6-15-2019

# The effect of “gender fit” on fitness app engagement

Qihui Xia
*School of Management, Xi'an Jiaotong University, China*, xiaqh20@hotmail.com

Xi Zhao
*School of Management, Xi'an Jiaotong University, China*, zhaoxi1@mail.xjtu.edu.cn

Wei Huang
*School of Management, Xi'an Jiaotong University, China*, waynehw21st@gmail.com

Atreyi Kankanhalli
*Department of Information Systems and Analytics, National University of Singapore*, atreyi@comp.nus.edu

# The effect of "gender fit" on fitness app engagement

*Completed Research Paper*

Qihui Xia, School of Management, Xi'an Jiaotong University, Xi'an, China, xiaqh20@hotmail.com

Xi Zhao*, School of Management, Xi'an Jiaotong University, Xi'an, China, zhaoxi1@mail.xjtu.edu.cn

Wei Huang, School of Management, Xi'an Jiaotong University, Xi'an, China, waynehw@qq.com

Atreyi Kankanhalli, Department of Information Systems and Analytics, National University of Singapore, Singapore, atreyi@comp.nus.edu.sg

## Abstract

*The effectiveness of fitness app in health promotion success has been observed and may greatly contribute to large health improvements of the public. However, the actual engagement with these interventions is unexpectedly low. Fitness app developers design Behavior change techniques (BCTs) to enhance user engagement with fitness apps. Although several studies have examined the effectiveness of some BCTs in encouraging user engagement with Internet-based applications in general, investigations remain underspecified. Based on the theory of psychological fit, we focus on the gender boundary condition of the effectiveness of BCTs on user engagement with fitness apps. The purpose of this research is twofold. First, we aim to explore gender differences in preferences for BCTs. The second purpose is to investigate whether there exists "gender fit" effect on user engagement of fitness apps with different BCTs.*

**Keywords:** Fitness app, Gender, User engagement, Behavior change techniques

## Introduction

Physical inactivity has been extensively recognized a factor of many chronic illnesses (Paffenbarger Jr & Hyde, 1984; Siscovick, Laporte, & Newman, 1985) and is associated with almost 5.3 million deaths per year (Lee et al., 2012). With the prevalence of smartphones, mobile applications (Apps) are prevalent as tools to help users monitor and manage their health. The market share of mobile health has greatly increased. By 2012, almost one in five U.S. smartphone owners download at least one health app from the app market (Fox & Duggan, 2012), and fitness apps are the most popular health apps (Statista, 2017). Fitness apps can monitor health behavior in a noninvasive manner via smartphone built-in accelerometers and interact with users in real time, making them a powerful tool to promote physical activity levels on a large scale. The effectiveness of fitness app in health promotion success has been observed (Serrano, Yu, Coa, Collins, & Atienza, 2016) and may greatly contribute to large health improvements of the public. However, the actual engagement with these interventions is unexpectedly low (Brouwer et al., 2011), which is a concern of both researchers and health service practitioners.

Behavior change techniques (BCT) taxonomy is a systematic and rigorous method to characterize complex persuasive interventions in a consistent format (Michie et al., 2013). BCT refers to a component of an intervention (e.g., using a mobile fitness apps) designed to facilitate health behaviors, e.g., self-monitoring, feedback, and reward (Yang, Maher, & Conroy, 2015). Fitness app developers design BCTs to motivate user engagement with fitness apps (Serrano, Coa, Yu, Wolff-Hughes, & Atienza, 2017). Many studies have examined the effectiveness of some BCTs in encouraging user engagement with Internet-based applications in general (see review Brouwer et al., 2011). However, the boundary conditions of the effectiveness of BCTs receive scant attention. Research about psychological fit suggests that better psychological fit between persuasive interventions and individual preferences achieves better performance (S. Matz, Kosinski, Nave, & Stillwell, 2017). However, little is known about the effect of psychological fit between BCTs and users on user engagement with fitness apps.

This research focuses on preferences driven by gender differences. There are two reasons. First, gender information is a common factor of consumer segmentation. Gefen and Straub (1997) suggest that the development of a more favorable communication environment should take gender into consideration. This attempt may enrich the theory of health behavior change by identifying the gender boundary condition. Second, gender information can be identified easily in practice. As males and females have different preferences for mobile apps or services (Weiser, 2000; Xu, Frey, & Ilic, 2016), a closer look at the gender boundary of the effectiveness of BCTs can provide useful guidance for fitness app users to find a more suitable health behavior intervention tool, and for fitness app developers to attract higher user engagement according to the gender distribution of their users.

