Examining construct and predictive validity of the Health-IT Usability Evaluation Scale: confirmatory factor analysis and structural equation modeling results

Po-Yin Yen,[1] Karen H Sousa,[2] Suzanne Bakken[3]

[1]Department of Biomedical Informatics, The Ohio State University, Columbus, Ohio, USA
[2]College of Nursing, University of Colorado, Aurora, Colorado, USA
[3]School of Nursing and Department of Biomedical Informatics, Columbia University, New York, New York, USA

Correspondence to
Dr Po-Yin Yen, Department of Biomedical Informatics, The Ohio State University, Columbus, 333 W 10th Ave, Columbus, OH 43210, USA; Po-Yin.Yen@osumc.edu

ABSTRACT

Background In a previous study, we developed the Health Information Technology Usability Evaluation Scale (Health-ITUES), which is designed to support customization at the item level. Such customization matches the specific tasks/expectations of a health IT system while retaining comparability at the construct level, and provides evidence of its factorial validity and internal consistency reliability through exploratory factor analysis.

Objective In this study, we advanced the development of Health-ITUES to examine its construct validity and predictive validity.

Methods The health IT system studied was a web-based communication system that supported nurse staffing and scheduling. Using Health-ITUES, we conducted a cross-sectional study to evaluate users' perception toward the web-based communication system after system implementation. We examined Health-ITUES's construct validity through first and second order confirmatory factor analysis (CFA), and its predictive validity via structural equation modeling (SEM).

Results The sample comprised 541 staff nurses in two healthcare organizations. The CFA (n=165) showed that a general usability factor accounted for 78.1%, 93.4%, 51.0%, and 39.9% of the explained variance in 'Quality of Work Life', 'Perceived Usefulness', 'Perceived Ease of Use', and 'User Control', respectively. The SEM (n=541) supported the predictive validity of Health-ITUES, explaining 64% of the variance in intention for system use.

Conclusions The results of CFA and SEM provide additional evidence for the construct and predictive validity of Health-ITUES. The customizability of Health-ITUES has the potential to support comparisons at the construct level, while allowing variation at the item level. We also illustrate application of Health-ITUES across stages of system development.

INTRODUCTION

Health information technology (health IT) is 'the application of information processing involving both computer hardware and software that deals with the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making.'[1] It offers important benefits to healthcare, including decision support, knowledge management, improved communication, effective resource management, reduction of medical errors, time saving, and paperwork reduction.[2] However, usability factors are one of the major obstacles to health IT adoption. These factors include ease of use, usefulness, flexibility, relevancy, and completeness.[3 4] If usability is not considered, health IT could introduce unintended, negative consequences, such as increased medical errors,[5 6] and problems with communication between healthcare providers.[5]

The International Organization for Standardization (ISO) defines usability as 'the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.'[7] A key indicator of technology usability is user satisfaction, which is 'the opinion of the user about a specific computer application, which they use'[8] and is a critical measure of IT success.[9] 'Ease of use' and 'usefulness' are two constructs commonly included in several user satisfaction instruments.[8 10 11] This overlaps with the main concept of technology acceptance, in which a technology is perceived to be easy to use and useful.[12 13] Studies have demonstrated a close relationship between user satisfaction and technology acceptance,[14–17] and their roles in the evaluation of health IT usability.[18–20] In summary, user satisfaction and/or technology acceptance can be considered valid measures for usability evaluation.

A number of instruments have been designed and applied to measure user perceptions of health IT usability, such as the IBM Computer System Usability Questionnaire,[21] Technology Acceptance Model (TAM) Perceived Usefulness/Ease of Use,[12] Unified Theory of Acceptance and Use of Technology (UTAUT),[22] Questionnaire for User Interaction Satisfaction (QUIS),[10] Physician Order Entry User Satisfaction and Usage Survey,[11] and End-User Computing Satisfaction.[8]

Although validated instruments exist, a mismatch between study needs and concepts measured in the questionnaires often requires item addition, deletion, or modification without standardization.[23] While such an approach is useful in meeting the needs of a particular study, it limits aggregation of findings across studies. In addition, questionnaires frequently fail to explicate tasks in the questionnaire items, even though the relationship between the task and the health IT is essential for health IT usability evaluation.[24 25] Failure to consider tasks or level of expectations may lead to poor technology adoption, while the concept of job performance varies across authors. Holden and Karsh[23] found that 'it was unclear from the definition whether usefulness referred to enhanced performance process (eg, fewer steps, more

information for decision making) or enhanced performance outcomes (eg, faster care, more accurate decisions).' Several usability methods and human factors approaches have emphasized that 'task' is essential in usability testing and should be considered during IT implementation or evaluation.[24 26–30]

To address these knowledge gaps, we developed a customizable Health Information Technology Usability Evaluation Scale (Health-ITUES), which explicitly considers task by addressing various levels of expectation of support for the task by the health IT, and provided evidence of its factorial validity and internal consistency through exploratory factor analysis (EFA).[31] In this study, we advanced the development of Health-ITUES by applying the methods of confirmatory factor analysis (CFA) and structural equation modeling (SEM) to examine its construct validity and predictive validity. We also illustrated its application across stages of system development.

