Telematics and Informatics 59 (2021) 101552

Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com/locate/tele

Systematic evaluation of mobile fitness apps: Apps as the Tutor, Recorder, Game Companion, and Cheerleader

Yunwen Wang [a,*], William B. Collins [b]

[a] University of Southern California, Annenberg School for Communication and Journalism, Los Angeles, CA, USA
[b] Purdue University, Brian Lamb School of Communication, West Lafayette, IN, USA

A R T I C L E I N F O

Keywords:
Mobile app
mHealth
User experience
Technology adoption
Gamification
Tailored design

A B S T R A C T

Mobile fitness applications are innovating the ways in which smartphone users self-manage their health. Prior research found that app functions may impact app efficacy. However, research to date has not sought to systematically investigate how different combinations of app functions impact user response to apps, especially adoption intent. This article describes two studies on mobile fitness app characteristics and user attitudes. Study One used content analysis and hierarchical cluster analysis on 98 iPhone fitness apps and identified four app clusters: "Tutor", "Recorder", "Game Companion", and "Cheerleader." Tracking was the predominant function in current market, but tracking-focused Recorder apps received lowest user ratings among all app clusters. Users favored Tutor apps that combine exercise education and tracking, and Game Companion apps that combine gamification, tracking, and social functions. Function combinations, rather than standalone functions, impact app success. Following a Reasoned Action Approach, Study Two found various effects of individual differences (age, gender, BMI, eHealth literacy, smartphone experience, function preference) on user attitude toward different fitness app types. A comparison between two studies demonstrated a mismatch between market offerings and user needs regarding app functions. Implications of results for mobile fitness app design to improve consumer health and for theories are discussed.

1. Introduction

Since the introduction of mobile phones in the 1970s and smartphones in the 1990s, ownership of mobile devices has exploded worldwide. The trend in the U.S. has been especially rapid. In 2019, 96% of U.S. adults owned a mobile phone (i.e. a cellphone), among which 81% were smartphones (Pew Research Center, 2019). The prevalence of mobile phones promises to leverage mobile health technologies (mHealth) to enhance public health. The U.S. government's Healthy People 2020 initiative, in this context, sought to "use health communication strategies and health information technology to improve population health outcomes and health care quality, and to achieve health equity" by providing all people "personalized self-management tools and resources" (Healthy People 2020, n.d.-a).

Mobile device-based health apps, particularly, are revolutionizing the ways in which smartphone users self-manage their health. In 2020 first quarter, the numbers of iOS and Android mHealth apps are respectively 45,478 and 43,285 (Statista, 2020a, 2020b). These apps aim at disease management (e.g., diabetes, Goyal and Cafazzo, 2013) and prevention (e.g., promoting healthy food and physical

* Corresponding author.
E-mail address: yunwenwa@usc.edu (Y. Wang).

https://doi.org/10.1016/j.tele.2020.101552
Received 16 July 2020; Received in revised form 23 October 2020; Accepted 7 December 2020
Available online 14 December 2020

activities, Fanning et al., 2012). Previous research has evaluated the mobile app market (e.g. Azar et al., 2013; Sama et al., 2014; Vickey et al., 2012), but findings have been mostly descriptive, and they do not link measures to users' behavioral response to app features or compare market offerings with consumer needs. The fast-paced changes in the mobile environment require updated research to survey user response to mobile health apps, e.g. mobile fitness apps, to evaluate and optimize app efficacy.

Mobile fitness apps are intended to support physical activity. For app users, the support can be instrumental (problem-solving knowledge and skills) and emotional (encouragement), namely informational and emotional-social support (Peng et al., 2016). Despite numerous health benefits (e.g. preventing cardiovascular diseases and diabetes, Reiner et al., 2013), mobilizing physical activity nationally has been unattainable. From 1996 to 2017, the percentage of U.S. adults who do not engage in an adequate level of physical activity *increased* from 60% to 80% (U.S. CDC, 1996; Healthy People, n.d.-b, 2020). While lack of motivation is a major barrier to physical activity (Coll et al., 2017; Sharifi et al., 2013; Trost et al., 2002), mobile motivational tools, when tailored to user needs, may help build exercise habits.

This paper reports two studies on mobile fitness app characteristics and user attitudes. Study One used content analysis and hierarchical cluster analysis to identify fitness app features and prototypes. Study Two used the Reasoned Action Approach (Fishbein and Ajzen, 2010) to examine how individual differences, especially app function preference, influence attitudes toward various fitness apps and adoption intent. The paper also compares the current market (Study One) with consumer needs (Study Two). Results provide implications for mHealth literature as well as the design of mobile fitness app interventions to impact consumer health.

2. Study One

2.1. Literature review and hypotheses development

Previous research in mobile fitness apps has assessed app efficacy and user feedback. For instance, Kranz et al. (2013) examined 15 fitness apps that perform GPS tracking, workout planning, and workout performance assessment. The study highlighted user-app interaction through playful, social, and long-term motivation. Meanwhile, interactive and gamification designs were observed in health and wellness apps. Popular gamification features include badges, leaderboards, points, levels, challenges, quest, and social engagement (Miller et al., 2016). These social and gamification designs satisfy users' gratification need, which is the desire of people to stay connected, enjoy life, and have fun (Lim et al., 2015). Gratification is an important theoretical base for app design (Leung and Wei, 2000; Yang, 2013). Users adhere to a fitness app more if they find it fun, interesting, trendy, cool (Lee and Cho, 2016).

Besides social and gamification functions, literature discussed education, tracking, and motivation functions in mHealth apps. Mohr et al. (2014) identified education, goal-setting, monitoring, feedback, and motivation enhancement as behavioral change strategies. They define education as providing requisite knowledge to enable a behavioral change, and monitoring/tracking as recording past and current states. In online physical education, mobile fitness apps host instructional resources such as demonstration videos (Hung et al., 2018), connecting students with physical educators through assignments and assessments (Goad et al., 2019). The array of educational features in mobile fitness apps often occur simultaneously with monitoring functions, which track users' exercise and biometric data over time to inform personalized feedback (Papastergiou et al., 2020). Heart rate, GPS traces, acceleration data are example sensor data (Kranz et al., 2013; Papastergiou et al., 2020). Apart from education, mobile fitness apps also address motivation enhancement. They use positive reinforcement, contingent rewards, behavioral contracts, incentives, and social support to promote physical activity (Mohr et al., 2014). Kranz et al. (2013) believe that when an app diversifies training experience and provides social incentives and feedback on progress, it can help maintain extrinsic motivation and engagement. In comparison, the gamification app designs mentioned earlier tackle intrinsic motivation (Farrow et al., 2019; Xi and Hamari, 2019).

Based on existing literature, we conceptually categorize mobile fitness app functions into five themes: Education (Goad et al., 2019; Mohr et al., 2014; Papastergiou et al., 2020), Tracking (Kranz et al., 2013; Mohr et al., 2014), Social (Kranz et al., 2013; Lee and Cho, 2016), Gamification (Farrow et al., 2019; Lee and Cho, 2016; Miller et al., 2016), and Motivation (Mohr et al., 2014). Some apps are minimalist; others are complex. Despite a possible negative correlation between app function variety and expert-reported usability (Arnhold et al., 2014), most mobile fitness apps in practice combine several functions that app developers believe will provide an optimal combination of capabilities for a subset of the market. Previous studies (e.g. Sama et al., 2014) only examined apps' functional tools individually without systematically evaluating their collective, interactive outcomes. The rapid change in app development requires an updated investigation to answer the following two questions.

RQ1: What major functions and features do current mobile fitness apps have?
RQ2: What are the common combinations of functions in current mobile fitness apps?

This study also aims to test the link between user outcomes and specific app functions or combination of app functions. Previous research on apps for the care/prevention of STIs (Muessig et al., 2013) has measured app outcomes by number of downloads (for app popularity) and user rating (for satisfaction), but rarely are correlations explored between these outcomes and app functions or functional combinations. Mendiola et al. (2015) do explore the relationship between app features and ratings. Coding 12 predetermined features in 234 medial, health and fitness apps, they find tracking features lower app ratings – though popular. Their study reveals links between app features and outcomes but presents limitations. First, it contains apps of three topics (medial, health, and fitness) in one regression analysis, disregarding topic-specific data variances. Second, the 12 app features are non-paralleled concepts; their typology appears non-exhaustive and not mutually exclusive. To address gaps in research, this study seeks to create an improved feature typology specific to mobile fitness apps. We ask two additional questions.

RQ3a: How does variation in functions impact app popularity?
RQ3b: How does variation in functions impact user satisfaction toward the app?

2.2. Method of Study One

2.2.1. Sample and coding scheme development

Prior to sampling, an exploratory search in iPhone app store (iTunes version: 12.4.1.6) used 12 keywords (i.e., fitness, exercise, exercise/fitness/walk/running/cycling tracking, exercise/fitness/walk/running/cycling monitor) to generate 464 fitness apps solely for developing a coding scheme. The coding scheme has an outline of the five app functional themes: education, tracking, social, gamification, and motivation (see Section 2.1). An iterative coding approach on the 464 apps identified specific features under each functional theme. When no new features appeared, codes reached saturation. Afterwards, another search used two search terms ("fitness", "exercise") to improve relevancy of apps solicited. Apps designed by a same developer but only varied in price (free or paid) were considered different apps in iTunes store; they had to have different features to be both sampled. If a paid app has identical features as its free version, only different as ad-free, one of them was randomly excluded from the sample. The sample excluded duplicate and off-topic apps, retaining 98 apps for content analysis.

2.2.2. Measures

App popularity. We used the number of user reviews of an app as a proxy measure of app popularity because Apple does not provide a direct indictor of downloads. The presumption was that apps with more ratings were likely downloaded more.

