# Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology

CrossMark

William Brown III [a,b], Po-Yin Yen [c], Marlene Rojas [d], Rebecca Schnall [d,*]

[a] HIV Center for Clinical and Behavioral Studies, New York State Psychiatric Institute and Columbia University, New York, NY, United States
[b] Department of Biomedical Informatics, Columbia University, New York, NY, United States
[c] Department of Biomedical Informatics, The Ohio State University, Columbus, OH, United States
[d] School of Nursing, Columbia University, 617 West 168th Street, New York, NY 10032, United States

## ABSTRACT

*Background:* Over two decades of research has been conducted using mobile devices for health related behaviors yet many of these studies lack rigor. There are few evaluation frameworks for assessing the usability of mHealth, which is critical as the use of this technology proliferates. As the development of interventions using mobile technology increase, future work in this domain necessitates the use of a rigorous usability evaluation framework.

*Methods:* We used two exemplars to assess the appropriateness of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating the usability of mHealth technology. In the first exemplar, we conducted 6 focus group sessions to explore adolescents' use of mobile technology for meeting their health Information needs. In the second exemplar, we conducted 4 focus group sessions following an Ecological Momentary Assessment study in which 60 adolescents were given a smartphone with pre-installed health-related applications (apps).

*Data analysis:* We coded the focus group data using the 9 concepts of the Health-ITUEM: Error prevention, Completeness, Memorability, Information needs, Flexibility/Customizability, Learnability, Performance speed, Competency, Other outcomes. To develop a finer granularity of analysis, the nine concepts were broken into positive, negative, and neutral codes. A total of 27 codes were created. Two raters (R1 and R2) initially coded all text and a third rater (R3) reconciled coding discordance between raters R1 and R2.

*Results:* A total of 133 codes were applied to Exemplar 1. In Exemplar 2 there were a total of 286 codes applied to 195 excerpts. Performance speed, Other outcomes, and Information needs were among the most frequently occurring codes.

*Conclusion:* Our two exemplars demonstrated the appropriateness and usefulness of the Health-ITUEM in evaluating mobile health technology. Further assessment of this framework with other study populations should consider whether Memorability and Error prevention are necessary to include when evaluating mHealth technology.

© 2013 Elsevier Inc. All rights reserved.

## 1. Introduction

Mobile technology has become almost ubiquitous among millions of Americans. At the same time, mobile health (mHealth) technology applications (apps) have also increased in availability and popularity [1]. For example, 13,000 health-related iPhone apps were available for consumer use in 2012 [2]. mHealth interventions are increasingly important instruments in the toolkit of public health professionals and researchers [3–5]. For example, an increasing number of mobile disease management programs for chronic diseases such as diabetes and heart disease [6] are being developed. Additional examples of mobile health apps include interventions that help people quit smoking [7] or lose weight, or mental health apps to address depression, and/or anxieties [8]. Moreover, patient access to their health records via personal health records and patient portals make it increasingly appealing for patients to view their personal health information in real-time via mobile devices.

Despite the growth and popularity of mHealth apps, more than 95% have not been tested [9]. For instance, a recent systematic review revealed that there are only 42 controlled trials of mobile technology interventions for all disease processes and the effects that were demonstrated are only modestly beneficial [10]. In another recent study, Whitlock and McLaughlin studied three apps for tracking blood glucose and found that each product may pres-

* Corresponding author.
  *E-mail address:* rb897@columbia.edu (R. Schnall).

ent a number of usability issues, including small text, poor color contrast and scrolling wheels, especially for older adults [11]. As a result, further research needs to be done to ensure that mobile health technologies are appropriately designed and targeted to the end-users' needs before they are used as health interventions [12]. Prior to the trial of mHealth technologies for improving clinical outcomes, it is imperative that IT designers pay close attention to the usability of these technologies.

Usability factors are a major obstacle to health information technology (IT) adoption. While health IT such as mHealth tools can offer potential benefits, they can also interrupt workflow, cause delays, and introduce errors [13–15]. Lack of attention to health IT evaluation may result in dissatisfied users, decreased effectiveness, and increases in error costs [16]. While the promise of mHealth is that we can leverage the power and ubiquity of mobile and cloud technologies to monitor and manage side effects and treatment outside the clinical setting, it is essential to be attentive to usability, keeping in mind its intended users, task and environment.

Many health IT usability studies have been conducted to explore usability requirements, discover usability problems, and design solutions, but few of these studies have evaluated the usability of mobile technologies. Currently there are few evaluation frameworks for these technologies making rigorous evaluation a challenge. In this paper, we seek to demonstrate through the use of two data sources, the applicability of the Health-ITUEM usability evaluation framework – which addresses gaps in existing usability models – for evaluating mobile technologies [17].