BACKGROUND

We designed and evaluated the Health-ITUES items within the context of a web-based communication system for scheduling nursing staff that supports both nurse manager and staff nurse tasks. The web-based communication system is designed to improve the efficiency and effectiveness of the staffing and scheduling processes. The system provides functionality for nurse managers to announce open shifts throughout their organization and for staff nurses to request shifts for which they are qualified based upon their profile. If more than one nurse requests the same open shift, nurse managers are able to select a nurse based on her/his experience or working hours (not exceeding hospital overtime policy) for patient safety purposes.

The iterative development of Health-ITUES included conceptual mapping of the proposed items to the subjective measures of the Health IT Usability Evaluation Model (Health-ITUEM). Previously, we developed this integrated model based on multiple theories to include both subjective and objective measures for usability evaluation,[32] and consideration of items from existing questionnaires including TAM measurements of perceived usefulness and perceived ease of use[12] and the IBM Computer System Usability Questionnaire.[21] This process is described in detail elsewhere.[31]

Based upon the principle that usability is measured through the interaction of *user*, *tool*, and *task* in a *specified setting*,[7 33] we modified items to address the web-based communication system and specific user tasks. For example, to modify an original TAM question, 'Using [system] is useful in my job,' we identified the system by name and also specified user tasks. The resulting questions were 'Using *Bidshift* (system) is useful for *requesting shifts* (task)' for staff nurses. Also of note, in contrast to most satisfaction measures that report general information which cannot identify specific usability problems,[34] Health-ITUES items address different levels of expectation. These include: (1) task level ('I am satisfied with Bidshift for requesting open shifts'), (2) individual level ('The addition of BidShift has improved my job satisfaction'), and (3) organizational level ('BidShift technology is an important part of our staffing process').

The original Health-ITUES consisted of 36 items[35] rated on a 5-point Likert scale from strongly disagree to strongly agree: 'actual usage' (2 items), 'intention to use' (1 item), 'satisfaction' (5 items), 'perceived usefulness' (6 items), 'perceived ease of use' (3 items), 'perceived performance speed' (2 items), 'learnability' (2 items), 'competency' (2 items), 'flexibility/customizability' (3 items), 'memorability' (2 items), 'error prevention' (2 items), 'information needs' (3 items), and 'other outcomes' (3 items). A higher scale value indicates higher perceived usability of the technology.

Table 1 The 20 items of the Health IT Usability Evaluation Scale (Health-ITUES)

Item
Quality of work life (Cronbach's α=.94)
34   I think [BidShift] has been [a positive addition to nursing]
33   I think BidShift has been [a positive addition to our organization]
35   [BidShift technology] is an important part of [our staffing process]
Perceived usefulness (Cronbach's α=.94)
29   Using [Bidshift] makes it easier to [request the shift I want]
26   Using [Bidshift] enables me to [request shifts] more quickly
28   Using [Bidshift] makes it more likely that I [will be awarded a shift that I request]
30   Using [Bidshift] is useful for [requesting open shifts]
25   I think [Bidshift] presents a more equitable process for [requesting open shifts]
31   I am satisfied with [Bidshift] for [requesting open shifts]
21   I [am awarded shifts] in a timely manner because of [Bidshift]
27   Using [Bidshift] increases [requesting open shifts]
14   I am able to [find shifts that I am qualified to work] whenever I use [Bidshift]
Perceived ease of use (Cronbach's α=.95)
5   I am comfortable with my ability to use [Bidshift]
4   Learning to operate [Bidshift] is easy for me
6   It is easy for me to become skillful at using [Bidshift]
22   I find [Bidshift] easy to use
10   I can always remember how to log on to and use [Bidshift]
User control (Cronbach's α=.81)
12   [Bidshift] gives error messages that clearly tell me how to fix problems
13   Whenever I make a mistake using [Bidshift], I recover easily and quickly
16   The information (such as on-line help, on-screen messages, and other documentation) provided with [Bidshift] is clear

The EFA process revealed the four-factor structure and resulted in a 20-item Health-ITUES (table 1): 'Quality of Work Life' (QWL) (3 items), 'Perceived Usefulness' (PU) (9 items), 'Perceived Ease of Use' (PEU) (5 items), and 'User Control' (UC) (3 items).[31] Internal consistency reliabilities for the four factors ranged from .81 to .95.[31]

In this study, we tested the 20-item Health-ITUES through CFA and SEM, which are well-established methods for model testing and scale development.[36 37] CFA is used to verify the measurement model found from EFA. SEM is a combination of factor analysis and path analysis.[38] It is performed after the measurement model is confirmed through CFA. SEM evaluates a complex model with more than one linear equation and supports model comparisons by evaluating the fit of alternative models to identify the best model.[39]

METHODS

We conducted a cross-sectional study design to evaluate users' perception toward the web-based communication system after system implementation using Health-ITUES. We examined the construct validity of Health-ITUES through first and second order CFA, and its predictive validity via SEM.