App satisfaction/ratings. App ratings from iTunes served as the proxy measure of app satisfaction. Any iPhone user can rate an app after downloads but not all users would rate apps.

Functional theme scores. We counted the absence and presence of app sub-features using 0 and 1. A theme score is the total number of sub-features present in an app on each of the five functional themes. For example, theme Gamification had five sub-features such as leaderboards, role-playing avatar. Apps thus range 0 to 5 on Gamification theme scores, 0 to 9 on Education, 0 to 8 on Tracking, 0 to 3 on Social, 0 to 5 on Gamification, and 0 to 5 on Motivation. The differences in scale magnitudes require data standardization. We calculated z scores for each of the five theme scores, using z scores in subsequent clustering and regression procedures.

2.3. Results of Study One

2.3.1. Content analysis

RQ1: What major functions and features do current mobile fitness apps have?

Five app functional themes served as high-level codes. Each theme contains several sub-features. For example, theme Social consists of three sub-features: (1) allowing users to connect with real-life networks like families and friends, (2) having an in-app community of users, and (3) enabling sharing of workout accomplishments to social media through the fitness app. Preliminary analysis determined: Tracking (84.7%) was the most frequent functional theme, followed by Education (71.4%), Social (59.2%), Motivation (38.8%), and Gamification (35.7%). See Table 2 column "market" for feature frequencies per theme.

One author coded all apps; results entered subsequent analyses. To assess reliability, a second coder coded a random subset of the sample, 22 (22.4%) apps. We calculated Krippendorff's alpha at the functional theme level comparing the aggregated numbers of features coders each identified. Reliability reached tentative to good ($\geq$0.667) for four themes: education ($\alpha = 0.719$), social ($\alpha = 0.854$), gamification ($\alpha = 0.809$), and motivation ($\alpha = 0.669$), except for tracking theme ($\alpha = 0.475$). After retraining, the second coder coded different 21 (21.4%) apps on tracking features; alpha increased to 0.721.

2.3.2. Cluster analysis

RQ2: What are the common combinations of functions in current mobile fitness apps?

To answer RQ2, we performed a cluster analysis. Cluster analysis is a method of identifying the category structure or membership of a collection of observations (Anderberg, 2014). Observations, which can be individuals or objects, are assigned to homogeneous groups called clusters (Norusis, 2010). Within a cluster, cases are conceptually similar to each other, but a case from one cluster should differ from cases assigned to any other clusters (Duran and Odell, 2013). Among several clustering approaches, hierarchical cluster analysis is appropriate in this study given the non-predetermined number of clusters and a relatively small sample size (Norusis, 2010). We chose an agglomeration schedule (Yim and Ramdeen, 2015) for hierarchical cluster analysis in IBM SPSS Statistics 27 to identify natural grouping of mobile fitness apps.

The clustering criteria were the five functional theme scores, which signify the richness of feature affordance on Education, Tracking, Social, Gamification, and Motivation respectively. See Section 2.2.2. for the measure. We used z-score standardization in cluster analysis (Mooi and Sarstedt, 2011). The scores are ratio variables with true zeroes, meaning no sub-feature of a given function is observed in an app. Thus, we selected interval measure and squared Euclidean distance following the Ward's method to cluster standardized data (Murtagh and Legendre, 2014).

A dendrogram was generated to suggest clustering at different levels of abstraction. We closely examined five clustering solutions each containing two, three, four, five, or six clusters. Among all solutions, the four-cluster solution had the most even case numbers across clusters (34.69%, 24.49%, 15.31%, 25.51%), and relatively more distinct clusters. See Fig. 1. A bar of a positive or near-zero z score means that apps in that cluster include features of the corresponding functional theme in a rich or average amount.

We further validated the four-cluster solution through a one-way multivariate analysis of variance (MANOVA) by confirming clusters distinctiveness. There was a statistically significant difference in functional theme scores per cluster membership, $F(15, 248.852) = 30.670$, $p < .0001$; Wilk's $\Lambda = 0.056$, partial $\eta^2 = 0.618$. Variances were detected among clusters for each theme score: Education, $F(3, 94) = 42.675$, $p < .0001$; Tracking, $F(3, 94) = 18.697$, $p < .0001$; Social, $F(3, 94) = 7.747$, $p < .0001$; Gamification, $F(3, 94) = 38.198$, $p < .0001$; Motivation, $F(3, 94) = 22.305$, $p < .0001$. See Table 1 for post-hoc pairwise cluster comparisons.

Based on the unique function combinations of each cluster, we labeled the four clusters with meaningful personified names. Cluster 1, named **Tutor** apps, are rich in education features but mix education with social, tracking, and/or gamification. They typically do not have motivational features such as music or voice cheer. Example Tutor apps include *Fitness Trainer - Exercise & Workout Guide* and *8fit – Workouts*. Cluster 2, **Recorder** apps, are single-purpose tracking apps lacking other functions especially Education. Representative Recorder apps are *Fitbod Planner*, *Metric Me*, and *Tap & Track* are. Cluster 3, **Game Companion** apps, are mixtures of gamification, tracking, and social functions – occasional motivational features – but not education. Examples are *Fitbit* and *Nike + Running*. Cluster 4, **Cheerleader** apps, are heavy in motivation, with occasional social tools but below average in tracking, education, and gamification. Representative apps are *Interval Timer*, *myWOD* and *Relax Melodies*.

2.3.3. Functional themes and clusters on app popularity and satisfaction

RQ3a: How does variation in functions impact app popularity?

A multiple regression tested whether functional theme scores impact app popularity. There was no significant overall effect, Multiple $R = 0.247$, $F(5, 92) = 1.201$, $p = .315$, indicating that the presence or absence of functional themes (Education, Tracking, Social, Gamification, Motivation) did not predict app popularity using number of reviews as the proxy measure. Additionally, a one-way analysis of variance (ANOVA) did not detect significant differences on number of reviews among four app clusters (Tutor, Recorder, Game Companion, Cheerleader), $F(3, 94) = 0.86$, $p = .47$. Cluster membership was unrelated to app popularity.

RQ3b: How does variation in functions impact user satisfaction toward the app?

We examined the relationship between app functions and user satisfaction. A multiple regression, using the five theme scores as predictors of app satisfaction, yielded an overall significant effect, Multiple $R = 0.373$, $F(5, 92) = 2.97$, $p < .05$. Two of the five themes were significant predictors of user satisfaction: Education, $\beta = 0.31$, $p < .01$; Social, $\beta = 0.21$, $p < .05$. Mobile fitness apps with Educational/Social features received higher ratings.

A one-way ANOVA compared app satisfaction among four app clusters. It found an overall between-group effect, $F(3, 94) = 8.51$, $p < .001$, $\eta^2 = 0.213$. Post-hoc Tukey's HSD found that **Recorder** apps ($M = 3.89$, $SD = 0.92$) were rated significantly lower than **Tutor** apps ($M = 4.51$, $SD = 0.26$, $p < .001$), **Game Companion** apps ($M = 4.42$, $SD = 0.23$, $p = < 0.05$), and **Cheerleader** apps ($M = 4.35$, $SD = 0.24$, $p < .001$). Tutor, Game Companion, and Cheerleader Apps were not significantly different from each other in terms of app ratings.

2.4. Discussion of Study One

Among five functions, tracking was predominant in current mobile fitness apps, whereas gamification had minimal presence (tracking > education > social > motivation > gamification). Apps with gamification feature (35.7%) were fewer than half of apps with tracking features (84.7%). However, tracking-only Recorder apps received lower user ratings than other app clusters which prioritize education (Tutor apps), social (Cheerleader apps), or a combination of tracking, social, and gamification (Game Companion apps). Arnhold et al. (2014) found, "documentation" functions negatively impact app usability. What we found was more complex: it

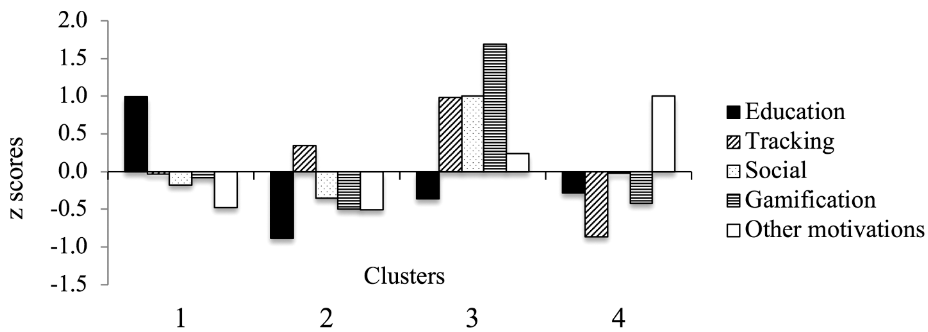

Fig. 1. Cluster profiles for the four-cluster solution.

Table 1
Cluster Means, Standard Deviations, and z Scores for the Four-Cluster Solution.