## 2. Background

### 2.1. Usability

Understanding deficiencies in HIT systems is critical to understanding why these technologies fail [18]. Usability evaluation is a method for identifying specific problems with IT products and specifically focuses on the interaction between the user and task in a defined environment [19]. Usability of a technology is determined by user-computer interactions and the degree to which the technology can be successfully integrated to perform a task in the intended work environment [20]. Effective usability evaluation improves predictability of products and saves development time and cost [21].
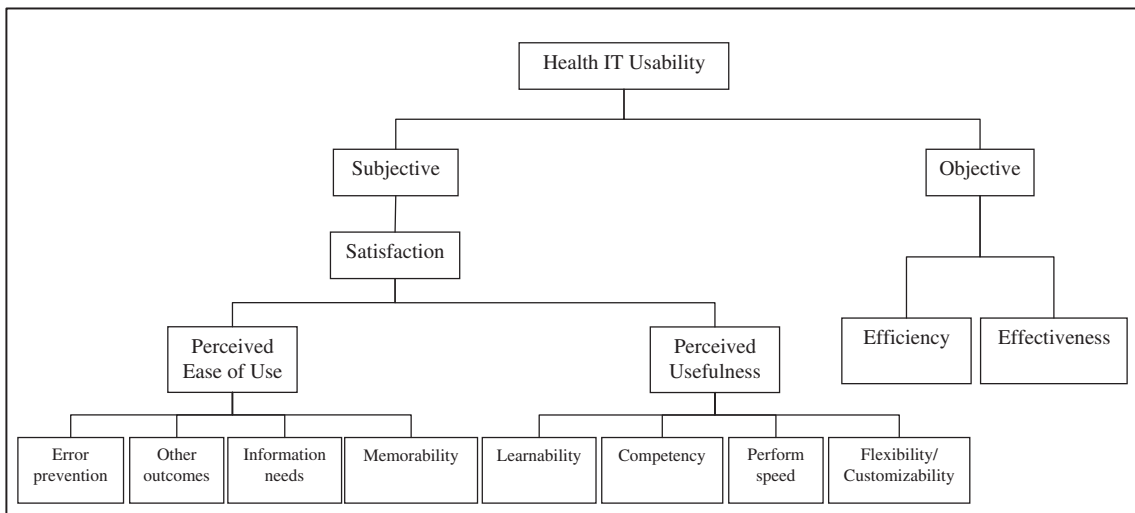
### 2.2. Health-ITUEM

The Health IT Usability Evaluation Model (Health-ITUEM) (Fig. 1) was developed in response to the current gaps in the existing usability models which had previously been developed [17]. The Health-ITUEM was developed as an integrated model of multiple theories as a comprehensive usability evaluation framework. The Health-ITUEM was developed by assessing the usability of a web-based communication system for scheduling nursing staff. The model has not been previously tested in other studies.

Definitions of usability from the technology acceptance model (TAM) [22] and ISO 9241-11 provide the fundamental constructs of the Health-ITUEM, while concepts were identified from usability decompositions [23], Nielsen's ten heuristics [24], Shneiderman's eight rules for user interface design [25], and Norman's seven principles for design [26]. Discussion with potential system users and developers also recognized additional concepts. Health-ITUEM concepts include: Error prevention, Completeness, Memorability, Information needs, Flexibility/Customizability, Learnability, Performance speed, Competency and Other outcomes.

### 2.3. Usability of mHealth technology

Currently there is a dearth of literature on the usability of mHealth technology. There are a number of studies focused on mHealth apps [27–29] and others focused on the usability of the devices [30,31] but usability evaluations of mHealth technology have not yet reached the level of rigor of web-based electronic health apps evaluation. There are two, somewhat overlapping challenges, which may have delayed the development of rigorous usability testing for mHealth technology. First, mobile devices present several unique challenges: small, low resolution screens; no mouse or keyboard); a slow operating system, and variable connectivity [32]. As a result, it is particularly important to develop and adapt frameworks for mHealth technology since the usability issues can be very different from those of web-based technology. A concurrent challenge for developing and refining mHealth technology is that while technology is rapidly advancing, end-user testing equipment and software is making much slower progress. For instance, traditional desktop screen-capture software still cannot record touch interactions on mobile devices and tablets, and so usability practitioners use strategically placed cameras for tablet



**Fig. 1.** Health-ITUEM.

testing but are still limited in how much usability data they can capture [33].

## 3. Methods

We used two sources of data to assess the applicability of the Health-ITUEM to mHealth technology. All study activities were approved by the Columbia University Medical Center IRB. Written informed assent was obtained from all study participants. The focus group sessions were recorded with two digital recorders to safeguard against mechanical failure. Each focus group session lasted approximately one hour. Data were generated using the basic focus group session format. Audio tape recordings and field notes were used for data collection. Data collection continued until saturation of themes was reached.

### 3.1. Exemplar 1

For our first exemplar, we conducted six focus groups sessions from November to December 2011. Focus groups had between 3 and 8 participants in each group. Participants were initially recruited on-site from a public high school in the Bronx, NY, followed by snowball sampling, where existing study subjects recruited their friends to participate. Our participants included 32 adolescents age 14–18 years ($M = 16$, $SD = 1.16$). We had 14 females and 17 male participants. Eighty-four percent of participants classified themselves as being Hispanic. All of our participants were mobile phone users and 81.3% of participants used a mobile device at least once a day. The mean self-reported physical and mental health status scores for our participants were 76.32 ($SD = 30.87$) and 75.19 ($SD = 9.80$), respectively. These values are significantly above the established national mean, which is 50 ($SD = 10.25$), as measured by the Medical Outcomes Study (MOS), 36-Item Short Form Health Survey (SF-36) [34], demonstrating a healthy participant cohort.

Participants were informed that the purpose of this study was to understand their health information seeking behavior and that their personal information and verbalizations would not be identifiable. We asked the following focus group questions: (1) Tell us a little about how you use mobile technology. What type of personal health information have you viewed using mobile health technology? (2) What are some of the reasons that have motivated you to use mobile technology? (3) What were some barriers you encountered when using mobile technology to meet your health Information needs? (4) What were some of the strategies you used to overcome these barriers?