Setting and sample

The sample for the CFA was recruited from an academic medical center with approximately 1200 staff nurses. At the time of questionnaire distribution, the web-based communication system had

been implemented for 8 months. The sample for the SEM analysis was comprised of both the CFA staff nurses and the sample from the previously reported EFA which was conducted in a community hospital in Philadelphia with approximately 1500 staff nurses.[31] The web-based communication system had been implemented in the community hospital for 2 years at the time of data collection. Staff nurses who had used the web-based communication system met the inclusion criterion for study participation.

Data collection procedures

Questionnaires were electronically distributed to eligible participants via email. An announcement regarding the opportunity to participate in the study was also posted on the system login page. The period of data collection was 8 weeks for the academic medical center sample and 4 weeks for the community hospital sample. As an incentive for participation, the academic medical center respondents were entered into a lottery with 1 in 50 chance of winning $100. Because the community hospital nurses were surveyed on a regular basis regarding use of the web-based communication system, no compensation for time to complete the questionnaire was provided. Questionnaires were considered complete when the amount of missing data was <20%. Demographic characteristics were collected only from the academic medical center sample. Both samples provided data on self-reported internet competency.

Data analysis

We tested two CFA models, the first order CFA model and the second order CFA model. The first order CFA model testing aimed to verify the model found in the EFA[31] using a new sample and assumed that all four factors correlated with each other. Once the first order CFA model testing was confirmed, we further tested the second order CFA model to see if the four factors could be explained by a broader general factor, which we assumed to be 'usability.' In other words, we hypothesized that a usable health IT system influences user perceptions of QWL, PU, PEU, and UC.

We used SEM to examine if Health-ITUES predicts 'intention to use' and/or 'actual usage.' We considered three dependent variables: Q1 (actual usage on my campus), Q2 (actual usage outside my campus), and Q36 (intention to use).[35] We ruled out use of Q2 ('I am more likely to bid on shifts outside of my campus since Bidshift') as its mean (2.40) was much lower than Q1 (4.10) and Q36 (4.14) because nurses at the academic medical center did not work at other campuses. Q1 and Q36 were highly correlated (r=0.704). Since 'intention to use' strongly predicts 'actual usage' and is used extensively in health IT evaluations,[12 22 40] we selected Q36 as the dependent variable. In other words, the SEM hypothesized that the general factor, 'usability', is able to predict a measured item, 'intention to use.'

Data analyses were conducted using SPSS V.16.0, SAS V.9.2, and Mplus V.5.21. SPSS V.16.0 was used for descriptive analysis. Statistical power was estimated based on the model structure and calculated using SAS V.9.2. The analysis program syntax was provided by MacCallum.[41] CFA and SEM were performed using Mplus V.5.21. We used the maximum likelihood (ML) robust extraction method (also called the Satorra-Bentler method) as the estimator. It is recommended for non-normal distributed data.[42] Five indices were used to assess model fit:

- Normed χ^2 (χ^2/df): reduces the sensitivity of χ^2 to sample size. Values <3.0 are considered reasonable fit.
- RMSEA: the root mean square error of approximation (RMSEA) measures the error of approximation. Values <0.05

indicate close approximate fit; values between 0.05 and 0.08 suggest reasonable fit; values >1.0 suggest poor fit.[43]
- SRMR: the standardized root mean square residual (SRMR) measures the mean absolute correlation residual. Values <0.08 are considered good fit.[44]
- CFI and TLI: the comparative fit index (CFI) and the Tucker-Lewis index (TLI) compare the researcher's model with a baseline model. Values >0.90 indicate reasonably good fit.[45]

RESULTS

Results are presented in the following order: descriptive analysis, power analysis, construct validity, and predictive validity.

Descriptive analysis

A total of 222 staff nurses from the academic medical center responded. After exclusion of duplicate entries, incomplete data (more than 20% of data missing), and self-reported 'never use the system' answers, there were 176 valid responses, which corresponded to approximately 18% of the staff nurses in the organization. Demographic characteristics are summarized in table 2.

We used listwise deletion for missing data, therefore only valid and completed responses without any missing information (n=165) were used in the CFA. The SEM sample combined the responses from the academic medical center (n=165) and the community hospital (n=377),[31] hence a total of 541 (377+165) responses were used.

The perceived internet competency of the respondents was high, with 82.9% of academic medical center respondents and 95.5% of community hospital respondents somewhat agreeing or strongly agreeing that they were competent.

Table 2 Demographics of CUMC-NYP staff nurse participants

Variable	n*	Mean (SD)
Age (years)	165	40.90 (11.06)
Gender		
Female	153	87.4%
Male	22	12.6%
Ethnicity		
Caucasian	43	24.4%
African American	32	18.2%
Hispanic	13	7.4%
Asian	71	40.3%
Other	17	9.7%
Education		
Associate	36	20.5%
Bachelor	121	68.8%
Master	18	10.2%
Working experience (years)		
<1	16	9.1%
1–3	34	19.3%
3–5	41	23.3%
>5	83	47.2%
Bidshift experience (months)		
<3	30	17.0%
3–6	55	31.3%
>6	90	51.1%

*n ≠ 176 because of missing data.
CUMC -NYP, Columbia University Medical Center-New York Presbyterian Hospital.

Power analysis

Given the sample sizes, power analysis for first order CFA, second order CFA, and SEM indicates powers of 0.98, 0.98, and 1.00, respectively, to detect RMSEA between 0.05 and 0.08 (table 3).