	Cluster 1 (n = 34) "Tutor"			Cluster 2 (n = 24) "Recorder"			Cluster 3 (n = 15) "Game Companion"			Cluster 4 (n = 25) "Cheerleader"		
	M	SD	z	M	SD	z	M	SD	z	M	SD	z
Clustering Variables												
1. Education (scale: 0–9)	3.71	0.906	0.990_a	0.79	0.833	-0.880_b	1.60	1.121	$-0.392_{b,\ c}$	1.275	2.16	-0.285_c
2. Tracking (scale: 0–8)	2.32	1.364	-0.034_a	2.92	0.881	$0.341_{a,\ b}$	3.93	1.668	0.983_b	1.00	1.190	-0.870_c
3. Social (scale: 0–3)	0.79	0.946	-0.183_a	0.63	0.711	-0.349_a	2.00	0.845	1.005_b	0.96	1.098	-0.019_a
4. Gamification (scale: 0–5)	0.53	0.788	-0.084_a	0.13	0.338	-0.497_a	2.27	1.033	1.687_b	0.20	0.408	-0.420_a
5. Other motivation (scale: 0–5)	0.15	0.359	-0.483_a	0.13	0.338	-0.511_a	0.73	1.033	0.236_b	1.36	0.810	1.006_c

Note. Subscripts, as a form of grouping labels, symbolize the results of post hoc pairwise differences for each cluster within each theme score. The group with a subscript "*a*" is statistically distinct from that with "*b*" or "*c*." However, the group with a subscript "*b*" is statistically not distinct from that with "*b, c*" since they share the subscript "*b*." Differed grouping labels indicate the heterogeneity across app clusters.

was the combinations of app functions – not standalone function – that impact app satisfaction.

Namely, tracking features have to coexist with other functional themes in a meaningful combination to achieve efficacy. One valuable functionality addition is gamification: users responded favorably to Game Companion apps. Gamification applies game-like features to the design of systems in traditionally non-game contexts (Huotari and Hamari, 2017) and has demonstrated some efficacy in promoting physical activity (Koivisto and Hamari, 2019).

Study One is not immune to limitations. First, controlling for the effect of search results ranking on app sampling was not possible because Apple keeps secret their ranking algorithm. A search simulation found that apps that updated more often would appear more front among search results (Lim and Bentley, 2013). Future research can explore ways to sample more representatively. Second, we only sampled iOS apps. Android and iPhone users may have fundamental differences that influence iTunes and Google Play Store. Given our focus on investigating app functional combinations as opposed to individual app functions and multiple research questions, we chose to focus on iTunes store for the present research. Future research may sample both iOS and Android apps and compare the two platforms.

3. Study Two

Study One found that users responded differently to various combinations of app functions, and we identified four typical app clusters: Tutor, Recorder, Game Companion, Cheerleader. These clusters, though very different in feature sets, tend to be similar in popularity and in user satisfaction ratings, with an average app store rating of 3.87/5. This led us to wonder the different user response to apps at an individual level. Study Two investigated: (1) whether the market meets user needs by comparing Studies One (current market) and Two (user needs); (2) factors associated with the intent to adopt different types of fitness apps.

3.1. The reasoned action approach

Systematic reviews have found numerous effective mHealth apps in interventions, on issues like diabetes, obesity, cancer, chronic obstructive pulmonary disease, and mental health (Donker et al., 2013; Wang et al., 2014). Users have diverse app choices, but little is known about their decision-making considering app functions or functional combinations. Proxy measures for app popularity and satisfaction from iTunes are insufficient for studying individual variances. We designed a user survey to address this gap in research.

Study Two drew upon the Reasoned Action Approach (RRA, Fishbein and Ajzen, 2010). RAA is a synthesized theory set from historical discussions on the Theory of Reasoned Action (TRA), Integrative Model of Behavioral Prediction (IMBP), and the Theory of Planned Behavior (TPB, Ajzen, 1991). As an expectancy-value theory, RAA posits that individuals make mental math to weigh the benefits and costs of incurring a behavioral change. If benefits are perceived to overweigh costs of performing a behavior, a change is more likely to happen. Such decision-making involves physical, emotional, social, and financial factors (DiClemente et al., 2013).

Behavioral intent. Among determinants that facilitate or hinder behavioral changes, behavioral intent is the most immediate predictor (Ajzen, 1991). To observe behaviors is a challenge; behavior intent is often a proxy measure and strong predictor for behavior (Gibbons, 2006). This study examined how individual differences influence intent to adopt mobile fitness apps, which made RAA an appropriate framework. For a given behavior, "attitude," "subjective norm," and "perceived behavioral control" are three immediate antecedents to behavioral intent (Fishbein and Ajzen, 2010).

Attitudes are psychological tendencies that one expresses upon evaluating an entity with some degree of favor or disfavor (Eagly and Chaiken, 1993). Though one's attitudes are generally consistent with behaviors, stronger attitudes are better predictors of behaviors (Holland et al., 2002). In the context of mobile fitness apps, attitudes toward mobile fitness apps – in general or toward particular functions/functional combinations – may influence adoption intent. However, we do not know whether consumers hold strong or weak attitudes on this topic, and whether app attitudes indeed predict adoption intent.

Subjective norms are perceived norms. They describe the extent one considers a behavior as expected by society (Fishbein, 2000). Unlike "social norms" (Cialdini et al., 1990), subjective norms underscore the subjective perception. They influence behavioral intent and behaviors (Cialdini et al., 1990; Fishbein, 2000), including health-related behaviors (Rimal and Real, 2005) and technology adoption (Teo et al., 2012). One may perceive adopting fitness apps as a social norm, but the degree of adoption intent is dependent on the degree of subjective norm.

Perceived behavioral control (PBC) is the extent to which one feels control over a given behavior, or capable of performing it (Fishbein and Ajzen, 2010). Users' incapability of downloading a fitness app to their smart phone can be a constraint of adoption.

RAA also posits, background factors at individual, social, and information levels influence behavioral, normative, and control beliefs, each of which respectively predicts behavior-related attitude, subjective norm, and perceived behavioral control. This study adapted RAA to mobile fitness app adoption. The first hypothesis is based on RAA.

H1: Consumers' mobile fitness app adoption intent is predicted by their attitudes toward the app, subjective norms of being expected to adopt a fitness app, and perceived behavioral control over app adoption.

Studies consistently indicated that "perceived value" especially "perceived usefulness" influences technology adoption intent (Teo et al., 2012). Consumers seek to maximize value by choosing behaviors that will lead to the most optimistic outcomes (Kim et al., 2007). This underlies the assumption of all value-based behavioral theories: human behaviors are the tradeoff of a complex evaluation of benefits and barriers. Mobile social apps, for instance, have three dimensions of perceived value: utilitarian motivation (perceived usefulness), hedonic motivation (perceived enjoyment), and social influence (social ties) (Hsiao et al., 2016). As Study One showed, these dimensions were also common in mobile fitness apps. With various function combinations, fitness apps carry perceived value with various emphases. A study on middle-aged individuals in Taiwan (Liao et al., 2017) found that fitness apps did not entail functions that met consumer demands. Study Two started with updated research on the same vein (RQ4, RQ5).

RQ4: What types of fitness apps are most appealing to consumers considering features?
RQ5: Do the types of fitness apps that consumers prefer align with the four app prototypes (Tutor, Recorder, Game Companion, Cheerleader) in current iPhone app market?

Study One found that fitness apps with certain functional combinations received better consumer evaluation, judging app store ratings. However, it remained unknown whether there exists an optimum app prototype that can appeal to the majority of consumers. Alternatively, different segments of consumers may have varied attitudes toward different types of mobile fitness apps. As each type of fitness apps entail unique feature affordances in bundles, consumer attitudes about each type of apps may be influenced by their app feature/function preference. Again, one of our main objectives is to systematically evaluate and understand the combinations of app functionality – as opposed to standalone functions. We thus tested in Study Two the associations between individual differences and consumer attitudes toward fitness apps in general, and toward the four app clusters identified in Study One (Tutor, Recorder, Game Companion, Cheerleader). Each of these four app clusters afford unique combinations of functional features on five themes: Education, Tracking, Social, Gamification, and Motivation (see Fig. 1 and Table 1). According to RAA, background factors predict behavioral attitudes (Fishbein and Ajzen, 2010). In mHealth, individual differences such as age, gender, Body Mass Index (BMI, weight-height ratio), health status, and current exercise behaviors are popular variables because they are closely tied to health behaviors (Ernsting et al., 2017; Deng et al., 2014; Leonard et al., 2013; Shin et al., 2001). A previous study using RAA framework, for example, found an age effect on the relationship between mHealth services adoption intent and related attitudes, perceived value, and perceived behavioral control (Deng et al., 2014). A German survey ($n = 4{,}144$) additionally reported associations between more health app usage and more Internet use, exercising, chronic diseases, higher health literacy and a low-fat diet (Ernsting et al., 2017). Namely, demographic variables and technology proficiency seem to underlie mHealth experiences. We speculated, however, a relevance of eHealth literacy, rather than health literacy, to mHealth adoption. To test this speculation and expand current theoretical discussion, we asked whether preference of fitness app functional features strongly predict user attitudes toward each app cluster, controlling for two sets of individual background factors (RQ6). See Fig. 2 for the research model of Study Two.

RQ6: Controlling for the effects of demographic and mHealth variables, does users' app function preference predict their attitudes toward the four mobile fitness app clusters (Tutor, Recorder, Game Companion, Cheerleader) and toward mobile fitness apps in general? If so, how?

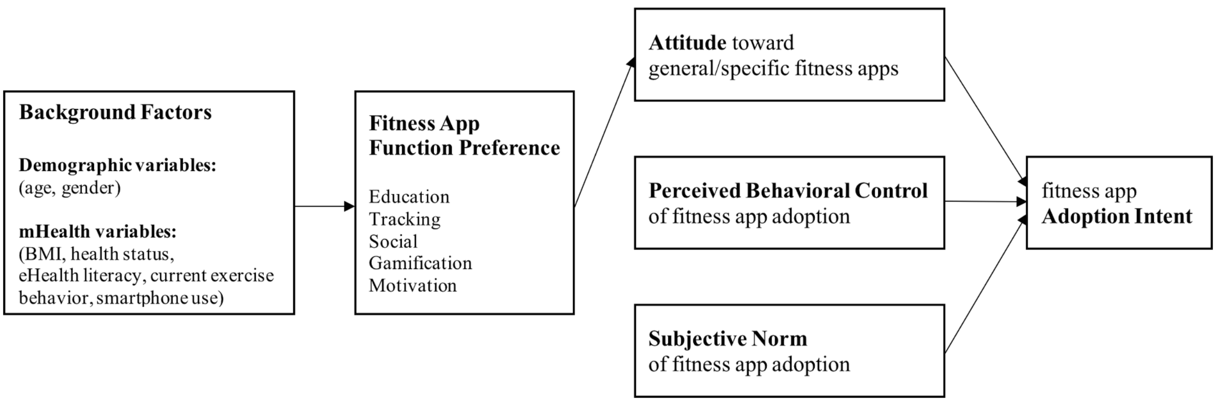

Fig. 2. Research model.