The moderator summarized the key points at the end of each question, which served as member checks [35]. The focus group included food and drinks appropriate for the time of day and reimbursement for time ($20). Peer debriefing among the members of the research team occurred immediately following the session.

### 3.2. Exemplar 2

For our second exemplar, we also conducted focus group sessions with a group of adolescents. Prior to our focus group sessions, we distributed Smartphones with the following pre-loaded apps: MyFitnessPal, SparkPeople, NIH obesity, and Asthma Check, for use during a 30-day ecological momentary assessment [36]. We had 38 participants in 4 focus group sessions. During these sessions, we asked participants to reflect upon their past 30 days of using the Smartphone and specifically answer the following questions: (1) What are some of the reasons that have motivated you to use the apps on your mobile device? (2) What were some barriers you encountered when using the mobile health apps on your phone? (3) What were some of the strategies you used to overcome these barriers? Participants for this exemplar ranged in age from 13 to 18 years old and 70.5% identified themselves as Hispanic.

## 4. Analysis

### 4.1. Health-ITUEM concepts and code development

The Health-ITUEM concepts guided our analysis of the two exemplars. Based on the Health-ITUEM concepts we developed the following data analysis codes: Error prevention, Completeness, Memorability, Information needs, Flexibility/Customizability, Learnability, Performance speed, Competency, and Other outcomes, which were used to categorize the data (Table 1). To develop a finer granularity of analysis, the nine concept codes were broken into positive, negative, and neutral codes. Concept codes for identifying positive sentiment were designated with a plus sign (+). Negative sentiment concept codes were designated using a minus sign (−), and neutral codes had no sign at all. The result was three codes for each of the nine Health-ITUEM concepts, for a total of 27 codes.

### 4.2. Data analysis

Free text was excerpted from the focus group transcripts and was coded using one or more of the 27 codes (Table 1) from the Health-ITUEM.

### 4.3. Dedoose QDA

Both data sources (Exemplars 1 and 2) were analyzed using Dedoose Qualitative Data Analysis (QDA) software, a web-based qualitative and mixed methods data analysis program. Dedoose was chosen because of its highly visual interactive data analysis interface, and its ability to analyze both qualitative and quantitative data when necessary [37]. This software is also cross platform, which allowed us to more easily analyze the data [38].

### 4.4. Inter-rater reliability analysis

The data was categorized according to the nine Health-ITUEM concepts. Free texts were excerpted from the focus group transcripts and each excerpt was coded using one or more of the 27 codes (Table 2). Two raters independently coded the data. We calculated the percent of agreement between the raters. Excerpts of compound statements that reflected more than one code were given multiple codes as appropriate; thus, standard percent agreement statistics was the most accurate and effective method of analysis. Moreover, our reliability check used 100% of excerpts and produced 10% or fewer disagreements; thus, the agreement achieved is improbably the result of chance agreement [39]. A third rater resolved any discrepant coding.

## 5. Results

For each excerpt, the rater could select one or more of 27 code categories. Excerpts that reflected more than one code were given multiple codes as appropriate. Each code, its definition and a sample excerpt are included in Table 2. All excerpts were first rated by two raters (R1 and R2), and disagreement was resolved by a third rater (R3). Some excerpts with overlapping speakers received one or more codes. Exemplar 1 data had 91 coded excerpts when saturation was achieved. A total of 133 codes were applied to this data set. In Exemplar 2 there were a total of 286 codes applied to 195 excerpts. The frequency of use of each code appears in Fig. 2.

**Table 1**
Codes and definitions derived from the Health-ITUEM.

| Title | Description |
|---|---|
| Error prevention | System offers error management, such as error messages as feedback, error correction through undo function, or error prevention, such as instructions or reminders, to assist users performing tasks |
| + Error prevention | Positive occurrence or response related to Parent Code Error prevention |
| − Error prevention | Negative occurrence or response related to Parent Code Error prevention |
| Completeness | System is able to assist users to successfully complete tasks. This is usually measured objectively by system log files for completion rate |
| + Completeness | Positive occurrence or response related to Parent Code Completeness |
| − Completeness | Negative occurrence or response related to Parent Code Completeness |
| Memorability | Users can remember easily how to perform tasks through the system |
| + Memorability | Positive occurrence or response related to Parent Code Memorability |
| − Memorability | Negative occurrence or response related to Parent Code Memorability |
| Information needs | The information content offered by the system for basic task performance, or to improve task performance |
| + Information needs | Positive occurrence or response related to Parent Code Information needs |
| − Information needs | Negative occurrence or response related to Parent Code Information needs |
| Flexibility/ Customizability | System provides more than one way to accomplish tasks, which allows users to operate system as preferred |
| + Flexibility/ Customizability | Positive occurrence or response related to Parent Code Flexibility/Customizability |
| − Flexibility/ Customizability | Negative occurrence or response related to Parent Code Flexibility/Customizability |
| Learnability | Users are able to easily learn how to operate the system |
| + Learnability | Positive occurrence or response related to Parent Code Learnability |
| − Learnability | Negative occurrence or response related to Parent Code Learnability |
| Performance speed | Users are able use the system efficiently |
| + Performance speed | Positive occurrence or response related to Parent Code Performance speed |
| − Performance speed | Negative occurrence or response related to Parent Code Performance speed |
| Competency | Users are confident in their ability to perform tasks using the system, based on Social Cognitive Theory |
| − Competency | Negative occurrence or response related to Parent Code Competency |
| + Competency | Positive occurrence or response related to Parent Code Competency |
| Other outcomes | Other system-specific expected outcomes representing higher level of expectations. (uses of non-phone app technology, non-mobile resources, other health related entities outside of study protocol) |
| + Other outcomes | Positive occurrence or response related to Parent Code Other outcomes |
| − Other outcomes | Negative occurrence or response related to Parent Code Other outcomes |

