ANNALS OF THE INTERNATIONAL COMMUNICATION ASSOCIATION
2022, VOL. 46, NO. 3, 135–158
https://doi.org/10.1080/23808985.2022.2096662

Routledge
Taylor & Francis Group

Self-transformation online through alternative presentations of self: a review, critique, and call for research

Joseph B. Walther [a,b] and Zijian Lew [a,c]

[a]Department of Communication, University of California, Santa Barbara, CA, USA; [b]Center for Information Technology and Society, University of California, Santa Barbara, CA, USA; [c]Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore, Singapore

ABSTRACT

Communication technologies such as text-based chat, blogs, virtual reality, and avatars allow people to present deviations from their offline personality or appearance, causing changes to their personality perceptions and social behavior. This review of self-transformation through online self-presentation explores two major paradigms – the discursive approach and the embodiment approach – in terms of the theories on which they draw. It also examines differences and inconsistencies within and between these paradigms with respect to six factors that different studies suggest as critical for self-transformation. It reviews empirical research examining these critical factors, and provides a roadmap for future research testing competing explanations and their generality or specificity across digital platforms.

KEYWORDS

Self-presentation; self-perception; identity workshop; identity shift; Proteus effect; avatars; communication technology

Numerous studies suggest the beguiling possibility that when we act online as if we are someone or something else, or exaggerate aspects of ourselves for others, we change our own minds, alter our own behaviors, and modify our own self-concepts. Reports of such phenomena go back to some of the earliest pre-Internet stories such as 'The Strange Case of the Electronic Lover' (Van Gelder, 1991; see also Stone, 1995), and the questions they raised remain important as new technology and new affordances arise. This review addresses how digital technology affords or facilitates online performances that result in a *transformation of self* (or *self-transformation*), defined as a change in individuals' self-concepts, personalities, or specific attitudes and behaviors pertaining to individuals' own attributions about themselves.

In a variety of online settings – in online role-playing, in the way we describe ourselves to others in blogs or online chats, in avatar-oriented interaction spaces, in virtual reality labs, on social media, and elsewhere – individuals often present an alternative or distorted version of themselves to others. An *online* self-presentation is the presentation of oneself via some digital, mediated platform; we refer to *offline* experiences as those that take place face-to-face (or in anticipation of face-to-face encounters) without electronic mediation. Because online platforms offer various degrees of plasticity, online self-presentations can take many forms and can be more or less veridical to one's self-presentation offline. For the present purposes we are concerned with *alternative or distorted self-presentations*, that is, the actors' portrayal, depiction, or embodiment of themselves online in a way that systematically differs from their offline self-presentation and their normal concept of their personality

CONTACT Joseph B. Walther jwalther@ucsb.edu 4007 Social Sciences & Media Studies Building, Santa Barbara, CA 93106-4020 USA

and/or appearance, from subtle accentuations of a personality trait or physical appearance characteristic, to outright fictive personae. Lines of research concerning the links between distorted self-presentation and self-transformation have focused on two distinctly different forms of online performance: self-presentation through on-screen or immersive virtual reality avatars, which we refer to in this essay as the *embodiment* paradigm; or via text-based messaging, which we refer to as the *discursive* paradigm.

Despite the similarity of their effects on self-transformation, different theoretical explanations and causal mechanisms have arisen between these paradigms, typically without consideration of whether the explanations that are associated with one paradigm generalize to the other (see Valkenburg, 2017). For instance, whether the discursive paradigm's identity shift phenomenon is mutually exclusive from embodied avatars' Proteus effect (both of which are detailed below) has seldom been addressed.[1] Some pre-digital theories that are commonly cited *across* paradigms (e.g. self-perception theory; Bem, 1967) are appropriated quite differently, or in some cases, gratuitously, in attempts to explain online transformations of self. Moreover, *within* embodied self-transformation contexts, debate has arisen over the most appropriate theoretical explanation for the same research results. Both paradigms now sport their own home-grown theories, as well, that have not (yet) been applied to the other paradigm. Additionally, research practices have evolved within each paradigm that hamper the possibility of making comparisons, unless these practices were to be surfaced and implemented across the paradigms.

It is no mystery how this theoretical and empirical quagmire developed. Original research studies in related areas tend to spotlight individual mechanisms (see Bunge, 2004) in an isolated manner in order to test novel and unique explanatory approaches. Research designs that simultaneously test rival hypotheses are rare (cf. Kocur et al., 2020; Van Der Heide et al., 2013). As a result, any given research study is likely to focus on the operation of some factor in a given way (a tall avatar vs. a short avatar, for instance), with little consideration whether alternative operations of that factor render similar or different results (tall and short avatars vs. verbal self-description of being tall or short). Moving forward, to detect whether disparate self-transformation approaches reflect the same or different phenomena, and what theoretical principles best explain them, researchers must synthesize, scrutinize, and operationalize a variety of factors that have been hidden or dismissed, that have heretofore been encoded into their respective research designs.

In an attempt to identify these potentially illuminating or confounding influences, the present work nominates a number of elements that appear in various approaches to self-transformation, with systematic yet unintentional inconsistency. We refer to them throughout this essay as *critical factors*. These factors are only sometimes linked to a theoretical construct that a given study, or a line of studies, involves. Quite often they appear as operational choices in research designs that may activate hypothesized reactions, or isolate or control hypothesized effects from extraneous, potentially confounding influences. Nevertheless, these critical factors may represent potentially necessary and/or sufficient conditions that cause or moderate self-transformation through alternative self-presentations. Due to their inconsistent treatment across all studies of self-presentation and self-transformation (within but primarily between paradigms), they may comprise confounds hindering the ability to synthesize or generalize across the body of research as a whole.

The first two of these critical factors pertain to the most superficial, yet essential difference between the two paradigms: the form and focus of distorted online self-presentations. The second two critical factors pertain to participants' cognitive awareness of and volition over the distorted self-presentation they enact online. The final two critical factors focus on participants' awareness of, and interaction with, other social actors, as they present themselves online. All six factors appear critical to consider in order in order to rule out rival hypotheses to various empirically-based theoretical claims, or offer generalization across paradigms.

This essay proceeds in several steps. First, it introduces the critical factors that have emerged in self-transformation research, and it discusses their importance. Next, it reviews each paradigm and the theories they have appropriated to explain self-transformation effects, as well as inconsistencies

or complexities with those appropriations, and any conflicts among these theoretical accounts. These discussions include particular attention to how empirical studies have treated the critical factors and the implications of those factors for the theories' robust support. The next section examines how research in each tradition has empirically examined, measured, manipulated, controlled, or neglected, each of the critical factors, and how these results inform the status of theories to which they relate and our ability to generalize across approaches. Finally, it recommends a research agenda for exploring parsimonious and robust explanations of the phenomena and testing whether these explanations pertain across various online media platforms. By providing a framework for analyzing self-transformation phenomena through the prism of causal or facilitating factors, we provide greater conceptual clarity to the different manifestations of self-transformation and offer methods for integrating otherwise disparate areas of research.

Critical factors

Across different approaches to self-transformation, several practices appear with varying consistency. Some lines of research consider their presence essential to the respective phenomena. Others assert that they should be avoided or controlled. Yet others ignore them. For each critical factor, we will discuss how it is regarded and treated within each respective paradigm.

Form and focus factors

The first critical factor focuses on whether online self-presentation involves (1) a *physical or personality alteration*. Is there a difference depending on whether a self-presentation alteration involves an individual's personality, appearance, or other characteristic? The embodiment paradigm largely focuses on alterations to one's visually-perceived, physical appearance. In the text-based discursive paradigm, alterations can be made to both personality and physical appearance. Are these various aspects of alteration merely different pathways to the same self-transformation outcomes, or do they involve different processes and therefore affect outcomes differently?

The second critical factor is whether online self-presentation appears as a (2) *verbal, graphical, or embodied representation of self*. Whereas the embodiment paradigm focuses on embodied/visual self-representations, the discursive paradigm explores text-based self-presentation. Is there a difference whether participants self-present through verbiage alone, a verbal and graphical/visual medium, or an embodied environment? Do written media have more or less influence than embodied virtual reality or graphical environments?

Volitional factors

The next critical factors pertain to individuals' consciousness of their distorted self-presentations online. (3) *Awareness of alteration:* Must an individual be overtly aware of their altered online self-presentation, completely unaware, or possess some tacit awareness for self-transformation most strongly to occur? Early studies on the Proteus effect within the embodiment approach (e.g. Yee & Bailenson, 2009) claimed that participants must not be aware of (experimental) alterations in their appearance, so that participants' responses cannot be attributed to experimenter demand characteristics. Some level of tacit awareness seems necessary nevertheless for the 'body ownership illusion' upon which the embodiment approach relies (Kilteni et al., 2015), but the kind of and degree of awareness that is necessary appears inconsistently across studies and theories in this tradition. In contrast, participants in the discursive and identity shift paradigms (e.g. Gonzales & Hancock, 2008; Tice, 1992) are overtly aware of the alternative self-presentations they chose or that they were explicitly instructed to perform. Do awareness or unawareness of the alternative self-presentations effect self-transformation processes outcomes differently?

The fourth factor focuses on the (4) *agency of choice or customization over one's online presentation*. In many instances, researchers specified the alternative self-presentations participants were required to enact, either generally (e.g. instructed participants' discourse to reflect extraversion or introversion; Gonzales & Hancock, 2008) or with specific computer-generated embodiment alterations applied by researchers (e.g. modifications in one's apparent height or physical attractiveness; Yee & Bailenson, 2007). In the discursive paradigm (e.g. Bruckman, 1992) and in some embodiment studies, people made their own choices (e.g. Aymerich-Franch et al., 2014) or modifications (e.g. Ratan & Sah, 2015) of their self-presentations. Is there a difference whether the alteration is selected by the participant or assigned by researchers? Different lines of research seem to require entirely opposite approaches on this issue.

