



Investigation of the moderating effect of race-based personalization of behavior model design in fitness application

Kiemute Oyibo¹ · Julita Vassileva¹

© Springer Nature Switzerland AG 2019

Abstract

Persuasive technologies are more likely to be successful if personalized to the target audience. However, in the context of exercise behavior modeling in fitness applications, there is limited research on the moderating effect of personalization based on the race characteristic of the user. To bridge this gap, we carried out an empirical study ($n = 669$) to examine how race-based personalization moderates the relationships between the *perceived persuasiveness* of behavior model design and three key social-cognitive determinants of behavior: *outcome expectations*, *self-efficacy beliefs*, and *self-regulation beliefs*. We based our investigation on user evaluation of bodyweight-exercise behavior models (videos) featured in a fitness app prototype based on all four constructs. Our results showed that, in both tailored and untailored contexts, behavior models (i.e., their *perceived persuasiveness*) have a significant influence on both black/brown and white users' *outcome expectations* and *self-regulation beliefs*. However, they are more likely to be effective in influencing black/brown observers than white observers in the respective contexts. Moreover, black/brown models are more likely to have a significant influence on users' *self-efficacy beliefs* than white models. Specifically, it is only when the black/brown models are tailored to black/brown observers and contra-tailored to white observers that the *perceived persuasiveness* of behavior model design has a significant effect on *self-efficacy beliefs*. Finally, among black/brown observers, our results indicate that behavior model designs are more likely to be effective when tailored than when contra-tailored or untailored, but this is hardly the case among white users. Our study contributes to the body of knowledge by being one of the first to demonstrate the potential effectiveness of race-based personalization of exercise behavior models in the domain of persuasive technologies aimed at motivating behavior change.

Keywords Persuasive technology · Fitness application · Bodyweight exercise · Video-based behavior modeling · Social cognitive model · Race-based personalization

1 Introduction

Owing to the limited information on how to correctly perform exercise behavior in the fitness domain, behavior modeling has become a popular persuasive technique used to motivate behavior change. Behavior modeling is a persuasive strategy employed by domain experts and coaches in a physical or virtual setting to demonstrate how to correctly perform a given behavior to an observer [1, 2]. The behavior modeling strategy can be considered

successful when the behavior of the observer is visibly changed in the future in line with the observed behavior [3]. In a systematic review of persuasive strategies used in top-ranked fitness apps in the marketplace, Conroy et al. [4] found that exercise behavior modeling (such as instructions, videos and simulations) is the most commonly used persuasive technique in motivating behavior change in the health domain. For example, in virtual (video- or simulation-based) exercise behavior modeling, users can visualize the steps involved in performing a given exercise

✉ Kiemute Oyibo, kiemute.oyibo@usask.ca; Julita Vassileva, jiv@cs.usask.ca | ¹Multi-User Adaptive Distributed Mobile and Ubiquitous Computing Laboratory, Department of Computer Science, University of Saskatchewan, Saskatoon S7N 5C9, Canada.



behavior, just as computer-based learners in the education domain can visualize abstract cognitive processes or concepts of a given subject simulated by the computer [3]. Moreover, video-based behavior models are employed to visualize the positive outcome expectations of certain health behaviors, for example, in exercise behavior, specific muscles that are targeted by specific body movements. Virtual behavior models have some advantages over real-life human models such as gym personal trainers and coaches. They include: (1) they are more affordable; and (2) they are more convenient, i.e., they enable the target users to observe and perform the modeled behavior at any time and wherever that is most convenient for them, e.g., at home, in the hotel, etc. [3, 5].

However, in the fitness domain, there is limited research on the moderating effect of personalization of exercise behavior models on their target users. Personalization is defined as the alignment of the features of a persuasive system to the characteristics of the end-user by the designer of the system [7]. It can be implemented at various levels—individual and group—with user characteristics being defined based on demographic variables such as age, gender, race, ethnicity, etc. In the context of this paper, we focused on group-based personalization based on the race of the end-user, also known as tailoring [6]. Research in the persuasive technology domain has shown that persuasive systems are more likely to be successful if tailored to the user by the designer. Thus, there have been several calls in the persuasive technology domain on designers to tailor or personalize persuasive systems aimed at behavior change to the characteristics of the users, as they are more likely to be successful [8, 9]. However, there is limited research in the area of race-based personalization of persuasive systems to the target user, especially in the health and fitness domain. Although, personalization has been applied in persuasive technology design in different health contexts, e.g., substance use [10], medication adherence [11], etc., there is limited evidence of its moderating effect in the fitness domain in the context of behavior modeling [7].

In this paper, in the context of user perception, we investigated the effect of race-based tailoring of exercise behavior model design on the target users using path analysis. Specifically, we examined the relationship between the *perceived persuasiveness* of bodyweight-exercise behavior model design and three key social-cognitive determinants of behavior (*self-efficacy beliefs*, *self-regulation beliefs* and *outcome expectations*) [5, 12] in a tailored, contra-tailored and untailored context. Primarily, our investigation is based on two same-race (tailored) observer-model dyads and two opposite-race (contra-tailored) observer-model dyads. They include: (1) black/brown observers who evaluated black/brown behavior

models (BB); (2) black/brown observers who evaluated white behavior models (BW); (3) white observers who evaluated white behavior models (WW); and (4) white observers who evaluated black/brown behavior models (WB). In addition, we investigated four other (untailored) dyads which resulted from the combination (summation) of two of the aforementioned dyads: (1) the group of random observers (BB and WB) that evaluated black/brown models (BM); (2) the group of random observers (WW and BW) that evaluated white models (WM); (3) the group of black/brown observers that evaluated random models (BR); and (4) the group of white observers that evaluated random models (WR).

The results of our study showed that the *perceived persuasiveness* of behavior model design has the potential to influence social-cognitive determinants of behavior, with the relationship between the persuasive design and social-cognitive constructs being moderated by race. Our path models showed that exercise behavior models are more likely to be effective in influencing *outcome expectations* and *self-regulation beliefs* than *self-efficacy beliefs*. In tailored (BB vs. WW) and untailored (BR vs. WR) contexts, behavior models are more likely to be effective in influencing *outcome expectations* and *self-regulation beliefs* of black/brown observers than those of white observers. For example, in the untailored context (BR vs. WR), the influence of *perceived persuasiveness* on *self-regulation beliefs* and *outcome expectations* is significantly stronger for black/brown users [$(\beta = 0.52, p < 0.001)$ and $(\beta = 0.59, p < 0.001)$, respectively] than for white users [$(\beta = 0.38, p < 0.001)$ and $(\beta = 0.39, p < 0.001)$, respectively]. Similarly, in the tailored context (BB vs. WW), the influence of *perceived persuasiveness* on *self-regulation beliefs* and *outcome expectations* is significantly stronger for black/brown users [$(\beta = 0.59, p < 0.001)$ and $(\beta = 0.62, p < 0.001)$, respectively] than for white users [$(\beta = 0.41, p < 0.001)$ and $(\beta = 0.35, p < 0.001)$, respectively]. However, in the contra-tailored (BW vs. WB) context, though the path coefficients for the three investigated relationships turned out to be higher for the black/brown observers than for the white observers, we found no significant differences between both user groups (BW vs. WB).

In sum, our findings revealed that behavior models, tailored and untailored, have the potential of being more effective among the black/brown user population than among the white user population, especially with respect to *self-regulation beliefs* and *outcome expectations*. Moreover, we found that behavior models, among black/brown users, are more likely to be effective, especially with respect to influencing *self-efficacy beliefs*, if tailored than otherwise. Therefore, we argue that, given that most existing fitness apps on the market are targeted at the western population (especially white users), there is a need for

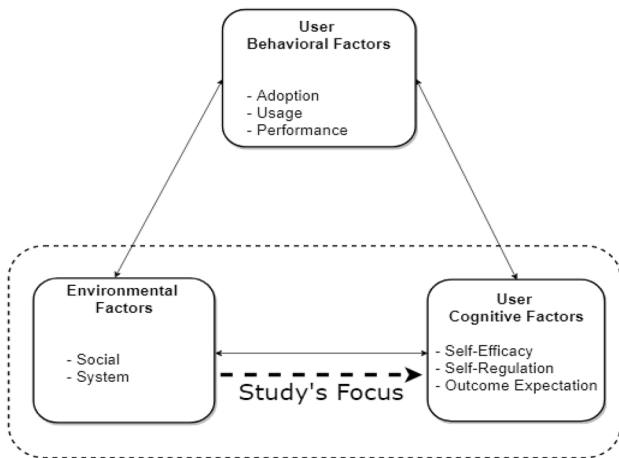


Fig. 1 SCT framework depicting the triad of reciprocal determinism [2]

designers to begin tailoring their apps to the racial characteristics of the target populations. Tailoring behavior models to the target users based on race will allow designers to deliver more effective fitness apps that users can identify with and use to support their behavior change.

The rest of this paper is organized as follows. Section 2 focuses on background information. Section 3 dwells on related work, while Sect. 4 on research method. Finally, Section 5 focuses on the result, while Sects. 6 and 7 on the discussion and conclusion, respectively.

2 Background

In this section, we provide an overview of the Social Cognitive Theory (SCT) and behavior modeling in the context of observational learning.

2.1 Social cognitive theory

The SCT is one of the most established behavior theories in the field of Psychology. It was put forward by Bandura [13] to explain human behaviors. As shown in Fig. 1, the SCT holds that *personal* factors, *environmental* factors, and the *target behavior* itself reciprocally influence one another.