3.2. Method of Study Two

3.2.1. Sample

Participants in this study were 351 Amazon Mechanical Turk (MTurk) Workers (48.72% female, 49.86% male, 1.14% transgender, 0.28% other gender). The mean age was 37.90 ($SD = 11.72$, $Min = 19$, $Max = 72$). Participants self-identified as Native American (1.42%), Hispanic/Latino (3.42%), Asian (5.41%), African American (5.70%), White/Caucasian (76.92%), and interethnic (7.12%). Majority (95.70%) reported English as their primary language. Participants reported their highest level of education as four-year degrees or Bachelor's degrees (37.6%), some college education but no degree (23.9%), two-year degrees or Associate's degrees (15.7%), graduate or professional degrees (10.8%), high school graduates or equivalency (10.5%), and doctorate (1.1%). The sample were 86.45% of all 406 participants who finished the survey; responses from 55 (13.55%) participants who failed the instructional manipulation check (IMC) or attention check were eliminated. The IMC was adapted from Oppenheimer et al. (2009). Participants were instructed to select a designated, fake smartphone operating system from a list of real smartphone operating systems. See Fig. 3 for our attention check question tailored to smartphone studies.

3.2.2. Procedure

The study was approved by the Institutional Review Board. A pilot study was conducted to ensure data integrity and clarity. Afterwards, a refined questionnaire was posted on MTurk at different times in two days to reach more diverse Workers. The compensation was one U.S. dollar. Median survey completion time was 12.22 mins ($M = 14.81$, $SD = 10.71$). Based on Study One findings, Study Two manipulated app descriptions and screenshots of one general fitness app and four fitness app archetypes (Tutor, Recorder, Game Companion, Cheerleader) to survey consumer adoption intent and feature preference. We provided visual representations of the apps to assist the textual descriptions in the online questionnaire. See Fig. 4.

3.2.3. Measures

BMI. Height and weight were surveyed to calculate Body Mass Index (BMI). Average BMI in this study was 26.87 ($SD = 6.64$), which was close to U.S. CDC's 2010 census data (26.6 for the average adult man and 26.5 for the average adult woman) (U.S. CDC, Centers for Disease Control and Prevention, n.d.).

Smartphone use. We measured smartphone use by frequency and history. Most survey participants (93%) own a smartphone.

Smartphones rely on operating systems (OS) and developed software to supply user needs. Individual and environmental factors can impact the decision process on which OS to adopt. This study is interested in knowing certain factors about you, the decision maker. Specifically, we are interested in whether you actually take the time to read the directions; if not, results to some of our questions that rely on instructions can be invalid. To confirm that you have read the instructions, please answer only "UltraTino" on the question below. Thank you very much.

What operating systems have you ever used? (select all that apply)

- ☐ Apple iOS
- ☐ Samsung Bata
- ☐ Google Android
- ☐ RIM BlackBerry
- ☐ Symbian
- ☐ Windows Mobile
- ☐ Window 7 Phone
- ☐ HP Palm OS (Garnet OS)
- ☐ Open WebOS
- ☐ UltraTino
- ☐ Maemo
- ☐ MeeGo
- ☐ Verdict
- ☐ Others:

Fig. 3. Adapted instructional manipulation check (IMC), or attention check.

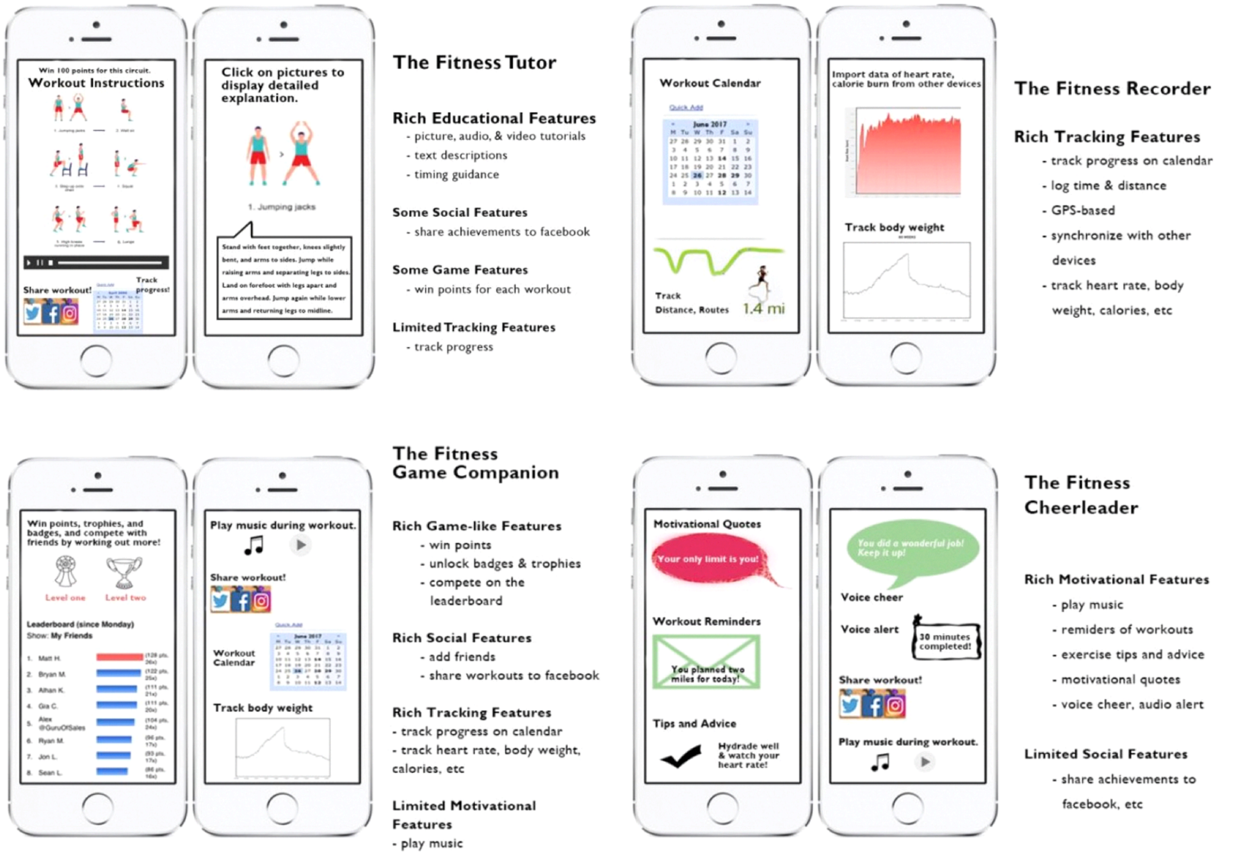

Fig. 4. Survey manipulations: screen shots of four hypothetical apps.

Majority use smartphones multiple times a day (74.4%) or at least daily (16.2%). Smartphone use history ranges 0–18 years ($M = 6.57$, $SD = 3.46$). About 31.3% participants have never used mobile apps to support exercise, physical activities or fitness, while two-thirds (68.7%) have. An average participant has downloaded three ($SD = 4.67$) fitness apps. Main devices participants used to access the Internet were laptops (50.4%), desktop computers (38.2%), and mobile phones or tablets (11.4%).

eHealth literacy. The eHealth Scale, or eHEALS, consists eight items that measure people's electronic literacy level for health purposes: navigating resources, skills to use resources, usefulness for oneself, and critical evaluation (Cameron and Harvey, 2006; Richtering et al., 2017). One sample item is "*I know what health resources are available on the Internet*" on a 5-point Likert-type scale, 1 = strongly disagree and 5 = strongly agree. Scores of all items were averaged at individual level to produce eHealth literacy scores ($M = 4.37$, $SD = 0.53$). On this dataset, eHealth Scale was internally reliable ($\alpha = 0.89$).

App function preference index. A checkbox question "*Create Your App*" instructed participants to "*select just the features you think you would use. (Check all that apply)*" from a randomized list of 28 mobile fitness app features that Study One identified. The number of features selected under each of the five functional themes (Education/ Tracking/ Social/ Gamification/ Motivation) was respectively summed to yield five function preference indexes.

Intent, attitude, subjective norms, and perceived behavioral control. Referencing Theory of Planned Behavior (TPB) sample questionnaire, 16 statements measured app adoption intent, attitude, subjective norms, and PBC on each of the five fitness apps (a general app, Tutor, Recorder, Game Companion, and Cheerleader). Participants reported how accurate statements describe their situations on a 5-point scale from "strong disagree" to "strongly agree". We averaged item scores for each multiple-item scale of behavioral constructs to produce a construct score. For example, the intent score for the general app was computed as the average points for four intent statements about the general app. The measurement statements for the general app on attitude are "*For me to download and try out a fitness app is interesting/ good/ pleasant/ valuable.*" Statements on perceived behavioral control include "*I am confident that if I wanted to I could download and try out a fitness app.*" Language was adjusted for specific apps, e.g. "*For me to download and try out this fitness app is interesting.*" Scales for all four behavioral constructs reached internal reliability on five apps: Cronbach's alphas range 0.73 to 0.95 for the general app, 0.79 to 0.96 for Tutor App, 0.78 to 0.96 for Recorder App, 0.85 to 0.97 for Game Companion App, and 0.80 to 0.97 for Cheerleader App.