## 5.1. Exemplar 1 data – Health information seeking

The Health-ITUEM successfully identified the mHealth usability issues that existed in this data set. Codes in this data set focused around three major concepts: Information needs, Performance speed, and Other outcomes. Moreover, the granularity of codes were overwhelmingly positive (103 instances), with instances of negative codes occurring only 25 out of 133 instances (18.8%), and instances of neutral codes occurring only 5 out of 133 instances (3.8%).

The most commonly occurring code in this data set was Other outcomes, which was used to code 46 (34.6%) of the 91 excerpts. Other outcomes captured system-specific or expected outcomes that represent higher-level user expectations. This included, uses of non-phone app technology (i.e. phone, books), non-mobile resources (i.e. parents, teachers, siblings), and other health related entities not directly related to the usability of mHealth. Information needs was the second most frequently occurring concept code. It was used to code 38 (28.6%) of the 91 excerpts from this data set. Similarly, the concept code Performance speed was a frequently occurring code, accounting for 28 (21.1%) excerpts in this data set. Performance speed captures instances where users are able to use the system efficiently. Instances in which participants discussed their use mobile technology, or any technology with which they are familiar, was captured using the Performance speed code. Each of the following codes: Flexibility/Customizability, Completeness, Competency also appeared in this data set, but to a small degree. Completeness was used 15 (11.3%) times, and both Flexibility/Customizability and Competency were used 3 (2.3%) times each.

## 5.2. Exemplar 2 – Reported usability of mobile devices during an ecological momentary assessment

The Health-ITUEM allowed us to identify the usability concepts, which were discussed by participants in this exemplar. The 27 positive, negative, and neutral codes based on the 9 Health-ITUEM concepts were captured in the data from these focus groups.

The most frequently applied codes were Performance speed, Other outcomes, and Information needs. These three concepts accounted for 74.1% of the excerpts. Eighty-seven (30.4%) excerpts were coded as Performance speed. Sixty-five (22.7%) excerpts were coded as Other outcomes. These included topics such as use of Google Search, CDC.gov, and other Internet-based health information query tools. Lastly, 60 (21.0%) excerpts were coded as Information needs.

Subsequently, Learnability, Completeness, Flexibility/Customizability were the next most frequently applied codes. Learnability was used for 22 (7.7%) excerpts and Completeness was used to code 21 (7.3%) excerpts. Flexibility/Customizability was used to code 15 (5.2%) excerpts. Competency, Error prevention and Memorability were the least used codes and were only applied to 8 (2.8%) 7, (2.5%) and 1 (0.5%) excerpt(s) respectively. There were only 19 (6.6%) neutral codes applied to the excerpts in Exemplar 2. There were a total of 137 (47.9%) positive codes applied to the excerpts and a total of 130 (45.5%) negative codes applied to excerpts.

## 5.3. Inter-rater reliability results

Exemplar 1 data was coded to saturation, resulting in 91 excerpts. A total of 133 codes were applied to this data set, and 94% of the raters' coding (R1 and R2) had strong agreement. Strong agreement was defined as raters agreeing on one or more concepts per-excerpt. The third rater resolved the discrepant 6% of disagreement. In Exemplar 2 there were a total of 286 codes applied to 195 excerpts. In Exemplar 2 the observed percentage for strong agreement was 89%. Rater R3 resolved the remaining 11% of discrepant excerpt coding. The combined total strong agreement between R1 and R2 for Exemplar 1 and 2 was 90%. Inter-rater reliability was also calculated for each code and for each exemplar (Table 3).

