One explanation is that the commercial software used in the studies examined did not actually have much higher quality than the custom-built software, while the custom-built software was not actually much better at targeting study goals than the commercial software. Although commercial software is presumably developed by professional teams with advanced tools, the skills and tools available to research teams may be advanced enough that the difference in quality is negligible. Conversely, although software custom built by researchers could theoretically include more options tailored to the study at hand, due to the popularity and breadth of commercial software, researchers could also locate portions of a commercial game that highlight their study hypotheses and goals. However, the lack of an observed difference could be due to the presumed higher quality of commercial software balancing out the influence of custom-built software. Future research could examine this potential by specifically comparing the strength of Proteus effects in higher- and lower-quality versions of the same software platform.

6.1 Limitations and Future Directions

There may not yet be enough published Proteus effect research to support a meta-analysis of many moderating factors. This method generally involves subgrouping studies by moderator, but as subgroup sample sizes shrink, confidence intervals increase, thereby decreasing the likelihood of significant differences between groups. Considering that the remaining unexplained variance in the observed Proteus effect, after accounting for modality, is .01, it is possible that there are currently not enough Proteus effect studies for a meta-analysis to be the appropriate tool for further moderator exploration.

The present study was able to analyze differences in hardware and software used because there were sufficient sample sizes within subgroups, but there are a number of other potential moderating factors that future research could consider. For example, the Proteus effect may depend on temporal elements of avatar exposure, such as the duration of avatar use, length of delay between avatar exposure and Proteus effect measurement, and frequency of reminders about avatar characteristics. Longer use durations, shorter measurement delays, and more frequent reminders of avatar identity characteristics may increase the strength and duration of the Proteus effect, but these temporal facets of study design were difficult to infer from the study descriptions included in this research. Regarding avatar identity reminders, if they do indeed strengthen the Proteus effect, this might suggest that avatar use in the third-person perspective is more influential than in the first-person perspective. However, the present study found VR—in which first-person avatar use is more common—to be associated with stronger effects than flat-screen technologies, in which third-person avatar use is more common. Hence, future research should be carefully designed to tease apart these influences, for example, by including both first- and third-person conditions in both VR and flat-screen technologies.

Qualitative elements of avatar exposure may influence Proteus effect strength as well. For example, the depth of character development—defined as the amount of detail provided about an avatar's identity characteristics—may influence user perceptions of the avatar's personality and likability, thus increasing the likelihood of behavioral conformity. Avatar–self similarity (e.g., likes, dislikes, opinions) increases presence (Rahill & Sebrechts, 2021) and could thereby augment conformity as well.

Likewise, perceived avatar personalities that include improved versions of the participants' self-attributes (e.g., social skills, intelligence) could similarly augment the effect, as participants have been found to favor "better" self-representations through avatar personalization (Fokides, 2021). We recommend that future studies explore this route and provide details about the breadth and extent of interactions presented to participants. Likewise, other elements of software type may influence the Proteus effect, such as specifications of the scene, context, or situation. Past familiarity with the software may also play a role in promoting the Proteus effect.

Similarly, differences in perceptions of avatar appearance may play a role. Positive perceptions, such as avatar humanness (Hamilton & Nowak, 2010), attractiveness (Liao et al., 2019), and realism (Garau et al., 2003; Kang & Watt, 2013; Latoschik et al., 2017), could enhance the Proteus effect by motivating avatar-characteristic internalization. Negative perceptions, such as avatar eerieness and uncanniness (Shin et al., 2019), could hinder the Proteus effect by discouraging user attention to and identification with avatar characteristics. Examining such expectations within a meta-analytical framework (e.g., by having new participants rate avatar images from previous Proteus effect studies) would be difficult due to the large differences in the experience of original and new participants. However, new experimental research could be designed to manipulate and/or measure avatars' visual characteristics and compare these to Proteus effect sizes.