3.3. Results of Study Two

RQ4: What types of fitness apps are most appealing to consumers considering features?

As Section 3.2.3. explained, participants individually "designed" their ideal fitness apps given a randomized list of 28 features. The 28 features unevenly fall under the five functional themes (Education, Tracking, Social, Gamification, Motivation). A participant's number of preferable features on each functional theme constitutes their "function preference index", or the extent to which they prefer that theme such as Education as opposed to Tracking. To understand the types of mobile fitness apps that appeal to participants, we replicated the clustering methods of Study One in Study Two: hierarchical clustering using standardized function preference indexes as the clustering criterion. We clustered survey participants, or essentially their ideal apps ($n = 351$). Clustering results represent homogeneous groups of mobile fitness apps which share the features that a group of users prefer, answering RQ4.

Among all clustering solutions, the five-group solution was most plausible because it offers best interpretability and least differences in cluster sizes. Clusters A (36.47%) and B (14.81%) are lean-feature apps. Cluster A contains apps with average tracking features, lacking other four themes in a moderate amount. Cluster B lacks almost all functional themes but have some social features. Cluster C (12.25%) addresses Motivation, Education, and Tracking. Clusters D (13.39%) and F (23.08%) are comprehensive apps: Cluster D contains all themes in a rich amount; Cluster F contains all themes with a moderate amount.

RQ5: Do the types of fitness apps that consumers prefer align with the four app prototypes (Tutor, Recorder, Game Companion, Cheerleader) in current iPhone app market?

A comparison of the clustering results between Studies One and Two answered RQ5. Clusters across both studies match at some point but not closely. Cluster A in Study Two is similar to the Recorder App in Study One, both focusing on Tracking. Cluster B is close to the Cheerleader App from Study One, but the apps in current market use Social functions secondary to the primary focus of Motivation. Cluster C is like the Game Companion App in terms of the richness of theme combination. They both combined three themes. However, although Cluster C and the Game Companion App both include Tracking, Cluster C additionally includes Motivation and Education, while Game Companion App contains Gamification and Social. Cluster D and F, based on their comprehensive nature, are distinct from any clusters from Study One regarding function combinations.

The comparison of clustering results implies, fitness apps in the current iOS app market may be related but not perfectly reflecting consumer needs in respect to app functions. To test this observation, we conducted a Pearson's r correlation test between market offerings and consumer needs. The frequency of app features that designers/market built into mobile fitness apps ($M = 22.24$, $SD = 16.95$, $n = 30$) significantly correlated ($r = 0.38$, $p = .046$) with the frequency of features that survey participants prefer ($M = 40.75$, $SD = 19.06$, $n = 28$). Among survey participants, 96.6% would want to have at least one tracking feature in their fitness apps, while 84.7% of sampled fitness apps offer at least one tracking feature. As 92.6% participants liked educational features, 71.4% fitness apps in Study One sample contain educational features. Namely, tracking is the most popular functional theme for both app designers and app users, closely followed by education. However, market offerings and user needs diverged on social features: the market (59.2%) overrepresented social features, given only 37.6% survey participants would like social features in their fitness apps. The current market, on the contrary, underrepresented gamification and motivational features. See Table 2 for comparisons between market offerings and user needs in regard to mobile fitness app features per theme.

H1: Consumers' mobile fitness app adoption intent is predicted by their attitudes toward the app, subjective norms of being expected to adopt a fitness app, and perceived behavioral control over app adoption.

Five multiple linear regressions were calculated to predict the intent to download mobile fitness apps, possibly by three behavioral variables (Attitude, Perceived Behavioral Control, and Subjective Norm). One model tested general fitness apps; another four models respectively tested Tutor, Recorder, Game Companion, and Cheerleader apps. Variables were mutually correlated in each model, *p values* < 0.001. A significant regression equation was found for **general fitness apps**, $F(3, 347) = 344.35$, $p < .001$, $R^2 = 0.75$. Participants' predicted intent to download fitness apps is equal to -0.109 + 0.94(Attitude) + 0.21(Norm), where intent, attitude, and norm are measured on 5-point scales (see Section 3.2.3 for measures). Intent to download fitness apps increased 0.94 point for each point of attitude ($\beta = 0.74$) and 0.21 point for each point of norm ($\beta = 0.16$).

A significant regression equation was found for **Tutor** app adoption, $F(3, 347) = 327.01$, $p < .001$), $R^2 = 0.74$. Participants' predicted intent to download Tutor apps is equal to – 0.31 + 0.75(Attitude) + 0.39(Norm). Intent to download Tutor apps increased 0.75 point for each point of attitude ($\beta = 0.65$) and 0.39 point for each point of norm ($\beta = 0.30$). A significant regression equation was found for the **Recorder** app, $F(3, 347) = 254.42$, $p < .001$, $R^2 = 0.69$. Participants' predicted intent to download Tutor apps is equal to -0.60 + 0.80(Attitude) + 0.32(Norm). Intent to download Recorder apps increased 0.80 point for each point of attitude ($\beta = 0.65$) and 0.32 point for each point of norm ($\beta = 0.25$). A significant regression equation was found for the **Game Companion** app, $F(3, 347) = 405.27$, $p < .001$, $R^2 = 0.78$. Participants' predicted intent to download Game Companion apps is equal to – 0.36 + 0.77(Attitude) – 0.10(Perceived Behavioral Control) + 0.40(Norm). Intent to download Recorder apps increased 0.77 point for each point of attitude ($\beta = 0.66$) and 0.40 point for each point of norm ($\beta = 0.31$) but decreased 0.10 point for each point of PBC ($\beta = -0.06$). We also found a significant regression equation for the **Cheerleader** app, $F(3, 347) = 429.17$, $p < .001$, $R^2 = 0.79$. Participants' predicted intent to download Cheerleader apps is equal to −0.27 + 0.70(Attitude) – 0.12(Perceived Behavioral Control) + 0.45(Norm). Intent to download Cheerleader apps increased 0.70 point for each point of attitude ($\beta = 0.62$) and 0.45 point for each point of norm ($\beta = 0.33$) but decreased 0.12 point for each point of PBC ($\beta = -0.08$).

Attitude and Subjective Norm were significant predictors of participants' intent to download all five apps. Perceived Behavioral

Table 2
Comparing Mobile Fitness App Features between the Market (98 apps) and Users (n = 351).

Features	Market % (n)	User % (n)	% Delta
Education	**71.4 % (70)**	**92.6 % (325)**	**21.2**
text tutorial	28.6 % (28)	22.2 % (78)	-6.4
photo tutorial	17.3 % (17)	31.9 % (112)	14.6
audio tutorial	21.4 % (21)	22.8 % (80)	1.4
video tutorial	52.0 % (51)	40.2 % (141)	-11.8
timing/interval guidance	25.5 % (25)	47.0 % (165)	21.5
real-time audio feedback	6.1 % (6)	27.9 % (98)	21.8
custom training	43.9 % (43)	55.3 % (194)	11.4
goal-setting	19.4 % (19)	72.4 % (254)	53
Tracking	**84.7 % (83)**	**96.6 % (339)**	**11.9**
users log data	27.6 % (27)	61.5 % (216)	33.9
app logs workout progress	69.4 % (68)	69.8 % (245)	0.4
sensor log data	55.1 % (54)	49.3 % (173)	-5.8
GPS based	17.3 % (17)	51.0 % (179)	33.7
Biometrics: Heart Rate	10.2 % (10)	64.7 % (227)	54.5
Biometrics: Weight	22.4 % (22)	69.5 % (244)	47.1
Biometrics: Calories	32.7 % (32)	76.4 % (268)	43.7
Social	**59.2 % (58)**	**37.6 % (132)**	**-21.6**
add friends	25.5 % (25)	20.5 % (72)	-5
community among app users	35.7 % (35)	22.2 % (78)	-13.5
social media sharing	36.7 % (36)	19.4 % (68)	-17.3
Gamification	**35.7 % (35)**	**63.0 % (221)**	**27.3**
badges and trophies	13.3 % (13)	35.6 % (125)	22.3
leaderboards	12.2 % (12)	21.1 % (74)	8.9
points and values	6.1 % (6)	35.0 % (123)	28.9
challenges and quests	25.5 % (25)	42.7 % (150)	17.2
role playing video game	4.1 % (4)	14.8 % (52)	10.7
Motivation	**38.8 % (38)**	**77.5 % (272)**	**38.7**
music	28.6 % (28)	52.7 % (185)	24.1
reminder/notification	8.2 % (8)	44.2 % (155)	36
motivational quotes	3.1 % (3)	23.4 % (82)	20.3
tips, advice	8.2 % (8)	32.8 % (115)	24.6
voice cheer, audio alert	6.1 % (6)	14.8 % (52)	8.7

Note. Bars represent differences between percentages of apps containing the feature and percentages of users who favor a feature. Bars to the left indicate that the current market overrepresented the features; bars to the right indicate that the current market underrepresented the features.

Control was an additional significant predictor of the intent to download **Game Companion** and **Cheerleader** apps, where users with lower PBC are more likely to adopt fitness apps that require less sophisticated smartphone skills. H1 on RAA was, by and large, supported – with an interesting exception about PBC.

RQ6: Controlling for the effects of demographic and mHealth variables, does users' app function preference predict their attitudes toward the four mobile fitness app clusters and toward mobile fitness apps in general? If so, how?