Construct validity

We verified the first order CFA model (figure 1) found in the EFA. Further, we also confirmed the second order CFA model (figure 2), which assumes that the four factors (the first order latent variables) can be explained by a broader dimension of

general factor (the second order latent variable). In other words, our hypothesis, user perceptions of QWL, PU, PEU, and UC can be explained by a general factor ('usability'), was accepted. Model indices demonstrate adequate fit in both models (figures 1 and 2).

Predictive validity

The SEM hypothesis (figure 3) was accepted. In other words, the general factor, 'usability', is able to predict 'intention to use.' These results support the predictive validity of the Health-ITUES for users' intention to use health IT. The model estimated that the second order factor, health IT usability, accounted for 78.1%, 93.4%, 51.0%, and 39.9% of the variance in QWL, PU, PEU, and UC, respectively. Model indices demonstrate adequate fit and the model explained 63.7% of the variance in intention to use (figure 3).

DISCUSSION
Construct and predictive validity of Health-ITUES

The four-factor model structure identified in the EFA[31] was confirmed in the CFA with adequate model fit, thus providing additional evidence for the construct validity of Health-ITUES.

Figure 1 First order confirmatory factor analysis model. The measurement model variables (eg, Q34, Q33) are presented in squares; the first order latent variables (eg, Quality of Work Life, Perceived Usefulness) are presented in ovals. The numbers on arrows between variables are coefficients. The number at the far right (eg, .146, .057) represents the residual variance not explained by the latent variables.
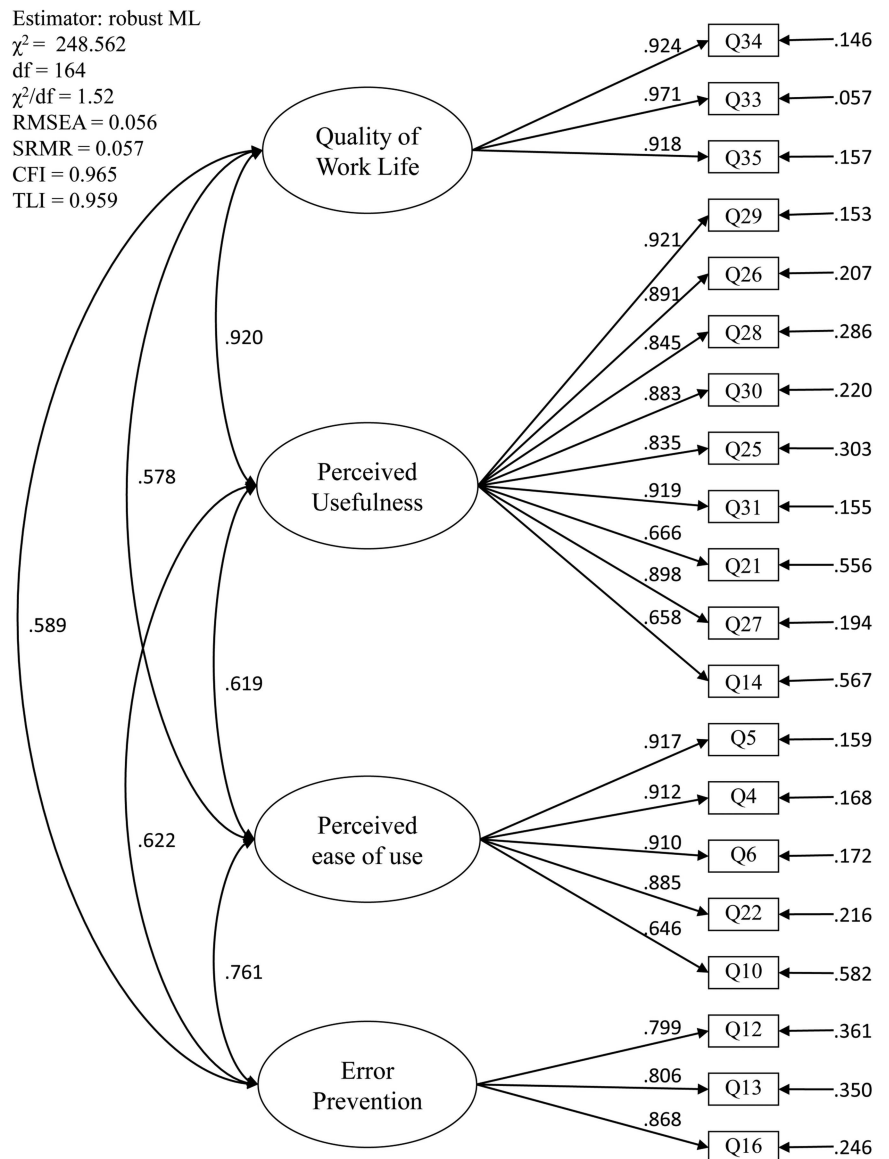
Figure 2 Second order confirmatory factor analysis model. The four factors (the first order latent variables) are explained by a broader dimension of general usability (the second order latent variable). In other words, we hypothesized that a usable health information technology would influence user perceptions of Quality of Work Life, Perceived Usefulness, Perceived Ease of Use, and User Control, and the hypothesis was accepted. The number that points to each latent variable represents the residual variance not explained by the latent variables.
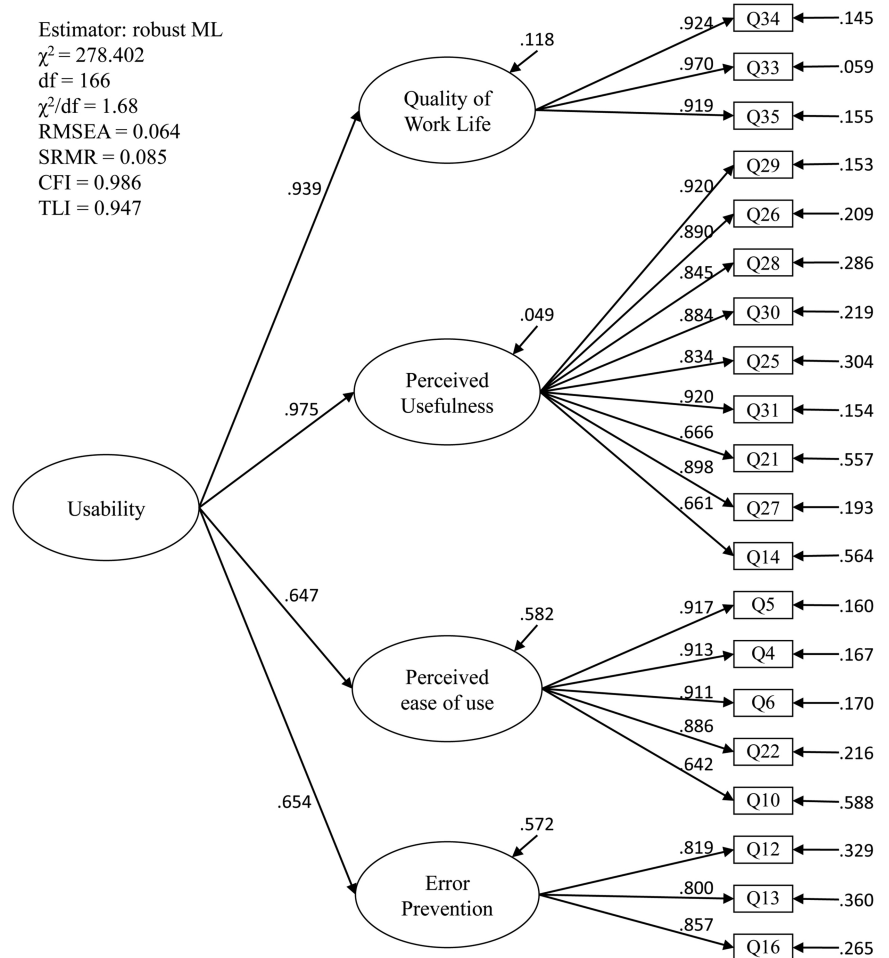