**Table 2**
Health-ITUEM concepts and representative quotes from the exemplars.

| | | |
|---|---|---|
| Error prevention | System offers error management, such as error messages as feedback, error correction through undo function, or error prevention, such as instructions or reminders, to assist users performing tasks | R: Oh no. When the phone first came to me, I did not know how to turn it off. I took out the battery and put it back in, and still said emergency, and I thought it was something wrong with the phone. So, I pushed the reset and it just restarted everything<br>M: It erased everything that was on there.<br>R: Yes. (Exemplar 2) |
| Completeness | System is able to assist users to successfully complete tasks. This is usually measured objectively by system log files for completion rate | R: Did you guys put an app of books; like that you could look up books?<br>R: I saw that<br>R: Yes, when I first…<br>M: It was already on there for the phone, yes<br>R: That came in handy<br>M: The books, the list of books?<br>R: Yes. I was reading a lot<br>M: Really?<br>R: Yes. (Exemplar 2) |
| Memorability | Users can remember easily how to perform tasks through the system | M: You forgot about your diet or you forgot about the app?<br><br>R: Both. (Exemplar 2) |
| Information needs | The information content offered by the system for basic task performance, or to improve task performance [26,27] | M: And now do you find those answers on your phone?<br><br>R: Well, simply type in the answer. They have like the Yahoo answer. And I can see if anybody else is going through the same problems I am. (Exemplar 1) |
| Flexibility/Customizability | System provides more than one way to accomplish tasks, which allows users to operate system as preferred | "Voice to text," is an important feature (Exemplar 1) |
| Learnability | Users are able to easily learn how to operate the system | M: So, did anyone look at any of the apps or some of the information online and just not understand what it was saying? Like the obesity app, was it confusing?<br>R: No (Exemplar 2) |
| Performance speed | Users are able use the system efficiently | M: Okay, so what do you use it for?<br>R: Facebook and (inaudible), mostly. Also check my email there because I do not feel like going on a computer to check the mail. (Exemplar 1) |
| Competency | Users are confident in their ability to perform tasks using the system, based on Social Cognitive Theory [28,29] | So like it depends on if you feel you can trust it, or if it matches up with what's going on with you. (Exemplar 1) |
| Other outcomes | Other system-specific expected outcomes representing higher level of expectations. | R: If I have any health-related problems I usually just go on Google and it takes to me some doctor website, where real doctors answer the questions (Exemplar 2) |

# 6. Discussion

mHealth technology is growing exponentially [1,40]. New wireless technologies and optimized mobile interfaces are adding to the ubiquity of mobile device use. As a result, users are increasingly using mHealth technology to meet their health information needs, self-management of their health and as communication tool with their providers. Researchers and interventionists are finding ways to integrate mobile technologies in public health and clinical practice. Thus, it is fundamental to evaluate the usability of mHealth technology before interventions are put into practice. A number of challenges occur in the development of mHealth technologies. As noted in earlier studies, the length of time required to develop content, complete usability testing and iteratively refine systems is a barrier to development [6,41]. Moreover, there are few usability frameworks, which have been developed or evaluated for mHealth technologies, which is an impediment to rigorously evaluating these technologies. Findings from this study fill a gap in the literature by assessing the use of the Health-ITUEM for mHealth technology. Specifically, this study contributes to the literature on the evaluation of mobile health tools by testing the stability, flexibility, and concept permanence [42] of the Health-IT concept models within the framework.

Rarely does a framework fit all study designs and data types. Nonetheless, the Health-ITEUM can be appropriately applied to the usability evaluation of mHealth technology as is demonstrated in our two exemplars. Memorability and Error prevention were the least often used codes. Further assessment of this framework with other study populations should consider whether these codes are necessary to include when evaluating mHealth technology.

Information needs was the second most frequently occurring concept code. It was used to code 38 (28.6%) of the 91 excerpts in Exemplar 1. The high occurrence of the "Information needs" code in Exemplar 1 can be attributed to the "information seeking" nature of the focus group. Performance speed was also a frequently used code. Performance speed captures instances where users are able use the system efficiently. In Exemplar 1, users were asked to refer to tools of which they were already familiar; this would likely make them effective users of the technology.