Social factors

The fifth critical factor pertains to the (5) *publicness of the performance setting*. Different paradigms feature different assumptions about the necessity of an audience or observers to one's self-presentation in order to stimulate self-transformation. Within the discursive approach, social factors are critical, and the identity shift approach explicitly hypothesizes that altered self-presentations occur before a virtual, public audience, rather than in private, for transformation of self to occur (Carr & Foreman, 2016; Gonzales & Hancock, 2008). In the embodiment paradigm, some studies featured public self-presentations (e.g. Fox et al., 2013), some featured private performances (e.g. Peña et al., 2009), and other studies provided an illusion of publicness (e.g. Peña & Hill, 2020). Is self-presentation in a public setting essential for self-transformation outcomes?

The final factor is generally an essential aspect of communication, but the two major paradigms differ markedly in terms of its role in self-transformation: (6) *social interaction*. The discursive approach has been most explicit on the importance of social interaction to self-transformation (e.g. Parks & Roberts, 1998; Walther et al., 2011). In the embodiment approach, social interaction has generally been controlled and scripted when it appeared at all (e.g. Peña, 2011; Yee & Bailenson, 2007), and is not generally theorized to be critical to self-transformation. Is social interaction or feedback critical to self-transformation outcomes?

In addition to these six critical factors, our research also examines differences in observed outcomes: What outcomes are specified, among self-report measures of self-concept or behavioral measures that are interpreted to represent transformation of self?

Self-transformation research paradigms

In this section, we discuss two major paradigms of research: the embodiment paradigm in which people appear on screens or in virtual space in a visually altered representation of themselves, and the discursive paradigm, in which individuals generate altered self-presentations through writing. Within each paradigm, we focus on the theoretical issues and debates regarding its key mechanisms that can potentially be clarified by research that explicitly surfaces and examines one or more of our six critical factors. We start with the embodiment paradigm, which is accorded greater discussion because the greater number of inconsistencies/disagreements therein require more nuanced treatment in terms of theories, and how different theoretical approaches affect the relevant critical factors. Following that, we present the discursive paradigm in a comparatively briefer manner, because there appears to be less dissensus about its theoretical statements and, at the same time, more empirical research that begins to examine its links to our six critical factors.

The embodiment paradigm

Graphical digital environments, such as those found in online games or virtual reality, provide avenues through which individuals can be represented in alternative ways. Research in this area

explores how the mutability of digital self-representations affects users' behaviors and self-concepts (see Ducheneaut et al., 2009; Van Looy et al., 2012). Graphical means of alternative *physical* self-presentations also appear to result in behavioral and self-perception changes, as primary studies and a meta-analysis show (Ratan et al., 2020). Studies of self-transformation in graphical digital environments often examine how the appearance of an *avatar* – a graphical embodiment that is operated, typically in real-time, by a specific human being that it resembles – can influence the thoughts and behaviors of its user (Bailenson & Blascovich, 2004). These studies differ from the discursive paradigm with respect to several critical factors (e.g. verbal/graphical/embodied environment, publicness, and social interaction). Additional variation also exists among studies within this paradigm in terms of how various critical factors are acknowledged and implemented. Different theoretical explanations appear describing the underlying mechanisms responsible for the embodiment effect: The Proteus effect, relying to some extent on self-perception theory (Yee & Bailenson, 2007), priming (Peña et al., 2009), and avatar self-relevance (Ratan & Dawson, 2016).

The Proteus effect

Among various graphical environments, the use of embodied virtual reality to alter an individual's appearance and trigger changes in behavior and/or self-concept has been called the *Proteus effect* (Yee & Bailenson, 2007). Participants wear eye goggles and apparatus that allow remote tracking of their major limb and head movements, creating an imaginary environment in which they appear to control life-size replicas of themselves. Experiments have modified various physical dimensions of users' avatars (e.g. physical attractiveness, relative height), finding that such modifications can stimulate changes in user's behaviors. Empirical data generally show consistent support for the outcomes of the Proteus effect.

In the original work on the Proteus effect, Yee and Bailenson (2007) describe the Proteus effect using self-perception theory (Bem, 1967). Self-perception theory argues that as individuals observe their own behavior, they cognitively adopt the same perspective as a third-person observer, and infer – from their behavior – their own internal attitudes and beliefs (Bem, 1972). According to Yee and Bailenson (2007), in the digital realm, individuals observe their avatars as the embodiment of themselves, and make inferences about appropriate behavior based on the appearance of their avatars. That is, individuals 'may conform to the expectations and stereotypes of the identity of their avatars' insofar as behavior is concerned (Yee & Bailenson, 2007, p. 274).

In a series of experiments conducted using embodied, immersive virtual reality, Yee et al. (2009) randomly assigned participants to embodied avatars that looked like themselves, but who appeared to be either 10 cm taller or shorter than a confederate. Users of taller avatars negotiated more assertively, consistent with stereotypes equating height to assertiveness and dominance (Young & French, 1996). This effect also occurred with respect to physical attractiveness: Avatars' faces comprised the participants' own photos, digitally morphed to appear more attractive, or less attractive, than the participants' natural faces. Participants with more attractive avatars behaved in a more sociable manner toward a confederate than did users of unattractive avatars, following the stereotype that physically attractive people tend to also be sociable and friendly (Yee & Bailenson, 2007). Lee et al. (2014) also found that participants who were represented by a male avatar – regardless of their actual gender – showed the strongest performance when competing against two female avatars (which were in reality controlled by scripts) in solving math equations, compared to other combinations of opponent gender, consistent with normative stereotypes linking females to lack of math ability. There are sometimes confederates in these studies but in most cases their conversations are tightly scripted.

Despite the reference to self-perception theory, its application of the Proteus effect remains unclear. There is a discrepancy between the process predicted by Bem's (1967, 1972) articulation of self-perception theory and the process described in the Proteus effect literature (Yee & Bailenson, 2007). Theoretically, self-perception theory (Bem, 1972) begins with observation of one's *behavior* (the causal factor), which leads to a perception of oneself in terms of self-concept, opinions,

preferences, and other attitudes (the measured outcome). The Proteus effect (Yee & Bailenson, 2007) begins with observation of an avatar's *appearance* (the causal factor), which leads to the activation of certain self-concept perceptions (based on the stereotypes associated with the avatar's appearance), which then leads to stereotype-consistent *behavior* (the measured outcome).

This raises two problematic issues. First, behavior is a causal influence in self-perception theory (Bem, 1972) but it is an outcome in the self-perception explanation for the Proteus effect (Yee & Bailenson, 2007), at least in early Proteus effect studies. To the extent that Proteus effect research does not measure attitudinal/self-concept outcomes (e.g. Lee et al., 2014; Yee & Bailenson, 2007, 2009), there is no evidence that self-perception theory is the mechanism driving the Proteus effect. This first issue is a matter of measurement: It does not invalidate the arguments made by Yee and Bailenson (2007), and it can be resolved (see e.g. Fox et al., 2013). The second issue is whether it makes sense to accept the self-perception explanation for the Proteus effect when the Proteus effect predicts that people change their behaviors (Yee & Bailenson, 2007) or attitudes (Fox et al., 2013) based on the *appearance* of their avatars, despite the fact that self-perception theory states unequivocally that people's inference about their *behavior* is what influences self-concept (Bem, 1972). The act of merely observing an avatar and perceiving that the avatar represents oneself is not consistent with self-perception *theory*, although some self-perception (as an independent concept) can occur. Making inferences from observed behavior (self-perception as a theory) and activating stereotypes/expectations based on appearances (self-perception as a concept) are two very different mechanisms.

Priming

Peña (2011), in contrast, rejects the application of self-perception theory to explain embodiment effects on self-transformation. He contends that the effect is better understood through priming theory. The priming argument posits that an avatar's appearance serves as a situational cue that activates (i.e. primes) neurological networks of associated concepts (i.e. nodes) in users' minds. These primed networks then render users more likely to think and act in ways consistent with the activated concepts (see Peña, 2011).

To test this explanation, Peña et al. (2009, experiment 2) examined embodiment effects using non-immersive two-dimensional avatars to represent participants. Participants toured a virtual museum while controlling avatars that either dressed as a Ku-Klux-Klan (KKK) member or a doctor. They were told to write stories for two pictures in the virtual museum that were designed to allow several equivocal interpretations. Their stories were later rated by judges for various themes like aggression. Results showed that participants whose avatars dressed as KKK members wrote stories with more aggressive and antisocial themes than participants whose avatars dressed as doctors. Another study reported also found users of black-robed avatars reported more aggressive intentions than users of white-robed avatars in group discussions of how to punish a hypothetical other (Peña et al., 2009, experiment 1). The authors link the observed increase in aggression to stereotypical associations between the color black and themes of evil and hostility (Adams & Osgood, 1973). In all, Peña et al. (2009) argued that KKK gowns and black robes primed participants to be more aggressive, and medical gowns and white robes primed others to be less aggressive.

Peña et al. (2009) also argued that the priming explanation is more parsimonious than Yee and Bailenson's (2007) self-perception explanation. The priming approach avoids the theoretical requirement, central to the application of self-perception theory, for self-perception or self-concept change to operate as an intervening construct. It also implies that after the priming effect wears off, people should not experience any long-term behavioral change. It is not clear that self-perception has been empirically demonstrated to mediate between alternative self-presentation and behavior change, in the research that advocates this position, nor has it been empirically excluded in research that challenges the self-perception approach. Such tests may prove difficult to do, however, as the self-perception process is not posited to be a conscious one. Indeed, some Proteus effect research attempts to demonstrate that the effects of visual alterations to avatars that lead to behavioral changes

operate below the level of consciousness (Yee & Bailenson, 2009). In terms of the critical factors, this means that participants have no awareness of the alternative self-presentation. The potential necessity that self-perception must be unconscious, however, also remains untested.