Figure 3 Structural equation model to predict intention to use. The model hypothesizes that the four factors (the first order latent variables, Quality of Work Life, Perceived Usefulness, Perceived Ease of Use, and User Control) are explained by a broader dimension of the general factor, 'Usability' (the second order latent variable), and the general factor, 'Usability' is able to predict a measured item, 'Intention to Use' and the hypothesis was accepted. The number that points to each latent variable represents the residual variance not explained by the latent variables.
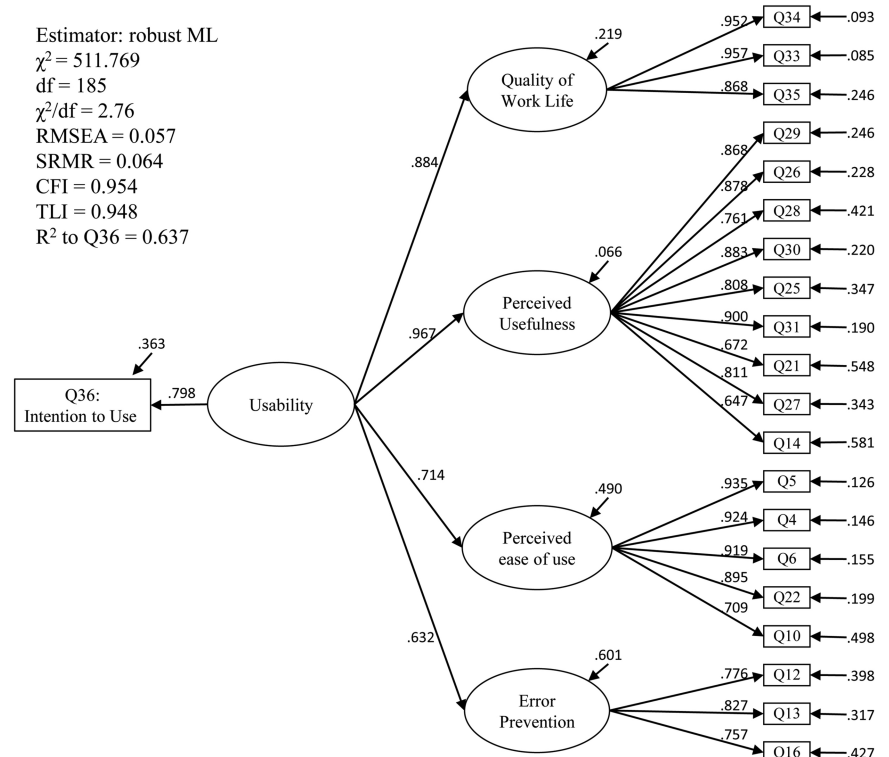
In addition, the final model estimated that a general usability factor accounted for 78.1%, 93.4%, 51.0%, and 39.9% of the variance for QWL, PU, PEU, and UC, respectively. This suggests that a usable system greatly influences users' perceptions of their QWL as well as the PU of the system. The somewhat smaller influence on PEU is consistent with prior research that PU is a more important influence on behavioral intention than PEU.[12] Moreover, user expectations influence user satisfaction, which further contributes to intention to continue system use.[46] From our study findings, we can additionally interpret the relationships as system interactions that supporting task/goal accomplishment is more of a determinant of behavioral intention for system use than simple user–system interaction.