We calculated inter-rater reliability using 100% of our sample, most often inter-rater reliability is calculated from 10% to 20% of the sample, thus we expect the utility of these codes to be reliable and translate to other mHealth studies. Error Prevention (100%), Memorability (100%), Competency (95%), Completeness (93%), Other Outcomes (93%), and Information Needs (91%) were codes with the highest (ninetieth percentile) inter-rater agreement. The inter-rater agreement of Learnability (88%) and Performance Speed (87%) were in the fair range (eightieth percentile). Despite Learnability and Performance Speed falling in the fair range, eight out of nine codes achieved almost 90% agreement. On the other hand, Flexibility/Customizability had the lowest inter-rater agreement in Exemplar 2 (59%), and the lowest agreement overall (66%). Low agreement for this concept may be due to conflicting interpretations of "Flexibility" and "Customizability". It is suggested that
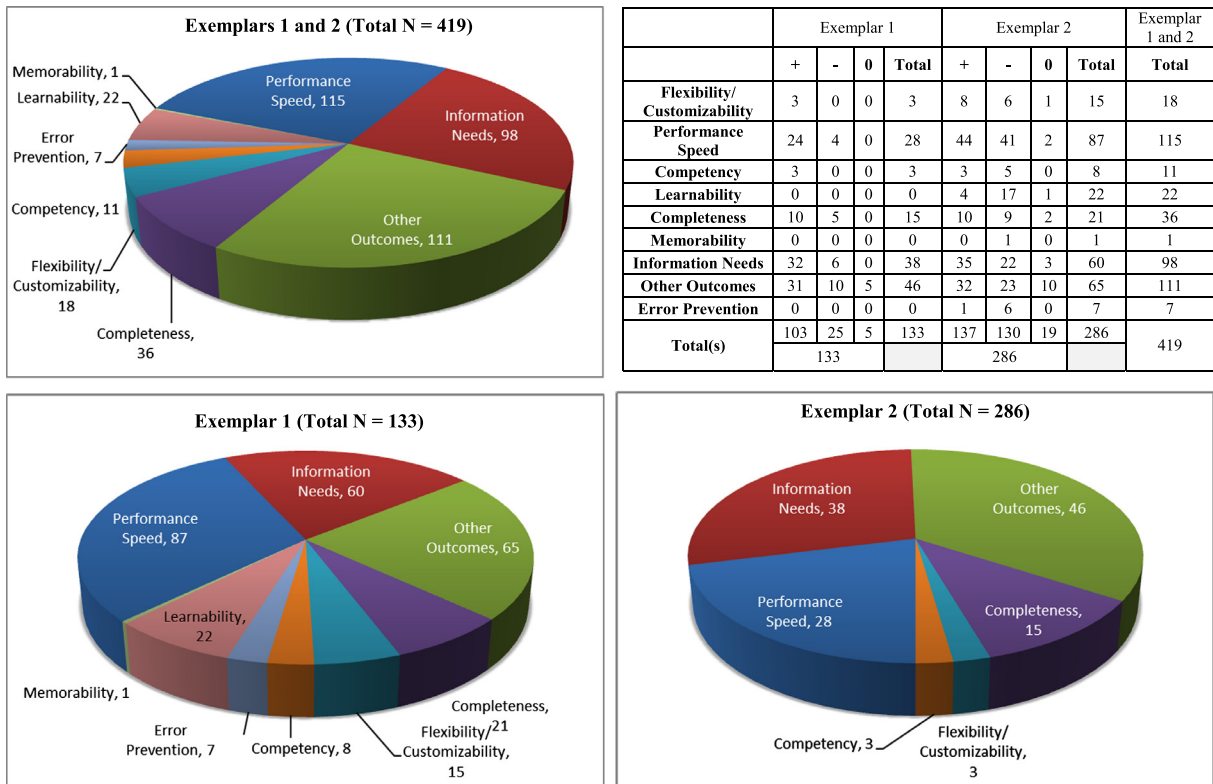
|  | Exemplar 1 | | | | Exemplar 2 | | | | Exemplar 1 and 2 |
|---|---|---|---|---|---|---|---|---|---|
|  | + | - | 0 | Total | + | - | 0 | Total | Total |
| Flexibility/ Customizability | 3 | 0 | 0 | 3 | 8 | 6 | 1 | 15 | 18 |
| Performance Speed | 24 | 4 | 0 | 28 | 44 | 41 | 2 | 87 | 115 |
| Competency | 3 | 0 | 0 | 3 | 3 | 5 | 0 | 8 | 11 |
| Learnability | 0 | 0 | 0 | 0 | 4 | 17 | 1 | 22 | 22 |
| Completeness | 10 | 5 | 0 | 15 | 10 | 9 | 2 | 21 | 36 |
| Memorability | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| Information Needs | 32 | 6 | 0 | 38 | 35 | 22 | 3 | 60 | 98 |
| Other Outcomes | 31 | 10 | 5 | 46 | 32 | 23 | 10 | 65 | 111 |
| Error Prevention | 0 | 0 | 0 | 0 | 1 | 6 | 0 | 7 | 7 |
| Total(s) | 103 | 25 | 5 | 133 | 137 | 130 | 19 | 286 | 419 |
|  | | | 133 | | | | 286 | | |





**Fig. 2.** Frequency of code use in exemplars.

**Table 3**
Inter-rater reliability by Health-ITUEM concept.

|  | Exemplar 1 IRR (%) | Exemplar 2 IRR (%) | IRR Total (%) |
|---|---|---|---|
| Flexibility/Customizability | 100 | 59 | 66 |
| Performance speed | 82 | 88 | 87 |
| Competency | 83 | 100 | 95 |
| Learnability | 100 | 75 | 88 |
| Completeness | 87 | 100 | 93 |
| Memorability | 100 | 100 | 100 |
| Information needs | 100 | 86 | 91 |
| Other outcomes | 97 | 91 | 93 |
| Error prevention | 100 | 100 | 100 |
| Total | 94 | 89 | 90 |

future use of this concept code be clearly defined by either "Flexi-bility" or "Customizability" so that nomenclature does not nega-tively impact coding or rater agreement.

We adjusted the Health-ITEUM major concept codes into sub-code triplets (i.e. negative, positive, neutral) to assess our data at a finer level of granularity. Fig. 2 indicates that the ratio of + to − sub coding is greatly varied between Exemplar 1 and Exemplar 2. Exemplar 1 had overwhelmingly positive responses (103 positive, 25 negative, and 5 neutral). In contrast Exemplar 2 was more bal-anced between positive and negative responses (133 positive, 137 negative, and 15 neutral). There are important distinctions be-tween the two exemplars, which may have contributed to these differences. In Exemplar 1 participants were asked to reflect on mobile tools and apps that they were already using prior to partic-ipating in the study. Participants in Exemplar 2 did not select their mobile devices or apps since they were provided by the researcher team. Given these differences, it is likely that participants in Exem-plar 1 would speak more favorably about mobile tools and apps since they were of their own choosing. These findings are congruent with earlier models of technology acceptance which po-sit that use of technology is influenced by perceived usefulness and ease of use [22]. Findings similar to ours have been reported in the development mobile applications for wellness management [8]. There were few excerpts that were coded as neutral. Use of sub code triplets (+, −, neutral) was useful in understanding some of the facilitators and barriers of mHealth technology.