User engagement of apps is a multi-facetted concept, which emphasizes on users' positive experiences of interacting with the app (Lalmas, O'Brien, & Yom-Tov, 2014). In accord with the web analytics-based approach (Lalmas et al., 2014), we measure user engagement of fitness apps by the frequency of access to the app. The purpose of this research is accordingly twofold. First, we aim to examine whether there is any gender difference in preference for BCTs. The second purpose is to investigate whether there exists "gender fit" effect on user engagement with fitness apps. In this research in progress paper, we selected two typical fitness apps with BCTs that were more preferred by men and women respectively. Then, we assign them a gender of male or female, namely app gender. We also refer to the gender of user as user gender. Research questions are:

(1) Whether female and male users are more likely to adopt fitness apps with BCTs that are more preferred?

(2) Whether the fit between app gender and user gender may improve user duration use of fitness app?

## Theoretical background

Psychological fit is a concept from psychology literature that indicates the extent to which people feel the environment fit their preferences (Mazt, Gladstone, & Stillwell, 2016). It is well-established that better psychological fit of the environment makes people feel more comfortable to express themselves and reinforces their self-concepts (Grubb & Grathwohl, 1967), which leads to higher user satisfaction and better persuasion effectiveness (Jokela, Bleidorn, Lamb, Gosling, & Rentfrow, 2015; S. Matz et al., 2017). For example, a social incentive, such as the verbal praise, may be more persuasive in changing behaviors to the people who are interdependent than others who are independent because the behavior may fulfill their need of being cared by others and reinforce their self-expression of being part of the team. This mechanism was referred to as the theory of psychological fit by Mazt, Gladstone, & Stillwell (2016).

The theory of psychological fit has been widely used by consumer psychologists and in marketing practices where self-expression behaviors may occur (S. C. Matz, Gladstone, & Stillwell, 2016). Dubois, Rucker, and Galinsky (2016) report a power fit effect on the persuasive effectiveness and suggest that the interaction between communicator power and audience power should be considered to achieve a successful persuasive communication. Cesario, Higgins, and Scholer (2008) propose several methods to increase the effectiveness of changing behavior by inducing regulatory fit. In the age of computational social science (Lazer et al., 2009), more recent studies apply this theory to real-world persuasion on a large-scale population. S. Matz et al. (2017) find that the fit of advertisements with individual preferences driven by personality significantly change consumer behavior in terms of online ad clicks and purchases.

We apply the theory of psychological fit to this research, as user engagement with fitness apps is a form of self-expression. When BCTs used by fitness apps fit the preferences of users, users are likely to be more motivated to engage with the intervention and to regulate their physical activity behaviors following the guides. In our paper, user preferences for BCTs are driven by gender differences.

## Hypotheses Development

In accord with the theory of psychological fit, we posit that female and male users are more likely to adopt fitness apps that provide their preferred BCTs. By reviewing relevant major theories and literature about gender differences, Meyers‐Levy and Loken (2015) conclude five major differences. Our hypotheses about gender differences in preferences for BCTs are built upon three of their conclusions that are relevant to our research.

First, males focus on instrumentality and independence, whereas females focus on inclusiveness and interdependence (Meyers‐Levy and Loken, 2015). Males tend to view themselves as separate from others, whereas females feel more connected to others (Guimond, Chatard, Martinot, Crisp, & Redersdorff, 2006). In this regard, Peterson, Lawman, Wilson et al. (2013) suggested that males and females have different preferences for social support on physical activity. Females prefer emotional social support that helps them to disclose themselves to others. For instance, users are allowed to post their exercise data, which can been seen and liked by their friends (coded as the BCT 3.3 social support (emotional), 3.3 is the BCT's coding number consistent with Michie et al., 2013). By contrast, males prefer instrumental social support. For instance, users are allowed to join groups and seek technical support on exercise from others (coded as 3.1 social support (unspecific)). Accordingly, we expect that males prefer BCTs about personal goal achievement, whereas females prefer BCTs about social and affective supports.