To answer RQ6, we ran five hierarchical regression analyses. Dependent variables were attitudes toward mobile fitness apps in general, and toward each app cluster (Tutor, Recorder, Game Companion, Cheerleader apps). Independent variables were entered in three steps. Step 1 included demographic variables: age, gender. Step 2 included mHealth variables: BMI, health status, eHealth literacy, exercise time, smartphone use history. Step 3 included users' App Function Preference Indexes on five functional themes: Education, Tracking, Social, Gamification, and Motivation. Section 3.2.3 explained how we measured these variables.

Table 3
Results of Hierarchical Regression Analyses for Tutor, Recorder, Game Companion, Cheerleader Apps, and Fitness Apps in General.[†]

	Tutor Apps			**Recorder Apps**			**Game Companion Apps**			**Cheerleader Apps**			**General Fitness Apps**		
Independent Variables	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3	Model 1	Model 2	Model 3
Step 1: Demographic variables															
Age	–0.016	–0.023	–0.020	0.011	0.000	–0.001	–0.136*	–0.152**	–0.090	0.034	0.025	0.011	–0.038	–0.051	–0.038
Gender (female = 1, male = 2)	–0.090	–0.075	–0.008	–0.109*	–0.102	–0.072	–0.125*	–0.106*	–0.056	–0.150**	–0.144**	–0.073	–0.159**	–0.146**	–0.099
Step 2: mHealth variables															
BMI		0.046	0.026		0.049	0.016		0.139*	0.091		0.166**	0.097		0.122*	0.081
Health status		–0.063	–0.058		–0.032	–0.030		–0.068	–0.083		–0.057	–0.082		0.002	–0.006
eHealth literacy		0.193***	0.168**		0.150**	0.120*		0.172**	0.151**		0.029	0.019		0.227***	0.206***
Exercise in hours		0.044	0.055		0.052	0.070		0.011	0.028		0.085	0.062		0.032	0.038
Smartphone use history		0.202***	0.168**		0.088	0.049		0.144**	0.108*		0.137*	0.119*		0.104	0.072
Step 3: App feature preference															
Educational			0.155**			–0.015			–0.017			–0.044			0.006
Tracking			–0.061			0.201***			–0.012			–0.119*			0.027
Social			–0.015			0.054			0.127*			0.205***			0.104
Gamification			0.110			–0.014			0.270***			–0.034			0.070
Motivation			0.121			0.135*			0.016			0.313***			0.131*
F	1.375	5.251***	5.583***	2.055	2.491*	3.679***	5.466**	5.873***	7.705***	4.206*	4.122***	7.764***	4.467*	5.734***	5.366***
R^2	0.008	0.100***	0.171***	0.012	0.050*	0.120***	0.032**	0.111***	0.221***	0.024*	0.080**	0.223***	0.026*	0.108***	0.165***
ΔR^2	0.008	0.092***	0.071***	0.012	0.038*	0.069***	0.032**	0.079***	0.111***	0.024*	0.056**	0.142***	0.026*	0.082***	0.057***

[†] Dependent variables: attitudes toward each of the five clusters of mobile fitness apps. All regression coefficients are standardized β.
Total $n = 351$; effective n was around 340 in each model as missing cases in respective analysis were deleted listwise.
* $p < .05$, ** $p < .01$, *** $p < .001$.

In all five regression analyses, plausibility of models increased as steps increased because variable blocks incrementally added significant R^2 to the regressions. For example, at Step 1, age and gender explained 3.2% variance in attitude toward Game Companion apps, $F(2, 335) = 5.466$, $p < .01$; at Step 2, BMI, eHealth literacy, and smartphone use history explained 7.9% additional variance in outcome, $F(7, 330) = 5.873$, $p < .001$; at Step 3, users' preference for Social and Gamification features significantly explained 11.1% additional variance in outcome, $F(12, 325) = 7.705$, $p < .001$. Together variables from three steps explained 47.1% of the variance in attitude toward Game Companion apps. Regressions on other app clusters as well as fitness apps in general found different sets of predictors for app attitudes, all including feature preferences – though different features – as strong predictors. See Table 3 for regression statistics: coefficients β, F values, F test significance, R, R significance, ΔR^2, ΔR^2 significance.

Hierarchical regression results show that people are not equally attracted to all types of mobile fitness apps. People's attitudes for general or specific apps are strongly predicted by their preference of the specific combination of features, which define the key functionality of the apps. For example, user preference of Educational features (i.e. the defining function of Tutor apps) strongly predicts favorable attitudes toward Tutor apps, $\beta = 0.155$, $p < .01$. Similarly, preference of Tracking ($\beta = 0.201$, $p < .001$) and Motivational ($\beta = 0.135$, $p < .05$) features predicts favorable attitudes toward Recorder apps. Preference of Social ($\beta = 0.127$, $p < .05$) and Gamification ($\beta = 0.270$, $p < .001$) features predicts favorable attitudes toward Game Companion apps. Preference of Tracking ($\beta = -0.119$, $p < .05$), Social ($\beta = 0.205$, $p < .001$) and Motivation ($\beta = 0.313$, $p < .001$) features predicts favorable attitudes toward Cheerleader apps. Attitudes toward mobile fitness apps in general are predicted by preference of Motivational app features, $\beta = 0.131$, $p < .05$.

While younger/female users seem to like Game Companion apps more than their older/male counterparts, age and gender effects will be lessened to non-significant when user preference of Tracking and Motivational features comes into play. Notably, Age only slightly predicts the attitude toward Game Companion apps but no other app clusters or fitness apps in general; but female users do seem to welcome fitness apps more than male across app types.

As for mHealth variables: (1) People with higher BMI may favor Game Companion and Cheerleader apps more than people with lower BMI; (2) eHealth literacy facilitates favorable attitudes toward fitness apps in general and toward almost all app clusters except for Cheerleader apps; (3) Smartphone proficiency, measured by usage history in years, facilitates favorable attitudes toward Tutor, Game Companion and Cheerleader apps, but not Recorder apps or fitness apps in general; (4) Health status and current exercise behaviors are not significant predictors for app attitude, but future research should still test their associations with app adoption.

3.4. Discussion of Study Two in comparison with Study One

Mediated experience enabled by mobile apps is proven promising to persuade behavioral intent change, through meaningful human–computer interaction and tailored user experience (Wang, 2020). Our research contributes to literature in persuasive technologies, where interactive systems target attitude or behavior change (Oinas-Kukkonen and Harjumaa, 2009). Through social, communicative persuasion and attitude change, persuasive technologies can foster desired behavioral outcomes (Hamari et al., 2014) such as mobile fitness apps for physical activity.

In our evaluation, the current mobile fitness app market turned out not closely aligned with user needs, according to the comparisons between Studies One and Two. Survey data of user preferences for certain features differ substantially from the presence of those features as identified in actual fitness apps available in the market, with some features being overrepresented (e.g. social media sharing) and other features being underrepresented (e.g. reminders/ notifications) relative to consumer preferences. We found only 14.44% shared feature representation between market and consumer demands, which requires more tailored design.

Relatively consistent evidence in the literature supports the benefits of tailored approaches to health behavior change over untailored approaches (Bull et al., 1999; Kreuter et al., 2000; Lustria et al., 2009), especially when tailored to address identified barriers to change (Baker et al., 2010). As we found, people with higher BMI liked fitness apps entailing Gamification and Motivational tools more than Educational or Tracking tools. None of our data suggests, nor was it the goal of our study, to examine the actual efficacy associated with app use. However, our presumption is that fitness app adoption is a necessary prerequisite for any real or perceived efficacy. Any fitness app designed to improve health outcomes should attempt to maximize adoption through *tailoring* of features aligned with specific consumer preferences. Formative research on these preferences should be a critical aspect of the app design process.

There is no "perfect" or "best" app for all. Our study indicates that a "kitchen sink" approach to app feature inclusion is unwarranted. In some cases, the inclusion of additional unwanted features had a negative impact on fitness app attitudes even though other desired features were present. However, the presences of some features may generally augment a wider range of app types without impacting attitudes. Given that we also saw evidence of factors such as eHealth literacy and smartphone experience being predictors of fitness app attitudes, the addition of more features may be overwhelming to some users and reduce perceived behavioral control associated with app adoption or utilization. Hence, app developers should strongly reconsider the potential negative impact of unprincipled additions in functionality even though it may be well within their power to do so.

Consistent with previous literature (Hsiao et al., 2016; Kim et al., 2007; Teo et al., 2012), Study Two found that attitudes toward different types of fitness apps were consistently predicted by individual beliefs regarding the value of specific fitness app features. Consumer app feature preferences significantly improved the prediction of fitness app attitudes after controlling for basic demographics and mHealth related individual differences. Specifically, consumer feature preferences corresponded to the defining characteristic of the fitness app type under consideration. As Table 3 reported, the feature preference most predictive of attitudes toward Tutor apps was about Educational features; for Recorder apps, preference for Tracking features was the strongest attitude predictor. This was true for all app types. Although Study One found that the presence of Educational and Social features in mobile fitness apps were overall predictive of higher user satisfaction ratings, Study Two revealed individual nuances on app attitude per various feature preference. Notably, users who liked Game Companion and Cheerleader apps were fond of Social features in the apps. We confirmed the significance of social emotional support in mHealth interventions beyond instrumental support (Peng et al., 2016), which gradually fosters exercise habits. Bridging the gap between market offerings and user needs may help optimize and tailor user experience in mobile fitness apps, which subsequently increases app efficacy to improve consumer health.

Like the RAA posits, attitudes, subjective norms, and perceived behavior control were very strong predictors of adoption intent, accounting for 69% to 79% of the variation in adoption intent across all app types, with attitude consistently being the strongest predictor, and behavioral control being the weakest predictor. Furthermore, the theory presumes that underlying beliefs are direct predictors of attitudes; other factors, such as demographics and other individual difference variables only influence attitudes, behavioral intent, and actual behavior indirectly. In other words, specific beliefs are presumed to mediate relationships between individual differences and attitudes. Though not a specific goal of the study, our models show that when demographic and mHealth variables have effects associated with attitudes, those effects were reduced or became non-significant when app feature preferences were subsequently included in the model. This finding is consistent with Fishbein and Ajzen's presumed mediating role of attitudes (2010).