Avatar self-relevance theory

A third theory has been developed and applied to the accumulated results of embodiment self-transformation research. Ratan and Dawson (2016) argue that the degree to which a user perceives an avatar as relevant to the individual's self moderates the effect of avatar use on self-transformation. Relevance can be related to the similarity of an avatar's appearance and/or behavior to the avatar's operator, and the theory proposes that a number of antecedent factors can increase avatar self-relevance. Three such specific factors have been advanced: The first factor is users' ability to customize their own avatars, which can take the form of selecting an avatar from a set of options or by altering the features of an avatar. Another factor is the existing similarity between the avatar and the user's self; in one study, Ratan and Dawson (2016) operationalized similarity as the match between an avatar's apparent gender and the user's gender. A third factor is a control over an avatar's actions (rather than merely observing it).

The authors of avatar self-relevance theory (Ratan & Dawson, 2016) argue that the framework 'is situated between the self-perception and priming-oriented explanations for the Proteus effect' (Ratan & Dawson, 2016, p. 1066). They describe the application of the self-perception process, in virtual reality, as conforming to the behavioral expectations associated with one's avatar's characteristics (rather than, as Bem argued, observing oneself and inferring attitude from observed behavior). Priming also comes into play, they suggest: Through avatar use, 'the schema of self-related concepts becomes associated with the schema of avatar-related concepts, and thus the priming of one schema potentially activates the other' (p. 1066).

The evidence supporting the self-relevance model is mixed, although the model itself is relatively young. One experiment (Ratan & Dawson, 2016) found inconsistent associations between both customizing an avatar, and between gender similarity among an avatar and its user, with various physiological measures predicted to change as a result of variations in the antecedent conditions. Ratan et al. (2020) suggest that meta-analytic results also support the theory post hoc, in that Proteus effect studies that examined interaction effects (including such variables as user control, customization, user gender, and user race) accounted for greater variance than did studies that examined only main effects. Nevertheless, effect sizes among these particular studies were not uniformly greater than many others in their meta-analysis, leaving the conclusion questionable.

Regardless of whether the effects of embodiment through some kind of alteration to one's avatar rely on an application of self-perception theory, priming theory, or self-relevance, empirical studies tend not to differ in terms of most of the critical factors their experiments reflect, the exception being users' awareness of deliberate changes to their self-presentation. While the original Proteus effect studies were adamant that research participants were not overtly aware of alterations to their avatars (e.g. Yee & Bailenson, 2007), later studies focused on the effects of participants' active manipulation thereof (e.g. Ratan & Sah, 2015). Across studies, though, alterations appear in the physical appearance of avatars (rather than users' personality), that are hypothetically linked to stereotypical personality or social category characteristics. The environments are graphical, although they may be immersive (that is, involving the use of virtual reality headsets and first-person visual perspective) or not (e.g. avatars on a computer screen). The agency of the choice of alterations is generally the researchers' rather than the users', except in cases where customization is a variable. Performances are considered private rather than public, despite the presence of a lone confederate in some cases. However, it is possible that participants experience some sense of audience, since the extensive apparatus and laboratory settings involved in these studies may plausibly lead participants to believe they are being observed by researchers. Dynamic social interaction is not hypothesized to have an impact; confederates are often scripted, and spontaneous social interaction has been treated as a potential confound, experimentally controlled by researchers.[2] In some cases, there

are ad hoc conversations in group experiments, but they are not considered essential to the theoretical mechanism (see Lee et al., 2014). The outcome measures are most often behavioral, yet self-concept change is implicitly assumed to occur according to some accounts, even if self-concept change is not empirically measured.

The discursive paradigm

A second important tradition in the study of online self-presentation and its effects on self-concept appears in research that examined settings in which people represent some altered or fictive version of themselves through discourse, that is, through electronic written communication, without embodiment in any visual form. This section briefly described some of the earliest ethnographies of Internet culture that explored these themes, largely through the prism of postmodern identity issues or psychoanalytic approaches to identity, collectively referred to as *identity workshops*. Second, we examine at greater length more recent work, in which social psychological theories and experimental methods have advanced the understanding of self-transformation through discursive self-presentation as seen in a body of research referred to as *identity shift*.

Identity workshops

Prior to the Internet boom in the late 1990s, a number of treatises exposed and explored the use of online social interaction in text-based virtual reality systems and the effects of online role-playing on individuals' online and offline identities. These studies generally asserted that individuals' online pretensions had profound effects on their self-concepts. The studies themselves had a large impact on the emerging understanding of cyberpsychology and computer-mediated communication and society.

Many such studies reported interactions that took place in real-time, multi-user conversation spaces variously known as MOOs and MUDs, the entire representation of which was textual. These environments allowed individuals to create characters online (specifying a name, gender, and textual self-description that other users could see). Characters chatted in the form of statements as well as emotes (descriptions of physical or mental actions that appeared as if narrated; Curtis, 1992). Users forged relationships with other characters/users that developed over time, many of which became intensely intimate (Parks & Roberts, 1998). Research on these domains relied primarily on ethnographic observation, participant-observer accounts, and interviews, and cited various humanistic psychology theories, sociological approaches, and postmodern philosophers. Specific theoretical mechanisms or propositions were seldom advanced.

Bruckman (1992, 1993) was the first to describe text-based virtual worlds as 'identity workshops.' She described how creating and enacting characters whose gender or other aspects of identity differed from their users' offline characteristics provided lessons to users about how gender, power, and hierarchy operate in social interaction. Turkle (1995) extended this approach, and suggested psychoanalytic benefits from inventing a character in a virtual space and observing others' reactions to its behavior, potentially improving one's real-life perspective-taking abilities and empathy.

In terms of critical factors, Bruckman's work strongly emphasizes participants' voluntarily self-selected alteration of traits, personalities, and even physical characteristics as implied by textual descriptions. The communication environment was discursive, participants were aware of their manipulations, and the public aspects could include intimate one-on-one interactions in private conversation spaces, to interactions involving dozens of other characters. Dynamic social interaction was present and it constituted a matter of critical conceptual importance in this research: How other actors behaved in response to one's alternative self-presentations provided the experiences from which to learn. In Bruckman's work, users were self-aware and learning was overt, supplemented for some participants through other online discussions about what people experienced in the

virtual environment. Outcomes were related to self-concept and insights into social dynamics at a larger scale.

As a whole, the identity workshop research has several implications for the transformation of self and a number of critical factors. First, the text-based virtual reality chats described in these early studies were voluntary and naturally occurring, unlike the more contemporary approaches such as identity shift (discussed below) or embodiment experiments in which researchers induced self-presentation distortions. More recent approaches do not typically consider the role of volition (cf. Aymerich-Franch et al., 2014). The older identity workshop reports suggest that agency and volition are worthwhile and potentially important factors that have been dormant in the research until their rediscovery by Ratan and Dawson's (2016) work (in the embodiment domain) on avatar self-relevance. Second, to the extent that alterations of self that are voluntarily chosen, priming can be ruled out as a causal explanation. Third, these older approaches show that transformations of self can occur even in very rudimentary online textual interfaces. The assumption may be quite false that highly graphical or embodied digital environments exclusively or more strongly facilitate the body ownership illusion, compared to older and allegedly more primitive text-based environments. Anecdotal reports suggest alarmingly potent body ownership even in 1990s-era text-based virtual environments (e.g. 'A Rape in Cyberspace,' Dibbel, 1993).

Identity shift

Concurrent with the initial publications on the Proteus effect, seemingly unrelated research started on selective self-presentation in text-based environments and its effects on one's self-concept (Gonzales, 2006). The research was premised on pre-digital social psychological studies that examined how the publicness (i.e. public or private context) of self-presentational performances influences one's perception of self. Specifically, public performances of self-presentations, in which participants deliberately accentuate some trait, have a greater impact on individuals' subsequent self-perceptions on that trait than do equivalent performances that are private (Schlenker et al., 1994; Tice, 1992). Original research in online self-presentation drew on these prior notions in an effort to expand the hyperpersonal model of computer-mediated communication (CMC), focusing on the transformational effects of selective self-presentation (see Walther, 1996; Walther & Whitty, 2021).

Gonzales (2006; Gonzales & Hancock, 2008) adapted the public presentation research, described above, to CMC contexts. The original experiment induced participants to respond to questions as if they were extraverted or introverted, via typed text, in either a public blog or in a private document. Participants' subsequent ratings of their own personalities differed in extraversion/introversion when participants believed they were self-presenting in a public blog rather than a private document. Gonzales and Hancock labeled this phenomenon *identity shift*. In terms of critical factors, the original identity shift research involved an online alteration of a personality trait, of which participants were overtly aware, for a public audience specifically compared to a private setting, without social interaction, ultimately affecting offline personality.

Since the introductory identity shift research, a number of studies have accomplished much: increased its scope to include other text-based CMC platforms, its openness to other foci besides personality traits, its consideration of secondary influences other than publicness effects, its precision, as well as its boundaries. A summary of both published and unpublished findings recently appeared, referring to identity shift as an independent theory (Carr et al., 2021). Despite this recent explication, the theoretical basis of the identity shift phenomenon remains somewhat conflicted, complicating the facility with which to compare its premises and predictions to those theories and practices of the embodiment paradigm.

One theoretical conflict pertains to the role of biased scanning, which Gonzales (2006) asserted to be an important mechanism in the identity shift process. Biased scanning involves searching one's autobiographical memory for instances consistent with a particular self-presentation goal (Tice, 1992), in order to draw on them to describe oneself in ways that exemplify the trait requested by the researchers. It has been instantiated by researchers' instructions to participants to 'think of

experiences in both their past and present that were consistent with (extraversion or introversion) and to rely on those examples when answering questions' in their discursive self-presentations (Gonzales & Hancock, 2008, p. 173).