Perceived stereotypes of avatars could also lead to Proteus effect studies that are difficult to interpret. The Proteus effect is a phenomenon driven by the stereotypes as well as other identity characteristics of an avatar (Ratan et al., 2020). As such, it can be difficult to interpret some Proteus effect studies if relevant stereotypes are not fully understood and examined as part of the study. The attitude shift due to avatar usage should be in the direction of the stereotyped attitude; however, stereotypes are subjective. If a researcher misinterprets or makes a misassumption about a relevant stereotype within the population, study results could show an attitude shift in a different direction than predicted. To address this in the present meta-analysis, we examined each study in terms of the effect size of the attitude shift due to the usage of an avatar in experimental studies, without interpreting the direction of the shift. Such an interpretation of the direction of the shift would require data on the stereotypes held by study participants, to which we did not have access. Future studies should not assume that participants hold specific stereotypes. Rather, researchers should take care to measure or manipulate stereotypes of interest.

We have offered these potential future directions based on our understanding of the Proteus effect literature, but we should also note that Walther and Lew (2022) offer a systematic roadmap for research on self-transformation online that is highly relevant, although slightly broader than the Proteus effect (e.g., including effects of text-based communication). They specify six factors of mediated self-presentations that potentially facilitate self-transformation: physical or personality alteration (e.g., avatar appearance), verbal, graphical, or embodied environment (e.g., text-based versus virtual avatar), awareness of alteration (e.g., overt or tacitly manipulated avatar characteristics), agency of choice/customization (e.g., avatar customization system), publicness of setting (e.g., perceived public audience of avatar use), and social interaction (e.g., social VR chat room). Future Proteus effect research should work to situate findings within these six frames.

Finally, there is a concern that the lack of preregistration in Proteus effect research could have inflated the effect sizes found in this meta-analysis. Methodological challenges resulting from a lack of preregistration, as well as possible demand characteristics in experimental studies, can drive spurious statistically significant findings and effect sizes (Drummond et al., 2020; Ferguson et al., 2022). To mitigate this concern about finding upwardly biased mean effect sizes due to non-preregistered studies (Drummond et al., 2020), we chose conservative approaches to our analyses whenever possible. Foremost, we selected the lowest effect size if a study contained multiple Proteus effect observations. We were also conservative in our decision about final analyses, following the assumption that there is one true effect size underlying observations of the Proteus effect and that any differences in effect sizes between studies is due to error. For this reason, we ran our meta-analysis using a fixed-effect model rather than a random-effects model. While the random-effects model is often considered the appropriate meta-analysis model to use in most cases, the fixed-effect model provides a greater weight to larger studies (Dettori et al., 2022). As the possibility for publication bias in Proteus effect studies favored larger effect sizes with smaller sample sizes, the fixed-effect model allowed us to be as conservative as possible in our results. While this decision does not fully eliminate the concerns about inflated effect sizes in non-preregistered Proteus effect studies, our approach mitigates this concern to the greatest extent possible within the constraint of the existing publications in the field.

7 Conclusion

These limitations notwithstanding, the present study offers a notable contribution to the Proteus effect literature, highlighting that VR technologies produce stronger effects. Despite the lack of other differences found, the paper also offers novel theoretical perspectives on potential moderators of the Proteus effect for future research to examine. Future researchers should also replicate previous Proteus effect studies with new preregistered experiments so that the field can develop a better understanding of this phenomenon.

Acknowledgments

We would like to thank Connor Bird and Justin Duby for their research assistance on this project, and the AT&T endowment to the Media & Information Department at MSU, which supports Dr. Ratan's AT&T Scholar position.

REFERENCES

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. 10.1177/1555412014568870

Aviles, J. A. (2017). *Paths to prejudice reduction utilizing virtual avatars and agents*. Retrieved from https://etda.libraries.psu.edu/catalog/14363jaa313

Aymerich-Franch, L., Kizilcec, R. F., & Bailenson, J. N. (2014). The relationship between virtual self similarity and social anxiety. *Frontiers in Human Neuroscience*, *8*, 944. 10.3389/fnhum.2014.00944

Banakou, D., Groten, R., & Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *PNAS*, *110*(31), 12846–12851. 10.1073/pnas.1306779110

Banakou, D., Hanumanthu, P. D., & Slater, M. (2016). Virtual embodiment of White people in a Black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in Human Neuroscience*, *10*, 601. 10.3389/fnhum.2016.00601