Second, males show more risk seeking tendency, whereas females behave with more cautiousness (Meyers-Levy and Loken, 2015). Males tend to take a risk to pay for more gains in return (Buchan, Croson, & Solnick, 2008). For instance, male users are more likely to participate in paid programs where they cannot get the payment back if they don't complete exercise goals (coded as 14.3 Remove reward). Moreover, males prefer competitive activities than females who prefer a variety of activities (Wright, Wilson, Griffin, & Evans, 2008), leading to a higher preferences for BCTs about getting material rewards or incentives for goal achievement (e.g., 10.1 Material incentive (behavior), 10.2 Material reward (behavior)). By contrast, females may prefer BCTs about reducing costs and loss. Social incentives or rewards, such as earning a virtual metal that can be seen by others or verbal praise (10.5 Social incentive, 10.4 Social reward), may be more attractive to females. 6.3 Information about others' approval may also satisfy females' preference for reducing costs by giving them more information.

Third, females exhibit more comprehensiveness and males are more selective in information processing (Meyers-Levy and Loken, 2015). Females tend to elaborate more message cues than males (Meyers-Levy & Sternthal, 1991). Thus, we posit that males prefer BCTs that require less cognitive process and help them achieve behavior goals straightway, whereas females may prefer BCTs that provide rich and diverse information about the behavior and the process. For example, the BCT 5.4 Monitoring of emotional consequences, which enables users to record their feelings after exercises, may be more preferred by females than males. Additionally, the BCT of frequent reminders on doing exercises (i.e., 8.3 Habit formation) or motivational cues (i.e., 7.1 Prompts/cue) may attract more female users because it requires considerable cognitive efforts to deal with the frequent contacts.

H1: More male users than female user use BCTs that involve (a) more instrumentality and independence, (b) more risks and material rewards and (c) more simplicity.

H2: More female users than male users use BCTs that involve (a) more inclusiveness and interdependence, (b) less punishments and (c) more variety.

Considerable previous studies suggest that the persuasive effectiveness is particular high when the persuasive communication fits the psychological preferences of individuals (Dubois et al., 2016; S. Matz et al., 2017; S. C. Matz et al., 2016). Fitness apps implement various BCTs to motivate their users to follow their persuasive interventions on health behavior change. The level of user engagement with the app is a reflection of the persuasive effectiveness of the app because users are more likely to change behaviors when they have frequent access to the persuasive interventions (Brouwer et al., 2011). As argued above, there are gender differences in preferences for BCTs. BCTs that match individual psychological preferences tend to be more persuasive (Yardley, Morrison, Bradbury, & Muller, 2015). Hence, we expect that the fit between gender preferences and BCTs used by fitness apps may elicit higher engagement.

H3: User engagement with fitness apps is higher in congruent conditions where user gender match app gender than in incongruent conditions.

## Methodology

In this section, we describe the methodology we utilize to test hypotheses. First, we discuss the dataset and the preprocessing of the dataset. Then we describe the statistical analysis method to analyze the data.

### *Dataset and Preprocessing*

The data used in this research were derived from an anonymized mobile app usage database of 300,000 mobile phone users from a province of China. We made a list of 59 fitness apps, which were relevant to physical activity and top-ranked on two major application marketplaces in China: Apple iTunes (iPhone operating system [iOS]) and Ying Yong Bao (Android application platform of Tencent Inc.). Users were extracted from the whole database who used at least one of the 59 apps in a week from Jan 7, 2018 to Jan 13, 2018, resulting in 5821 users (3990 males and 1831 females).

An app used by a user in the week was saved as a record. A user might use more than one app, leading to 6403 records. Then, we calculated the number of records for each fitness app and observed that nine apps took up 91% cumulative share of total engagement records. Thus, we only included the users of the top nine fitness apps in the analysis, resulting in 5440 users (3715 males and 1725 females) and 5830 engagement records. The coverage rate of our sample was 93.5% (5440/5821). The age distribution of the sample was in consistent with Ernsting, Dombrowski, Oedekoven, and LO (2017)'s population-based survey in German that fitness app users were young. More than 50 percent of our sample users were below 35 years old.