Despite its clear explication in the original identity shift experiment, empirical research has yet to isolate the effect of biased scanning as a mediator between altered self-presentation in public rather than private settings, and its effects on self-transformation. In fact, the construct seems to have disappeared from all published identity shift studies. It may be theoretically unnecessary, and detract from the parsimony of a model that does not otherwise need it. However, other issues in the present state of identity shift research call this into question. From a biased scanning perspective, one might argue that if individuals lack the relevant experience they are triggered to scan, or if memories of relevant experience are unavailable, biased scanning might not take place, commuting the identity shift process. Given that a recent review concluded that 'that there may be some facets of one's personality not susceptible to identity shift' (Carr et al., 2021, p. 204), it may be that it is not certain traits, but the lack of personal experience and memory of enacting those traits, that renders them unavailable to biased scanning and therefore to identity shift. This notion would be consistent with Carr et al.'s assertion that self-presentations must be authentic, not fabricated or invented: 'participants must draw from personal experiences to experience self-transformation' (p. 205; see also propositions 4 and 5 in Carr et al., 2021).

Indeed, the Carr et al. review went so far as to assert that the one's online self-portrayal of alternative social categories, deceptive performances, an animal, or any fabricated characteristic that is untrue to one's authentic self 'may not experience any transformation' (p. 209). Although such a view may be consistent with the 'lost' biased scanning notion, it is inconsistent with Gonzales and Hancock's (2008, p. 170) contention that identity shift effects occur even for self-presentations that are 'artificially induced' by researchers. It would be further countered by the anecdotal and ethnographic accounts of transformation effects emanating from gender-swaps and other pretensions as were described in the identity workshop research that was summarized earlier in this essay. If, on the other hand, it is true that individuals cannot experience identity shift for fictive self-presentations, then however identity shift operates, it does so differently than the Proteus effect and priming explanations that have been applied to the embodiment approach to self-transformation effects, yet with potentially greater resemblance to the avatar self-relevance approach within that paradigm.

Despite this theoretical ambiguity, the identity shift research has expanded its analytic precision with respect to a number of the critical factors discussed in this essay. For instance, although identity shift was initially introduced as an intrapersonal phenomenon independent of social interaction, subsequent research (Walther et al., 2011) found that a single instance of (prefabricated) feedback confirming an individual's selective performance magnified the identity shift effect. This adds an explicit test of variations within the critical factor of social interaction (specifically adding controlled social interaction versus no social interaction). The same study also found that the language styles used across treatment groups differed in ways prior research was associated with introversion/extraversion, thus adding a behavioral outcome to identity shift research.

Further research examined disconfirmatory feedback's potential to attenuate self-transformation in relation to attitudes toward a product brand. An experiment by Carr and Hayes (2017) induced participants to present themselves online as either a highly loyal brand advocate or not, to which they received either confirmatory or disconfirmatory feedback. Significant differences in attitude toward the brand resulted from both the nature of the feedback and the initial self-presentation. Turning back to the critical factors, identity shift research as a whole can be said to have isolated and tested variations related to the critical factors of social interaction and publicity: Audiences have been public, absent, or intimate (relationally close), and social interaction in the form of feedback has been present yet controlled, or absent.

Research in the public self-presentation and identity shift traditions involves changes to behavior, personality, or self-concept as a result of alternative self-presentations through sending verbal (i.e.

oral or written) messages during which participants are entirely aware of the alteration that they enact (by choice or by experimental induction). In all cases, the actual or anticipated performance before some sort of audience enhances the self-transformation effect. Social interaction or anticipated social interaction with others was assumed to be a part of the mechanism in most cases, although the role of public responses in the accentuation of self-transformation effects was only rarely demonstrated empirically.

Empirical evaluations of critical factors in the process of self-transformation

Our introduction of various approaches to self-transformation through alternative online self-presentation nominated a variety of critical factors that potentially affect how individuals' modified self-presentations affect their self-concepts. The necessity or potency of any of these factors across platforms is difficult to assess, because of the considerable variation with which these factors are implicated in different theories, and especially, as variables or constants in different research settings. Not all factors have been directly tested. Some factors may often (but not always) be implicitly active when another factor also is (for example, there is usually some degree of publicness of one's behavior in research in which the variable under study is feedback; how can there be feedback when one's behavior is not observable to others?).

Looking across paradigms, research must ask why some studies implement critical research factors one way, while other studies implement them differently, and others, not at all. For example, consider whether participants should or should not be consciously aware of an alteration in their self-presentation, reflected in their avatar. Yee and Bailenson (2007) argued that they should be unaware, because otherwise, responses could reflect demand characteristics rather than the results of self-perception-based Proteus effects. Alternatively, in identity shift studies, participants get overt instruction to construct verbal self-presentation, actively and consciously, that exaggerate some aspect of their personalities. Is the latter approach necessary? These are the kinds of questions to which a review of research examining critical factors in self-transformation effects can lead.

In order to chart a course for future research, it is appropriate to assess which of the critical factors accompanying online self-transformation has been isolated and tested, rather than assumed, in research. In some cases, a critical factor has received considerable empirical study, refining what we know about its operation in one paradigm, raising further questions about its potential application across paradigms. In the following we examine evidence from previous online and offline research that may help to address the paradoxes presented by the operation of critical factors across online self-transformation research paradigms. The remainder of this section on critical factors may be read while referring to the roadmap for future research (Figure 1), which gives a macroscopic summary of the most cardinal issues that will be discussed.

Physical or personality alteration

The difference between studies involving explicit distortions of personality characteristics, or of physical characteristics (that may be associated with personality and its associated behavior) is one major distinction between research paradigms. For example, discursive, identity shift research most frequently involved altered self-presentations of personality characteristics, while embodiment research typically involved altered physical characteristics. Both research traditions posit changes to individuals' self-concept as a result. To complicate matters, the critical factor of physical versus personality alteration is often confounded with another critical factor: conscious awareness of alterations. In some Proteus effect studies, participants were not overtly aware of modifications to their (avatars') physical appearance. However, in discursive environments, participants are always aware of the exaggeration of personality traits because they, themselves, enact them. Therefore, the awareness or unawareness of distorted self-presentation factor provides a rival explanation to the personality/physicality distinction.

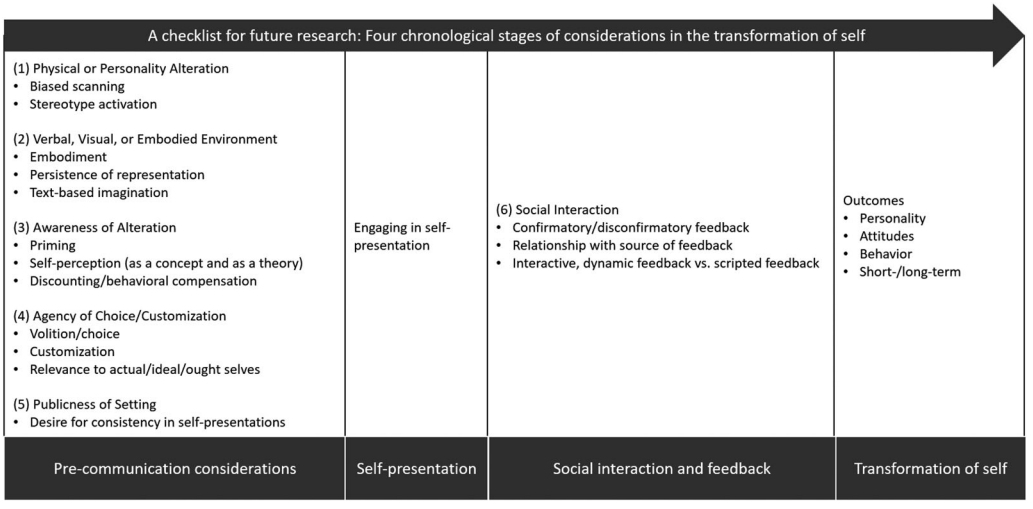

Figure 1. A roadmap for future research on the transformation of self.

Can we infer that discursive personality self-presentations and embodied physical alterations, which produce similar outcomes that are commensurate with the alterations, represent the same phenomenon? The theoretical mechanisms underlying personality or physical alterations have been described quite differently, but whether the phenomena occur due to biased scanning (Gonzales, 2006), or self-perception (Yee & Bailenson, 2007) or something else is not yet clear. Biased scanning has not been examined in responses to physical appearance alterations in embodiment studies. Although self-perception is mentioned in discussions of identity shift (Carr et al., 2021), it does not drive hypotheses in identity shift research. If the mechanism behind physical alterations is conformity to stereotypes and the mechanism behind personality alterations is biased scanning, then physical alterations and personality alterations involve completely different pathways to self-transformation. However, research has yet to try a verbal self-descriptive induction in an embodied environment. Few studies examined how characteristics of a visual avatar affected subsequent discursive messages. In one exception, women who were assigned male-looking avatars to represent them on computer screens apologized less, use tentative language less, and referenced their own emotional states less – linguistic behaviors normatively linked to femininity – when communicating with others online (Palomares & Lee, 2010; see also Beyea et al., 2021).

Verbal, graphical, or embodied environment

Research involving the presentation of alternative personality or physical characteristics have not been compared across various graphical or text-based environments using consistently focused or themed experimental manipulations. It is possible that embodied virtual reality, graphical avatars presented on a screen, and text-based discursive virtual realities all facilitate self-transformation by the same mechanism(s) with similar potency.

One distinction between 2D graphical versus immersive embodiment is the persistence of the representation. When research employs a consistent representation of a user over some episode, it may be unclear whether priming or self-perception explains effects. A typical immersive embodiment study begins with a brief glimpse by a participant into a virtual mirror, to establish a 'body ownership illusion' (Kilteni et al., 2015). In contrast, 2D avatar studies involve constant observation of one's avatar's appearance over the duration of the activity. Research by Lim and Reeves (2009) indicates that a third-person point of view of one's avatar magnifies the effect of choosing one's avatar,

on players' level of arousal in a game, compared to a first-person point of view, due to the persistent visual reminder of the avatar's appearance.

