INTERNATIONAL JOURNAL OF ACADEMIC RESEARCH IN BUSINESS & SOCIAL SCIENCES

Vol 14, Issue 1, (2024) E-ISSN: 2222-6990

Instrument's Reliability and Validity of Continuance Intention toward Using Mobile Fitness Apps Model: Pilot Study

Rasha Najib Aljabali and Norasnita Ahmad
Information Systems Department, Faculty of Management, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
Corresponding Author Email: rashagabaly@gmail.com

To Link this Article: http://dx.doi.org/10.6007/IJARBSS/v14-i1/20521 DOI:10.6007/IJARBSS/v14-i1/20521

Published Date: 12 January 2024

Abstract
The prevalence of mobile fitness apps has led to a substantial increase in their market size and download rates. However, in practical situations, consistently high dropout rates and limited usage of mobile fitness apps continue to pose significant challenges for successful post-adoption. Therefore, the current study aimed to evaluate the instrument's reliability and validity for the proposed model of users' continuance intention toward using mobile fitness apps. The model comprises 10 constructs derived from the Expectation Confirmation Model (ECM), guilt of temporary discontinuance of the mobile fitness app, four Neutralization Techniques, and switching costs. The pilot study was conducted in Malaysia through social media platforms. The data analysis was conducted using the Partial Least Squares-Structural Equation Modelling (PLS-SEM) technique for 65 valid respondents to assess the reliability and validity of the questionnaire. SmartPLS 4 software was utilized to perform the analysis. The findings indicated good reliability and construct validity. This pilot study provides a foundation for further research on continuance intention to use mobile fitness apps and validates the measurement instrument for future large-scale studies.
Keywords: Construct Validity, Construct Reliability, Continuance Intention, Instrument Development, Mobile Fitness Apps

Introduction

Approximately 656 million worldwide downloads of Mobile Health (M-health) and fitness apps were recorded in 2021, with expectations of further growth in the future (Statista, 2021). Similarly, the market for M-health apps reached a value of USD 43.5 billion in 2022 and is projected to experience a compound annual growth rate (CAGR) of 11.6% from 2023 to 2030 (Research, 2023). Mobile fitness apps are among the top popular categories of M-health apps and have been found to offer social and emotional utilitarian benefits such as competition and a sense of accomplishment which can serve as incentives for users to boost their levels of physical activity, particularly among younger and technologically savvy generations (Cheng et al., 2021; Yousaf et al., 2021). Additionally, the pandemic of COVID-19 has prompted a shift to home-based workouts, leading to a spike in the utilization of mobile fitness apps that offer precise exercise measurements (Kwon et al., 2022).

Despite the large number and the increased market size of M-health apps, the dropout rate of these apps remains a challenge (Amagai et al., 2022; Huang & Ren, 2020). Research indicates that despite mobile fitness apps primarily targeting the younger generation, a substantial number of users discontinue their usage within 90 days (Ding, 2019; Yin et al., 2021). This phenomenon is compounded by low app engagement rates and remarkably high uninstallation rates. (Vinnikova et al., 2020; Zhang & Xu, 2020). Statistics show that the daily active users of M-health apps are substantially lower than those of other apps such as social media, instant messaging, and gaming apps (Li et al., 2021). This discrepancy in user engagement highlights the need for a deeper understanding of the post-adoption behavior of Information Systems (IS).

Theoretically, the adoption lifecycle of an Information Systems (IS) has recently been reconceptualized to elaborate on the adoption stages. According to Soliman and Rinta-Kahila (2020), there are three main key lifecycle stages of IS use: exposure, adoption, and continued use. Therefore, the continued usage stage represents a dynamic phase in which individuals continue to use the IS, subject to changes in their environment, psychological perception, and satisfaction status, contrasting with the relatively stable pre-adoption phase. In the context of physical activity and exercise (e.g., jogging, running, cycling, and swimming), it is widely supported by evidence that the benefits gained from exercise are closely tied to both the intensity (low or high) and the duration, which is typically longer. Practically, users' fitness and health goals are facilitated by the usage of the mobile fitness app which acts as a tracker of physical activity and exercise and motivates users to continue or repeat their workouts. As a result, the continued use stage occurs after a long period of usage where benefits and expectations are achieved. Nevertheless, individuals may develop discontinuance behaviors based on the changes that occur during continued use, such as quitting, temporary discontinuance, and switching to another mobile fitness app (Soliman & Rinta-Kahila, 2020).

In this study, the factors that influence mobile fitness app users to continue using the app have been proposed. These factors will prevent users from either temporarily abandoning it, permanently quitting, or switching to another mobile fitness app. Hence, this paper presents the instrument's reliability and validity for continuance intention to use mobile fitness apps through a pilot study. In the next section, the model development and methodology employed to conduct the pilot study are discussed.