Banakou, D., Kishore, S., & Slater, M. (2018). Virtually being Einstein results in an improvement in cognitive task performance and a decrease in age bias. *Frontiers in Psychology*, *9*, 917. 10.3389/fpsyg.2018.00917

Bartoš, F., Maier, M., Quintana, D. S., & Wagenmakers, E.-J. (2022). Adjusting for publication bias in JASP and R: Selection models, PET-PEESE, and robust Bayesian meta-analysis. *Advances in Methods and Practices in Psychological Science*, *5*(3). 10.1177/25152459221109259

Beaudoin, M., Barra, J., Dupraz, L., Mollier-Sabet, P., & Guerraz, M. (2020). The impact of embodying an "elderly" body avatar on motor imagery. *Experimental Brain Research*, *238*(6), 1467–1478. 10.1007/s00221-020-05828-5

Beyea, D., Van der Heide, B., Ewoldsen, D., Eden, A., & Meng, J. (2021). Avatar-based self-influence in a traditional CMC environment. *Journal of Media Psychology: Theories, Methods, and Applications*. 10.1027/1864-1105/a000309

Bian, Y., Zhou, C., Tian, Y., Wang, P., & Gao, F. (2015). The Proteus effect: Influence of avatar appearance on social interaction in virtual environments. *HCI International—Posters' Extended Abstracts*, 78–83.

Borenstein, M. (2005). Software for publication bias. In H. R. Rothstein, A. J. Sutton, & M. Borenstein (Eds.), *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*, 193–220. Wiley.

Buisine, S., & Guegan, J. (2020). Proteus vs. social identity effects on virtual brainstorming. *Behaviour & Information Technology*, *39*(5), 594–606.

Buisine, S., Guegan, J., Barré, J., Segonds, F., & Aoussat, A. (2016). Using avatars to tailor ideation process to innovation strategy. *Cognition, Technology & Work*, *18*(3), 583–594.

Carpenter, C. J. (2012). A trim and fill examination of the extent of publication bias in communication research. *Communication Methods and Measures*, *6*(1), 41–55. 10.1080/19312458.2011.651347

Chang, F., Luo, M., Walton, G., Aguilar, L., & Bailenson, J. (2019). Stereotype threat in virtual learning environments: Effects of avatar gender and sexist behavior on women's math learning outcomes. *Cyberpsychology, Behavior and Social Networking*, *22*(10), 634–640. 10.1089/cyber.2019.0106

Chen, G. M., Schweisberger, V. N., & Gilmore, K. (2012). "Conductor effect": Violent video game play extends anger, leading to triggered displaced aggression among women. *62nd Annual Conference of the International Communication Association*.

Christou, C., & Michael, D. (2014). Aliens versus humans: Do avatars make a difference in how we play the game? *6th International Conference on Games and Virtual Worlds for Serious Applications*. 10.1109/vs-games.2014.7012029

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, *19*(2), 272–309. 10.1080/15213269.2015.1015740

De Rooij, A., van der Land, S., & van Erp, S. (2017). The creative Proteus effect: How self-similarity, embodiment, and priming of creative stereotypes with avatars influences creative ideation. *Proceedings of the ACM SIGCHI Conference on Creativity and Cognition*, 232–236.

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2022). Fixed-effect vs random-effects models for meta-analysis: 3 points to consider. *Global Spine Journal*, *12*(7), 1624–1626. 10.1177/21925682221110527

Drummond, A., Sauer, J. D., & Ferguson, C. J. (2020). Do longitudinal studies support long-term relationships between aggressive game play and youth aggressive behaviour? A meta-analytic examination. *Royal Society Open Science*, *7*(7), 200373. 10.1098/rsos.200373

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. 10.1111/j.0006-341X.2000.00455.x

Ferguson, C. J., Gryshyna, A., Kim, J. S., Knowles, E., Nadeem, Z., Cardozo, I., Esser, C., Trebbi, V., & Willis, E. (2022). Video games, frustration, violence, and virtual reality: Two studies. *The British Journal of Social Psychology*, *61*(1), 83–99. 10.1111/bjso.12471