These differences in persistence have implications for different theoretical explanations of embodiment effects, as well. Traditional priming research suggests that a sequential experimental manipulation – whereby participants are first primed, then instructed to do a second task in the absence of the priming material – is sufficient to influence behavior (see for review Bargh & Ferguson, 2000). Peña et al. (2009) claimed that priming was the mechanism in their study, which found that people expressed more aggressive thoughts when they had Klan-like avatars than when they did not. In that study, participants *always* saw their avatars while they expressed their thoughts. As such, the persistence of the Klan avatar in the Peña et al. study raises the possibility that participants were engaging in self-perception; the study does not eliminate that possibility.

The degree of 'body ownership illusion' afforded by immersive rather than on-screen avatars is also unsettled. Some research shows that the kinesic one-to-one body-to-avatar mapping in immersive embodiment is more effective at changing people's attitudes. For example, a study showed that participants who were embodied in a Black avatar in an embodied virtual reality environment showed less post-induction racial prejudice than participants in the same virtual reality environment using a non-embodied Black avatar whose appendages did not match the participants' physical movements (Peck et al., 2013).

Other research, however, suggests that immersive embodiment may not be necessary for self-transformation, and that similar outcomes occur whether experimental inductions employ immersion or if users control two-dimensional graphic avatars on computer monitors or TV screens. For instance, women represented on screens embodied as average-sized avatars exerted more energy while playing a tennis video game on a Nintendo Wii compared to women embodied as overweight avatars (Peña & Kim, 2014). One study directly comparing 2D and immersive versions of avatars playing a violent videogame found no difference on post-game implicit attitudes toward aggression (Zumbach et al., 2015). Another experiment explored whether choosing a 2D avatar, and/or developing a sense of control over the avatar, led to greater self-transformation responses; avatar control (and not choice) affected transformation (Beyea et al., 2021). As a whole, extant research cannot rule out the possibility that it is control over an avatar, rather than the form of embodiment, that drives changes in self-perception and behavior. Klimmt et al. (2009, p. 354) suggest that controlling a 2D avatar in online games leads to identification with that avatar, in a 'nondyadic, or *monadic* user-character relationship' (emphasis in original). Controlling an avatar using a mouse and four-key controls (e.g. WASD keys), among other methods of avatar control, may be as effective as being immersively embodied in a virtual reality avatar in causing self-transformation.

Ultimately, much of the research makes no theoretical distinction between the two forms of embodiment insofar as transformation effects are concerned (e.g. Ratan & Dawson, 2016), and seems to assume that individuals identify with their avatars whether from the inside-out (immersively, with first-person perspective) or the outside-in (observing onscreen avatars as if a spectator of oneself). A good deal of the virtual reality literature seems premised on technology's provision of an immersive self-representation, through visual perspective, proprioceptive control, and potentially multi-sensory experience; yet, as Peck and Gonzalez-Franco (2021, p. 1) note, the body ownership illusion may be triggered by 'imaginative suggestion' and experimental demand characteristics, even in less immersive technologies.

Theoretically speaking, why may first-person and third-person views of an avatar differentially influence people's self-perception? If self-perception (Bem, 1972) is cardinal to self-transformation, must one literally see how oneself looks or see what oneself does? When people in text-based environments are unable to see a digital representation of themselves, yet are able to see what they write, how critical is first or third-person perspective and embodiment? In the original self-perception theory, 'seeing' what oneself does is a metaphor; in some of these online applications, it is literal. If self-perception is a mechanism for self-transformation, then additional research should address the on-screen and immersion characteristics as variables to see how these factors facilitate or inhibit self-perception and self-transformation (e.g. Walther et al., 2018).

Awareness of alteration

Awareness of alteration in self-presentation refers to whether participants are overtly aware that they present themselves in a way that differs from, or accentuates, their normal offline characteristics, or, alternatively, that they are induced by researchers to self-present in a certain way. The concern over participants' awareness of experimental manipulations stems from the possibility of a discounting effect: Participants may attribute their behaviors to others' experimental manipulations or instructions, recognize its fiction, and therefore not experience any transformation of self (see Festinger & Carlsmith, 1959).

In some research, participants were aware of the induction, because they were given explicit instructions to act in a certain way. For example, Gonzales (2006; Gonzales & Hancock, 2008), following Tice (1992), explicitly prompted participants to portray themselves either as introverted or extroverted. In other studies, participants were unaware of the manipulation (see, e.g. Chaiken & Baldwin, 1981; Fazio et al., 1981). Early Proteus studies argued participants' overt awareness of experimenters' alterations to their virtual appearance led to demand characteristics (Yee & Bailenson, 2007, 2009), whereas more recent interpretations of embodiment effects (Ratan & Dawson, 2016; Ratan et al., 2020) argue the opposite: Overt awareness of the manipulation of self-presentation actually strengthens self-transformation effects, when an avatar is overtly customized by a user (Ratan & Dawson, 2016).

Other effects of the alteration of one's appearance rely on one's overt recognition of one's self-presentation online. Van Der Heide et al. (2013) explored how the attractiveness of an avatar may affect one's expectation of others' responses, and whether these expectations in turn affect the way an avatar user communicates. Drawing on behavioral compensation mechanisms (see Tong & Walther, 2015), Van Der Heide et al. offered alternative hypotheses about the influence on a user from being embodied as an unattractive avatar: While the Proteus effect posits users adopt the behaviors associated with the avatar's appearance (i.e. behave in a way consistent with an unattractive appearance), behavioral compensation predicts that users behave in ways that counteract this appearance (i.e. behave more attractively). To test these alternative predictions, the researchers randomly assigned female dyad members to an attractive, unattractive, or no avatar representation of themselves, as they interacted online with a male partner. Females who saw unattractive avatars of themselves behaved more positively toward their partners than those in other conditions, lending credence to more complex behavioral compensation patterns rather than the main effects of most Proteus studies (Van Der Heide et al., 2013). Continued exploration of how Proteus-type effects interact with other interpersonal processes remains an area ripe for future investigation.

Despite the claims of early Proteus studies that participants lacked overt awareness of the alterations between their avatars' appearances and their own native appearance, participants involved in these studies must have experienced a tacit awareness of the self-representational alteration. That is, while participants may not be able to articulate the precise physical alteration they self-presented, participants were nevertheless exposed to the experimental manipulation at the self-perception level when they saw themselves in a virtual mirror, to establish the body ownership illusion described above (e.g. Peña et al., 2009; Yee & Bailenson, 2007, 2009). Indeed, failure to do so would render the experimental design ineffectual by virtue of not providing participants with the visual stimulus necessary to activate the Proteus effect, regardless of whether the effect is rooted in self-perception or priming.

Additional research has capitalized on participants' overt awareness that their embodied avatars' appearance differs from their normal appearance. Slater and colleagues, for instance, often employ multiple embodiment episodes but with alternating appearances that induce different perspectives. For example, Slater et al.'s (2019) two-stage experiment initially placed participants in an avatar of their own likeness, in which they described some personal problem they were having to another avatar that resembled Sigmund Freud. Then, participants in the control condition received scripted responses from Freud, while other participants swapped avatars with Freud. Embodied in Freud's

avatar, they give problem-solving advice to their own avatars. Although the outcome of this research strays somewhat from the transformation of self paradigm (did they become more Freud-like?), the treatment significantly increased participants' perspectives about and understanding of their own problem. Another two-stage, embodiment-switching study involved police officers who, in their own avatars, initially participated alongside another officer in an 'intimidatory interrogation clearly based on racial bias' of a Black suspect (Kishore et al., 2021, p. 8). In a second-stage, half the police participants were interrogated, themselves, embodied as the suspect. Those participants more often intervened, several weeks later, in a virtual café when another officer exhibited aggression toward an innocent Black customer. In this case, the effect did not make participants more like the suspect but rather, as intended, helped transform the self by engendering empathy and a third-person perspective about one's own behavior. In any case, Slater's approach in these and other studies has relied on participants' explicit awareness that their second-stage avatars are not themselves, but are rather seeing themselves through another perspective, through embodiment.

Evidence across research paradigms does not seem to reconcile these competing opinions whether ignorance or knowledge of the alteration of self-presentation produces self-transformation. Identity shift and public self-presentation studies that involve self-reports, linguistic analyses, and behavioral measures (sitting distance from a confederate and initiating conversation with the confederate) show that participants do become more introverted or extraverted after their self-presentations, even if they are aware of the manipulation (Gonzales & Hancock, 2008; Tice, 1992, experiment 2). In other research, participants who were unaware of the manipulation also showed similar shifts in their introversion or extroversion, according to self-reports and observer ratings (Fazio et al., 1981). As a whole, awareness of a manipulated self-presentation is important to the extent that self-transformation effects are caused by priming (for which there should not be overt awareness; Peña et al., 2009), negated by discounting (Festinger & Carlsmith, 1959), or enhanced due to avatar self-relevance (Ratan & Dawson, 2016).

Agency of choice/customization

One critical factor closely linked to awareness of alternative self-presentations is the extent to which alternative self-presentations are experimentally assigned by researchers or volitionally selected by participants. This factor can also be viewed as a continuum of user-customization, which at one end contains self-presentations that are completely researcher-assigned, and the other end complete liberty for people to customize physical features (e.g. skin tone, facial structure, body shape, accessories) and personality attributes (e.g. introversion/extraversion). If people select an avatar or enact certain text-based self-presentations, then they are necessarily aware of the alternative self-presentation. But if an alternative self-presentation is assigned by researchers, then participants may or may not be consciously aware of the manipulation.

Even within self-transformation studies using graphical environments, the results are inconsistent with respect to the extent to which alternative presentations of the self are assigned by researchers or selected by the participants themselves. Unlike early Proteus effect research, more recent research required participants to select or customize their avatar at the inception of virtual experiences (Ratan & Dawson, 2016; Ratan & Sah, 2015). Another study required participants consciously to alter the appearance of their avatars during a virtual experience where people's avatars got fatter the more junk food they ate (Hwang & Mamykina, 2017). Simply choosing one's avatar seems to increase the potential for its actions to affect users, in online game environments, compared to one's avatar being assigned by another agent (Lim & Reeves, 2009).