Model Development

This research proposes a model based on the three discontinuance scenarios according to IS discontinuance Soliman & Rinta-Kahila (2020): mobile fitness app quitting, switching to another mobile fitness app, and temporary discontinuance of the mobile fitness app. Each scenario comprised the constructs that contribute to exploring the individual's intention to continue using the mobile fitness app. In addition, this study utilized a theoretical foundation to establish the hypothesis such as the Expectation Confirmation Model (ECM) Bhattacherjee (2001) and neutralization techniques (Sykes & Matza, 1957). Therefore, according to the M-health app, marketing, and sociology literature, the authors were able to expand the previous models by adding new factors: switching costs, guilt of temporary discontinuance of the mobile fitness app, neutralization techniques and emotional satisfaction. The model contains ten constructs: Expectation Confirmation (COM), Emotional Satisfaction (ESAT), Cognitive Satisfaction (CSAT), Guilt of Temporary Discontinuance of the mobile fitness apps (GUL), Switching Costs (SWI), Denial of Injury (DOI), Appeal to Higher Loyalties (AHL), Metaphor of Ledger (MOL), Condemning the Condemners (CTC), and Continuance Intention to Use Mobile Fitness App (CON). Some of these constructs have been replicated in the context of this study, and we need to confirm their reliability and validity. A questionnaire has been developed based on prior literature and has undergone content validity following best practices (Yusoff, 2019).

Research Methodology

A questionnaire is a highly utilized tool in quantitative approaches in Information Systems research for collecting data. It is designed to gather information from participants through a structured set of measurement items. The primary objective of using a questionnaire in research is to obtain relevant information in the most reliable and valid manner possible (Taherdoost, 2016). Validity refers to the extent to which an instrument accurately measures the construct or concept that intends to measure. There are four main types of validity: construct validity, face validity, content validity, and criterion validity. Construct validity examines the extent to which a measurement instrument measures the theoretical construct or concept it intends to assess. Therefore, validity is a judgment made about how well the scores obtained from an instrument can be interpreted to represent the construct being measured (Flake et al., 2022). As the same construct may be replicated in heterogeneous contexts and settings, it is important to continually gather evidence to support the validity of the scores in these new contexts. This ongoing validation process ensures that the scores obtained from the instrument remain valid and meaningful in various settings. Meanwhile, constructs' reliability indicates that measurement items consistently produce similar results when administered under similar conditions. Reliability is enhanced through standardized question formats, clear instructions, and well-defined response options (Taherdoost, 2016).

A pilot study is a small feasibility study that tests different methods for a larger research study to ensure that it will be successful (Lowe, 2019). A pilot study is not meant to answer research questions, but to help researchers avoid costly mistakes in a larger study. In the quantitative research approach, the pilot study should confirm the validity and reliability of the instrument used (Van Teijlingen & Hundley, 2001). After confirming the content validity of the questionnaire, the instrument should be pre-examined with a representative sample of the respondents prior to the main data collection. Conducting a pilot study is crucial for enhancing the quality and effectiveness of the primary research study.

Instrument Development

The research instrument was developed based on previous literature; refer to Table 2. The instrument was divided into four sections: (1) respondent's information sheet and informed consent, (2) demographic and screening Information, (3) Type of Mobile fitness app, Physical Activity Behavior, and Usage Characteristics, and (4) questionnaire items.
The measurement includes 42 items corresponding to the model's constructs (refer toTable 2) to explore the continuance intention to use mobile fitness apps. The 5-point

Likert scale was applied to collect the data to measure the 10 constructs (i.e., Continuance Intention to Use Fitness App, Confirmation, Emotional Satisfaction, Cognitive Satisfaction, Appeal to Higher Loyalties, Denial of Injury, Metaphor of the Ledger, Condemning the Condemners, Guilt of Temporary Discontinuance of the mobile fitness apps, and Switching costs) of the research model. The 5-point Likert scale is a valid and reliable scale for collecting data in information systems research (Taherdoost, 2019). The scale varied between low 1 (Strongly Disagree) to high 5 (Strongly Agree) in agreement with the items' statement. However, the guilt associated with temporary discontinuance of the mobile fitness app was measured on a 5-point Likert scale, ranging from 1 (Not at All) to 5 (Extremely).

Respondents of the Study

The respondents in this study are current users of mobile fitness apps in Malaysia. According to Aljabali and Ahmad (2022), mobile fitness app users in Malaysia use the apps mostly for monitoring their physical activity and frequently engage in online social communities (e.g., Facebook, and Instagram) dedicated to sport and fitness. Malaysian users of the apps also tend to have higher levels of education. Moreover, they maintain a regular exercise regimen and typically have a high monthly income. Therefore, the pilot study was conducted in Malaysia, and respondents were aged 18 and above. The questionnaire was administered online. Respondents were recruited from sport and fitness public and private groups on social media platforms such as Facebook, and Instagram. Questionnaire links were provided in English and Malay languages.

The respondents who filled out the questionnaire were 107. The analysis was performed on only 65 respondents who were considered valid, after excluding those who do not use mobile fitness apps, those who did not give consent, and those who answered the attention question incorrectly. The participants from sports, fitness and running Facebook groups were invited to participate in the research questionnaire. The Google Form questionnaire links have been posted and the respondents voluntarily access the preferred language: Malay version or English version.

Data Analysis

Items were coded with 3 or 4 letters of the construct, followed by an underscore (_), and the item number to create a unique code for each item (see Table 2). The measurement model can be either reflective or formative, depending on the conceptualization of the Latent Variables (LVs). In this study, the reflective measurement model has been utilized and the measures: reliability parameters (e.g., indicator reliability, Internal consistency reliability) and validity parameters (e.g., convergent, discriminant validity) were performed. SmartPLS 4 version 4.0.9.6 (Ringle et al., 2022) was used for data analysis. Therefore, numerous studies have highlighted the recommendations, guidelines, thresholds, and activities within PLS-SEM for rigorous presentation (Ali et al., 2018; Hair et al., 2019). This study undertook the R statistical package version 4.1.2 to perform the descriptive statistics. Additionally, the PLS-SEM algorithm in SmartPLS 4.0.9.6 was used to report the results, which constitute a measurement model (see Table 2).