Fokides, E. (2021). My avatar and I. A study on avatars, personality traits, self-attributes, and their perceived importance. *Journal of Ambient Intelligence and Humanized Computing*, *12*(1), 359–373. 10.1007/s12652-020-01977-1

Fox, J., Bailenson, J., & Binney, J. (2009). Virtual experiences, physical behaviors: The effect of presence on imitation of an eating avatar. *Presence: Teleoperators and Virtual Environments*, *18*(4), 294–303. 10.1162/pres.18.4.294

Fox, J., & Bailenson, J. N. (2009). Virtual self-modeling: The effects of vicarious reinforcement and identification on exercise behaviors. *Media Psychology*, *12*(1), 1–25. 10.1080/15213260802669474

Fox, J., Bailenson, J. N., & Tricase, L. (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. 10.1016/j.chb.2012.12.027

Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 529–536.

Gomes, S. B. (2012). *Avatars and self-concept change: The effects of embodiment on identity shift in computer-mediated environments*. Retrieved from https://search.proquest.com/openview/5e05f98f8e56e36f7b024b5e43c02000

Guegan, J., Buisine, S., Mantelet, F., Maranzana, N., & Segonds, F. (2016). Avatar-mediated creativity: When embodying inventors makes engineers more creative. *Computers in Human Behavior*, *61*, 165–175. 10.1016/j.chb.2016.03.024

Hamilton, M. A., & Nowak, K. L. (2010). Advancing a model of avatar evaluation and selection. *PsychNology Journal*, *8*(1). http://www.psychnology.org/File/PNJ8(1)/PSYCHNOLOGY_JOURNAL_8_1_HAMILTON.pdf

Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research*, *48*, S23–S37. 10.1509/jmkr.48.SPL.S23

Hilgard, J. (2019). Comment on Yoon and Vargas (2014): An implausibly large effect from implausibly invariant data. *Psychological Science*, *30*(7), 1099–1102. 10.1177/0956797618815434

Kang, S.-H., & Watt, J. H. (2013). The impact of avatar realism and anonymity on effective communication via mobile devices. *Computers in Human Behavior*, *29*(3), 1169–1181. 10.1016/j.chb.2012.10.010

Kaye, L. K., Pennington, C. R., & McCann, J. J. (2018). Do casual gaming environments evoke stereotype threat? Examining the effects of explicit priming and avatar gender. *Computers in Human Behavior*, *78*, 142–150. 10.1016/j.chb.2017.09.031

Kilteni, K., Bergstrom, I., & Slater, M. (2013). Drumming in immersive virtual reality: The body shapes the way we play. *IEEE Transactions on Visualization and Computer Graphics*, *19*(4), 597–605. 10.1109/TVCG.2013.29

Kocur, M., Kloss, M., Schwind, V., Wolff, C., & Henze, N. (2020). Flexing muscles in virtual reality: effects of avatars' muscular appearance on physical performance. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 193–205.

Kocur, M., Schwind, V., & Henze, N. (2019). Utilizing the Proteus effect to improve interactions using full-body avatars in virtual reality. *Mensch Und Computer 2019-Workshipband*. 10.18420/muc2019-ws-584

Koda, T., & Oguri, R. (2019). Analysis of the effects of appearances of avatars on user's self-evaluation of extroversion. *ICAART* (1) (pp. 232–237). https://pdfs.semanticscholar.org/c003/d108beb0d8e9f2d9791c1778d4ec96d269e6.pdf

Koster, R. (2018). The cost of games. *VentureBeat*. https://venturebeat.com/2018/01/23/the-cost-of-games/

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., & Botsch, M. (2017). The effect of avatar realism in immersive social virtual realities. *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Article 39.

Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *British Medical Journal, 333*(7568), 597–600. 10.1136/bmj.333.7568.597

Lee, J.-E. R., Nass, C. I., & Bailenson, J. N. (2014). Does the mask govern the mind? Effects of arbitrary gender representation on quantitative task performance in avatar-represented virtual groups. *Cyberpsychology, Behavior and Social Networking*, *17*(4), 248–254. 10.1089/cyber.2013.0358

Lee-Won, R. J., Tang, W. Y., & Kibbe, M. R. (2017). When virtual muscularity enhances physical endurance: Masculinity threat and compensatory avatar customization among young male adults. *Cyberpsychology, Behavior and Social Networking*, *20*(1), 10–16. 10.1089/cyber.2016.0418

Leung, G. Y. S., Ng, A. K. T., & Lau, H. Y. K. (2021). Effect of height perception on state self-esteem and cognitive performance in virtual reality. *Engineering Psychology and Cognitive Ergonomics*, 172–184. 10.1007/978-3-030-77932-0_15

Levine, T. R., Asada, K. J., & Carpenter, C. (2009). Sample sizes and effect sizes are negatively correlated in meta-analyses: Evidence and implications of a publication bias against nonsignificant findings. *Communication Monographs*, *76*(3), 286–302. 10.1080/03637750903074685

Liao, G.-Y., Cheng, T. C. E., & Teng, C.-I. (2019). How do avatar attractiveness and customization impact online gamers' flow and loyalty? *Internet Research*, *29*(2), 349–366. 10.1108/IntR-11-2017-0463

Li, B. J., Lwin, M. O., & Jung, Y. (2014). Wii, myself, and size: The influence of Proteus effect and stereotype threat on overweight children's exercise motivation and behavior in exergames. *Games for Health Journal*, *3*(1), 40–48. 10.1089/g4h.2013.0081

McCain, J., Ahn, S. J. (Grace), & Keith Campbell, W. (2018). Is desirability of the trait a boundary condition of the Proteus effect? A pilot study. *Communication Research Reports*, *35*(5), 445–455. 10.1080/08824096.2018.1531212

Navarro, J., Cebolla, A., Llorens, R., Borrego, A., & Baños, R. M. (2020). Manipulating self-avatar body dimensions in virtual worlds to complement an internet-delivered intervention to increase physical activity in overweight women. *International Journal of Environmental Research and Public Health*, *17*(11). 10.3390/ijerph17114045

Ogawa, N., Narumi, T., Kuzuoka, H., & Hirose, M. (2020). Do you feel like passing through walls? Effect of self-avatar appearance on facilitating realistic behavior in virtual environments. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–14.

Oyanagi, A., & Ohmura, R. (2019). Transformation to a bird: Overcoming the height of fear by inducing the Proteus effect of the bird avatar. *Proceedings of the 2nd International Conference on Image and Graphics Processing*, 145–149.

Palomares, N. A., & Lee, E.-J. (2010). Virtual gender identity: The linguistic assimilation to gendered avatars in computer-mediated communication. *Journal of Language and Social Psychology*, *29*(1), 5–23. 10.1177/0261927X09351675

Peck, T. C., & Gonzalez-Franco, M. (2021). Avatar embodiment. A standardized questionnaire. *Frontiers in Virtual Reality*, *1*, 44. 10.3389/frvir.2020.575943

Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a Black avatar reduces implicit racial bias. *Consciousness and Cognition*, *22*(3), 779–787. 10.1016/j.concog.2013.04.016

Peña, J. (2011). Integrating the influence of perceiving and operating avatars under the automaticity model of priming effects. *Communication Theory*, *21*(2), 150–168. 10.1111/j.1468-2885.2011.01380.x

Peña, J., Hancock, J. T., & Merola, N. A. (2009). The priming effects of avatars in virtual settings. *Communication Research*, *36*(6), 838–856. 10.1177/0093650209346802

Peña, J., Khan, S., & Alexopoulos, C. (2016). I am what I see: How avatar and opponent agent body size affects physical activity among men playing exergames. *Journal of Computer-Mediated Communication*, *21*(3), 195–209. 10.1111/jcc4.12151

Peña, J., & Kim, E. (2014). Increasing exergame physical activity through self and opponent avatar appearance. *Computers in Human Behavior*, *41*, 262–267. 10.1016/j.chb.2014.09.038

Rahill, K. M., & Sebrechts, M. M. (2021). Effects of avatar player-similarity and player-construction on gaming performance. *Computers in Human Behavior Reports*, *4*, 100131. 10.1016/j.chbr.2021.100131

Ratan, R., Beyea, D., Li, B. J., & Graciano, L. (2020). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the Proteus effect. *Media Psychology*, *23*(5), 651–675. 10.1080/15213269.2019.1623698

Ratan, R., & Dawson, M. (2016). When Mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research*, *43*(8), 1065–1093. 10.1177/0093650215570652

Ratan, R., Rikard, R., Wanek, C., McKinley, M., Johnson, L., & Sah, Y. J. (2016). Introducing avatarification: An experimental examination of how avatars influence student motivation. *49th Hawaii International Conference on System Sciences*, 51–59.