This paper presents both theoretical and practical implications for app evaluation and mHealth intervention. We stopped at studying app adoption intent. Future research can explore the continued use of mobile fitness apps, i.e. engagement and adherence, to answer "why would consumers stop using a fitness app which they initially believed to be helpful for achieving their fitness goals?" Future research should also evaluate app feature implementation and track longitudinally the actual health outcomes that fitness apps bring to users.

Limitations. MTurk is a considerably optimistic channel to obtain relatively representative data, but it does not perfectly represent the entire population. Also, findings of both studies are based on iPhone apps. Conclusions should be interpreted without over-extrapolating to all platforms such as Android. In a future study, we hope to test findings of the present research on app ecosystems supported by other smartphone operating systems.

4. Conclusion

Mobile fitness apps tend to represent distinct feature combinations that emphasize different strategies for supporting consumer exercise goals. We synthesized current literature on mHealth to propose an updated typology of mobile fitness app features, organized around five functional themes: Education, Tracking, Social, Gamification, and Motivation, each including several sub-features. The typology demonstrated legitimate applicability. Future research should cast more attention to the interplay between diverse app features in shaping app functionality.

The Reasoned Action Approach provided a meaningful theoretical framework to evaluating adoption intent of fitness apps through surveying attitudes, subjective norms, and perceived behavioral control. A comparison between current market offerings (Study One) and consumer preference (Study Two) indicates a mismatch and room for market improvement. Adding to tailored design literature, we found various user feature preference meaningfully associated with attitudes toward different types of apps. This corresponded with recommendations in the academic literature that stress the efficacy of theoretically-driven physical activity interventions (Gourlan, et al., 2016). Though it is unknown the extent to which fitness apps in this study utilized theory to guide application development, our results demonstrated the ability of theoretical concepts to predict app adoption intent, satisfaction, and attitude. Hence, app developers are likely to benefit from the use of theoretical developments associated with fitness-related behavior change. Some highlights of this paper's contribution to mHealth and persuasive technology literature are as follows.

- The different combinations of app functions – rather than standalone functions – contribute to app success.
- There are four typical clusters of mobile fitness apps in current iPhone app store: Tutor, Recorder, Game Companion, Cheerleader. They afford unique combinations of functionality amongst Education, Tracking, Social, Gamification, and Motivation features – varying in both feature types and feature richness. See Section 2.3.2.
- Motivation features in mobile fitness apps appeal to all user groups, e.g. music, reminder/notification, motivational quotes, tips/advice, voice cheer, audio alert.
- User attitude toward a fitness app is predicted by both how much they value the defining functionality of the app and how much they tolerate a feature they dislike.
- Smartphone proficiency and eHealth literacy (i.e. the skills to navigate, use, and critically evaluate health resources online) underpin mHealth adoption.
- Younger/female/higher-BMI users welcome more gamification features in mobile fitness apps than their older/male/lower-BMI counterparts.

We provided theoretical and practical implications for future efforts on mobile fitness apps, mHealth, and persuasive technology to improve the health outcomes of unique subsets of the entire population through tailored design.

CRediT authorship contribution statement

Yunwen Wang: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing - original draft, Visualization, Funding acquisition, Project administration. **Bart Collins:** Supervision, Conceptualization, Methodology, Formal analysis, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The project was funded by Purdue University through the PROMISE Award. The authors thank Dr. Maria Venetis, Dr. Marifran Mattson, Dr. Randy Lake, Dr. Lynn Miller, and colleagues Natalie Jonckheere and Junyi Lv, as well as anonymous reviewers for their feedback on earlier versions of the manuscript. We also thank Ashleigh Shields for her help on inter-coder reliability testing.

References

Ajzen, I., 1991. The theory of planned behavior. Org. Behav. Hum. Decis. Process. 50 (2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.
Anderberg, M.R., 2014. Probability and mathematical statistics: A series of monographs and textbooks, vol. 19. Academic Press, New York, NY.
Arnhold, M., Quade, M., Kirch, W., 2014. Mobile applications for diabetics: A systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J. Med. Internet Res. 16 (4) https://doi.org/10.2196/jmir.2968.
Azar, K.M.J., Lesser, L.I., Laing, B.Y., Stephens, J., Aurora, M.S., Burke, L.E., Palaniappan, L.P., 2013. Mobile applications for weight management. Am. J. Preventive Med. 45 (5), 583–589. https://doi.org/10.1016/j.amepre.2013.07.005.
Baker, R., Camosso-Stefinovic, J., Gillies, C., Shaw, E.J., Cheater, F., Flottorp, S., Robertson, N., 2010. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst. Rev. 3 https://doi.org/10.1002/14651858.CD005470.pub2.
Bull, F.C., Kreuter, M.W., Scharff, D.P., 1999. Effects of tailored, personalized and general health messages on physical activity. Patient Educ. Counsel. 36 (2), 181–192.
Cameron, D., Harvey, A., 2006. eHEALS: the ehealth literacy scale. J Med Internet Res 8 (4). https://doi.org/10.2196/jmir.8.4.e27.
Cialdini, R.B., Reno, R.R., Kallgren, C.A., 1990. A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. J. Personality Soc. Psychol. 58, 1015–1026. https://doi.org/10.1037/0022-3514.58.6.1015.
Coll, C.V.N., Domingues, M.R., Gonçalves, H., Bertoldi, A.D., 2017. Perceived barriers to leisure-time physical activity during pregnancy: a literature review of quantitative and qualitative evidence. J. Sci. Med. Sport 20 (1), 17–25. https://doi.org/10.1016/j.jsams.2016.06.007.
Deng, Z., Mo, X., Liu, S., 2014. Comparison of the middle-aged and older users' adoption of mobile health services in China. Int. J. Med. Inf. 83 (3), 210–224. https://doi.org/10.1016/j.ijmedinf.2013.12.002.
DiClemente, R.J., Salazar, L.F., Crosby, R.A., 2013. Health Behavior Theory for Public Health. Jones & Bartlett Publishers.
Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.R., Christensen, H., 2013. Smartphones for smarter delivery of mental health programs: a systematic review. J. Med. Internet Res. 15 (11) https://doi.org/10.2196/jmir.2791.
Duran, B.S., Odell, P.L., 2013. Cluster analysis: A survey, vol. 100. Springer Science & Business Media.
Eagly, A.H., Chaiken, S., 1993. The Psychology of Attitudes. Harcourt brace Jovanovich college publishers.
Ernsting, C., Dombrowski, S.U., Oedekoven, M., Kanzler, M., Kuhlmey, A., Gellert, P., 2017. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J. Med. Internet Res. 19 (4) https://doi.org/10.2196/jmir.6838.
Fanning, J., Mullen, S.P., McAuley, E., 2012. Increasing physical activity with mobile devices: a meta-analysis. J. Med. Internet Res. 14 (6) https://doi.org/10.2196/jmir.2171.
Farrow, M., Lutteroth, C., Rouse, P.C., Bilzon, J.L.J., 2019. Virtual-reality exergaming improves performance during high-intensity interval training. Eur. J. Sport Sci. 19 (6), 719–727.
Fishbein, M., 2000. The role of theory in HIV prevention. AIDS Care 12 (3), 273–278. https://doi.org/10.1080/09540120050042918.
Fishbein, M., Ajzen, I., 2010. Predicting and changing behavior: The reasoned action approach.
Gibbons, F.X., 2006. Behavioral intentions, expectations, and willingness. Health Behavior Constructs: Theory, Measurement, and Research. National Cancer Institute, Bethesda, MD.
Goad, T., Towner, B., Jones, E., Bulger, S., 2019. Instructional tools for online physical education: using mobile technologies to enhance learning. J. Phys. Educ., Recreation Dance 90 (6), 40–47.
Gourlan, M., Bernard, P., Bortolon, C., Romain, A.J., Lareyre, O., Carayol, M., Ninot, G., Boiché, J., 2016. Efficacy of theory-based interventions to promote physical activity. A meta-analysis of randomised controlled trials. Health Psychol. Rev. 10 (1), 50–66.
Goyal, S., Cafazzo, J.A., 2013. Mobile phone health apps for diabetes management: current evidence and future developments. QJM: Int. J. Med. 106 (12), 1067–1069.
Hamari, J., Koivisto, J., Pakkanen, T., 2014. Do persuasive technologies persuade?-a review of empirical studies. In: International Conference on Persuasive Technology. Springer, Cham, pp. 118–136. https://doi.org/10.1007/978-3-319-07127-5_11.
Healthy People 2020, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, n.d.-a. Health communication and health information technology. Retrieved from <https://www.healthypeople.gov/2020/topics-objectives/topic/health-communication-and-health-information-technology>.
Healthy People 2020, U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, n.d.-b. Physical activities. Retrieved from <https://www.healthypeople.gov/2020/topics-objectives/topic/physical-activity>.
Holland, R.W., Verplanken, B., Van Knippenberg, A.d., 2002. On the nature of attitude-behavior relations: the strong guide, the weak follow. Eur. J. Soc. Psychol. 32 (6), 869–876.
Hung, H.-C., Shwu-Ching Young, S., Lin, K.-C., 2018. Exploring the effects of integrating the iPad to improve students' motivation and badminton skills: a WISER model for physical education. Technol., Pedagogy Educ. 27 (3), 265–278.
Huotari, K., Hamari, J., 2017. A definition for gamification: anchoring gamification in the service marketing literature. Electr. Markets 27 (1), 21–31. https://doi.org/10.1007/s12525-015-0212-z.
Hsiao, C.-H., Chang, J.-J., Tang, K.-Y., 2016. Exploring the influential factors in continuance usage of mobile social Apps: satisfaction, habit, and customer value perspectives. Telematics Inf. 33 (2), 342–355. https://doi.org/10.1016/j.tele.2015.08.014.
Kim, H.-W., Chan, H.C., Gupta, S., 2007. Value-based Adoption of mobile internet: an empirical investigation. Decis. Support Syst. 43 (1), 111–126. https://doi.org/10.1016/j.dss.2005.05.009.
Koivisto, J., Hamari, J., 2019. Gamification of physical activity: A systematic literature review of comparison studies. In: 3rd International GamiFIN Conference, GamiFIN 2019. CEUR-WS. <http://ceur-ws.org/Vol-2359/paper10.pdf>.
Kranz, M., MöLler, A., Hammerla, N., Diewald, S., PlöTz, T., Olivier, P., Roalter, L., 2013. The mobile fitness coach: Toward individualized skill assessment using personalized mobile devices. Pervasive Mobile Comput. 9 (2), 203–215. https://doi.org/10.1016/j.pmcj.2012.06.002.
Kreuter, M.W., Oswald, D.L., Bull, F.C., Clark, E.M., 2000. Are tailored health education materials always more effective than non-tailored materials? Health Educ. Res. 15 (3), 305–315.
Lee, H.E., Cho, J., 2016. What motivates users to continue using diet and fitness apps? Application of the uses and gratifications approach. Health Commun. 32 (12), 1445–1453.