Results and Discussions

Demographic Information

As shown in

Table **1**, the table describes the demographic data of the respondents of the pilot study. Overall, the pilot study was based on 65 valid respondents who filled out the questionnaire, 30 (46%) were aged between 30 to 39, 33 (51%) were males, and 31 (48%) held bachelor's degree. Additionally, 55 (85%) of the respondents were employed, with 39 (61%) having a monthly salary of more than 4000MYR. For the race, 45 (69%) were Malay ethnic group. The physical activity behavior description of the respondents is shown in the table too. While 22 (33%) were exercising 5 times a week, 14 (22%) exercised 3 times a week. The duration of the exercise ranged from 31 minutes to 60 minutes for 36 (55%) of the respondents. For the mobile fitness app characteristics, 61(94%) of the respondents use the free version of the mobile fitness app. 35% of the respondents use the mobile fitness app from 4 to 5 times per week. Furthermore, 35 (54%) of the respondents have never temporarily discontinued the mobile fitness app in the last 6 months.

Table 1
Respondents' Demographic Information for the Pilot Survey

Demographic	**Frequency(n) N=65**	**Percentage (%)**
Age		
18-23	2	3%
24-29	5	8%
30-39	30	46%
40-49	20	31%
>=50	8	12%
Gender		
Male	33	51%
Female	32	49%
Educational Level		
Bachelor's degree	31	48%
Diploma	11	17%
Graduate degree	17	26%
High school degree	3	4%
Some college/vocational school/ apprenticeship	3	4%
Employment Status		
Employed	55	85%
Retired	2	3%
Self-employed	4	6%
Student	3	5%
Unemployed	1	2%
Monthly Income		
Below MYR 1000	1	9%
More than MYR 4000	39	60%
MYR 1001- MYR 2000	2	3%

MYR 2001- MYR 3000	5	8%
MYR 3001- MYR 4000	12	18%
Not Mentioned	1	2%
Race/Ethnicity		
Chinese	10	15%
Indian	5	8%
Malay	45	69%
Other	5	8%
Physical Activity Behavior		
Weekly Exercise		
1 time	1	2%
2 times	8	12%
3 times	14	22%
4 times	10	15%
5 times	22	33%
6 times	4	6%
7 times	3	5%
More than 7 time	3	5%
Duration of Exercise per Day		
15 mins to 25 mins	3	5%
30 mins	4	6%
31 mins to 60 mins	36	55%
Less than 15 mins	3	5%
More than 60 mins	19	29%
Paid		
Use only free version of app	61	94%
<$5	1	2%
$10 – $20	2	3%
>$20	1	2%
Mobile fitness app Experience		
1-2 years	10	15%
2-3 years	15	23%
3-4 years	13	20%
Less than a year	5	8%
More than 5 years	22	33%
Mobile fitness app Usage Frequency		
2 times to 3 times	14	22%
4 times to 5 times	23	35%
6 times to 7 times	14	22%
More than 7 times per week	13	20%
Once	1	2%
Temporary discontinuance of the mobile fitness app in the last 6 months		
1-2 weeks	7	11%
1-3 months	3	5%
2-4 weeks	5	8%
Less than a week	15	23%
Never	35	54%

Construct Reliability

The data of the pilot study has been analysed based on Structural Equation Modelling (SEM) with Partial Least Squares (PLS). The SmartPLS version 4.0.9.2 was employed for the data analysis. In the pilot study, the validity and reliability of the questionnaire were assessed. In order to assess a measurement model, it's important to evaluate the reliability of the measures being applied (including indicator reliability and internal consistency reliability) as well as their validity (including convergent and discriminant validity) (Hair et al., 2017).

The initial step in assessing the measurement model involves examining the loadings of the indicators. Ideally, the loadings should be above 0.708, indicating that the construct being measured explains more than 50% of the variance in the indicator, which suggests good item reliability (Ali et al., 2018; Hair et al., 2019). As shown in Table 2, the indicator loadings of all the indicators are higher than 0.708, except for COM_4 and CTC_3. COM_4 and CTC_3 items have indicator loadings of 0.60, and 0.57, respectively, which do not meet the criterion of the reliable indicator (Hair et al., 2019). That gives the indication of the low reliability of these items, so COM_4 and CTC_3 should be considered for removal.

After that, internal consistency reliability was assessed using composite reliability, and Cronbach alpha. High values indicate a higher level of reliability. Therefore, the values above 0.7 are "good satisfactory" (Fornell & Larcker, 1981). According to the analysis, the Composite Reliability (CR) of the constructs are above 0.7. Moreover, Cronbach alpha has been assessed for the constructs, all values ranged from 0.7 to 0.948 which indicates "good satisfactory" reliability.