Ratan, R., & Sah, Y. J. (2015). Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior*, *50*, 367–374. 10.1016/j.chb.2015.04.010

Reeves, B., Yeykelis, L., & Cummings, J. J. (2016). The use of media in media psychology. *Media Psychology*, *19*(1), 49–71. 10.1080/15213269.2015.1030083

Reinhard, R., Shah, K. G., Faust-Christmann, C. A., & Lachmann, T. (2020). Acting your avatar's age: Effects of virtual reality avatar embodiment on real life walking speed. *Media Psychology*, *23*(2), 293–315. 10.1080/15213269.2019.1598435

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*(1), 59–82. 10.1146/annurev.psych.52.1.59

Sah, Y. J., Ratan, R., Tsai, H.-Y. S., Peng, W., & Sarinopoulos, I. (2017). Are you what your avatar eats? Health-behavior effects of avatar-manifested self-concept. *Media Psychology*, *20*(4), 632–657. 10.1080/15213269.2016.1234397

Sherrick, B., Hoewe, J., & Waddell, T. F. (2014). The role of stereotypical beliefs in gender-based activation of the Proteus effect. *Computers in Human Behavior*, *38*, 17–24. 10.1016/j.chb.2014.05.010

Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, *78*, 64–73. 10.1016/j.chb.2017.09.012

Shin, M., Kim, S. J., & Biocca, F. (2019). The uncanny valley: No need for any further judgments when an avatar looks eerie. *Computers in Human Behavior*, *94*, 100–109. 10.1016/j.chb.2019.01.016

Slater, M. D., Johnson, B. K., & Cohen, J. (2014). Temporarily expanding the boundaries of the self: Motivations for entering the story world and implications for narrative effects. *Journal of Communication, 64*(3), 439–455. 10.1111/jcom.12100

Stavropoulos, V., Rennie, J., Morcos, M., Gomez, R., & Griffiths, M. D. (2021). 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.

Sylvia, Z., King, T. K., & Morse, B. J. (2014). Virtual ideals: The effect of video game play on male body image. *Computers in Human Behavior*, *37*, 183–188. 10.1016/j.chb.2014.04.029

Van Der Heide, B., Schumaker, E. M., Peterson, A. M., & Jones, E. B. (2013). The Proteus effect in dyadic communication: Examining the effect of avatar appearance in computer-mediated dyadic interaction. *Communication Research*, *40*(6), 838–860. 10.1177/0093650212438097

Via, C. M. (2016). *The Proteus Effect and Gaming: The Impact of Digital Actors and Race in a Virtual Environment* [Doctoral dissertation. Embry-Riddle Aeronautical University]. https://commons.erau.edu/edt/291

Walther, J. B., & Lew, Z. (2022). Self-transformation online through alternative presentations of self: a review, critique, and call for research. *Annals of the International Communication Association*, *46*(3), 135–158. 10.1080/23808985.2022.2096662

Won, A. S., Bailenson, J., Lee, J., & Lanier, J. (2015). Homuncular flexibility in virtual reality. *Journal of Computer-Mediated Communication, 20*(3), 241–259. 10.1111/jcc4.12107

Yee, N. (2007). *The Proteus effect: Modification of social behaviors via transformations of digital self-representation*. Stanford University.

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*(3), 271–290. 10.1111/j.1468-2958.2007.00299.x

Yee, N., & Bailenson, J. N. (2009). The difference between being and seeing: The relative contribution of self-perception and priming to behavioral changes via digital self-representation. *Media Psychology*, *12*(2), 195–209. 10.1080/15213260902849943

Yee, N., Bailenson, J. N., & Ducheneaut, N. (2009). The Proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research*, *36*(2), 285–312. 10.1177/0093650208330254