Moreover, the Health-ITUEM addressed challenges, which com-monly contribute to the complexity of analyzing IT related qualita-tive data. For instance, identifying important units of analysis are critical to extrapolating from data. The Health-ITUEM was devel-oped by integrating concepts from The Technology Acceptance Model [TAM] [22] and International Organization for Standards 9241-11 [ISO 9241-11] [43] which are well-studied usability mod-els. The integrative approach of the Health-ITUEM builds on the strengths and addresses weaknesses of both the TAM and ISO 9241-11. This model produces a robust usability evaluation frame-work which includes nine concepts which are informed by four overarching constructs: Quality of (Work) Life, Perceived Useful-ness, Perceived Ease of Use, and User Control [44]. Our data dem-onstrates the usefulness of this model in understanding the usability issues in mHealth technology

### 6.1. Limitations

Our study seeks to evaluate the use of mHealth technology, which encompasses both devices and apps. However, our results indicate that there may be a confounding effect between the usability of the mobile device and the mobile application embed-ded in the device. One future solution may include assessing the usability of the mobile device (hardware) first and then assessing the usability of the mobile application (software). While qualita-

tive methods have been a popular data collection type in informatics research [33,45,46], there are also unique data analysis challenges, such as unit of analysis identification, construct development, within-group dynamics, between-group variations, and data inconsistency [47,48]. These limitations can be exacerbated in focus group data collection with adolescents. To mitigate these issues, the research team trained focus group leaders on youth moderation techniques, and took care to focus directly on topics related to the research question. Finally our study population was largely Hispanic, which limits the generalizability of our findings.

### 6.2. Future work

Our study population only included adolescents and was largely Hispanic. Future work should include a more generalizable sample of mobile technology users. Moreover, given the breadth of mobile technology that is currently being used for health, the Health-ITUEM should be evaluated in the context of a wider range of mobile devices (e.g. iPad, Tablets, etc.). Lastly, since Error prevention, Learnability, and Memorability were the least frequently used codes; the Health-ITUEM should be assessed on mobile devices that address IT cognitive load. Future studies should aim to address the limitations that exist in our work.

## 7. Conclusion

The Health-ITUEM is evidence-based and draws its concepts, constructs, and items from widely used and tested usability frameworks. The aim of this study was to elucidate the usefulness of the Health-ITUEM for evaluating the usability of mHealth technology. The framework was applied to two unique data sets. Moreover, we wanted to be able to capture the data within the concepts to a greater granularity. Having adjusted the concept codes to account for positive, negative, and neutral response types on a code-by-code basis, we were able to identify the most commonly occurring concept codes, which allowed us to discern if the tone of participants eschewed negatively, positively or stayed neutral for that concept. Ultimately, we found that the Health-ITUEM offers a new framework for understanding the usability issues related to mHealth technology. This study demonstrated the flexibility, robustness, and limitations of this model. In our estimation, the Health-ITUEM framework advances the science of mHealth technology evaluation and supports the effective use of these tools.

## Acknowledgments

## References

[1] Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008;35(2):177–81.
[2] Dolan B. Report: 13K iPhone consumer health apps in 2012; 2011. <http://mobihealthnews.com/13368/report-13k-iphone-consumer-health-apps-in-2012/> [cited 03.18.13].
[3] Ritterband LM, Andersson G, Christensen HM, Carlbring P, Cuijpers P. Directions for the international society for research on Internet interventions (ISRII). J Med Internet Res 2006;8(3):e23 [Epub 2006/10/13].
[4] Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004;6(4):e40 [Epub 2005/01/06].
[5] Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [Epub 2006/07/27].
[6] Pfaeffli L, Maddison R, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mHealth cardiac rehabilitation exercise intervention: findings from content development studies. BMC Cardiovasc Disorders 2012;12:36.
[7] Abroms LC, Ahuja M, Kodl Y, Thaweethai L, Sims J, Winickoff JP, et al. Text2Quit: results from a pilot test of a personalized, interactive mobile health smoking cessation program. J Health Commun 2012;17(Suppl 1):44–53.
[8] Mattila E, Parkka J, Hermersdorf M, Kaasinen J, Vainio J, Samposalo K, et al. Mobile diary for wellness management–results on usage and usability in two user studies. IEEE Trans Inform Technol Biomed: Publ IEEE Eng Med Biol Soc 2008;12(4):501–12.
[9] Furlow B. mHealth apps may make chronic disease management easier; 2012. <http://www.clinicaladvisor.com/mhealth-apps-may-make-chronic-disease-management-easier/article/266782/> [cited 03.17.13].
[10] Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362.
[11] Whitlock LA, McLaughlin AC, editors. Usability of blood glucose tracking apps for older users. Human factors and ergonomics society 56th annual meeting. Santa Monica, CA; 2012.
[12] Wolf JA, Moreau J, Akilov O, Patton T, English JC, Ho J, et al. Diagnostic inaccuracy of smartphone applications for melanoma detection. JAMA Dermatol 2013:1–4.
[13] Patterson ES, Cook RI, Render ML. Improving patient safety by identifying side effects from introducing bar coding in medication administration. J Am Med Inform Assoc: JAMIA 2002;9(5):540–53.
[14] Kossman SP, Scheidenhelm SL. Nurses' perceptions of the impact of electronic health records on work and patient outcomes. Comput Inform Nurs: CIN 2008;26(2):69–77.
[15] Kushniruk A, Triola M, Stein B, Borycki E, Kannry J. The relationship of usability to medical error: an evaluation of errors associated with usability problems in the use of a handheld application for prescribing medications. Studies Health Technol Inform 2004;107(Pt 2):1073–6.
[16] Kaufman D, Roberts WD, Merrill J, Lai TY, Bakken S. Applying an evaluation framework for health information system design, development, and implementation. Nurs Res 2006;55(2 Suppl):S37–42.
[17] Yen PY. Health information technology usability evaluation: methods, models and measures. New York: Columbia University; 2010.
[18] Dumas JS, Redish JC. A practical guide to usability testing. Intellect Ltd.; 1999.
[19] Abran A, Khelifi A, Suryn W, Seffah A. Usability meanings and interpretations in ISO standards. Softw Qual J 2003;11(4):325–38.
[20] Shackel B. Usability—context, framework, definition, design and evaluation. In: Shackel B, Richardson SJ, editors. New York (NY): Cambridge University Press; 1991.
[21] Nielsen J. Heuristic evaluation. In: Nielsen J, Mack RL, editors. Usability inspection methods. New York (NY): John Wiley & Sons; 1994.
[22] Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart 1989;13(3):319–40.
[23] Folmer E, Bosch J. Architecting for usability: a survey. J Syst Softw 2004;70(1–2):61–78.
[24] Nielsen J, Mack RL. Usability inspection methods, vol. xxiv. New York: Wiley; 1994. 413 p..
[25] Shneiderman B, Plaisant C. Designing the user interface: strategies for effective human–computer interaction, vol. xviii. 5th ed. Boston: Addison-Wesley; 2010. 606 p..
[26] Norman DA. The design of everyday things, vol. xxi. 1st Basic paperback ed. New York: Basic Books; 2002. 257 p..
[27] Amith M, Loubser PG, Chapman J, Zoker KC, Rabelo Ferreira FE. Optimization of an EHR mobile application using the UFuRT conceptual framework. In: AMIA annual symposium proceedings/AMIA symposium AMIA symposium; 2012. p. 209–17.
[28] Luxton DD, Mishkind MC, Crumpton RM, Ayers TD, Mysliwiec V. Usability and feasibility of smartphone video capabilities for telehealth care in the US military. Telemed J e-Health: Official J Am Telemed Assoc 2012;18(6):409–12.
[29] Burnay E, Cruz-Correia R, Jacinto T, Sousa AS, Fonseca J. Challenges of a mobile application for asthma and allergic rhinitis patient enablement-interface and synchronization. Telemed J e-Health: Official J Am Telemed Assoc 2013;19(1):13–8.