Table 2
Measurement Items and Reliability for Constructs

NO	Construct	Item Code	Items	Indicator Loading	References	Assessment Results	
						Inclusion	Exclusion Criteria
1	**CON** (4 items), CA: 0.89, CR: 0.923, AVE:0.856	CON_1	I intend to continue using Fitness app rather than discontinue its use.	0.73	(Bhattacherjee, 2001) (Cheng et al., 2021) (Wang, 2015)	✔	
2		CON_2	My intentions are to continue using Fitness app than use any alternative means	0.88		✔	
3		CON_3	If possible, I would like to continue my use of Fitness app	0.95		✔	
4		CON_4	In the near future, I intend to continue using Fitness app.	0.91		✔	

5	**COM** (5 items), CA: 0.874, CR: 0.905, AVE: 0.667	COM_1	My experience with using Fitness app was better than what I expected.	0.86	(Bhattacherjee, 2001) (Huang et al., 2019) (Zheng, 2019)	✔	
6		COM_2	The features provided by Fitness app were better than what I expected.	0.92		✔	
7		COM_3	The benefits provided by the Fitness app were better than what I expected.	0.90		✔	
8		COM_4	The problems encountered with Fitness app were less serious than I expected	**0.60***		✗	Loading <0.7
9		COM_5	Overall, most of my expectations from using Fitness app were confirmed.	0.77		✔	
10	**CSAT** (5 items), CA: 0.851, CR: 0.872, AVE: 0.758	CSAT_1	My choice to use Fitness app was a wise one.	0.85	(Kim et al., 2018)	✔	
11		CSAT_2	I think I did the right thing in using the Fitness app.	0.88		✔	
12		CSAT_4	The Fitness app is close to my ideal app.	0.76		✔	
13		CSAT_5	The fitness app meets the functions that I needed.	0.84		✔	
14	**ESAT** (6 items), CA: 0.933, CR: 0.961, AVE: 0.747	ESAT_1	I feel interested in using the Fitness app.	0.80	(Kim et al., 2018) (Zheng, 2019)	✔	
15		ESAT_2	I feel cheerful in using the Fitness app	0.90		✔	
16		ESAT_3	I feel happy in using the Fitness app	0.94		✔	
17		ESAT_4	I feel excited in using the Fitness app	0.85		✔	

18		ESAT_5	I enjoy using the Fitness app.	0.91		✔	
19		ESAT_6	I feel active in using the Fitness app	0.79		✔	
20	**AHL** (4 items), CA: 0.944, CR: 0.944, AVE: 0.856	AHL_1	As long as I am spending good time with my family members and friends, it's OK to temporarily stop using a Fitness app.	0.94	(Siponen et al., 2020)	✔	
21		AHL_2	As long as I am carrying out important work, it's OK to temporarily stop using a Fitness app.	0.94		✔	
22		AHL_3	As long as I am in vacation, it's OK to temporarily stop using a Fitness app.	0.89		✔	
23		AHL_4	As long as I am studying, it's OK to temporarily stop using a Fitness app.	0.93		✔	
24	**CTC** (3 items), CA: 0.754, CR: 0.764, AVE: 0.674	CTC_1	It is OK to temporarily stop using a Fitness app because everyone does not use Fitness apps at times.	0.93	(Siponen et al., 2020)	✔	
25		CTC_2	It is OK to temporarily stop using a Fitness app because everybody else is doing it.	0.89		✔	
26		CTC_3	It is OK to temporarily stop using a Fitness app if the doctor/physician did not recommend it.	**0.57***		✗	Loading <0.7 and Discriminant Validity

27	**DOI** (4 items), CA: 0.948, CR: 0.953, AVE: 0.868	DOI_1	It is OK to temporarily stop using a Fitness app if no harm is done.	0.94	(Siponen et al., 2020)	✔	
28		DO_2	It is OK to temporarily stop using a Fitness app if no damage is done to my health	0.95		✔	
29		DOI_3	It is OK to temporarily stop using a Fitness app if no one gets hurt.	0.95		✔	
30		DOI_4	It is OK to temporarily stop using a Fitness app if I am young and healthy.	0.88		✗	Discriminant Validity
31	**MOL** (3 items), CA: 0.929, CR: 0.929, AVE: 0.876	MOL_1	I feel my general adherence to a healthy lifestyle compensates for occasionally stopping using Fitness apps.	0.91	(Silic et al., 2017)	✔	
32		MOL_2	I feel my good health compensates for occasionally stopping using Fitness apps.	0.95		✔	
33		MOL_3	I feel my daily exercise, compensates for occasionally stopping using Fitness apps.	0.94		✔	
34	**SWI** (3 items), CA: 0.7, CR: 0.72, AVE: 0.611	SWI_1	If I transfer to a new Fitness app, it will take me much more time and energy to learn and become familiar to the new Fitness App.	0.84	(Ye et al., 2019) (Aydin et al., 2005)	✔	
35		SWI_2	Switching to a new Fitness app causes monetary cost.	0.77		✔	
36		SWI_3	If I switch to a new Fitness	0.73		✔	

			app, the features offered by the new Fitness app might not work as well as expected.				
	GUL (6 items), CA: 0.938, CR: 0.967, AVE: 0.758		This list consists of a number of words and phrases that describe different feelings and emotions you may have had when you temporarily discontinue the Fitness app. Please reflect on your last temporary discontinuance of the Fitness app, and indicate to what extent you have felt this way:		(Turel, 2017)	✔	
37		GUL_1	Guilty	0.83		✔	
38		GUL_2	Ashamed	0.89		✔	
39		GUL_3	Blameworthy	0.91		✔	
40		GUL_4	Angry at self	0.87		✔	
41		GUL_5	Disgusted with self	0.87		✔	
42		GUL_6	Dissatisfied with self	0.87		✔	

AVE: Average variance extracted, ***CR***: Composite Reliability, ***CA***: Cronbach's alpha, ***CON***: Continuance Intention to Use Fitness App, ***COM***: Confirmation, ***CSAT***: Cognitive Satisfaction, ***ESAT***: Emotional Satisfaction, ***AHL***: Appeal to Higher Loyalties, ***DOI,*** Denial of Injury, ***CTC***: Condemning the Condemners, ***MOL***: Metaphor of Ledger, ***GUL***: Guilt of temporary Discontinuance the mobile fitness apps , ***SWI***: Switching Costs
*Factor loadings of items that are less than 0.7 which does not meet required criterion.