It is notable that our analyses below are based on BCTs, thus app names are not essential. Hence, we anonymized app names due to a confidentiality agreement. We sort the nine fitness apps by the number of users from most to least and refer to them as App1, App2, … , App9. At first, we clarify some definitions that are used in data processing. We assume that there are $n$ fitness apps (in this research, $n = 9$). $\boldsymbol{u}_i$ indicates a set of $n$ fitness apps that are used by user i, i.e., $\boldsymbol{u}_i = \{a_{i1}, a_{i2}, \dots, a_{in}\}$, where $a_{ik}$ represents whether the app $k$ is used by user $i$ (value=1) or not (value=0).

### *BCTs Coding Scheme*

We recruited and trained 3 coders to independently code the presence or absence of BCTs implemented in each app using the BCT taxonomy (v1), which proposed 93 different BCTs (Michie et al., 2013). They used the above nine fitness apps for two weeks. After that, they were asked to code all apps in a room independently (reviewing apps were allowed during the coding process). Coding discrepancies were discussed and resolved to achieve consensus. We screen out those BCTs that are unused in the analysis, and define fitness app $j$ as $\boldsymbol{a}_j$ with the remaining $m$ BCTs, i.e., $\boldsymbol{a}_j = \{b_{j1}, b_{j2}, \dots, b_{jm}\}, j \in \{1,2,\dots,n\}$, where $b_{jk}$ is the presence (value=1) or absence (value=0) of BCT $k$ in fitness app $j$. Thus, $\boldsymbol{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}^T$ indicates the BCT coding result of $n$ fitness apps.

## *Gender Differences in Preferences for BCTs Analyses*

We define another vector $\boldsymbol{c}_i = \{c_{i1}, c_{i2}, \dots, c_{im}\}$, which indicates a set of $m$ BCTs that are adopted by user $i$. $c_{ik} = 1$ means BCT $k$ is implemented in at least one fitness app that is used by user $i$, otherwise, $c_{ik} = 0$. We acknowledge that a user may prefer some of the BCTs implemented in a fitness app rather than all of them, but our proposition make sense in the big data context because we perform the calculation at the group level and individual biases may reduce owing to the aggregation effect.

**Measurement of gender differences in preferences for BCTs**. The preference for BCTs that we investigate is illustrated by the use of BCTs. We assume that a larger proportion of male (or female) users than female (or male) users who use a specific BCT indicates a higher preference for the BCT among males (or females). Hence, we measure gender preferences for a BCT with the proportion of users (*male* vs. *female*) who use a specific BCT $j$, i.e., $P_{Gender,j}$ (*Gender=male* or *female*).

$$P_{Gender,j} = \frac{\sum_{i \in I_{Gender}} c_{ij}}{Total_{Gender}},$$

where $I_{Gender}$ represents that a set of male users (*Gender=male*) or female users (*Gender=female*). $Total_{Gender}$ is the total number of male or female users. Gender differences $D$ in preferences for BCT $j$ is computed by

$$D_j = P_{Male,j} - P_{Female,j}, \qquad \text{Equation (1)}$$

where $D_j > 0$ indicates that BCT $j$ is more preferred by males, and $D_j < 0$ indicates that BCT $j$ is more preferred by females. Then we get two lists of BCTs. $L_{Male}$ contains a list of BCTs where $D_{j \in L_{Male}} > 0$, whereas $L_{Female}$ contains a list of BCTs where $\mathrm{D}_{j \in L_{Female}} < 0$.

## *Gender Fit Effect on Engagement Level Analyses*

Based on previous analysis, we obtain two lists of BCTs that are preferred by males and females respectively. In this part, we differentiate fitness apps by BCTs that they adopt. In congruent conditions, male and female users use fitness apps, which adopt more male- and female-preferred BCTs. In incongruent conditions, male and female users use fitness apps, which adopt more female- and male-preferred BCTs.

**Definition of app gender**. In accord with S. Matz et al. (2017) who refer to "the personality of the audience an ad[vertisement] is aimed at as ad personality", we refer to the gender of users by whom BCTs of a fitness app are preferred as app gender. App gender can be identified by two compatible dimensions, i.e., male and female. A fitness app which adopts both male- and female-preferred BCTs is valued high on both dimensions.

**Measurement of user engagement level with fitness apps**. Adapted from Serrano et al. (2017), we measure user engagement level with a fitness app using the number of days logged in the observation week (ranging from 1 to 7).