The bidirectional interplay among the three SCT factors is known as the triad of reciprocal determinism. In this paper, we are particularly concerned with the influence of the environmental factors on three social-cognitive factors of behavior: *self-efficacy beliefs*, *self-regulation beliefs*, and *outcome expectations*. Table 1 shows the definitions of the three factors. The environmental factor includes friends, family or other influential people in face-to-face conversations, mass communication media (e.g., newspaper, radio, television, etc.), electronic communication systems (e.g., mobile phone, computer, etc.) or social networks (e.g., Facebook, Twitter, etc.) where people engage online. Given the critical role mobile technology plays in our lives today, to the extent that mobile devices such as smartphones have become an extension of the human body and are becoming ubiquitous and invisible, it has become important to investigate how mobile technology affects our cognition in the context of beneficial behavior change. Thus, in our study, we are focused on investigating, at the perception level, the effect of mobile-technology-based artifacts on three key social-cognitive factors using bodyweight-exercise behavior models featured in a mobile fitness app as our case study. More specifically, we set out to investigate how the *perceived persuasiveness* of behavior model design influences the three SCT factors. All of the three SCT factors (*self-efficacy beliefs*, *self-regulation beliefs*, and *outcome expectations*) have been shown by prior research [14–19] to be significant determinants of physical activity behavior.

2.2 Behavior modeling

Research [22] shows that most behavior changes require new knowledge, though knowledge alone may not be sufficient to bring about the desired behavior change. According to Jimison et al. [22], human cognition (what people know and think) and behavior (people’s action) go hand in hand. Behavior modeling, in the context of SCT, is a behavior change technique through which a domain expert or coach teaches an observer (by way of demonstration) how to perform a given behavior correctly, in either physical setting (e.g., gym) or virtual setting (e.g., video). In this paper, we are concerned with the latter.

Table 1 Study’s constructs and definition

Construct	Definition
Self-efficacy belief	It is the belief in one’s ability to perform a given behavior [2, 20]
Self-regulation belief	It is the belief in one’s ability to regulate his/her behavior through planning, self-monitoring, goal-setting, and tracking of his/her behavior over time [2, 21]
Outcome expectation	It is the belief that engaging in a given behavior will result in certain positive outcomes [2, 21]
Perceived persuasiveness	It is the belief that a certain persuasive system has the motivational capacity to influence one to positively change his/her behavior [2]

Behavior modeling stems from the concept of observational learning, which is based on Bandura's [23] Social Learning Theory (SLT). The SLT, over time, evolved into the SCT. According to Wouters et al. [3], "*the social cognitive model of sequential skill acquisition [sic] describes how learners initially start with observing a model, but then start practicing and gradually learn how to self-regulate their own performance*" (p. 328). During the observation of the behavior model, Wouters et al. explain [3], the learners constructs a mental representation of the modeled behavior without actually performing it immediately. Eventually, the observed behavior is performed by the observer in different situations and contexts. Wouters et al. [3] argue that the transition from observing the model to self-regulated practice is "*accompanied by different cognitive processes and that the interaction takes different forms that all facilitate different cognitive processes*" (p. 328). Research has shown that behavior modeling—also regarded as vicarious modeling—can influence behaviors through mediating social-cognitive factors such as *self-efficacy*, *self-regulation*, *outcome expectation*, etc. [24, 25]. In behavior modeling, individuals observe the actions and consequences of the behavior of others and then determine through social-cognitive processes whether to imitate and/or adopt the newly learned behavior. Three reasons have been put forward to explain why behavior modeling has the potential to be effective. They include the following [3]:

1. By observing an expert performing a complex behavior, the observer can construct a sufficient cognitive representation, which enables him/her to rehearse and fine-tune the observed behavior mentally and physically.
2. Observational learning might be more effective than other forms of learning because the expert modeling the behavior does not only demonstrate to the observer *what* is happening but explains *why* it is happening as well.
3. While the performance of a complex behavior by novices might place high demands on memory resources such that vital information can no longer be processed, observing an expert performing the complex behavior, on the other hand, tends to free up cognitive resources that can be channeled to processing the vital information.

3 Related work

In this section, we present a review of the literature related to behavior modeling in the health domain. Jimison et al. [22] investigated the effectiveness of using interactive exercise videos for remote coaching of older adults in their

homes. They developed an online-based interactive video exercise system (composed of a human behavior model and Kinect camera's skeleton representation), which was used to demonstrate the performance of certain types of chair exercises to older adults residing in their homes. They found that the online interactive video exercise system can be used to foster adherence to exercise goals despite the challenge in finding adequate room in participants' homes. However, the authors did not investigate the perceived effect the interactive video exercise system has on participants' cognition such as *self-efficacy beliefs*. Oyibo et al. investigated the influence behavior model design has on three SCT factors. They found that the *perceived persuasiveness* of behavior model design affects *self-regulation beliefs*, *outcome expectations* and *self-efficacy beliefs* (the least affected). However, they did not investigate the moderating effect of race-based personalization of the behavior model design.

Moreover, Devi et al. [26] conducted a study to investigate the effect of breast-feeding behavior modeling and other media influences on postnatal women. The study used a video in which correct breastfeeding techniques and safe practices were taught to postnatal women by a behavior model. They found that the video-based model of breastfeeding—a cost-effective method of disseminating breastfeeding information—yielded positive results, including providing the target women with the right knowledge and fostering a prolonged breastfeeding period. However, apart from the study being in the breast-feeding subdomain of health, it is difficult to tell how much influence the video-based behavior model had on the target audience due to their exposure to other media (e.g., television, newspapers), health personnel, friends and family, etc.) as well during the same time period. Torres et al. [26] carried out an experimental study to evaluate the effects of a video-enhanced activity schedule on exercise behavior among adolescents with autism spectrum disorder (ASD) using an iPad®. They found that the intervention, consisting of a video-based model and graduated guidance, increased the independent schedule-following behavior and on-task behavior of adolescents with ASD, with participants' skills generalizing to a novel exercise and novel setting (such as a new fitness center). Specifically, they found that the intervention was socially acceptable to behavior analysts, instructors, and paraprofessionals. However, their study, though experimental, did not investigate the effect the video modeling had on cognitive factors. Our current study aims to bridge this gap in a large-scale study. Specifically, at the cognitive level, we set out to investigate the influence of exercise behavior model design on three important social-cognitive factors (*outcome expectations*, *self-efficacy* and *self-regulation beliefs*), which mediate behavior change, and the potential effectiveness of race-based personalization in fitness app design.

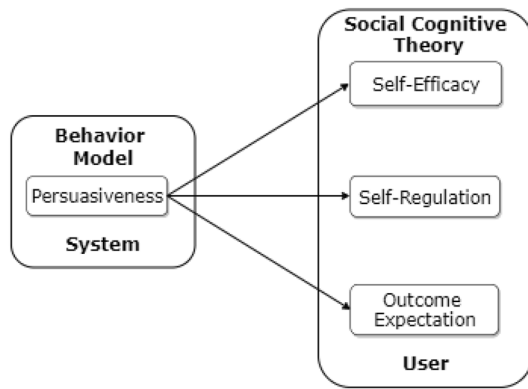


Fig. 2 Research model based on the SCT triad of reciprocal determinism [2]

4 Method

In this section, we present the research objective, empirical instruments used in measuring the constructs in our research model and demographics of participants who took part in the study.

4.1 Research objective

In this paper, we aimed to investigate the moderating effect of race-based tailoring of exercise behavior models by uncovering the differences that exist between two different groups of observers in tailored, contra-tailored and untailored contexts. This means, in all three contexts, the two comparative groups of interest evaluated behavior models of the same, opposite and random race, respectively. In a nutshell, we aimed to answer this overarching research question—“*is race-based tailoring of behavior model design more likely to be effective for both racial target groups than otherwise?*” To answer this question, we adopted a prior research model [2] shown in Fig. 2 to study the relationships between the *perceived persuasiveness* of behavior model design and the three social-cognitive factors of interest.

Using path analysis, we investigated whether behavior model designs are more likely to be effective if tailored to the target users or not. Owing to the limited studies in the context of tailoring exercise behavior models to the race of the target users in the fitness domain, we decided to adopt an exploratory approach using Partial Least Square Path Modeling (PLS-PM) [27]. We chose PLS-PM because it is a technique used for predicting response and not for explaining a phenomenon, for which covariance modeling is a better suited alternative modeling technique. While explanatory models may be “*perceived to be primarily concerned with testing the faithful representation of causal mechanisms by the statistical model and the estimation of true population parameter values from samples*” (p. 2) [28], predictive

models, though possibly based on causal mechanisms, are developed in a more exploratory and data-driven way. In other words, a PLS-based predictive model emphasizes the observed data over the underlying theory, which makes it preferable to covariance analysis for the prediction of target constructs¹ [28]. According to Tobias [29], PLS is a method for building predictive models which lays emphasis on “*predicting the responses and not necessarily on trying to understand the underlying relationship between the variables*” (p. 1) in the causal model [28]. In our case, we aim at investigating the ability of the *perceived persuasiveness* of behavior model design in fitness apps to predict SCT factors (*perceived self-efficacy*, *perceived self-regulation* and *outcome expectation*) in the context of tailoring, contra-tailoring and non-tailoring and not the interrelationships among the SCT variables in our path model. In a nutshell, our research design is based on the research model shown in Fig. 2 and the following four observer-model dyads:

1. Black/brown observers who evaluated black/brown behavior models (BB).
2. Black/brown observers who evaluated white behavior models (BW).
3. White observers who evaluated white behavior models (WW).
4. White observers who evaluated black/brown behavior models (WB).