Convergent Validity

The third step is to assess the convergent validity of each construct measure. Convergent validity refers to how well a construct can explain the variance shared by its items. To assess a construct's convergent validity, the average variance extracted (AVE) metric is employed, which involves squaring the loading of each indicator on the construct and then calculating the mean value. An AVE of 0.50 or higher is considered acceptable, as this suggests that the construct can explain at least 50% of the variance in its items (Hair et al., 2019). As shown in Table 2, AVE values are higher than 0.5 which means acceptable values.

Discriminant Validity

To assess the discriminant validity of the instrument, three crucial tests must be conducted. Firstly, cross-factor loadings are examined to ensure that all item loadings onto other latent variables are lower than the factor loading for their intended construct (Hair et al., 2011). Discriminant validity is established when items exhibit higher loadings on their intended construct compared to other constructs in the model. The difference between the factor loading for the intended construct and the second-highest cross-factor loading should be at least 0.10 (Vance et al., 2012). As shown in **Error! Not a valid bookmark self-reference.**, all factor loadings for their respective constructs exceed the cross-factor loadings. However, for measurement item CTC_3, the factor loading is 0.724 for the "Condemning the Condemners" construct, while the cross-factor loading for "Denial of Injury" is 0.662. Additionally, for DOI_4, the factor loading is 0.846, whereas the cross-factor loading for the "Condemning the Condemners" construct is 0.757. These findings indicate that CTC_3 and DOI_4 should be considered for removal from the instrument to enhance discriminant validity.

Table 3
Loading and Cross-Factor Loading of the Constructs

	AHL	COM	CON	CSAT	CTC	DOI	ESAT	GUL	MOL	SWI
AHL_1	**0.946**	-0.148	-0.245	-0.258	0.71	0.707	-0.23	-0.232	0.598	0.067
AHL_2	**0.94**	-0.165	-0.231	-0.257	0.69	0.708	-0.286	-0.22	0.586	0.01
AHL_3	**0.89**	-0.262	-0.294	-0.334	0.74	0.742	-0.277	-0.229	0.669	0.018
AHL_4	**0.922**	-0.055	-0.112	-0.193	0.79	0.779	-0.182	-0.288	0.701	0.008
COM_1	-0.097	**0.862**	0.536	0.634	-0.036	-0.074	0.462	-0.01	-0.045	-0.009
COM_2	-0.06	**0.913**	0.377	0.621	-0.038	-0.036	0.327	-0.102	-0.008	-0.094
COM_3	-0.113	**0.9**	0.424	0.574	-0.086	-0.055	0.426	0.011	-0.041	0.005
COM_5	-0.313	**0.778**	0.492	0.642	-0.296	-0.273	0.358	0.076	-0.3	-0.036
CON_1	-0.218	0.306	**0.725**	0.361	-0.215	-0.299	0.305	0.049	-0.218	-0.196
CON_2	-0.225	0.512	**0.876**	0.694	-0.129	-0.181	0.493	0.22	-0.174	-0.019
CON_3	-0.235	0.488	**0.945**	0.693	-0.153	-0.256	0.472	0.182	-0.293	-0.113
CON_4	-0.161	0.5	**0.91**	0.669	-0.129	-0.146	0.48	0.224	-0.196	-0.116
CSAT_1	-0.264	0.545	0.68	**0.845**	-0.25	-0.225	0.317	0.126	-0.244	-0.119
CSAT_2	-0.141	0.686	0.749	**0.875**	-0.14	-0.147	0.336	0.167	-0.178	-0.163
CSAT_3	-0.223	0.577	0.348	**0.757**	-0.091	-0.144	0.423	0.044	-0.115	0.063
CSAT_4	-0.335	0.564	0.532	**0.838**	-0.151	-0.232	0.372	0.126	-0.18	-0.047
CTC_1	0.596	0.019	-0.038	-0.129	**0.852**	0.736	-0.031	-0.334	0.554	0.08
CTC_2	0.682	-0.098	-0.211	-0.184	**0.88**	0.721	0.038	-0.265	0.728	0.2
CTC_3	0.678	-0.257	-0.162	-0.163	**0.724**	0.662	-0.107	-0.096	0.647	-0.021
DOI_1	0.774	-0.129	-0.231	-0.232	0.838	**0.955**	-0.171	-0.298	0.714	0.119
DOI_2	0.761	-0.147	-0.223	-0.233	0.798	**0.959**	-0.213	-0.293	0.74	0.056
DOI_3	0.752	-0.177	-0.24	-0.29	0.812	**0.961**	-0.146	-0.259	0.684	0.075
DOI_4	0.671	-0.01	-0.197	-0.062	0.757	**0.846**	-0.167	-0.393	0.687	0.128
ESAT_1	-0.168	0.501	0.628	0.492	0.024	-0.13	**0.799**	0.191	-0.043	-0.057
ESAT_2	-0.299	0.338	0.306	0.224	-0.104	-0.232	**0.895**	0.212	-0.017	0.063
ESAT_3	-0.276	0.428	0.389	0.313	-0.063	-0.184	**0.94**	0.206	0.006	0.067
ESAT_4	-0.173	0.298	0.204	0.205	0.013	-0.093	**0.845**	0.105	0.1	0.06
ESAT_5	-0.234	0.405	0.49	0.419	-0.073	-0.22	**0.907**	0.123	0.002	0
ESAT_6	-0.22	0.291	0.442	0.405	0.008	-0.097	**0.79**	0.067	0.002	0.134
GUL_1	-0.296	0.154	0.251	0.251	-0.327	-0.398	0.265	**0.834**	-0.294	-0.341
GUL_2	-0.212	-0.068	0.164	0.062	-0.25	-0.291	0.137	**0.887**	-0.265	-0.215