At first, we select two typical fitness apps that adopt more male- and female-preferred BCTs, namely male app and female app. The principle of selection is that these two apps use a relatively large number of BCTs in $L_{Male}$ and $L_{Female}$ lists respectively. This principle is to ensure that one of the two apps is more representative of male app, whereas another is more representative of female app. In doing so, we first pair two fitness apps, either of which is taken from male apps and female apps. Then, we calculate correlations for all pairs. The pair that are most uncorrelated are selected as our target male and female apps. Finally, we make a 2 (male app vs. female app ) $\times$ 2 (male user vs. female user) design to test the gender fit effect on engagement level.

# Preliminary Findings

## *BCT Coding Result*

The mean Cohen's Kappa of the initial coding result is 0.644, ranging from 0.621-0.677, indicating moderate to substantial agreement. After discussion, all the three coders agree with a consensus coding result, which is summarized in Table 1. Among the total 93 BCTs of the BCT taxonomy (v1), we observe 33 BCTs in the coded fitness apps (therefore, m = 3). Appropriate 18 BCTs on average with a range 6-23 (Mean=17.89, SD=5.34, median=19) are used by each fitness app. The number is relatively high in our research compared with in some western countries which is below 10 (Conroy, Yang, & Maher, 2014; Yang, Maher, & Conroy, 2015).

**Table 1. BCTs used in the selected nine fitness apps**

| BCT code | App1 | App2 | App3 | App4 | App5 | App6 | App7 | App8 | App9 | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 1.1 Goal setting (behavior) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 1.3 Goal setting (outcome) | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 5 |
| 1.4 Action planning | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 7 |
| 1.5 Review behavior goal(s) | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 3 |
| 1.6 Discrepancy between current behavior and goal | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 1.7 Review outcome goal(s) | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 2.2 Feedback and behavior | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 2.4 Self-monitoring of outcomes of behavior | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 6 |
| 2.6 Biofeedback | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 3 |
| 2.7 Feedback on outcome(s) of behavior | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 6 |
| 3.1 Social Support (unspecified) | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 4 |
| 3.3 Social Support (emotional) | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| 4.1 Instruction on how to perform a behavior | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| 5.1 Information about health consequences | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 4 |
| 5.4 Monitoring of emotional consequences | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 2 |
| 6.1 Demonstration of the behavior | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 8 |
| 6.2 Social comparison | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 6.3 Information about others' approval | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 2 |
| 7.1 Prompts/cue | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 4 |
| 8.3 Habit formation | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 |
| 8.7 Graded tasks | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 6 |
| 9.1 Credible source | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 10.1 Material incentive (behavior) | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 4 |
| 10.2 Material reward (behavior) | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 4 |
| 10.4 Social reward | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 3 |
| 10.5 Social incentive | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 3 |
| 10.11 Future punishment | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 3 |
| 11.3 Conserving mental resource | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| 13.5 Identity associated with changed behavior | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 14.1 Behavior cost | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 3 |
| 14.3 Remove reward | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 4 |
| 14.10 Remove punishment | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 2 |
| 16.3 Vicarious consequences | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 5 |
| **Total** | 23 | 22 | 19 | 23 | 16 | 17 | 22 | 6 | 13 | |

Note. The code numbers of the 33 BCTs that are used by the nine fitness apps are the same as they are in the BCT taxonomy (v1). 1 in the table represents the presence of the BCT in the app, whereas 0 represents the absence of the BCT in the app.

### *Hypotheses Test*

We measure gender differences in preferences for each BCT with the method described in methodology section. Figure 1 displays the result of $D_j$ (see Equation (1) for computation) for each BCT (coding numbers are derived from Michie et al. 2013). Coded BCTs are plotted on the horizontal axis, whereas $D_j$ for all BCTs are plotted on the vertical axis. BCTs with values above the horizontal axis are regarded more preferred by males, whereas those below the horizontal axis are regarded more preferred by females. As shown in Figure 1, females prefer BCTs including 6.3 Information about others' approval, 8.3 Habit formation, 5.4 Monitoring of emotional consequences, 10.5 Social incentive, 10.4 Social reward, 1.5 Review behavior goal(s), 7.1 Prompts/cue, etc. Males prefer BCTs including 10.1 Material incentive (behavior), 10.2 Material reward (behavior), 2.6 Biofeedback, 14.3 Remove reward, 3.1 Social Support (unspecified), etc. Most H1 and H2 are supported.