Furthermore, using exploratory analysis, we analyzed the path models for the following four observer-model dyads:

1. The group of random observers (BB and WB) that evaluated black/brown models (BM).
2. The group of random observers (WW and BW) that evaluated white models (WM).
3. The group of black/brown observers that evaluated random models (BR).
4. The group of white observers that evaluated random models (WR).

We examined in a multigroup analysis the statistical significant differences that exist in the respective path coefficients between pairs of path models, e.g., BB versus WW, BB versus BW, WW versus WB, BM versus WM, BR versus WR, etc. We also examined the numerical differences in the respective path coefficients between a tailored dyad and an untailored dyad (BB versus BR and WW versus WR) in order to see whether *perceived persuasiveness* is more

¹ Prediction has been recognized as an important aspect of information systems research [28].

Table 2 Research questions in words and observer-model dyadic symbols

No.	Research question (RQ) in words	RQ in symbols
RQ1	Overall, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger when observers evaluate black models or white models?	BM > WM? or WM > BM?
RQ2	For which of the groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger when they evaluate random behavior models (i.e., in an untailored context)?	BR > WR? or WR > BR?
RQ3	For black/brown observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	BB > BR?
RQ4	For white observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	WW > WR?
RQ5	For black/brown observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	BB > BW?
RQ6	For white observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	WW > WB?
RQ7	Are the relationships between <i>perceived persuasiveness</i> of black/brown behavior model design and the SCT factors more likely to be stronger for black/brown observers (tailored) than for white observers (contra-tailored)?	BB > WB?
RQ8	Are the relationships between <i>perceived persuasiveness</i> of white behavior model design and the SCT factors more likely to be stronger for white observers (tailored context) than for black/brown observers (contra-tailored context)?	WW > BW?
RQ9	For which of the two groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context?	BB > WW? or WW > BB?
RQ10	For which of the two groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a contra-tailored context?	BW > WB? or WB > BW?
RQ11	In general, which of the three relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors is more likely to be the strongest?	PERS → SE? or PERS → SR? or PERS → OE?

">": are the path coefficients in the left path model greater than those in the right path model? SE: Self-Efficacy Beliefs, SR: Self-Regulation Beliefs, OE: Outcome Expectations

BM = Random observers' evaluation of black/brown models, WM = Random observers' evaluation of white models

BR = black/brown observers' evaluation of random models, WR = white observers' evaluation of random models

BB = black/brown observers' evaluation of black models, BW = black/brown observers' evaluation of white models

WW = white observers' evaluation of white models, WB = white observers' evaluation of black models

likely to predict the three SCT factors in a tailored context than in an untailored context.

For the sake of proper organization and easy reference, we present all of our research questions in Table 2, in which the third column represents their symbolic representations. For example, the symbol, BM > WM? or WM > BM?, means: *are the relationships between the perceived persuasiveness of the behavior model design and the SCT factors stronger in the BM path model than those in the WM path model or vice versa?* Overall, our findings will provide empirical evidence on the moderating effect of tailoring exercise behavior models to the race of the target users.

4.2 Research design and measurement instruments

To answer our research questions, we designed a fitness app prototype we called "Homex App," aimed at encouraging

the performance of bodyweight exercise on the home front. The app features eight versions of exercise behavior model designs (see Fig. 6 in Appendix 1), each of which demonstrates to the target user (in text and graphics) how to correctly and successfully perform a given bodyweight exercise. The textual representation of the modeled exercise behavior (see the top-left corner of the application screen) is similar to the Tunneling strategy in the Persuasive System Design (PSD) model [30]. Tunneling is a persuasive strategy in which a user is guided in a step-by-step fashion to perform the target behavior [10, 31]. For example, the tunneling-based instructions for the squat exercise include: (1) *Stand upright, with your arms stretched forward;* (2) *Push your butt backward until hips are lower than knees;* (3) *Return to the starting position;* and (4) *Repeat as many times as possible.* Apart from focusing on two exercise-types (push-up and squat), in designing the behavior models, we considered

Table 3 Study's constructs and measurement items [20, 21]

Construct	Overall question and items
Perceived persuasiveness [0—strongly disagree to 7—strongly agree]	Assume the Homex App works as described at the beginning of the survey and seen in the previous video. Please answer the following general questions about yourself and regarding the app The app would influence me The app would be convincing The app would be personally relevant to me The app would make me reconsider my physical activity habits
Perceived self-efficacy [18] [0—not confident to 100—confident]	How confident are you that you can complete at home the proposed weekly number of push-ups (entered previously) for the next three months Even when you have worries and problems? Even if you feel depressed? Even when you feel tense? Even when you are tired? Even when you are busy?
Outcome expectation [32] [1—strongly disagree to 5—strongly agree]	The [name of exercise] will... Improve my ability to perform daily activities Improve my overall body functioning Strengthen my bones Increase my muscle strength Improve the functioning of my cardiovascular system Improve my social standing Make me more at ease with people Increase my acceptance by others
Perceived self-regulation [19] [1—strongly disagree to 5—strongly agree]	To achieve my proposed weekly average number of push-ups... I will set a goal I will develop a series of steps to reach my weekly goal I will keep track of my progress in meeting my goal I will endeavor to achieve the set goal for myself I will make my goal public by telling others about it

gender (male vs. female) and race (black/brown vs. white). However, in this paper, we are focused on investigating the moderating effect of race-based tailoring of behavior model designs to the target users (black/brown observers vs. white observers). In our questionnaire, to put our study in proper context, we described to participants the functionality of the fitness app prototype as shown in Appendix 1. Afterwards, we presented each participant with one of the randomized versions of the eight behavior model designs (a GIF-based video). Next, we requested the participant to answer questions relating to the behavior model design and SCT factors (see Table 3).

4.3 Participants

The survey questionnaire was submitted to and approved by the Behavior Research Ethics Board of our university. Afterward, we posted it on Amazon Mechanical Turk (AMT)

to recruit participants. Prior to answering the questionnaire, participants were presented with an informed consent form, which explained what the study was about, the potential risks, potential benefits, compensation, rights to withdraws, etc. Participants agreed to the information contained in the consent form before completing the survey. In appreciation of participants' time, each was compensated with US \$0.6. Table 4 shows the demographic information of the valid participants after data cleaning.

5 Results

In this section, we present the results of our path analysis and multigroup analysis (MGA). The analyses were conducted using PLS-PM library ("*plspm*") in R programming language [27]. Statistical significance tests were done using the bootstrapping method with 5000 resamples [33].

Table 4 Demographics of participants (n = 669)

Criterion	Subgroup	Number	Percent
Gender	Female	327	48.9
	Male	342	51.1
Age	18–24	126	18.8
	25–34	296	44.2
	35–44	155	23.2
	45–54	60	9.0
	54+	32	4.8
Education	Technical/trade school	86	12.9
	High school	136	20.3
	BSc	316	47.2
	MSc	96	14.3
	PhD	15	2.2
	Others	20	3.0
Country of origin	Canada	214	32.0
	United States	378	56.5
	Others	77	11.5
Race	Black	52	7.8
	Brown	99	14.8
	White	518	77.4
Observer-model race	BB	74	11.1
	BW	78	11.7
	WW	255	38.1
	WB	262	39.2
Years on the internet	0–3	4	0.6
	4–6	31	4.6
	7–9	60	9.0
	> 10	574	85.8

BB=black/brown observers who evaluated black models, WW=white observers who evaluated white models, BW=black/brown observers who evaluated white models; WB=white observers who evaluated black models

5.1 Measurement models

Before analyzing the structural models, we evaluated the measurement models and ensured the preconditions for the former were met [33, 34]. The definitions of the preconditions (criteria) and the result of the evaluation of the measurement models are shown in Table 5.

5.2 Structural models

We began our path analysis by building structural models that will help in addressing our first research question: “Overall, are the relationships between perceived persuasiveness of behavior model design and the SCT factors more likely to be stronger when observers evaluate black models or white models?” Figure 3 shows the path models for the evaluation of the two different race-based behavior model designs. The path models represent the relationships

Table 5 Evaluation of measurement models [33–36]

Criterion	Definition	Evaluation result
Indicator reliability	It is the degree to which the measure of a construct (indicator) is reliable. Thus, it is defined as the variance of an indicator that is not explained by measurement error [35]	About 95% of the indicators had an outer loading greater than 0.7, while the other 5% exceeded 0.6. See Table 8 in Appendix 2 for the outer loading of the indicators in each of the path models we presented
Internal consistency reliability	It is a measure of the extent to which a set of indicators that purport to measure a construct produces similar scores [36]	The Dillon-Goldstein metric (DG.rho) for each construct was greater than 0.7. See Table 9 in Appendix 2 for the respective values for each construct
Convergent validity	It is a measure of the extent to which the indicators that measure a construct are closely related [36]	The average variance extracted (AVE) for each construct was greater than 0.5. See Table 10 in Appendix 2 for the respective values for each construct
Discriminant validity	It is a measure of the extent to which the indicators that measure a given construct are unrelated to another construct [36]	The cross-loading criterion for each construct was used, and no indicator loaded higher on any other construct than the one it was designed to measure