	AHL	COM	CON	CSAT	CTC	DOI	ESAT	GUL	MOL	SWI
GUL_3	-0.271	-0.037	0.239	0.14	-0.246	-0.276	0.184	**0.915**	-0.249	-0.144
GUL_4	-0.103	-0.03	0.09	0.083	-0.16	-0.175	0.097	**0.86**	-0.149	-0.18
GUL_5	-0.196	-0.096	0.065	0.084	-0.211	-0.249	0.103	**0.86**	-0.263	-0.123
GUL_6	-0.199	-0.042	0.178	0.066	-0.226	-0.241	0.066	**0.867**	-0.178	-0.154
MOL_1	0.702	-0.041	-0.222	-0.164	0.725	0.731	-0.058	-0.251	**0.92**	-0.039
MOL_2	0.637	-0.139	-0.261	-0.236	0.75	0.701	0.034	-0.253	**0.956**	0.153
MOL_3	0.603	-0.144	-0.218	-0.217	0.731	0.695	0.024	-0.281	**0.932**	0.131
SWI_1	0.039	-0.053	-0.104	-0.047	0.03	0.088	-0.026	-0.359	0.058	**0.842**
SWI_3	-0.059	0.015	-0.098	-0.08	0.1	0.02	0.102	0.038	0.043	**0.771**
SWI_4	0.163	-0.076	-0.044	-0.134	0.209	0.194	0.008	-0.282	0.151	**0.729**

The second test for assessing discriminant validity employs the Fornell-Larcker criterion (Fornell & Larcker, 1981). In a PLS path model, the purpose of this test is to confirm that a reflective construct displays stronger associations with its indicators compared to indicators of other constructs, thus establishing its uniqueness from other constructs in the model (Jr et al., 2022). As shown in Table 4, the correlations between the constructs are consistently lower than the square root of the Average Variance Extracted (AVE) in the upper triangle values.

In partial least squares (PLS) and other variance-based structural equation modeling techniques, two primary methods are employed to assess discriminant validity: the Fornell-Larcker criterion and the evaluation of cross-loadings. While these approaches are widely used for assessing discriminant validity in PLS path models, it is important to note that in certain research scenarios, relying solely on these methods may not provide a comprehensive understanding of whether reflective constructs are genuinely distinct from one another in a given model (Henseler et al., 2015).

Table 4
Heterotrait-monotrait (HTMT) Scores (lower triangle) with the Square Root of Average Variance Extracted (AVE) on the Diagonal and Discriminant validity of the measurement model: Fornell-Larcker criterion upper triangle.

	AHL	CSAT	CTC	COM	CON	DOI	ESAT	GUL	MOL	SWI
AHL	**0.925**	-0.281	0.698	-0.169	-0.237	0.785	-0.254	-0.284	0.692	0.028
CSAT	0.324	**0.83**	-0.172	0.717	0.719	-0.259	0.48	0.16	-0.22	-0.096
CTC	0.791	0.203	**0.821**	-0.047	-0.142	0.771	0.012	-0.332	0.703	0.156
COM	0.197	0.82	0.143	**0.865**	0.533	-0.156	0.483	0.016	-0.115	-0.038
CON	0.265	0.772	0.177	0.583	**0.868**	-0.239	0.579	0.241	-0.251	-0.115
DOI	0.819	0.285	**0.867***	0.167	0.271	**0.932**	-0.184	-0.306	0.734	0.086
ESAT	0.289	0.549	0.083	0.526	0.619	0.197	**0.864**	0.197	-0.013	0.029
GUL	0.298	0.169	0.382	0.113	0.249	0.315	0.203	**0.871**	-0.285	-0.251
MOL	0.737	0.243	0.802	0.152	0.279	0.773	0.064	0.303	**0.936**	0.087
SWI	0.138	0.167	0.249	0.104	0.16	0.147	0.105	0.38	0.179	**0.782**

*Indicates HTMT> 0.85

To further assess discriminant validity in this study, the Heterotrait-Monotrait (HTMT) ratio of correlations was employed as the third stage of analysis. HTMT was initially introduced by Henseler et al (2015) as a means to address discriminant validity concerns. HTMT is calculated by averaging the correlations between items across different constructs and dividing this by the (geometric) mean of the average correlations for items measuring the same construct. High HTMT values indicate potential issues with discriminant validity (Hair et al., 2019).

In accordance with HTMT threshold guidelines, it's acceptable for HTMT values to exceed 0.90 in models where constructs are conceptually similar. However, when dealing with conceptually distinct constructs, it's advisable to adopt a more conservative approach, and a suggested threshold is set at 0.85 (Franke & Sarstedt, 2019). The choice of a specific HTMT cutoff value should depend on the similarity of the constructs under investigation and the researcher's desired level of caution in assessing discriminant validity.