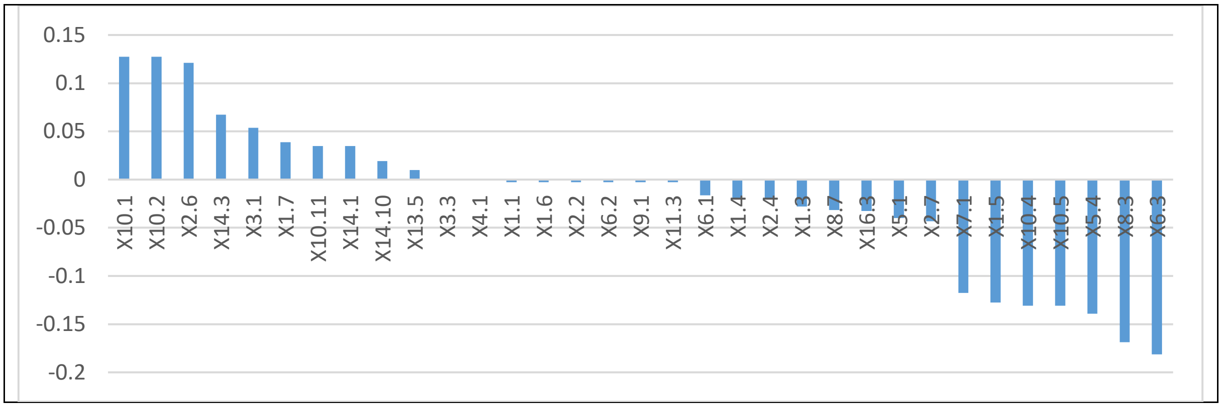


**Figure 1. Results of gender differences in preferences for BCTs analyses.**

Note. Each BCT has a coding number derived from (Michie et al., 2013). BCTs with $\boldsymbol{D_j}$ values above the horizontal axis are regarded more preferred by males, whereas those below the horizontal axis are regarded more preferred by females.

In testing H3, we first select two target apps following the method described in the methodology section. These two target apps distinct in app gender. We conduct hierarchical linear regression analyses for engagement level with fitness apps using user gender, app gender and their two-way interaction as predictors. The result is shown in Table 2. We observed the significant interaction effect of user gender and app gender on engagement level ($\mathrm{B} = 0.228, \mathrm{p} < 0.05$). The result is robust after controlling the effect of age and its interaction effect with app gender.

Further, we test whether or not an app has a higher engagement among users in congruent conditions where user gender and app gender fit. The engagement level with apps is redefined as low engagement level (using an app for up to three days, N=1360) and high engagement level (using an app for at least five days, N=955). Then we conduct Chi-square test in congruent group and incongruent group with low and high engagement levels. The result suggests that the engagement level is 1.447 times more likely to be higher in congruent group than in incongruent group ($\chi^2(1) = 18.649, \text{odds ratio (OR)} = 1.447\ [1.223 - 1.711], \mathrm{P} = 0.000 < 0.05$). This effect supports H3.

To be specific, we conduct Chi-square test for user engagement level in male and female apps respectively. As depicted in Figure 2, the gender fit effect is significant in using the female app, as females are 1.364 times more likely to increase engagement with the female app ($\chi^2(1) = 8.340, \text{odds ratio (OR)} = 1.364\ [1.105 - 1.686], \mathrm{P} = 0.004 < 0.05$). However, in using the male app, the gender fit effect on engagement level is not significant ($\chi^2(1) = 0.184, \text{odds ratio (OR)} = 1.077\ [0.766 - 1.514], \mathrm{P} = 0.668 > 0.05$). This result suggests that males and females engage equally with the male app and the engagement level is relatively higher than with the female app. However, females significantly engage more with the female app than males, indicating that the female app is more persuasive among females and less persuasive among males.