Recent research involving the avatar self-relevance theory (see for review Ratan et al., 2020) offers direct contradictions to Yee and Bailenson's (2007) original position on the potentially mitigating effects of users' awareness of manipulations to the presentation of self: The more participants choose their avatars' appearance through selection or modification, the stronger the effect of avatar use on self-transformation (and stronger Proteus effects are produced). However, the

concept of avatar self-relevance is also described as the similarity between an avatar and its user (Ratan & Dawson, 2016), and it is not clear, theoretically, how self-transformation may take place as users can customize their avatars to be relevant to different facets of their selves.

One theoretical opening left from Ratan and colleagues' research has to do with the theory's proposition regarding the effect of customization. The theory argues that the more one customizes one's avatar, the greater the relevance of the avatar to oneself, hence the greater its influence. But the theory does not specify that customization must make one's avatar more similar to one's offline self, only that customization (in an unspecified direction) increases relevance. Although one empirical study (Ratan & Dawson, 2016) did direct some participants to customize the avatar in the direction of similarity to self, research should explore whether customization of any nature generates greater self-relevance and greater transformation of self. A different approach to customization of one's avatar to reflect one's self follows from Higgins' (1987) notion of different selves: the actual self, the ideal self, and the ought self, which may or may not overlap with each other. So, rather than the notion that the more one modifies one's avatar the more one identifies with it, it may be that it entirely depends on the nature, not the effort, of the modification process; little self-transformation may be expected if avatars are modified to depict users' actual selves, but more self-transformation occurs if avatars are made to resemble ideal or ought selves. This very hypothesis was tested in an experiment by Sah et al. (2014), finding greater user identification with ideal- and ought-designed avatars than with actual-appearing avatars.

Future research should continue to explore the causal mechanisms of embodiment effects. Measures have not yet been applied in ways that could empirically disentangle the three potential explanatory mechanisms: self-perception, priming, and/or self-relevance. As such, existing attempts to validate one explanation over the other largely fall short of their goal. Yee and Bailenson (2009), for instance, attempted to discern whether priming was more effective as a result of embodiment as a Black or White person, than as a result of seeing photographs of Black or White persons; different embodied primes led to different attitudes, but no overall differences between embodiment and photo-based priming occurred. Such results are difficult to interpret as research generally continues to lack a clear theoretical understanding of how embodiment effects operate. As a consequence, it is not well understood how transformation of self may or may not take place as a result of choice or customization in graphical and/or embodied digital environments.

Future research can also consider several principles that appeared in the discursive paradigm's identity workshops research from text-based virtual reality and chats (Bruckman, 1992, 1993; Parks & Roberts, 1998). These older approaches assume that volition – and by extension, choice and/or customization – is the starting point of any transformation of self. They assume that long-lasting transformations of self are a result of people's desire to explore various aspects of their actual, ought, ideal, or repressed selves (Higgins, 1987; Turkle, 1995), and not a result of momentary experimental inductions.

Publicness of setting

Although publicness of the self-presentation has been a theoretical and empirical mainstay of the identity shift literature, it is not exactly clear from this literature why it enhances self-transformation. Earlier, offline research argued that public situations make individuals accountable to others and evokes a desire to portray a consistent self to others (Schlenker et al., 1994). Several offline experiments (Schlenker et al., 1994; Tice, 1992, Experiment 3) have determined that, among those who self-presented publicly, internalization of the self-presented trait was greater among those who anticipated meeting the person to whom they self-presented, compared to participants who did not anticipate such a meeting. While many online identity shift studies confirm that public self-presentations generate greater self-transformation than do private self-presentations, research has not ascertained whether the effect is due to being accountable to others, or if it is due to one of many concepts that may be associated with the phenomenon of publicness (e.g. Carr et al., 2021, proposition 2).

Offline research provides additional precedents that could account for publicness effects. In an offline study by Kelly and Rodriguez (2006), participants privately self-presented twice, once as introverted and once as extroverted. They recorded their self-presentations on video. However, one of the video recordings – either the one containing the introverted self-presentation or the one containing the extroverted self-presentation – was subsequently deleted in front of the participants. Participants were told that the undeleted video would be made public. Results indicated that participants experienced a transformation of self in line with the trait that was retained in the undeleted video. Had priming or self-perception been the more dominant mechanism for transformation of self than the need to appear consistent with one's public self-presentations, no significant differences would have been found regardless of whether the video was deleted or not. That is, participants would have been primed or would have self-perceived even if their self-presentations were subsequently erased. The results, therefore, suggest that public self-presentations, as well as the desire to appear consistent, can have an effect on self-concept change above and beyond the effect of priming or self-perception. Replication of these precedents in online self-transformation studies may be very telling.

The importance of perceived publicness of self-presentation also raises concerns for embodiment research. Different embodiment studies entail varying levels of publicness in their experimental designs, yet with little explicit consideration of the potential effects of publicness on the results. Yee and Bailenson (2007, 2009) examined the Proteus effect in relation to traits of extroversion and confidence using observational measures of conversations and negotiation tasks involving participants and confederates. Participants must have known they were being observed or recorded, or at least been in the presence of a confederate. In contrast, other virtual reality studies (e.g. Fox et al., 2013) examined the Proteus effect using self-reported attitudinal measures and did not involve any confederates or other elements of publicness in their protocols. Perhaps participants in the latter type of study do not suspect they are being observed by researchers.

This variation similarly applies to studies of the Proteus effect in desktop virtual environments. For instance, whereas some studies (e.g. Peña & Kim, 2014; Peña et al., 2009, experiment 1) involve highly public designs involving spontaneous interaction or outright competition between participants, other studies (e.g. Peña et al., 2009, experiment 2) do not involve any other human – participant or confederate – beyond the participant. In sum, much of the extant embodiment research does not adequately control for the role of publicness across experimental designs, resulting in a potential confound in the literature. That said, the theoretical frameworks – self-perception, priming, and avatar self-relevance – do not entail any element of public awareness for their activation. It is unknown whether participants in these studies had some sense of audience, even if it was only the researcher.

Social interaction

Feedback plays an important role in confirming or disconfirming a self-presentation performance. In the hyperpersonal model of computer-mediated communication, confirmatory feedback to an individual's self-presentation reinforces individuals' perceptions that they possess a self-presented characteristic, and leads them to continue to interact in ways that are consistent with the self-presented characteristic (Walther, 1996). Conversely, negative feedback undermines individuals' self-image, and causes them to behave in ways intended to restore self-image (Cialdini & Richardson, 1980).

The literature reviewed in the previous section revealed that the publicity of self-presentations is an important factor in facilitating transformation of self. However, one study found that even when self-presentations were private, participants nevertheless experienced transformation of self if they were given confirmatory feedback. Walther et al. (2011) found that regardless of whether participants self-presented publicly (on a blog) or privately (to a computer), they experienced a greater transformation of self (on introversion/extraversion) when they received confirmatory feedback

than when they received no feedback. Participants who self-presented to a computer also perceived that their self-presentations were more private than participants who self-presented on a blog, ruling out the possibility that participants may have perceived the computer as a social actor to whom they were publicly presenting (see Nass & Moon, 2000). This study raises the possibility that feedback may be an even more important mechanism than the publicness of self-presentation to explain the transformation of self.

Other aspects of feedback also play a role in modifying people's self-concepts. For instance, public feedback causes greater transformation of self than private feedback, as does feedback from relationally close ties compared to feedback from relationally distant ties (Carr & Foreman, 2016). An interaction effect found that public feedback from relationally close sources led to the greatest transformation of self, while private feedback from relationally distant sources led to the least transformation of self.

In the embodiment paradigm's Proteus effect research, initial studies sought to keep feedback and social interaction constant in an effort to demonstrate the uniquely intrapersonal nature of the effect. This was often achieved either through research designs that were devoid of any sort of interpersonal influence (e.g. Peña et al., 2009, experiment 2) or through use of confederates trained to provide scripted, uniform responses (Yee & Bailenson, 2007, 2009). Additional studies of embodiment effects that interact with various interpersonal processes (e.g. Van Der Heide et al., 2013) will help further illuminate the role of feedback and social interaction in influencing avatar-induced self-concept change.

Outcome measures: behaviors and attitudes

Although outcome measures do not constitute a critical factor, their comparability across different studies is essential to understand which critical factor is important to self-transformation under various conditions. If measured outcomes are solely behavioral (as is the case in many studies on the Proteus effect), researchers cannot know if there is any self-transformation the moment people are taken out of the environment in which they alter their self-presentations, be it embodied virtual reality, graphical avatars, or text-based chats. As people cease to be in the alternative self-presentation environment, they are no longer primed by the appearance of their avatars and they can no longer engage in self-perception of their avatars. Without measuring people's attitudes and personalities (or any other variable related to self-concept), there is little opportunity to show the plausibility of any long-term effects, even without longitudinal study designs.

If outcomes are solely perceptual (as in most identity shift research), researchers will not know if there are demand characteristics, especially if participants are instructed to enact specific self-presentations. A reasonable solution for future studies would be to use both attitudinal and behavioral measures of self-transformation (e.g. when Oh et al., 2016, virtually enhanced participants' smiles as they appeared on their embodied avatars, it led not only to participants' utterances of more positive language but also to more positive scores on self-reported affect measures; see also Beyea et al., 2021; Falconer et al., 2016; Walther et al., 2010). Longitudinal designs should also be welcomed, because there is relatively little existing evidence on how temporary or long-lasting self-transformation effects may be at either the perceptual or behavioral level. Although some longitudinal experiments exist in the literature, their results may conflict: Some multiple-exposure embodiment studies show longer-term self-transformations (e.g. Barberia et al., 2018) whereas others show a dissipation of effects when the novelty wears off after multiple exposures to embodiment (e.g. Bailenson & Yee, 2006).