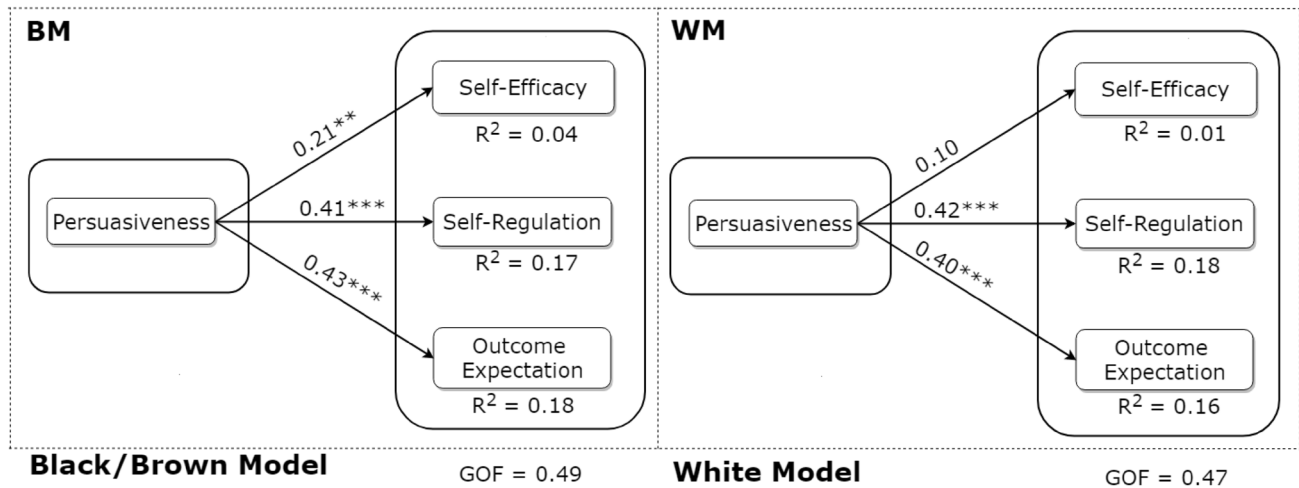


Fig. 3 Path models for the evaluation of black/brown and white behavior models

between the *perceived persuasiveness* of behavior model design and the three SCT factors for the two groups of random observers that evaluated the behavior models in an untailored context. In the survey, the first group evaluated black/brown models (BM) and the second group evaluated white models (WM). In the path models, the goodness of fit (GOF) represents how well the respective models fit their data, while the coefficient of determination (R^2) stands for the amount of variance of an SCT factor explained by *perceived persuasiveness*. Finally, the path coefficient (β) represents the strength of the relationship between *perceived persuasiveness* and an SCT factor. The GOF for BM and WM is 49% and 47%, respectively (roughly moderate), while the R^2 -value for *self-regulation* and *outcome expectation* is about 17%. In particular, the R^2 -value for *self-efficacy* is quite low (4% and 1%, respectively). Moreover, the β -value for *self-regulation* and *outcome expectations* is over ($\beta = 0.4$, $p < 0.001$). However, with respect to *self-efficacy*, while the relationship for the black/brown models is significant ($\beta = 0.21$, $p < 0.01$), that for the white models is non-significant ($\beta = 0.10$, $p = n.s$). The MGA will reveal whether the numerical differences between the corresponding path coefficients in the BM and WM path models are statistically significant at $p < 0.05$ or not.

Furthermore, to be able to answer the second to sixth research questions, we built the path models for the untailored (BR and WR), tailored (BB and WW) and contra-tailored (BW and WB) dyads we introduced earlier on (see Fig. 4). The left column represents the black/brown group of observers, while the right column represents the white group of observers. The GOF for the six dyads ranges from 46 to 56%—almost similar to those for BM versus WM path models. However, unlike in the BM versus WM path models, where the β - and R^2 -values, especially

for *self-regulation* and *outcome expectations*, are approximately equal for both path models, in the six dyads shown in Fig. 4, both parameters are higher for the black/brown observers than for the white observers. For example, regarding BB versus WW dyads, the β -values for *self-efficacy*, *self-regulation* and *outcome expectations* are higher for the black/brown observers ($\beta = 0.30$, $p < 0.05$; 0.59 , $p < 0.001$; and 0.62 , $p < 0.001$, respectively) than for the white observers ($\beta = 0.11$, $p = n.s$; 0.41 , $p < 0.001$; and 0.35 , $p < 0.001$, respectively). The same applies to the R^2 -values for the black/brown observers (9%, 35%, and 39%, respectively) compared with those for the white observers (1%, 17%, and 12%, respectively). Again, the MGA will reveal whether the numerical differences between the corresponding path coefficients in the respective comparative path models are statistically significant at $p < 0.05$ or not.

5.3 Multigroup analysis

To investigate the moderating effect of the race of the behavior models and that of the observers in the evaluation of the former, we conducted MGAs as shown in Table 6. The MGA is based on the following six dyadic pairs: (1) BM versus WM, (2) BR versus WR, (3) BB versus WB, (4) WW versus BW, (5) BB versus BW, (6) WW versus WB, and (7) BB versus WW. The results of the MGAs showed that, regarding BR versus WR, there are significant differences in two relationships: PERS \rightarrow SR ($p = 0.059$ —marginal) and PERS \rightarrow OE ($p < 0.05$). Secondly, regarding BB versus WB, there are significant differences in two relationships: PERS \rightarrow SR ($p < 0.05$) and PERS \rightarrow OE ($p < 0.05$). Finally, regarding BW versus WW, there are significant differences in only one relationship: PERS \rightarrow OE ($p < 0.05$). However, regarding BR versus WR, BB versus BW and WW versus WB, there are no

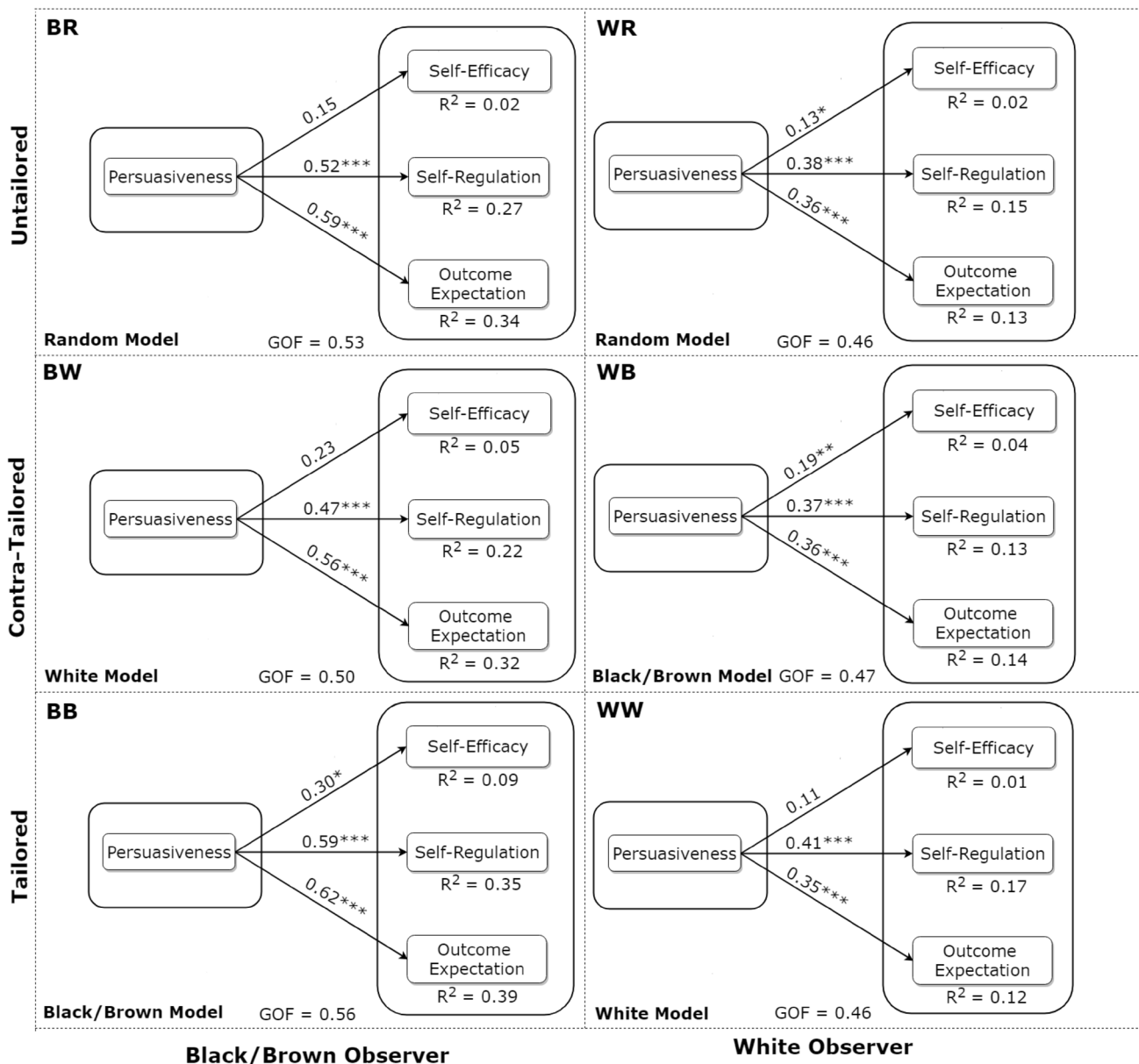


Fig. 4 Race-based tailored, contra-tailored and untailored path models

significant differences in the three investigated relationships. We discuss the implications of the significant differences between the respective dyadic pairs in Sect. 6.