**Table 1 Hierarchical linear regression analyses for engagement level**

| | Model 1 | Model 2 |
|---|---|---|
| User gender | 0.046 (0.022) | -0.113 (0.151) |
| App gender | -0.125 (0.000) | -0.290 (0.001) |
| Age | 0.013 (0.518) | -0.037 (0.596) |
| User gender × App gender | | 0.228 (0.035) |
| Age × App gender | | 0.063(0.442) |

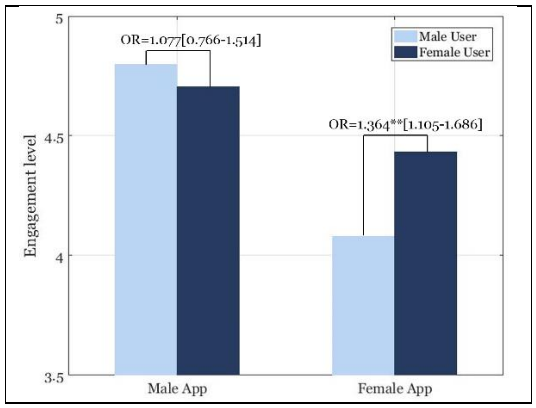


**Figure 2. Interaction effect of user gender and app gender on engagement level with fitness apps**

## Conclusions

Our research findings suggest that males and females have different preferences for BCTs. If users use the "right" fitness app which adopts their preferred BCTs, users tend to engage more with this app. This gender effect is significant for the female app, which indicates a better performance of the female app to enhance female user engagement than male user engagement. The reason why gender fit effect is not significant for the male app may be due to the ceiling effect. The engagement level with the male app is relatively high, making it difficult to facilitate significant increase of engagement level via matched BCTs.

Our research makes several initial theoretical contributions that deserve a deep investigation in future research. First, we make a complement to the BCT research by identifying gender differences in preferences for BCTs adopted by fitness apps using objective app usage data. To our best acknowledge, this research is the first to investigate gender differences in preferences for BCTs implemented in fitness apps in the context of real-life environment. A better understanding of which BCTs fit which users may shed light on designing personalized health behavior change interventions. In addition to gender differences, we call for future investigations on other characteristics of users.

Second, we observe a gender fit effect between user gender and app gender on user engagement with fitness apps. Gender is a personal trait that can be easily identified or predicted (Hu, Zeng, Li, Niu, & Chen, 2007). As most existing research on psychological fit emphasizes self-reported psychological traits, our investigation on gender preferences can be more easily applied in practice without the need of user participation. As an exploratory work, we only target at nine fitness apps that cover more than 93% of our sample records, and two typical apps that are distinct in male- and female-preferred BCTs. Our future research is to expand our sample and test the generalizability robustness of our findings.

For practice, our research may benefit both the users and fitness app developers. For app users, our research suggests them to choose the fit mobile fitness app from hundreds in the market that contains their preferred BCTs to maximize their engagement in future. For app developers, our research suggests them to provide different BCTs for female and male users in order to sustain user engagement. We have identified some BCTs that are more preferred by males and females for both of them to make decisions.

## References


Brouwer, W., Kroeze, W., Crutzen, R., de Nooijer, J., de Vries, N. K., Brug, J., & Oenema, A. (2011). Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. *Journal of Medical Internet Research, 13*(1).

Buchan, N. R., Croson, R. T., & Solnick, S. (2008). Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization, 68*(3-4), 466-476.

Cesario, J., Higgins, E. T., & Scholer, A. A. (2008). Regulatory fit and persuasion: Basic principles and remaining questions. *Social and Personality Psychology Compass, 2*(1), 444-463.

Conroy, D. E., Yang, C.-H., & Maher, J. P. (2014). Behavior change techniques in top-ranked mobile apps for physical activity. *American Journal of Preventive Medicine, 46*(6), 649-652.

Dubois, D., Rucker, D. D., & Galinsky, A. D. (2016). Dynamics of communicator and audience power: The persuasiveness of competence versus warmth. *Journal of Consumer Research, 43*(1), 68-85.

Ernsting, C., Dombrowski, S. U., Oedekoven, M., & LO, J. (2017). Using smartphones and health apps to change and manage health behaviors: a population-based survey. *Journal of Medical Internet Research, 19*(4).

Fox, S., & Duggan, M. (2012). Mobile health 2012. Retrieved from http://www.pewinternet.org/2012/11/08/mobile-health-2012/

Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the technology acceptance model. *MIS Quarterly*, 389-400.