Study results also provide evidence for the predictive validity of Health-ITUES. The 64% of explained variance in behavioral intention is similar to or higher than that in other studies.[23] The 36% of residual on intention to use could be potentially influenced by user variance (eg, age, gender, and education), organizational supports, or other non-system-design issues.

Application of Health-ITUES

The interest in user perceptions of health IT is on the rise. Davis' TAM[12] has been modified in several studies to include constructs or variables other than PU and PEU. Also, the definition of TAM constructs varies in different studies.[23] We argue that the modifications and diverse definitions in studies are due to the varied expectations associated with different types of health IT. For example, we might expect a clinical decision support system (CDSS) to improve patient safety, a picture archiving and communication system (PACS) to improve work productivity, or an information management system to improve evidence-based practice. Even though patient safety, work productivity, and improvement in evidence-based practice can all be considered measurable aspects of PU, the measures are appropriately tied to the characteristics of the health IT. This serves as another illustration of the interaction among user, task, system, and environment and emphasizes the need for explication of task in usability evaluation measures.

Health-ITUES varies from most traditional measurement scales in that it is designed to support customization at the item level to match the specific task/expectation and health IT system while retaining comparability at the construct level. For instance, an electronic health records system may offer multiple functionalities achieving diverse goals, such as order entry, data management and validation, workflow improvement, provider communication, or knowledge support. Customization can target tasks and expectations associated with these functionalities. The comparison could occur at the construct level to better identify usability issues related to these goals. Usability could be evaluated at the QWL level to compare user perceptions of system impact on work processes or at the PEU level to assess user–system interaction. Health-ITUES supports evaluation of three levels of task/expectation: user-system, user-system-task, and user-system-task environment. The approach implemented through Health-ITUES is designed to meet the need to compare across studies, similar to the needs identified by Holden and Karsh[23] in a review of TAMs.

Although some health IT usability models have identified mediating variables, most studies examining health IT usability are cross-sectional, thus limiting the ability to confirm directionality and causality among factors that influence usability. We believe that Health-ITUES can be useful across stages of system development and across time. This can be illustrated through the use of Health-ITUES at the five stages of system
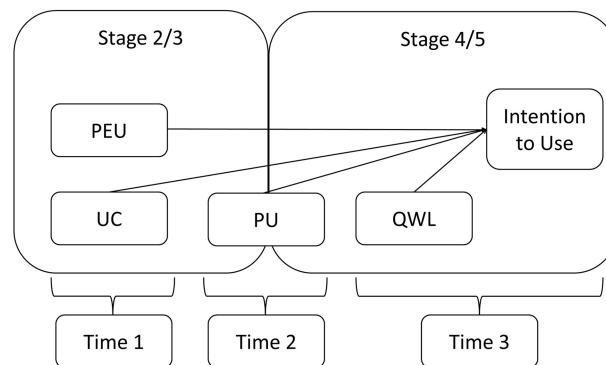

Figure 4 Longitudinal study plan for usability evaluation. PEU, Perceived Ease of Use; PU, Perceived Usefulness; QWL, Quality of Work Life; UC, User Control.

development proposed by Stead and colleagues[47]: (1) specification, (2) component development, (3) combination of components into a system, (4) integration of system into environment, and (5) routine use (figure 4). PEU and UC measures could be implemented in stages 2 and 3 to verify user perceptions of system operation. PU measures could be included in stages 3 and 4 to confirm the system effectiveness to accomplish tasks. QWL, a measure of system impact, could be included in stages 4 and 5 to assess the long-term effect. Also, time-series comparison of PEU, UC, and PU through stages 3–5 could be used to assess learnability and the system effectiveness. Lastly, longitudinal designs may overcome the overgeneralization that occurs in directional analysis in SEM with cross-sectional study designs.

Limitations

The study has some limitations. First, the response rate for the academic medical center survey used in CFA was only about 18%. Second, most participants were competent in internet use. The results may vary in users with low internet competency. Third, the study was conducted using only one system and one professional group (registered nurses), which may potentially limit broader applicability. Finally, Health-ITUES was designed to be customizable based on the user-system-task environment interaction, but the customization could vary due to the level of description of a task; this requires additional research. Our future work will include applying Health-ITUES to other health IT and other professions, and incorporating task complexity and user expectations to offer further guidance on the customization.