6 Discussion

We have presented a path model of the relationships between the *perceived persuasiveness* of behavior model design and three key social-cognitive factors (*perceived self-efficacy*, *perceive self-regulation* and *outcome expectations*) in tailored, contra-tailored and untailored contexts. Moreover, we showed the significant differences that exist

between key observer-model dyadic pairs in a number of MGAs. In this section, we discuss our key findings and recommend persuasive technology design guidelines accordingly.

6.1 Perceived effect of black/brown and white behavior model design on SCT factors: BM versus WM

Overall, as shown in Fig. 3, the *perceived persuasiveness* of behavior model design, irrespective of its race characteristic, has a positive influence on the three social-cognitive factors, especially on *perceived self-regulation* and *outcome*

Table 6 Multigroup analysis for untailored, contra-tailored and tailored dyadic pairs

Relationship between Model Constructs	R ² value		β value		Sig. Diff.
Black/brown model versus white model: BM versus WM (BM = 336, WM = 333)	BM	WM	BM	WM	<i>p</i> Value
PERS → SE	.04	.01	.21**	.10	n.s
PERS → SR	.17	.18	.41	.42	n.s
PERS → OE	.18	.16	.43	.40	n.s
Untailoring for black/brown observers versus untailoring for white observers: WR versus. BR (BR = 152, WR = 517)	BR	WR	BR	WR	<i>p</i> Value
PERS → SE	.02	.02	.15	.13*	n.s
PERS → SR	.27	.15	.52	.38	0.059
PERS → OE	.34	.13	.59	.36	0.015
Tailoring for black/brown observers versus contra-tailoring for white observers: BB versus WB (BB = 74, WB = 262)	BB	WB	BB	WB	<i>p</i> Value
PERS → SE	.09	.04	.30*	.19**	n.s
PERS → SR	.35	.13	.59	.37	0.043
PERS → OE	.39	.14	.62	.36	0.047
Contra-tailoring for black/brown observers versus tailoring for white observers: BW versus WW (BW = 78, WW = 255)	BW	WW	BW	WW	<i>p</i> Value
PERS → SE	.05	.01	.23	.11	n.s
PERS → SR	.22	.17	.47	.41	n.s
PERS → OE	.32	.12	.56	.35	0.041
Tailoring for black/brown observers versus tailoring for white observers: BB versus WW (BB = 74, WW = 255)	BB	WW	BB	WW	<i>p</i> Value
PERS → SE	.09	.01	.30*	.11	n.s
PERS → SR	.35	.17	.59	.41	0.059
PERS → OE	.39	.12	.62	.35	0.017
Contra-tailoring for black/brown observers versus contra-tailoring for white observers: BW versus WB (BW = 78, WB = 262)	BW	WB	BW	WB	<i>p</i> Value
PERS → SE	.05	.04	.23	.19**	n.s
PERS → SR	.22	.13	.47	.37	n.s
PERS → OE	.32	.14	.56	.36	n.s
Tailoring for black/brown observers versus contra-tailoring for black/brown observers: BB versus BW (BB = 74, BW = 78)	BB	BW	BB	BW	<i>p</i> Value
PERS → SE	.09	.05	.30*	.23	n.s
PERS → SR	.35	.22	.59	.47	n.s
PERS → OE	.39	.32	.62	.56	n.s
Tailoring for white observers versus contra-tailoring for white observers: WW versus WB (WW = 255, WB = 262)	WW	WB	WW	WB	<i>p</i> Value
PERS → SE	.01	.04	.11	.19**	n.s
PERS → SR	.17	.13	.41	.37	n.s
PERS → OE	.12	.14	.35	.36	n.s

PERS Persuasiveness, SE self-efficacy beliefs, SR self-regulation beliefs, OE outcome expectation

BM = Random observers' evaluation of black/brown models, WM = random observers' evaluation of white models

BR = black/brown observers' evaluation of random models, WR = white observers' evaluation of random models

BB = black/brown observers' evaluation of black models, BW = black/brown observers' evaluation of white models

WW = white observers' evaluation of white models, WB = white observers' evaluation of black models

β value in regular font: $p < 0.001$, β value in bold font: fully or marginally significant at $p < 0.05$, β value in italic font: non-significant (n.s)

expectations. In the WM path model, for example, the influence of *perceived persuasiveness* on *perceived self-regulation* ($\beta = 0.42, p < 0.001$) and *outcome expectations* ($\beta = 0.40, p < 0.001$) is stronger than that on *perceived self-efficacy* ($\beta = 0.10, p = n.s.$). The numerical difference between the first two path coefficients and the third is also evident in the tailored, contra-tailored, and untailored path models shown in Fig. 4. These results suggest that irrespective of the race characteristic of the behavior model, the higher users perceive behavior model design to be *persuasive*, the higher will be their *outcome expectations* and their perceived ability to *regulate* themselves to engage in the modeled behavior. Based on these findings, we can conclude that, *in the context of perception, irrespective of tailoring, exercise behavior models are more likely to be effective in increasing outcome expectations and self-regulation beliefs and less likely to be effective in increasing self-efficacy beliefs*. As a result, we recommend that fitness apps designers should leverage exercise behavior models to promote regular exercise, as their *perceived persuasiveness* has the potential of impacting *outcome expectations* and *self-regulation*, both of which mediate behavior change. For example, as seen in Fig. 6 (Appendix 1), positive outcome expectations (e.g., physical benefits of exercise) can be fostered by highlighting the body parts or groups of muscles affected when users perform a given bodyweight exercise. Moreover, users should be given the opportunity, as a way of self-regulation, to set goals, plan their workouts, track their activities, performance and progress.

Moreover, we found that it is only the *perceived persuasiveness* of the black/brown behavior model design that has a significant influence on users' *perceived self-efficacy* ($\beta = 0.21, p < 0.01$). That with regard to white behavior model design is non-significant ($\beta = 0.10, p = n.s.$). While the numerical difference between both path coefficients is not significant, in the light of our first research question, this result seems to suggest that the black/brown behavior models are more likely to influence users' *perceived self-efficacy* than white behavior models. This also seems to be evident in the BB dyad ($\beta = 0.30, p < 0.05$) versus WB dyad ($\beta = 0.19, p < 0.01$), in which the relationship PERS \rightarrow SE—unlike that of WW ($\beta = 0.11, p = n.s.$) and BW ($\beta = 0.23, p = n.s.$)—is statistically significant. However, as a result of the non-statistically significant difference between BM and WM path models with respect to PERS \rightarrow SE, there is a need for further research to answer our first research question, which, in other words, reads: *“in a one-size-fits-all fitness app, which of the behavior model designs (black/brown vs. white) is more likely to influence the social-cognitive factors of observers?”* Meanwhile, a possible answer to this question could be found in the qualitative data we collected (i.e., comments provided by participants on the respective race-based behavior model designs). Thus, we look forward, in future studies, to conducting a sentiment

analysis on participants' comments to uncover any possible explanations as to why the *perceived persuasiveness* of the black/brown models seems to have a stronger significant effect on observers' *perceived self-efficacy*, in particular, and social-cognitive factors, in general.

6.2 Perceived effect of untailored behavior model design on SCT factors: BR versus WR

Figure 4 (first row) shows the path model for the untailored dyads (BR and WR). Overall, the three model parameters (β , R^2 and GOF) are similar to those of BM versus WM in Fig. 3. In the BR and WR path models, the influence of *perceived persuasiveness* of behavior model design on *outcome expectations* and *perceived self-regulation* is stronger than on *perceived self-efficacy*. Moreover, the three relationships tend to be numerically stronger for the BR dyad than for the WR dyad. It is worthy of note that the PERS \rightarrow SE relationship ($\beta = 0.15, p = n.s.$) for BR might have been non-significant because of the relatively small sample size ($n = 152$) compared with the larger sample size of WR ($n = 517$) for which the relationship is significant ($\beta = 0.13, p < 0.05$). Despite that the PERS \rightarrow SE relationship is significant for WR and non-significant for BR, the MGA shows that there is no significant difference between both dyads regarding this relationship. However, for PERS \rightarrow OE and PERS \rightarrow SR, the MGA results, which are fully significant ($p < 0.05$) and marginally significant ($p = 0.059$), respectively, indicate the relationships are stronger in the BR path model than in the WR path model. These findings suggest that, in the light of our second research question, the influence of *perceived persuasiveness* on *outcome expectations* and *self-regulation* is more likely to be stronger among black/brown observers than among white observers. Moreover, the stronger influence of the untailored behavior models on black/brown observers than on white observers is reflected in the amount of variance of the target social-cognitive factors explained by *perceived persuasiveness*. For example, the variances of *outcome expectations* and *perceived self-regulation* (see Fig. 4) explained by *perceived persuasiveness* for the black/brown observers (BR) are 27% and 34%, respectively, while those for the white observers (WR) are 15% and 13%, respectively. These findings, coupled with those regarding BB versus WW and BW versus WB (see Fig. 4), suggest that *the use of behavior models in fitness apps, whether in a tailored, contra-tailored or untailored context, is more likely to be effective among black/brown users than among white users, especially regarding the effect of their perceived persuasiveness on outcome expectations and self-regulation beliefs*.