Conclusions and recommendations

A theorist may start from the assumption that the embodiment and discursive paradigms – that both involve distorted self-presentations – leading to self-transformations are distinct and little-

related; that theorist will not look far to find differences in the theories and research designs investigating these phenomena, in order to support the assumption. Alternatively, a theorist may start by saying, let us assume they are essentially the same; does the research prove that they are not? At present, it would be difficult to provide compelling proof of differences, because most any two studies that found oppositional results also are likely to feature too many other procedural and operational differences among them to be compared. These features might seem theoretically important in one study, but superfluous in another. These features, often activated in grossly different ways across paradigms and studies, inhibit our ability to make theoretical sense across the field of self-transformation studies. They are the six critical factors that have been the focus of this synthetic review.

The six critical factors are potentially necessary and/or sufficient to facilitate the effects of distorted online self-presentations on self-transformation: (1) physical or personality alteration, (2) verbal, graphical, or embodied environment, (3) awareness of alteration, (4) agency of choice/customization, (5) publicness of setting, and (6) social interaction (see Figure 1). Past research from the embodiment approach and the discursive approach has frequently articulated assumptions about the values of several of these factors that are presumed to be necessary to the processes that each hypothesizes. Yet associated research studies have rarely tested these critical factors in isolation from, or in explicit combination with, each other. This pattern has resulted in empirical confounds and conceptual conflicts across a potentially theoretically common field.

New directions

In order for future research efforts to address these inconsistencies, research designs should explicitly and systematically alternate and equalize the activation of these factors, one or several at a time, across other factors while pursuing theoretical hypotheses. For example, addressing the physical versus personality factor, as well as the verbal versus embodied factor, a study might ask some participants to create discursive self-descriptions online focusing on (different levels of) a personality trait, while other participants discursively self-describe variations on some physical characteristic that has stereotypical trait associations. Then repeat in virtual reality: Participants immersed in an avatar are prompted to describe their most shy and introverted moments or their most gregarious ones to a virtual confederate. What distance do they adopt from the virtual other in each variation, and how much mutual gaze? Then cross the personality/physicality and verbal/embodied factors with the publicness factor: the same instructions, but the embodied participant describes these events to a single virtual confederate, a virtual group of them, or a stadium filled with virtual confederates; what distances and gaze this time?

There are examples of attempts at such synthesis studies. In a study by Peña and Hill (2020), for instance, participants were instructed to self-present as either extraverted or introverted, in a virtual reality classroom that was either full of (static) students or completely empty. Peña and Hill (2020) found no identity shift effects between those who self-presented as extraverted/introverted, nor was there a difference due to public versus private conditions. This one study raises the possibility that embodiment does hinder identity shift. Or that feedback, missing from this particular study, may be critical in alternative environments. Or that static figures do not induce a feeling of a public setting. More research is needed, but it is an admirable attempt at the replication of a discursive effects study in an embodiment protocol.

Another hybrid case combined elements of discursive and embodied approaches to physical versus personality cues, in an unconventional offline experiment. Kashian et al. (2012) replicated Yee and Bailenson's (2007) original Proteus study but did not use virtual reality. Participants verbally described themselves to a confederate as being very physically attractive or very unattractive, while wearing paper bags over their heads in a physical room. Results matched Yee and Bailenson's Proteus outcomes with respect to how close participants stood to confederates, but not on the self-disclosure outcomes Yee and Bailenson (2007) reported.

Can people be induced to generate distorted self-descriptions verbally, as is typical in the discursive tradition, but without awareness of their manipulation? Possibly so, if they are subtly primed to enact some aspect of identity. Early CMC research on visual anonymity and social identification employed puzzles which, when solved, primed online group participants to favor different values (Postmes et al., 2001). Techniques such as these offer promise to exploring the question of volition and awareness across paradigms.

These are a few examples of a recommended strategy. To remove confounding inconsistencies across studies, it is important to surface the potentially confounding factors, and isolate their influences by employing and exploring rather than ignoring them across studies. Studying the online transformation of self requires the examination of technological affordances, social factors, and research practices themselves. It is easier to adopt research design choices intended to test one particular mechanism, without asking if it inadvertently evoked others, an unfortunately common problem in communication and technology research (see Walther, 2010). The aim of the present work to advocate, by urging greater focus on critical factors, for a more nuanced delineation of the different mechanisms that can drive self-transformation effects online.

Conclusion

The potential for media to alter our personalities is a common societal concern. It is usually expressed in fears of what technology does to people, and less often, what people do with technology to themselves. That media affordances and the way we employ them have profound effects on our perceptions and ourselves. For those who use media to present their better selves, to put their best foot forward, to experiment with identity; for those who learn new things and undertake therapy in virtual reality; for those who make and play games for hours and hours and years and years, whether the behavioral changes that take place *in situ* lead to transformations of self is a serious matter. As we have seen, our understandings of these phenomena develop and re-develop themselves with new digital platforms, potentially hiding common causal threads and facilitating the uneven application of critical factors. This essay attempts to present a roadmap for future research. A comprehensive understanding of the ways in which these phenomena are the same, and in what ways they critically differ, requires integrative research that tests competing hypotheses among multiple factors such as those uncovered in this review.

Notes

1. Valkenburg (2017) nominated *self-effects* as an organizing concept, following Pingree's (2007) proposition that, when an individual says something to another, the preparation and delivery of saying it steers that individual's own attitude about the topic in the direction of the utterance. Self-presentation/self-transformation effects differ, however, in that (a) the message senders engage in deliberately alternative representations of themselves as they generate messages, that (b) change how they perceive themselves rather than some attitude object. For instance, a self-effect may occur when individuals discuss cat foods in an online chat, and their comments affect their own feelings about cat food. When individuals discuss cat food while pretending to be a cat, however, and it leads them to feel more feline, it is a self-transformation, rather than a self-effect.
2. A major methodological exception appears in Yee et al.'s (2011) analysis of the prevalence and consequences of gender-swapping using avatars within the multiplayer game, World of Warcraft. In this case, the physical alteration in a graphical environment involved users' awareness of alteration, instigated by their own agency, in a public setting, with social interaction. Alternative genders affected in-game healing vs. killing ratios.

Acknowledgements

The authors express their gratitude to Justice Quick and Giselle Langley for their assistance at the inception of this work, and to two anonymous reviewers for their helpful suggestions.

Disclosure statement

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

ORCID

Joseph B. Walther http://orcid.org/0000-0003-2393-9208
Zijian Lew http://orcid.org/0000-0003-1769-7898

References

Adams, F. M., & Osgood, C. E. (1973). A cross-cultural study of the affective meanings of color. *Journal of Cross-Cultural Psychology*, *4*(2), 135–156. https://doi.org/10.1177/002202217300400201

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

Bailenson, J., & Blascovich, J. (2004). Avatars. In W. S. Bainbridge (Ed.), *Encyclopedia of human-computer interaction* (pp. 62-64). Berkshire.

Bailenson, J. N., & Yee, N. (2006). A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. *Presence: Teleoperators and Virtual Environments*, *15*(6), 699–716. https://doi.org/10.1162/pres.15.6.699

Barberia, I., Oliva, R., Bourdin, P., & Slater, M. (2018). Virtual mortality and near-death experience after a prolonged exposure in a shared virtual reality may lead to positive life-attitude changes. *PLoS ONE*, *13*(11), e0203358. https://doi.org/10.1371/journal.pone.0203358

Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*(6), 925–945. https://doi.org/10.1037/0033-2909.126.6.925

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 L. Berkowitz (Ed.), *Advances in experimental social psychology (Vol. 6)* (pp. 1–62). Academic Press.

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*. Advance online publication, https://doi.org/10.1027/1864-1105/a000309

Bruckman, A. (1992). Identity workshop: Emergent social and psychological phenomena in text- based virtual reality. https://www.cc.gatech.edu/~asb/papers/misc/identity-workshop.rtf.

Bruckman, A. S. (1993). *Gender swapping on the Internet*. Paper presented at the meeting of the Internet Society, San Francisco, CA. http://citeseerx.ist.psu.edu/viewdoc/download?doi = 10.1.1.56.4441&rep = rep1&type = pdf.

Bunge, M. (2004). How does it work? The search for explanatory mechanisms. *Philosophy of the Social Sciences*, *34*(2), 182–210. https://doi.org/10.1177/0048393103262550

Carr, C. T., & Foreman, A. C. (2016). Identity shift III: Effects of publicness of feedback and relational closeness in computer-mediated communication. *Media Psychology*, *19*(2), 334–358. https://doi.org/10.1080/15213269.2015.1049276

Carr, C. T., & Hayes, R. A. (2017). Identity shift effects of self-presentation and confirmatory and disconfirmatory feedback on self-perceptions of brand identification. *Media Psychology*, *22*(3), 418–444. https://doi.org/10.1080/15213269.2017.1396228

Carr, C. T., Kim, Y., Valov, J. J., Rosenbaum, J. E., Johnson, B. K., Hancock, J. T., & Gonzales, A. L. (2021). An explication of identity shift theory: Getting our shift together. *Journal of Media Psychology*, *33*(4), 202–214. https://doi.org/10.1027/1864-1105/a000314

Chaiken, S., & Baldwin, M. W. (1981). Affective-cognitive consistency and the effect of salient behavioral information on the self-perception of attitudes. *Journal of Personality and Social Psychology*, *41*(1), 1–12. https://doi.org/10.1037/0022-3514.41.1.1

Cialdini, R. B., & Richardson, K. D. (1980). Two indirect tactics of image management: Basking and blasting. *Journal of Personality and Social Psychology*, *39*(3), 406–415. https://doi.org/10.1037/0022-3514.39.3.406

Curtis, P. (1992). Mudding: Social phenomena in text-based virtual realities. In D. Schuler (Ed.), *DIAC-92: Directions and implications of advanced computing* (pp. 48–68). Computer Professionals for Social Responsibility.