In this pilot study, an HTMT analysis was conducted, and the results are detailed in the lower triangle values in Table 4. As displayed in Table 4, all HTMT values are below 0.85, indicating satisfactory discriminant validity. However, the HTMT value between the "Condemning the Condemners" and "Denial of Injury" constructs slightly exceeds this threshold, with an HTMT value of 0.861.

The concepts of "Condemning the Condemners" and "Denial of Injury" share a notable similarity, as both involve neutralization techniques. "Condemning the Condemners" entails criticizing those who disapprove of deviant behavior, while "Denial of Injury" involves minimizing or denying the harm caused by such behavior. Neutralization, in this context, refers to the cognitive strategies individuals employ to rationalize or excuse their deviant actions. Given their conceptual resemblance, it is reasonable to anticipate that these two constructs might exhibit a higher correlation in the model. Therefore, an HTMT value of 0.861 between these two constructs may not necessarily signify a discriminant validity issue in this particular case, as suggested by (Hair et al., 2019). Additionally, it is essential to consider that the study type and sample size can influence the results (Franke & Sarstedt, 2019). With the instrument's reliability and validity confirmed, the updated instrument is ready for use in the main data collection phase.

Conclusion

This pilot study aimed to assess the reliability and validity of a questionnaire designed to investigate factors influencing individuals' intention to continue using mobile fitness apps. The findings of this pilot study provided promising results, indicating good reliability and construct validity of the questionnaire. The pilot study sample consisted of 65 respondents from Malaysia, with a diverse range of demographic characteristics and mobile fitness app usage patterns. The results of the reliability analysis demonstrated the internal consistency of the questionnaire, suggesting that the items consistently measured the intended constructs. Additionally, the construct validity analysis provided evidence that the questionnaire items effectively captured the targeted aspects of the intended constructs. This pilot study serves as an important foundation for future research on the topic of continuance intention to use mobile fitness apps. The validated measurement instrument developed in this study can now be confidently employed in larger-scale studies to explore the factors influencing individuals' decisions to continue using mobile fitness apps. The findings of such studies could shed light on ways to improve user engagement and reduce dropout rates in the realm of health and mobile fitness app usage.

References

Ali, F., Rasoolimanesh, S. M., Sarstedt, M., Ringle, C. M., & Ryu, K. (2018). An assessment of the use of partial least squares structural equation modeling (PLS-SEM) in hospitality research. *International Journal of Contemporary Hospitality Management, 30*(1), 514-538. https://doi.org/10.1108/ijchm-10-2016-0568

Aljabali, R. N., & Ahmad, N. (2022). Profile of Fitness and Diet App Users in Online Wellness Social Communities in Malaysia: Cross-Sectional Pilot Study. The International Conference of Advanced Computing and Informatics.

Amagai, S., Pila, S., Kaat, A. J., Nowinski, C. J., & Gershon, R. C. (2022, Apr 26). Challenges in Participant Engagement and Retention Using Mobile Health Apps: Literature Review. *J Med Internet Res, 24*(4), e35120. https://doi.org/10.2196/35120

Aydin, S., Ozer, G., & Arasil, O. (2005). Customer loyalty and the effect of switching costs as a moderator variable. *Marketing Intelligence & Planning, 23*(1), 89-103. https://doi.org/10.1108/02634500510577492

Bhattacherjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly, 25*(3), 351-370. https://doi.org/10.2307/3250921

Cheng, L. K., Huang, H.-L., & Lai, C.-C. (2021). Continuance intention in running apps: the moderating effect of relationship norms. *International Journal of Sports Marketing and Sponsorship, ahead-of-print*(ahead-of-print). https://doi.org/10.1108/ijsms-08-2020-0143

Ding, Y. (2019). Looking forward: The role of hope in information system continuance. *Computers in Human Behavior, 91*, 127-137. https://doi.org/10.1016/j.chb.2018.09.002

Flake, J. K., Davidson, I. J., Wong, O., & Pek, J. (2022). Construct validity and the validity of replication studies: A systematic review. *American Psychologist, 77*, 576-588. https://doi.org/10.1037/amp0001006

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research, 18*(1), 39-50.

Franke, G., & Sarstedt, M. (2019). Heuristics versus statistics in discriminant validity testing: A comparison of four procedures. *Internet Research, 29*(3), 430-447. https://doi.org/10.1108/IntR-12-2017-0515

Hair, J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. SAGE Publications. https://books.google.com.my/books?id=JDWmCwAAQBAJ

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing theory and Practice, 19*(2), 139-152.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2-24. https://doi.org/10.1108/ebr-11-2018-0203

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2022). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer Nature.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135. https://doi.org/10.1007/s11747-014-0403-8

Huang, C.-K., Chen, S.-H., Tang, C.-P., & Huang, H.-Y. (2019). A trade-off dual-factor model to investigate discontinuous intention of health app users: From the perspective of information disclosure. *Journal of Biomedical Informatics, 100*. https://doi.org/10.1016/j.jbi.2019.103302

Huang, G., & Ren, Y. (2020). Linking technological functions of fitness mobile apps with continuance usage among Chinese users: Moderating role of exercise self-efficacy. *Computers in Human Behavior, 103*, 151-160. https://doi.org/https://doi.org/10.1016/j.chb.2019.09.013