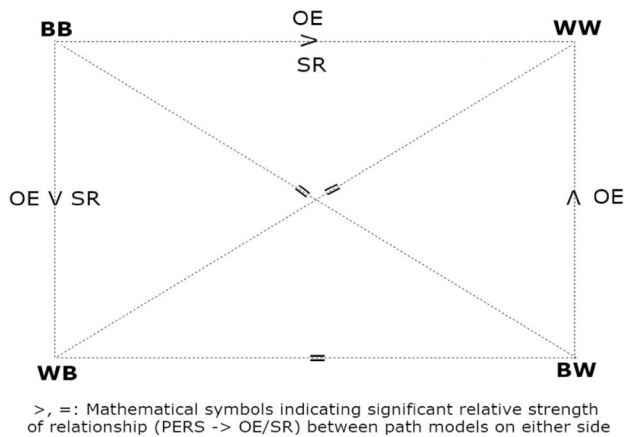


Fig. 5 Abstraction of pairwise comparisons of tailored and contra-tailored path models

6.3 Race-based tailoring and contra-tailoring of behavior models

In general, the results of our path analyses (Fig. 4) showed that exercise behavior model designs are more likely to be effective among black/brown users than among white users, especially in a tailored context in which the PERS \rightarrow SE relationship is significant for the former group. For easy visualization of the pairwise comparisons among the tailored and contra-tailored dyadic path models, we summarize the results of the MGAs (from Table 6) in Fig. 5. The results of the MGAs showed that there are significant differences between BB and WB, between BW and WW, and between BB and WW, at $p < 0.05$. However, there are no significant differences between BB and BW, and between WW and WB, and between WB and BW.

First, regarding BB versus WB, the MGA showed that the *perceived persuasiveness* of the black/brown model design has a significantly stronger effect ($p < 0.05$) on the *self-regulation beliefs* ($\beta = 0.59$, $p < 0.001$) and *outcome expectations* ($\beta = 0.62$, $p < 0.001$) of black/brown observers than those of white observers: ($\beta = 0.37$, $p < 0.001$) and ($\beta = 0.36$, $p < 0.001$), respectively. Overall, this suggests that, *in a real-life fitness app, black/brown models are more likely to be effective for black/brown users (tailored) than for white users (contra-tailored) in impacting outcome expectations and self-regulation.*

Second, with respect to BB versus WW, the MGA showed that, in a tailored context, the *perceived persuasiveness* of behavior model design is more likely to influence the *self-regulation beliefs* ($\beta = 0.59$, $p < 0.001$) and *outcome expectations* ($\beta = 0.62$, $p < 0.001$) of black/brown observers than white observers: ($\beta = 0.41$, $p < 0.001$) and ($\beta = 0.35$, $p < 0.001$), respectively. (Note: the significant difference between the two groups with respect to *self-regulation beliefs* is marginal:

$p = 0.059$.) Overall, this finding suggests that, *in a real-life fitness app, tailored behavior models are more likely to be effective among black/brown users than among white users in influencing outcome expectations and self-regulation.*

Third, with respect to BW versus WW, the MGA showed that the *perceived persuasiveness* of the white model design has a significantly stronger effect ($p < 0.05$) on the *outcome expectations* ($\beta = 0.56$, $p < 0.001$) of black/brown observers than those of white observers ($\beta = 0.35$, $p < 0.001$). The result suggests that, *in a real-life fitness app, regarding outcome expectations, white models are more likely to be effective among black/brown users (contra-tailored) than among white users (tailored).* This finding, coupled with that in the prior paragraph, suggests that, *regarding outcome expectations, behavior models—whether tailored or contra-tailored—are more likely to be effective among black/brown users than among white users.* This confirms the prior finding in Sect. 6.2, in which we saw that, in an untailored context (BR vs. WR), behavior models are more likely to influence the *outcome expectations* of black/brown users ($\beta = 0.59$, $p < 0.001$) than those of white users ($\beta = 0.36$, $p < 0.001$).

6.4 Summary of main findings

All of our findings based on the MGA results in Table 6 are summarized in Table 7 to provide answers to the research questions we posed earlier in Table 2.

As shown in Tables 6, 7 and Fig. 4, we see that, regardless of the race of the observers, behavior model design is more likely to be effective in enhancing *outcome expectations* and *self-regulation beliefs* than *self-efficacy beliefs*. Secondly, we see that, in general, behavior models, whether tailored (BB vs. WB/WW) or contra-tailored (BW vs. WW) or untailored (BR vs. WR), are more likely to have a greater effect on *outcome expectations* for black/brown users than for white users (see RQ2, RQ7, RQ8, and RQ9). Thirdly, as seen in RQ2, RQ7 and RQ9, behavior models, whether tailored (BB vs. WB/WW) or untailored (BR vs. WR), are more likely to have a greater effect on *self-regulation beliefs* among black/brown users than among white users.

To wrap up the discussion, we would like to revisit the overarching research question of this study—“*is race-based tailoring of behavior model design more likely to be effective for both racial target groups than otherwise?*”—in the light of our results. Going by the MGA results for BB versus BW (tailored vs. contra-tailored) and for WW versus WB (tailored vs. contra-tailored), the answer to our question seems to be “No,” as there are no significant differences between both comparative path models in either case regarding the three SCT relationships. However, for black/brown users, given the following:

Table 7 Summary of answers to research questions

No.	Research question (RQ) in words	RQ in symbols	Answer to RQ	Is the difference significant?
RQ1	Overall, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger when observers evaluate black models or white models?	BM or WM?	BM (based on PERS → SE)	No
RQ2	For which of the groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger when they evaluate random behavior models (i.e., in an untailored context)?	BR or WR?	BR (based on PERS → SR and PERS → OE)	YES
RQ3	For black/brown observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	BB > BR?	Yes (based on PERS → SE)	NA
RQ4	For white observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	WW > WR?	No, REV (based on PERS → SE)	NA
RQ5	For black/brown observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	BB > BW?	Yes (based on PERS → SE)	No
RQ6	For white observers, are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context than untailored context?	WW > WB?	No, REV (based on PERS → SE)	No
RQ7	Are the relationships between <i>perceived persuasiveness</i> of black/brown behavior model design and the SCT factors more likely to be stronger for black/brown observers (tailored) than for white observers (contra-tailored)?	BB > WB?	Yes (based on PERS → SR and PERS → OE)	Yes
RQ8	Are the relationships between <i>perceived persuasiveness</i> of white behavior model design and the SCT factors more likely to be stronger for white observers (tailored context) than for black/brown observers (contra-tailored context)?	WW > BW?	No, REV (based on PERS → OE)	Yes
RQ9	For which of the two groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a tailored context?	BB or WW?	BB (based on PERS → SR and PERS → OE)	Yes
RQ10	For which of the two groups of observers (black/brown vs. white) are the relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors more likely to be stronger in a contra-tailored context?	BW or WB?	BW (based on PERS → SE, PERS → SR, and PERS → OE)	No
RQ11	In general, which of the three relationships between <i>perceived persuasiveness</i> of behavior model design and the SCT factors is more likely to be the strongest?	SE or SR or OE?	PERS → OE and PERS → SR	NA

No, REV: means the reverse is true about the symbolic representation of the research question. NA: not applicable

1. the three relationships in the path models are numerically stronger when the behavior models are race-tailored (BB) than otherwise (Fig. 4);
2. the relationship between *perceived persuasiveness* and *self-efficacy beliefs* is only significant ($\beta = 0.30, p < 0.05$) when the behavior models are tailored (BB); and
3. there are significant differences between BB and WB (tailored vs. contra-tailored) but none between BW and WB (contra-tailored vs. contra-tailored).

we may conclude that *behavior model design in fitness apps are more likely to be effective if tailored to the race of black/brown users than otherwise*. However, this conclusion

needs further investigation in future research efforts to confirm it.

On the other hand, with respect to white users, we found no evidence that tailored behavior models are more likely to be effective than contra-tailored or untailored behavior models. Specifically, for the three relationships, we found the following:

1. There is no substantial numerical difference between WW and WR (see Fig. 4).
2. There is no statistically significant difference between WW and WB regarding all three relationships (see Table 6).

However, unexpectedly, we found that the relationship between *perceived persuasiveness* and *self-efficacy beliefs* for WB ($\beta = 0.19, p < 0.01$) is significant, but for WW, it is non-significant ($\beta = 0.11, p = n.s.$). These findings seem to suggest that white users are very likely to be indifferent to the race characteristic of the behavior model design. However, more investigation still needs to be carried out in the future to verify this hypothesis.

6.5 Contributions

Our main contribution to the body of knowledge is that our study provides empirical evidence on the potential effectiveness of the race-based personalization of behavior models in fitness apps aimed at exercise behavior change, especially for black/brown users. We showed that, in general—in an untailored context—behavior models, in fitness apps, are more likely to be effective among black/brown users than among white users. Secondly, we showed that race-based personalization is more likely to be effective among black/brown users than among white users. Finally, for black/brown users, we demonstrated that tailoring of behavior models is more likely to be effective than contra-tailoring and non-tailoring. In the context of fitness apps aimed at promoting regular bodyweight exercise, our study is one of the first to contribute these findings to the body of knowledge.

6.6 Limitations

There are several limitations to our study. The first limitation is that our study is based on user perceptions and not the actual use of fitness apps featuring exercise behavior models. Thus, our findings may not generalize to the context of actual use, in which users have to interact with the behavior models through monitored (logged) usage over a given period of time. For this reason, we recommend that, in future research efforts in the area of race-based personalization, researchers investigate our findings in an actual use context to uncover how far our findings generalize to or differ from a real-life setting.