Dibbel, J. (1993, Dec. 23). A rape in cyberspace: How an evil clown, a Haitian trickster spirit, two wizards, and a cast of dozens turned a database into a society. *The Village Voice*. Rpt 2005, https://www.villagevoice.com/2005/10/18/a-rape-in-cyberspace/.

Ducheneaut, N., Wen, M. H., Yee, N., & Wadley, G. (2009). Body and mind: A study of avatar personalization in three virtual worlds. *Paper presented at the CHI Conference on Human Factors in Computing Systems*, Boston, 2009. ACM. https://doi.org/10.1145/1518701.1518877.

Falconer, C. J., Rovira, A., King, J. A., Gilbert, P., Antley, A., Fearon, P., Ralph, N., Slater, M., & Brewin, C. R. (2016). Embodying self-compassion within virtual reality and its effects on patients with depression. *BJPsych Open*, *2*(1), 74–80. https://doi.org/10.1192/bjpo.bp.115.002147

Fazio, R. H., Effrein, E. A., & Falender, V. J. (1981). Self-perceptions following social interaction. *Journal of Personality and Social Psychology*, *41*(2), 232–242. https://doi.org/10.1037/0022-3514.41.2.232

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, *58*(2), 203–210. https://doi.org/10.1037/h0041593

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. https://doi.org/10.1016/j.chb.2012.12.02

Gonzales, A. L. (2006). *Self-concept change in text-based environments: Audience and selective self-presentation* [Unpublished master's thesis]. Cornell University. eCommons: Open scholarship at Cornell. https://ecommons.cornell.edu/handle/1813/3387.

Gonzales, A. L., & Hancock, J. T. (2008). Identity shift in computer-mediated environments. *Media Psychology*, *11*(2), 167–185. https://doi.org/10.1080/15213260802023433

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319–340. https://psycnet.apa.org/doi/10.1037/0033-295X.94.3.319

Hwang, M. L., & Mamykina, L. (2017). *Monster appetite: Effects of subversive framing on nutritional choices in a digital game environment*. Paper presented at the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO. https://doi.org/10.1145/3025453.3026052.

Kashian, N., Liang, Y., Jang, J.-W., Wisner, A., Falin, J., & Walther, J. B. (2012, November). *Masked interaction and interpersonal behaviors: A low-tech assessment of the Proteus Effect on proxemics and disclosures*. Paper presented at the annual meeting of the National Communication Association, Orlando, FL.

Kelly, A. E., & Rodriguez, R. R. (2006). Publicly committing oneself to an identity. *Basic and Applied Social Psychology*, *28*(2), 185–191. https://doi.org/10.1207/s15324834basp2802_8

Kilteni, K., Maselli, A., Kording, K. P., & Slater, M. (2015). Over my fake body: Body ownership illusions for studying the multisensory basis of own-body perception. *Frontiers in Human Neuroscience*, *9*, 141. https://doi.org/10.3389/fnhum.2015.00141

Kishore, S., Spanlang, B., Iruretagoyena, G., Szostak, D., & Slater, M. (2021). A virtual reality embodiment technique to enhance helping behavior of police towards a victim of police racial aggression. *Presence: Teleoperators and Virtual Environments*, *28*, 5–27. https://doi.org/10.1162/pres_a_00339

Klimmt, C., Hefner, D., & Vorderer, P. (2009). The video game experience as "true" identification: A theory of enjoyable alterations of players' self-perception. *Communication Theory*, *19*(4), 351–373. https://doi.org/10.1111/j.1468-2885.2009.01347.x

Kocur, M., Schauhuber, P., Schwind, V., Wolff, C., & Henze, N. (2020). The effects of self- and external perception of avatars on cognitive task performance on virtual reality. *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology (n.p.)*. https://doi.org/10.1145/3385956.3418969.

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. https://doi.org/10.1089/cyber.2013.0358

Lim, S., & Reeves, B. (2009). Being in the game: Effects of avatar choice and point of view on psychophysiological responses during play. *Media Psychology*, *12*(4), 348–370. https://doi.org/10.1080/15213260903287242

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Oh, S. Y., Bailenson, J., Krämer, N., & Li, B. (2016). Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PLoS ONE*, *11*(9), e0161794. https://doi.org/10.1371/journal.pone.0161794

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. https://doi.org/10.1177/0261927X09351675

Parks, M. R., & Roberts, L. D. (1998). Making MOOsic': The development of personal relationships on line and a comparison to their off-line counterparts. *Journal of Social and Personal Relationships*, *15*(4), 517–537. https://doi.org/10.1177/0265407598154005

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. https://doi.org/10.1016/j.concog.2013.04.016

Peck, T. C., & Gonzalez-Franco, M. (2021). Avatar embodiment. A standardized questionnaire. *Frontiers in Virtual Reality*, *1*, 575943. https://doi.org/10.3389/frvir.2020.575943

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

Peña, J., & Hill, D. (2020). Examining identity shift effects in virtual reality. *Cyberpsychology, Behavior, and Social Networking*, *23*(10), 697–701. https://doi.org/10.1089/cyber.2020.0010

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

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

Pingree, R. J. (2007). How messages affect their senders: A more general model of message effects and implications for deliberation. *Communication Theory*, *17*(4), 439–461. https://doi.org/10.1111/j.1468-2885.2007.00306.x

Postmes, T., Spears, R., Sakhel, K., & De Groot, D. (2001). Social influence in computer-mediated groups: The effects of anonymity on social behavior. *Personality and Social Psychology Bulletin*, *27*(10), 1243–1254. https://doi.org/10.1177/01461672012710001

Ratan, R. A., & Dawson, M. (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., 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. https://doi.org/10.1080/15213269.2019.1623698

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. https://doi.org/10.1016/j.chb.2015.04.010

Sah, Y. J., Ratan, R. A., Tsai, H. S., & Sarinopoulous, I. (2014, October). *Avatar projecting actual, ideal, or ought self: Which one do we feel close to?*. Paper presented at the Meaningful Play Conference, East Lansing, MI. http://meaningfulplay.msu.edu/proceedings2014/mp2014_submission_86.pdf.

Schlenker, B. R., Dlugolecki, D. W., & Doherty, K. (1994). The impact of self-presentations on self-appraisals and behavior: The power of public commitment. *Personality and Social Psychology Bulletin*, *20*(1), 20–33. https://doi.org/10.1177/0146167294201002

Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, M. Á. d. l. C., Alabèrnia-Segura, M., Spanlang, B., & Feixas, G. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Scientific Reports*, *9*(1), 10903. https://doi.org/10.1038/s41598-019-46877-3

Stone, A. R. (1995). *The war of desire and technology at the close of the mechanical age*. MIT.

Tice, D. M. (1992). Self-concept change and self-presentation: The looking glass self is also a magnifying glass. *Journal of Personality and Social Psychology*, *63*(3), 435–451. https://doi.org/10.1037/0022-3514.63.3.435

Tong, S. T., & Walther, J. B. (2015). The confirmation and disconfirmation of expectancies in computer-mediated communication. *Communication Research*, *42*(2), 186–212. https://doi.org/10.1177/0093650212466257

Turkle, S. (1995). *Life on the screen: Identity in the age of the Internet*. Simon & Schuster.

Valkenburg, P. M. (2017). Understanding self-effects in social media. *Human Communication Research*, *43*(4), 477–490. https://doi.org/10.1111/hcre.12113

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. https://doi.org/10.1177/0093650212438097

Van Gelder, L. V. (1991). The strange case of the electronic lover. In C. Dunlop, & R. Kling (Eds.), *Computerization and controversy: Value conflicts and social choices* (pp. 364–375). Academic Press.

Van Looy, J., Courtois, C., De Vocht, M., & De Marez, L. (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

Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, *23*(1), 3–43. https://doi.org/10.1177/009365096023001001

Walther, J. B. (2010). Computer-mediated communication. In C. R. Berger, M. E. Roloff, & D. R. Roskos-Ewoldsen (Eds.), *Handbook of communication science* (2nd ed., pp. 489–505). Sage.

Walther, J. B., Kashian, N., Jang, J.-W., Shin, S. Y., Dai, Y., & Koutamanis, M. (2018). The effects of message persistence and disclosure on liking in computer-mediated communication. *Media Psychology*, *21*(2), 308–327. https://doi.org/10.1080/15213269.2016.1247718

Walther, J. B., Liang, Y. J., DeAndrea, D. C., Tong, S. T., Carr, C. T., Spottswood, E. L., & Amichai-Hamburger, Y. (2011). The effect of feedback on identity shift in computer-mediated communication. *Media Psychology*, *14*(1), 1–26. https://doi.org/10.1080/15213269.2010.547832

Walther, J. B., Van Der Heide, B., Tong, S. T., Carr, C. T., & Atkin, C. K. (2010). Effects of interpersonal goals on inadvertent intrapersonal influence in computer-mediated communication. *Human Communication Research*, *36*(3), 323–347. https://doi.org/10.1111/j.1468-2958.2010.01378.x

Walther, J. B., & Whitty, M. T. (2021). Language, psychology, and new media: The hyperpersonal model of mediated communication at twenty-five years. *Journal of Language and Social Psychology*, *40*(1), 120–135. https://doi.org/10.1177/0261927X20967703

Yee, N., & Bailenson, J. N. (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., & 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. https://doi.org/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. https://doi.org/10.1177/0093650208330254

Yee, N., Ducheneaut, N., Yao, M., & Nelson, L. (2011). Do men heal more when in drag? Conflicting identity cues between user and avatar. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/1978942.1979054.

Young, T. J., & French, L. A. (1996). Height and perceived competence of US presidents. *Perceptual and Motor Skills*, *82*(3), 1002. https://doi.org/10.1177/003151259608200301

Zumbach, J., Seitz, C., & Bluemke, M. (2015). Impact of violent video game realism on the self-concept of aggressiveness assessed with explicit and implicit measures. Computers in Human Behavior, 53, 278–288. https://doi.org/10.1016/j.chb.2015.07.018