An additional limitation is that only one model was tested for predictive validity. We did not conduct comparisons to test competing theories. Testing only one model may identify a well-fitting model but also ignores other plausible models that may better account for the relationships among the data and decrease the probability of confirmation bias.[37]

Despite these limitations, our analyses addressed a number of methodological issues often associated with the model testing process. First, studies frequently fail to report power analysis to demonstrate adequacy of sample based on model structures.[48–53] A second methodological issue relates to failure to examine data distributions before proceeding to SEM.[54] Consequently, the typical default estimator in software packages, ML, may be applied even when the data do not meet ML's strong assumption of normally distributed data. A third issue relates to lack of attention to examining the measurement model by CFA prior to SEM. The two-step approach[55] and four-step approach[56] have been recommended as alternative

methods for finding an acceptable CFA model before examining the structural model.[57] To avoid these issues, we conducted a power analysis based upon model structures, examined data distribution for match to methods, and assessed the measurement model through CFA prior to SEM.

CONCLUSION

The results of CFA confirmed the factorial structure of Health-ITUES that was identified through the EFA and also demonstrated adequate model fit, thus providing additional evidence for the construct validity of the scale. SEM supported the predictive validity of Health-ITUES for behavioral intention for system use. The customizability (based on task-specific and level of expectation) of the Health-ITUES has the potential to support comparisons at the construct level, while allowing variation at the item level.