The second limitation of our findings is that we treated black/brown users as a monolithic group given that both subgroups belong to the collectivist culture based on Hofstede's framework of cultural classification [37, 38]. Similarly, in our study, we did not differentiate between black and brown behavior models in terms of their physical characteristics, e.g., skin color, hairstyle, etc. Rather, our black/brown behavior models seem to be more tailored to the physical characteristics of black users than brown users, e.g., based on skin color, hairstyle, and physique. As a result of these limitations, we recommend that future research efforts in this area should differentiate between

black and brown users to uncover the influence of the acknowledged limitation on our study and provide more nuanced findings related to race-based personalization of exercise behavior model design in the fitness domain.

The third limitation of our study is that our findings may neither generalize to the entire white race nor black/brown race, as our study mainly focuses on both groups of users from North America (Canada and United States). As a result of this limitation, we recommend that, in addition to investigating our findings in an actual application setting, future research efforts should focus on users from other countries and continents to uncover how our current findings generalize to them.

7 Conclusion

In this paper, we presented the path model of the relationships between *perceived persuasiveness* of behavior model design and three social-cognitive factors (*outcome expectations*, *self-regulation beliefs* and *self-efficacy beliefs*) in a tailored, contra-tailored and untailored contexts. In addition, we presented the results of our MGAs aimed at uncovering the potential effectiveness of tailoring exercise behavior model design in fitness apps to the target users. Overall, our path models showed that exercise behavior models are more likely to be effective in influencing users' *outcome expectations* and *self-regulation beliefs* than *self-efficacy beliefs*. Moreover, we found that, in a tailored context (BB vs. WW) and untailored context (BR vs. BR), behavior models are more likely to be effective in influencing the *outcome expectations* and *self-regulation beliefs* of black/brown observers than those of white observers. However, in the contra-tailored context (BW vs. WB), we found no significant differences between both user groups, though the path coefficients for the three relationships turned out to be higher for the black/brown observers than for the white observers. Overall, we did not find a statistically significant difference between tailored and contra-tailored behavior models for the black/brown observers (BB vs. BW) and the white observers (WW vs. WB). However, our path models indicated that race-based personalization is more likely to be effective for black/brown users as the path coefficients for the three relationships turned out to be greater in the tailored than in the contra-tailored and untailored context. Based on this finding and the finding that the relationship between *perceived persuasiveness* and *self-efficacy beliefs* is statistically significant when behavior models are tailored to black/brown users, we recommend that fitness apps featuring behavior models should be specifically tailored to this group of users. In future work, we look forward to conducting sentiment analysis on participants' comments on the behavior model designs

and studies among other populations to determine the generalizability of our findings.

Author’s contributors The first author designed and conducted the study. Both authors contributed in preparing the manuscript.

Funding We thank the Canadian Federal Government and the University of Saskatchewan for funding this research. The second authors’ research is funded through the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants Program, while the first author received the Saskatchewan Innovation and Opportunity Scholarship.

Compliance with ethical standards

Conflict of interests The author(s) declare(s) that there is no conflict of interest.

Research involving human participants and/or animals The studies complied with the research ethics guidelines provided the University of Saskatchewan.

Informed consent Prior to answering the questionnaire, participants were presented with an informed consent form, which explained what the study was about, the potential risk, compensation, right to withdraws, etc. Specifically, The informed consent to participate in the study, which participants agreed to prior to taking the survey, read: “By completing and submitting this questionnaire, your free and informed consent is implied and indicates that you understand the above conditions of participation in this study.”

Appendix 1: Description of fitness app and videos of race-based behavior models

Description of fitness app to participants prior to taking the survey

Kindly read the description of the above application before proceeding to the next page.

Research has shown about one-third of the world’s population are either obese or overweight as a result of a sedentary lifestyle, with about 6% dying as a result of physical inactivity.

Imagine you want to improve your personal health and fitness level. Given the challenges associated with going to the gym regularly, the “Homex App” has been developed to support you in performing exercises at home.

For example, it suggests to you to do three or more different home exercises three or more times a week. Additionally, “Homex” allows you to collaborate with a friend by setting collective goals, and alerts you when your friend achieves his/her part of the set goal.

After each exercise session, the application records data like duration of exercising, the intensity of exercising, and frequency of repetitions. “Homex” processes this data to

analyze your exercise program and adapt it to your personal goals.

A curve diagram informs you about your progress and steps back over time. Progress is rewarded with cheering up like “You did 10 more squats than the last time”. If you are losing sight of your goal, e.g., stopped exercising for some time, “Homex” encourages you not to give up and defines new interim goals to increase your motivation again.

Behavior model videos presented to participants prior to answering questionnaire

See Fig. 6.

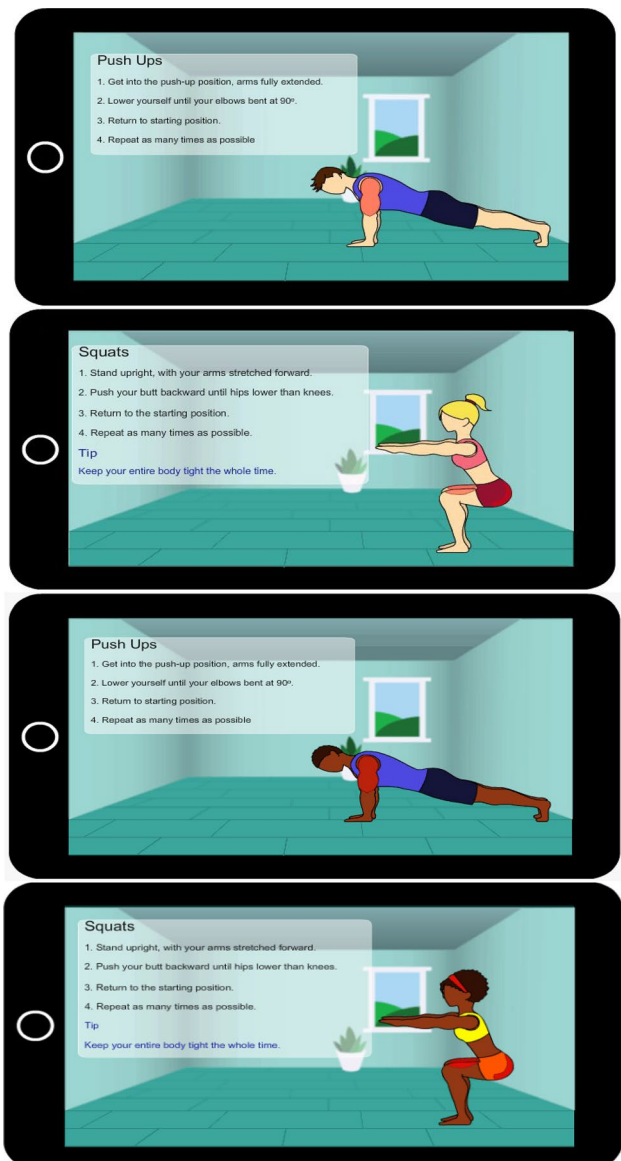


Fig. 6 Four of the eight race-based behavior models performing push-up and squat [2]

Appendix 2: Evaluation of measurement model criteria

See Tables 8, 9, 10.

Table 8 Outer loading of direct and second-order indicators on constructs

Indicator	Block	BM	WM	BR	WR	BW	WB	BB	WW
The app would be convincing	PERS	0.87	0.91	0.87	0.90	0.87	0.87	0.88	0.92
The app would be personally relevant for me	PERS	0.90	0.89	0.89	0.90	0.90	0.90	0.89	0.89
The app would influence me	PERS	0.91	0.93	0.89	0.93	0.91	0.93	0.87	0.93
The app would make me reconsider my physical activity habits	PERS	0.90	0.89	0.91	0.89	0.89	0.89	0.93	0.89
...even when you are busy?	SE	0.85	0.96	0.87	0.88	0.92	0.85	0.85	0.95
...even if you feel depressed?	SE	0.85	0.73	0.87	0.82	0.75	0.84	0.87	0.69
...even when you feel tense?	SE	0.87	0.87	0.87	0.90	0.75	0.88	0.85	0.90
...even when you are tired?	SE	0.90	0.88	0.92	0.89	0.96	0.90	0.88	0.81
...even when you have worries and problems?	SE	0.89	0.81	0.89	0.89	0.79	0.89	0.89	0.80
Second order indicator for physical outcome expectation (Phy OE)	OE	0.80	0.79	0.82	0.79	0.82	0.80	0.82	0.78
Second order indicator for social outcome expectation (Soc OE)	OE	0.78	0.78	0.81	0.77	0.80	0.77	0.83	0.78
...improve the functioning of my cardiovascular system	Phy OE	0.75	0.74	0.70	0.76	0.64	0.75	0.77	0.78
...improve my ability to perform daily activities	Phy OE	0.81	0.75	0.80	0.78	0.86	0.83	0.74	0.71
...improve my overall body functioning	Phy OE	0.84	0.76	0.83	0.80	0.78	0.83	0.88	0.75
...increase my muscle strength	Phy OE	0.74	0.67	0.71	0.71	0.70	0.75	0.70	0.65
...strengthen my bones	Phy OE	0.64	0.61	0.65	0.62	0.64	0.64	0.66	0.60
...improve my social standing	Soc OE	0.90	0.90	0.90	0.90	0.89	0.89	0.92	0.91
...increase my acceptance by others	Soc OE	0.90	0.89	0.90	0.89	0.89	0.89	0.91	0.89
...make me more at ease with people	Soc OE	0.86	0.88	0.88	0.87	0.88	0.85	0.88	0.88
I will endeavor to achieve the set goal for myself	SR	0.86	0.84	0.87	0.84	0.87	0.85	0.87	0.83
I will develop a series of steps to reach my weekly goal	SR	0.78	0.80	0.82	0.78	0.83	0.76	0.81	0.79
I will keep track of my progress in meeting my goal	SR	0.85	0.85	0.82	0.86	0.85	0.86	0.80	0.86
I will set a goal	SR	0.88	0.85	0.89	0.85	0.91	0.88	0.88	0.82