Kim, M.-K., Park, M.-C., Park, J.-H., Kim, J., & Kim, E. (2018). The role of multidimensional switching barriers on the cognitive and affective satisfaction-loyalty link in mobile communication services: Coupling in moderating effects. *Computers in Human Behavior, 87*, 212-223. https://doi.org/10.1016/j.chb.2018.05.024

Kwon, J. Y., Lee, J. S., & Park, T. S. (2022). Analysis of Strategies to Increase User Retention of Fitness Mobile Apps during and after the COVID-19 Pandemic. *Int J Environ Res Public Health, 19*(17). https://doi.org/10.3390/ijerph191710814

Li, A., Sun, Y., Guo, X., Guo, F., & Guo, J. (2021). Understanding how and when user inertia matters in fitness app exploration: A moderated mediation model. *Information Processing & Management, 58*(2), 102458. https://doi.org/10.1016/j.ipm.2020.102458

Lowe, N. K. (2019). What Is a Pilot Study? *J Obstet Gynecol Neonatal Nurs, 48*(2), 117-118. https://doi.org/10.1016/j.jogn.2019.01.005

Research, G. V. (2023). *mHealth Apps Market Size, Share & Trends Analysis Report By Type (Fitness, Medical), By Region (North America, Europe, Asia Pacific, Latin America, Middle East & Africa), And Segment Forecasts, 2023 - 2030*. Retrieved 25-03-2023 from https://www.grandviewresearch.com/industry-analysis/mhealth-app-market

Silic, M., Barlow, J. B., & Back, A. (2017). A new perspective on neutralization and deterrence: Predicting shadow IT usage. *Information & Management, 54*(8), 1023-1037. https://doi.org/10.1016/j.im.2017.02.007

Siponen, M., Puhakainen, P., & Vance, A. (2020). Can individuals' neutralization techniques be overcome? A field experiment on password policy. *Computers & Security, 88*. https://doi.org/10.1016/j.cose.2019.101617

Soliman, W., & Rinta-Kahila, T. (2020). Toward a refined conceptualization of IS discontinuance: Reflection on the past and a way forward. *Information & Management, 57*(2), 103167. https://doi.org/10.1016/j.im.2019.05.002

Statista. (2021). *Global health and fitness app downloads as of Q2 2020*. https://www.statista.com/statistics/1127248/health-fitness-apps-downloads-worldwide/

Sykes, G. M., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American sociological review, 22*(6), 664-670.

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *How to test the validation of a questionnaire/survey in a research (August 10, 2016)*.

Taherdoost, H. (2019). What is the best response scale for survey and questionnaire design; review of different lengths of rating scale/attitude scale/Likert scale. *Hamed Taherdoost*, 1-10.

Turel, O. (2017). Untangling the complex role of guilt in rational decisions to discontinue the use of a hedonic Information System. *European Journal of Information Systems, 25*(5), 432-447. https://doi.org/10.1057/s41303-016-0002-5

Van Teijlingen, E., & Hundley, V. (2001). The importance of pilot studies. *Social research update*(35), 1-4.

Vance, A., Siponen, M., & Pahnila, S. (2012). Motivating IS security compliance: Insights from Habit and Protection Motivation Theory. *Information & Management, 49*(3-4), 190-198. https://doi.org/10.1016/j.im.2012.04.002

Vinnikova, A., Lu, L. D., Wei, J. C., Fang, G. B., & Yan, J. (2020, Oct). The Use of Smartphone Fitness Applications: The Role of Self-Efficacy and Self-Regulation. *International Journal of Environmental Research and Public Health, 17*(20). https://doi.org/10.3390/ijerph17207639

Wang, K. (2015). Determinants of mobile value-added service continuance: The mediating role of service experience. *Information & Management, 52*(3), 261-274. https://doi.org/10.1016/j.im.2014.11.005

Ye, Q., Luo, Y., Chen, G., Guo, X., Wei, Q., & Tan, S. (2019, 2019/02/01). Users Intention for Continuous Usage of Mobile News Apps: the Roles of Quality, Switching Costs, and Personalization. *Journal of Systems Science and Systems Engineering, 28*(1), 91-109. https://doi.org/10.1007/s11518-019-5405-0

Yin, Q., Li, L., Yan, Z., & Guo, C. (2021). Understanding the effects of self-peer-platform incentives on users' physical activity in mobile fitness apps: the role of gender. *Information Technology & People, ahead-of-print*(ahead-of-print). https://doi.org/10.1108/itp-10-2020-0705

Yousaf, A., Mishra, A., & Gupta, A. (2021). 'From technology adoption to consumption': Effect of pre-adoption expectations from fitness applications on usage satisfaction, continual usage, and health satisfaction. *Journal of Retailing and Consumer Services, 62*, 102655. https://doi.org/10.1016/j.jretconser.2021.102655

Yusoff, M. S. B. (2019). ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal, 11*(2), 49-54. https://doi.org/10.21315/eimj2019.11.2.6

Zhang, X. X., & Xu, X. G. (2020, Sep). Continuous use of fitness apps and shaping factors among college students: A mixed-method investigation. *International Journal of Nursing Sciences, 7*, 580-587. https://doi.org/10.1016/j.ijnss.2020.07.009

Zheng, L. (2019). The role of consumption emotions in users' mobile gaming application continuance intention. *Information Technology & People, 33*(1), 340-360. https://doi.org/10.1108/itp-04-2018-0197