Grubb, E. L., & Grathwohl, H. L. (1967). Consumer self-concept, symbolism and market behavior: A theoretical approach. *The Journal of Marketing*, 22-27.

Guimond, S., Chatard, A., Martinot, D., Crisp, R. J., & Redersdorff, S. (2006). Social comparison, self-stereotyping, and gender differences in self-construals. *Journal of Personality and Social Psychology, 90*(2), 221.

Hu, J., Zeng, H.J., Li, H., Niu, C., & Chen, Z. (2007). Demographic prediction based on user's browsing behavior. In *Proceedings of the 16th international conference on World Wide Web* (pp. 151-160). ACM.

Jokela, M., Bleidorn, W., Lamb, M. E., Gosling, S. D., & Rentfrow, P. J. (2015). Geographically varying associations between personality and life satisfaction in the London metropolitan area. *Proceedings of the National Academy of Sciences, 112*(3), 725-730.

Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). Measuring User Engagement (Synthesis Lectures on Information Concepts, Retrieval, and Services). *San Rafael: Morgan & Claypool Publishers*.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., et al. (2009). Life in the network: the coming age of computational social science. *Science, 323*(5915), 721-723.

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., Katzmarzyk, P. T., & Group, L. P. A. S. W. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet, 380*(9838), 219-229.

Matz, S., Kosinski, M., Nave, G., & Stillwell, D. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the National Academy of Sciences*, 201710966.

Matz, S. C., Gladstone, J. J., & Stillwell, D. (2016). Money buys happiness when spending fits our personality. *Psychological Science, 27*(5), 715-725.

Meyers-Levy, J., & Sternthal, B. (1991). Gender differences in the use of message cues and judgments. *Journal of Marketing Research*, 84-96.

Meyers-Levy, J., & Loken, B. (2015). Revisiting gender differences: What we know and what lies ahead. *Journal of Consumer Psychology, 25*(1), 129-149.

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine*, 46(1), 81-95.

Paffenbarger Jr, R. S., & Hyde, R. T. (1984). Exercise in the prevention of coronary heart disease. *Preventive Medicine, 13*(1), 3-22.

Peterson, M. S., Lawman, H. G., Wilson, D. K., Fairchild, A., & Van Horn, M. L. (2013). The association of self-efficacy and parent social support on physical activity in male and female adolescents. *Health Psychology*, 32(6), 666-674.

Serrano, K. J., Coa, K. I., Yu, M., Wolff-Hughes, D. L., & Atienza, A. A. (2017). Characterizing user engagement with health app data: a data mining approach. *Translational Behavioral Medicine, 7*(2), 277-285.

Serrano, K. J., Yu, M., Coa, K. I., Collins, L. M., & Atienza, A. A. (2016). Mining health app data to find more and less successful weight loss subgroups. *Journal of Medical Internet Research, 18*(6).

Siscovick, D. S., Laporte, R. E., & Newman, J. M. (1985). The disease-specific benefits and risks of physical activity and exercise. *Public Health Reports, 100*(2), 180.

Statista. (2017). Which of these health app types have you used in the past 12 months? Retrieved 2018.04.27 https://participatorymedicine.org/wp-content/uploads/2016/04/Accenture-Survey-HIMSS-2016.pdf

Weiser, E. B. (2000). Gender differences in Internet use patterns and Internet application preferences: A two-sample comparison. *CyberPsychology and Behavior, 3*(2), 167-178.

Wright, M. S., Wilson, D. K., Griffin, S., & Evans, A. (2008). A qualitative study of parental modeling and social support for physical activity in underserved adolescents. *Health Education Research*, 25(2), 224-232.

Xu, R., Frey, R. M., & Ilic, A. (2016). Individual differences and mobile service adoption: an empirical analysis.In *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)* (pp. 234-243). IEEE.

Yang, C.-H., Maher, J. P., & Conroy, D. E. (2015). Implementation of behavior change techniques in mobile applications for physical activity. *American Journal of Preventive Medicine, 48*(4), 452-455.

Yardley, L., Morrison, L., Bradbury, K., & Muller, I. (2015). The person-based approach to intervention development: application to digital health-related behavior change interventions. *Journal of Medical Internet Research, 17*(1).