OE was a higher-order construct composed of two lower-order constructs (Phy OE and Soc OE)

Table 9 Internal consistency reliability based on Dillon–Goldstein metric (DG.rho)

Construct	BM	WM	BR	WR	BW	WB	BB	WW
PERS	0.94	0.95	0.94	0.95	0.94	0.95	0.94	0.95
SE	0.94	0.95	0.95	0.94	0.95	0.94	0.94	0.95
OE	0.77	0.76	0.80	0.76	0.80	0.76	0.80	0.76
Phy OE	0.87	0.83	0.86	0.86	0.86	0.86	0.87	0.83
Soc OE	0.92	0.92	0.92	0.92	0.92	0.92	0.93	0.92
SR	0.91	0.90	0.91	0.90	0.91	0.90	0.91	0.90

Table 10 Convergent validity based on average variance extracted (AVE)

Construct	BM	WM	BR	WR	BW	WB	BB	WW
PERS	0.80	0.82	0.79	0.82	0.79	0.82	0.80	0.83
SE	0.76	0.73	0.78	0.77	0.78	0.77	0.75	0.70
OE	0.63	0.62	0.67	0.61	0.67	0.61	0.67	0.61
Phy OE	0.58	0.50	0.55	0.54	0.55	0.54	0.57	0.49
Soc OE	0.78	0.80	0.80	0.79	0.80	0.79	0.82	0.80
SR	0.71	0.70	0.73	0.69	0.73	0.69	0.71	0.68

References

- Abraham C, Michie S (2008) A taxonomy of behavior change techniques used in interventions. *Health Psychol* 27:379–387. <https://doi.org/10.1037/0278-6133.27.3.379>
- Oyibo K, Adaji I, Orji R, et al (2018) Perceived persuasive effect of behavior model design in fitness apps. In: Proceedings of the 26th conference on user modeling, adaptation and personalization, Singapore, pp 219–228
- Wouters P, Tabbers HK, Paas F (2007) Interactivity in video-based models. *Educ Psychol Rev* 19:327–342. <https://doi.org/10.1007/s10648-007-9045-4>
- Conroy DE, Yang CH, Maher JP (2014) Behavior change techniques in top-ranked mobile apps for physical activity. *Am J Prev Med* 46:649–652. <https://doi.org/10.1016/j.amepre.2014.01.010>
- Oyibo K, Ifeoma A, Vassileva J (2018) Social cognitive determinants of exercise behavior in the context of behavior modeling: a mixed-method approach. *Digit Heal*. <https://doi.org/10.1177/2055207618811555>
- Sundar SS, Marathe SS (2010) Personalization versus customization: the importance of agency, privacy, and power usage. *Hum Commun Res* 36:298–322. <https://doi.org/10.1111/j.1468-2958.2010.01377.x>
- Van Dooren MMM, Visch VT, Spijkerman R (2017) Personalized design process for persuasive technologies. *Pers Persuas Technol Work Proc* 1833:26–31
- Kaptein M, De Ruyter B, Markopoulos P, Aarts E (2012) Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Trans Interact Intell Syst* 2:1–25. <https://doi.org/10.1145/2209310.2209313>
- Kaptein M, Markopoulos P, De Ruyter B, Aarts E (2015) Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. *Int J Hum Comput Stud* 77:38–51. <https://doi.org/10.1016/j.ijhcs.2015.01.004>
- Lehto T, Oinas-Kukkonen H (2011) Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *J Med Internet Res*. <https://doi.org/10.2196/jmir.1559>
- Stawarz K, Cox AL, Blandford A (2014) Don't forget your pill!: designing effective medication reminder apps that support users' daily routines. In: CHI'14 Proceedings of the SIGCHI conference on human factors computing system, pp 2269–2278. <https://doi.org/10.1145/2556288.2557079>
- Oyibo K (2016) Designing Culture-based Persuasive Technology to Promote Physical Activity among University Students. In: 24th conference on user modeling, adaptation and personalization (UMAP 2016). Halifax, Canada, pp 321–324
- Bandura A (2001) Social cognitive theory: an agentic perspective. *Annu Rev Psychol* 6:1–60. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Keller C, Fleury J, Gregor-Holt N, Thompson T (1999) Predictive ability of social cognitive theory in exercise research: an integrated literature review. *J Knowl Synth Nurs* 6:2
- Oyibo K, Orji R, Vassileva J (2018) Developing culturally relevant design guidelines for encouraging physical activity: a social cognitive theory perspective. *J Healthc Informatics Res*. <https://doi.org/10.1007/s41666-018-0026-9>
- Oyibo K (2016) Designing culture-based persuasive technology to promote physical activity among university students. In: UMAP 2016—proceedings of the 2016 conference on user modeling adaptation and personalization
- Anderson ES, Wojcik JR, Winett RA, Williams DM (2006) Social-cognitive determinants of physical activity: the influence of social support, self-efficacy, outcome expectations, and self-regulation among participants in a church-based health promotion study. *Health Psychol* 25:510–520. <https://doi.org/10.1037/0278-6133.25.4.510>
- Resnick B (2001) A prediction model of aerobic exercise in older adults living in a continuing-care retirement community. *J Aging Health* 13:287–310. <https://doi.org/10.1177/089826430101300207>
- Rovniak L, Anderson ES, Winett RA, Stephens RS (2002) Social cognitive determinants of physical activity in young adults: a prospective structural equation analysis. *Ann Behav Med* 24:149–156
- Bandura A (1997) *Self-efficacy: the exercise of control*. W H. Freeman, New York
- Anderson-Bill ES, Winett RA, Wojcik JR (2011) Social cognitive determinants of nutrition and physical activity among web-health users enrolling in an online intervention: the influence of social support, self-efficacy, outcome expectations, and self-regulation. *J Med Internet Res*. <https://doi.org/10.2196/jmir.1551>
- Jimison HB, Hagler S, Kurillo G, et al. (2015) Remote health coaching for interactive exercise with older adults in a home environment. In: Proceedings of the annual international conference of the IEEE engineering in medicine and biology society (EMBS) pp 5485–5488. <https://doi.org/10.1109/EMBC.2015.7319633>
- Bandura A (1971) *Social learning theory*. Stanford Univ. <https://doi.org/10.1111/j.1460-2466.1978.tb01621.x>
- Sheeshka JD, Woolcott DM, Mackinnon NJ (1993) Social cognitive theory as a framework to explain intentions to practice healthy eating behaviors. *J Appl Soc Psychol* 23:1547–1573. <https://doi.org/10.1111/j.1559-1816.1993.tb01047.x>
- Eyck A, Geerlings K, Karimova D, et al. (2006) Effect of a virtual coach on athletes' motivation. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). pp 158–161
- Devi B, Das M, Das M, Khandelwal B (2017) Application of Bandura's social cognitive theory in the technology-enhanced, blended learning environment. *Int J Appl Res* 3:721–724
- Sanchez G (2009) Introduction to the R package PLS-PM. 1–9
- Evermann J, Tate M (2012) Comparing the predictive ability of PLS and covariance models. *ICIS* 18:546–550
- Tobias RD (1995) An introduction to partial least squares regression. *Anal* 24:1250–1257
- Oinas-kukkonen H, Harjumaa M (2008) A systematic framework for designing and evaluating persuasive systems. In: International conference on persuasive technology, pp 164–176
- Oyibo K, Orji R, Vassileva J (2018) Developing culturally relevant design guidelines for encouraging physical activity: a social cognitive theory perspective. *J Healthc Informatics Res*. <https://doi.org/10.1007/s41666-018-0026-9>
- Wójcicki TR, White SM, Mcauley E (2009) Assessing outcome expectations in older adults: the multidimensional outcome expectations for exercise scale. *J Gerontol Ser B, Psychol Sci Soc Sci* 64:33–40. <https://doi.org/10.1093/geronb/gbn032>
- Sanchez G (2013) *PLS path modeling with R*. Berkley Trowchez Ed, Berkley
- Wong KK (2013) Partial least squares structural equation modeling (PLS-SEM) TECHNIQUES USING smartPLS. *Mark Bull* 24:1–32. <https://doi.org/10.1108/EBR-10-2013-0128>
- Raines-Eudy R (2000) Using structural equation modeling to test for differential reliability and validity: an empirical demonstration. *Struct Equ Model* 7:124–141. https://doi.org/10.1207/S15328007SEM0701_07
- Hair JF, Hult GTM, Ringle CM, Sarstedt M (2014) *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications Inc, Washington DC

37. Hofstede G (2011) Dimensionalizing cultures: the hofstede model in context. *Online Readings Psychol Cult* 2:1–26. <https://doi.org/10.9707/2307-0919.1014>
38. Hofstede G, Hofstede GJ, Minkov M (2010) *Cultures and organizations: software of the mind*, 3rd edn. McGraw Hill, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.