Leonard, T., Shuval, K., de Oliveira, A., Skinner, C.S., Eckel, C., Murdoch, J.C., 2013. Health behavior and behavioral economics: economic preferences and physical activity stages of change in a low-income African-American community. Am J Health Promot 27 (4), 211–221.

Leung, L., Wei, R., 2000. More than just talk on the move: uses and gratifications of the cellular phone. Journalism Mass Commun. Q. 77 (2), 308–320.

Liao, G.Y., Chien, Y.T., Chen, Y.J., Hsiung, H.F., Chen, H.J., Hsieh, M.H., Wu, W.J., 2017. What to build for middle-agers to come? Attractive and necessary functions of exercise-promotion mobile phone apps: a cross-sectional study. JMIR mHealth uHealth 5 (5). https://doi.org/10.2196/mhealth.6233.

Lim, S.L., Bentley, P.J., 2013. Investigating app store ranking algorithms using a simulation of mobile app ecosystems. In: Evolutionary Computation (CEC), 2013 IEEE Congress on Evolutionary Computation, pp. 2672-2679. https://doi.org/10.1109/CEC.2013.6557892.

Lim, S.L., Bentley, P.J., Kanakam, N., Ishikawa, F., Honiden, S., 2015. Investigating country differences in mobile app user behavior and challenges for software engineering. IIEEE Trans. Software Eng. 41 (1), 40–64. https://doi.org/10.1109/TSE.2014.2360674.

Lustria, M.L.A., Cortese, J., Noar, S.M., Glueckauf, R.L., 2009. Computer-tailored health interventions delivered over the web: review and analysis of key components. Patient Educ. Counsel. 74 (2), 156–173. https://doi.org/10.1016/j.pec.2008.08.023.

Mendiola, M.F., Kalnicki, M., Lindenauer, S., 2015. Valuable features in mobile health apps for patients and consumers: Content analysis of apps and user ratings. JMIR mHealth uHealth 3 (2). https://doi.org/10.2196/mhealth.4283.

Miller, A.S., Cafazzo, J.A., Seto, E., 2016. A game plan: Gamification design principles in mHealth applications for chronic disease management. Health Inf. J. 22 (2), 184–193.

Mohr, D.C., Schueller, S.M., Montague, E., Burns, M.N., Rashidi, P., 2014. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J. Med. Internet Res. 16 (6) https://doi.org/10.2196/jmir.3077.

Mooi, E., Sarstedt, M., 2011. In: A Concise Guide to Market Research. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 237–284. https://doi.org/10.1007/978-3-642-12541-6_9.

Muessig, K.E., Pike, E.C., LeGrand, S., Hightow-Weidman, L.B., 2013. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J. Med. Internet Res. 15 (1) https://doi.org/10.2196/jmir.2301.

Murtagh, F., Legendre, P., 2014. Ward's hierarchical agglomerative clustering method: which algorithms implement ward's criterion? J Classif 31 (3), 274–295. https://doi.org/10.1007/s00357-014-9161-z.

Norusis, M.J., 2010. In: Chapter 16: Cluster analysis. PASW Statistics 18 Statistical Procedures Companion. Prentice Hall, Upper Saddle River, NJ, pp. 361–391.

Oinas-Kukkonen, H., Harjumaa, M., 2009. Persuasive systems design: key issues, process model, and system features. Commun. Assoc. Inf. Syst. 24 (1), 28. https://doi.org/10.17705/1CAIS.02428.

Oppenheimer, D.M., Meyvis, T., Davidenko, N., 2009. Instructional manipulation checks: detecting satisficing to increase statistical power. J. Exp. Soc. Psychol. 45 (4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009.

Papastergiou, M., Natsis, P., Vernadakis, N., Antoniou, P., 2020. Introducing tablets and a mobile fitness application into primary school physical education. Educ. Inf. Technol. 1–18.

Peng, W., Kanthawala, S., Yuan, S., Hussain, S.A., 2016. A qualitative study of user perceptions of mobile health apps. BMC Publ. Health 16 (1), 1–11. https://doi.org/10.1186/s12889-016-3808-0.

Pew Research Center, 2019. Mobile Fact Sheet. Retrieved from <https://www.pewresearch.org/internet/fact-sheet/mobile/#:~:text=The%20vast%20majority%20of%20Americans,smartphone%20ownership%20conducted%20in%202011>.

Reiner, M., Niermann, C., Jekauc, D., Woll, A., 2013. Long-term health benefits of physical activity–a systematic review of longitudinal studies. BMC Publ. Health 13 (1), 813. https://doi.org/10.1186/1471-2458-13-813.

Richtering, S.S., Hyun, K., Neubeck, L., Coorey, G., Chalmers, J., Usherwood, T., et al., 2017. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum. Factors 4 (1). https://doi.org/10.2196/humanfactors.6217.

Rimal, R.N., Real, K., 2005. How behaviors are influenced by perceived norms: a test of the theory of normative social behavior. Commun. Res. 32 (3), 389–414.

Sama, P.R., Eapen, Z.J., Weinfurt, K.P., Shah, B.R., Schulman, K.A., 2014. An evaluation of mobile health application tools. JMIR mHealth uHealth 2 (2). https://doi.org/10.2196/mhealth.3088.

Sharifi, N., Mahdavi, R., Ebrahimi-Mameghani, M., 2013. Perceived barriers to weight loss programs for overweight or obese women. Health Promotion Perspect. 3 (1), 11. https://doi.org/10.5681/hpp.2013.002.

Shin, Y., Jang, H., Pender, N.J., 2001. Psychometric evaluation of the exercise self-efficacy scale among Korean adults with chronic diseases. Res. Nurs. Health 24 (1), 68–76. https://doi.org/10.1002/1098-240X(200102)24:1<68::AID-NUR1008>3.0.CO;2-C.

Statista, 2020a. Number of mHealth apps available at Google Play from 1st quarter 2015 to 1st quarter 2020. Retrieved <https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/>.

Statista, 2020b. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 1st quarter 2020. Retrieved from <https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/>.

Teo, A.C., Tan, G.W.H., Cheah, C.M., Ooi, K.B., Yew, K.T., 2012. Can the demographic and subjective norms influence the adoption of mobile banking? Int. J. Mobile Commun. 10 (6), 578–597. https://doi.org/10.1504/IJMC.2012.049757.

Trost, S.G., Owen, N., Bauman, A.E., Sallis, J.F., Brown, W., 2002. Correlates of adults' participation in physical activity: review and update. Med. Sci. Sports Exercise 34 (12), 1996–2001. https://doi.org/10.1249/01.MSS.0000038974.76900.92.

U.S. CDC, Centers for Disease Control and Prevention, n.d., Healthy weight, overweight, and obesity among U.S. adults (CDC Publication No. 03-0260 (7/03)). Retrieved from <https://www.cdc.gov/nchs/data/nhanes/databriefs/adultweight.pdf>.

U.S. CDC, Centers for Disease Control and Prevention, 1996. Physical activity and health: A report of the Surgeon General. Retrieved from <https://www.cdc.gov/nccdphp/sgr/pdf/sgraag.pdf>.

Vickey, T., Breslin, J., Williams, A., 2012. Fitness - there's an app for that: Review of mobile fitness apps. Int. J. Sport Soc. 3 (4) https://doi.org/10.18848/2152-7857/CGP/v03i04/53952.

Wang, Y., 2020. Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). Comput. Hum. Behav. 110, 106373. https://doi.org/10.1016/j.chb.2020.106373.

Wang, J., Wang, Y., Wei, C., Yao, N.(., Yuan, A., Shan, Y., Yuan, C., 2014. Smartphone interventions for long-term health management of chronic diseases: an integrative review. Telemedicine e-Health 20 (6), 570–583.

Xi, N., Hamari, J., 2019. Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. Int. J. Inf. Manage. 46, 210–221.

Yang, H.'., 2013. Bon appétit for apps: young american consumers' acceptance of mobile applications. J. Comput. Inf. Syst. 53 (3), 85–96.

Yim, O., Ramdeen, K.T., 2015. Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. Tutorials Quant. Methods Psychol. 11, 8–21. https://doi.org/10.20982/tqmp.11.1.p008.