REFERENCES

1 Thompson T, Brailer D. *Health IT strategic framework*. Washington, DC: DHHS, 2004.
2 Shortliffe EH, Perreault LE, Wiederhold G, *et al*. *Medical informatics: computer applications in health care and biomedicine*. 2nd edn. Springer, 2001.
3 Yusof MM, Stergioulas L, Zugic J. Health information systems adoption: findings from a systematic review. *Stud Health Technol Inform* 2007;129(Pt 1):262–6.
4 Miller RH, Sim I. Physicians' use of electronic medical records: barriers and solutions. *Health Affairs* 2004;23:116–26.
5 Ash JS, Berg M, Coiera E. Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc* 2004;11:104–12.
6 Kushniruk AW, Triola MM, Borycki EM, *et al*. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. *Int J Med Inform* 2005;74:519–26.
7 ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs) — Part 11: Guidance on usability, 1998.
8 Doll WJ, Xia WD, Torkzadeh G. A confirmatory factor-analysis of the end-user computing satisfaction instrument. *Mis Quarterly* 1994;18:453–61.
9 DeLone WH, McLean ER. The DeLone and McLean model of information systems success: a ten-year update. *J Manag Info Syst* 2003;19:9–30.
10 Norman K, Shneiderman B. Questionnaire for User Interaction Satisfaction (QUIS) 1997 [cited 2008]. http://lap.umd.edu/quis/ (accessed 13 Feb 2014).
11 Lee F, Teich JM, Spurr CD, *et al*. Implementation of physician order entry: user satisfaction and self-reported usage patterns. *J Am Med Inform Assoc* 1996;3:42–55.
12 Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 1989;13:318–40.
13 Dillon A, Morris MG. User acceptance of information technology: theories and models. *Annu Rev Info Sci Technol* 1996;31:3–32.
14 Wixom BH, Todd PA. A theoretical integration of user satisfaction and technology acceptance. *Info Syst Res* 2005;16:85–102.
15 Mather D, Caputi P, Jayasuriya R. Is the technology acceptance model a valid model of user satisfaction of information technology in environments where usage is mandatory? In: Wenn A, McGrath M, Burstein F, eds. *Enabling organisations and society through information systems*. Australia: ACIS 2002 School of Information Systems, Victoria University, 2002:1241–50.
16 Miyamoto M, Kudo S, Iizuka K. Measuring ERP success: integrated model of user satisfaction and technology acceptance. *An Empirical Study in Japan IPEDR Proceedings*. 2012.
17 Adamson I, Shine J. Extending the new technology acceptance model to measure the end user information systems satisfaction in a mandatory environment: a bank's treasury. *Technol Anal Strateg Manag* 2003;15:441–55.
18 Gadd CS, Ho YX, Cala CM, *et al*. User perspectives on the usability of a regional health information exchange. *J Am Med Inform Assoc* 2011;18:711–16.
19 Chau PYK, Hu PJH. Investigating healthcare professionals' decisions to accept telemedicine technology: an empirical test of competing theories. *Inform Manage-Amster* 2002;39:297–311.
20 Jaspers MWM, Peute LWP, Lauteslager A, *et al*. Pre-post evaluation of physicians' satisfaction with a redesigned electronic medical record system. *Stud Health Technol Inform* 2008;136:303–8.
21 Lewis JR. IBM computer usability satisfaction questionnaires—psychometric evaluation and instructions for use. *Int J Hum-Comput Interac* 1995;7:57–78.
22 Venkatesh V, Morris MG, Davis GB, *et al*. User acceptance of information technology: Toward a unified view. *Mis Quarterly* 2003;27:425–78.
23 Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. *J Biomed Inform* 2010;43:159–72.
24 Ammenwerth E, Iller C, Mahler C. IT-adoption and the interaction of task, technology and individuals: a fit framework and a case study. *BMC Med Inform Decis Mak* 2006;6:3.
25 Shackel B. Usability—context, framework, definition, design and evaluation. In: Shackel B, Richardson SJ, eds. *Human factors for informatics usability*. New York: Cambridge University Press, 1991:21–37.
26 Or CK, Valdez RS, Casper GR, *et al*. Human factors and ergonomics in home care: Current concerns and future considerations for health information technology. *Work* 2009;33:201–9.
27 Kushniruk A, Nohr C, Jensen S, *et al*. From usability testing to clinical simulations: bringing context into the design and evaluation of usable and safe health information technologies. Contribution of the IMIA human factors engineering for healthcare informatics working group. *Yearb Med Inform* 2013;8:78–85.
28 Borycki E, Kushniruk A. Identifying and preventing technology-induced error using simulations: application of usability engineering techniques. *Healthc Q* 2005;8 Spec No:99–105. PMID: 16334081
29 Carayon P. Human factors in patient safety as an innovation. *Appl Ergon* 2010;41:657–65.
30 Kushniruk AW, Patel VL. Cognitive and usability engineering methods for the evaluation of clinical information systems. *J Biomed Inform* 2004;37:56–76.
31 Yen P, Bakken S, Wantland D. Development of a customizable health information technology usability evaluation scale. *Proceedings of the American Medical Informatics Association Annual Symposium 2010*. PMID: 21347112
32 Schnall R, Hyun S, Yen P, *et al*. Using technology acceptance models to inform the design and evaluation of nursing informatics innovations (panels). *Proceedings of NI2009, the 10th International Congress in Nursing Informatics*. 2009 June 28—July 1.
33 Bennett J. *Visual display terminals: usability issues and health concerns*. EnglewoodCliffs, NJ: Prentice-Hall, 1984.
34 Hornbaek K. Current practice in measuring usability: challenges to usability studies and research. *Int J Hum-Comput Stud* 2006;64:79–102.
35 Yen P. *Health information technology usability evaluation: methods, models and measures*. Columbia University, 2010.
36 Noar SM. The role of structural equation modeling in scale development. *Struct Equation Model* 2003;10:622–47.
37 Martens MP. The use of structural equation modeling in counseling psychology research. *Couns Psychol* 2005;33:269–98.
38 Weston R, Gore PA. A brief guide to structural equation modeling. *Couns Psychol* 2006;34:719–51.
39 Tomarken AJ, Waller NG. Structural equation modeling: strengths, limitations, and misconceptions. *Annu Rev Clin Psychol* 2005;1:31–65.
40 Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. *Manag Sci* 2000;46:186–204.
41 MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods* 1996;1:130–49.
42 Satorra A, Bentler PM. A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika* 2001;66:507–14.
43 Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS.eds *Testing structural equation models*. Newbury Park, CA: Sage, 1993:136–62.
44 Hu LT, Bentler PM. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychol Methods* 1998;3:424–53.
45 Bentler PM. Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–46.
46 Bhattacherjee A. Understanding information systems continuance: an expectation-confirmation model. *Mis Quarterly* 2001;25:351–70.
47 Stead WW, Haynes RB, Fuller S, *et al*. Designing medical informatics research and library—resource projects to increase what is learned. *J Am Med Inform Assoc* 1994;1:28–33.
48 Yi MY, Jackson JD, Park JS, *et al*. Understanding information technology acceptance by individual professionals: toward an integrative view. *Inform Manage-Amster* 2006;43:350–63.

49 Wu JH, Wang SC, Lin LM. Mobile computing acceptance factors in the healthcare industry: a structural equation model. *Int J Med Inform* 2007;76:66–77.

50 Hulse NC, Del Fiol G, Rocha RA. Modeling end-users' acceptance of a knowledge authoring tool. *Methods Inf Med* 2006;45:528–35.

51 Gagnon MP, Godin G, Gagne C, *et al*. An adaptation of the theory of interpersonal behaviour to the study of telemedicine adoption by physicians. *Int J Med Inform* 2003;71:103–15.

52 Hu PJ, Chau PYK, Sheng ORL, *et al*. Examining the technology acceptance model using physician acceptance of telemedicine technology. *J Manage Inf Syst* 1999;16:91–112.

53 Tung F-C, Chang S-C, Chou C-M. An extension of trust and TAM model with IDT in the adoption of the electronic logistics information system in HIS in the medical industry. *Int J Med Inform* 2008;77:324–35.

54 Yuan KH, Bentler PM. Effect of outliers on estimators and tests in covariance structure analysis. *Br J Math Stat Psychol* 2001;54:161–75.

55 Anderson JC, Gerbing DW. Structural equation modeling in practice—a review and recommended 2-step approach. *Psychol Bull* 1988;103:411–23.

56 Mulaik SA, Millsap RE. Doing the four-step right. *Struct Equation Model* 2000;7:36–73.

57 Kline RB. *Principles and practice of structural equation modeling*. 2nd edn. New York: The Guilford Press, 2005.