[30] Sheehan B, Lee YJ, Rodriguez M, Tiase V, Schnall R. A comparison of usability factors of four mobile devices for accessing healthcare information by adolescents. Appl Clin Inform 2012;3(4):356–66.

[31] Sparkes J, Valaitis R, McKibbon A. A usability study of patients setting up a cardiac event loop recorder and BlackBerry gateway for remote monitoring at home. Telemed J e-Health: J Am Telemed Assoc 2012;18(6):484–90.

[32] Zhang DS, Adipat B. Challenges, methodologies, and issues in the usability testing of mobile applications. Int J Hum-Comput Int 2005;18(3):293–308.

[33] Schnall R, Smith AB, Sikka M, Gordon P, Camhi E, Kanter T, et al. Employing the FITT framework to explore HIV case managers' perceptions of two electronic clinical data (ECD) summary systems. Int J Med Inform 2012;81(10):e56–62.

[34] McHorney CA, Ware Jr JE, Lu JF, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32(1):40–66.

[35] Creswell JW. Qualitative inquiry and research design choosing among five traditions. Thousand Oaks (CA): Sage Publications; 1998.

[36] Schnall R, Okoniewski A, Tiase V, Low A, Rodriguez M, Kaplan S. Using text messaging to assess adolescents' health information needs: an ecological momentary assessment. J Med Internet Res 2013;15(3):e54.

[37] Lieber E, Weisner TS. Meeting the practical challenges of mixed methods research. Mix methods soc behav res. 2nd ed. Thousand Oaks (CA): SAGE Publications; 2010.

[38] Lieber E. Mixing qualitative and quantitative methods: insights into design and analysis issues. J Ethnogr Qual Res 2009;3(4):218–27.

[39] Birkimer JC, Brown JH. Back to basics – percentage agreement measures are adequate, but there are easier ways. J Appl Behav Anal 1979;12(4):535–43.

[40] Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012;12:CD007457.

[41] Whittaker R, Merry S, Dorey E, Maddison R. A development and evaluation process for mHealth interventions: examples from New Zealand. J Health Commun 2012;17(Suppl 1):11–21.

[42] Cimino JJ. Desiderata for controlled medical vocabularies in the twenty-first century. Methods Inf Med 1998;37(4–5):394–403 [Epub 1998/12/29].

[43] 9241-11. I. Ergonomic requirements for office work with visual display terminals (VDTs) ~ Part 11: Guidance on usability; 1998.

[44] Yen PY, Wantland D, Bakken S. Development of a customizable health IT usability evaluation scale. In: AMIA annual symposium proceedings/AMIA symposium AMIA symposium; 2010. p. 917–21 [Epub 2011/02/25].

[45] Ash JS, Guappone KP. Qualitative evaluation of health information exchange efforts. J Biomed Inform 2007;40(6 Suppl):S33–9.

[46] Kaufman DR, Starren J, Patel VL, Morin PC, Hilliman C, Pevzner J, et al. A cognitive framework for understanding barriers to the productive use of a diabetes home telemedicine system. In: AMIA annual symposium proceedings/AMIA symposium AMIA, symposium; 2003. p. 356–60.

[47] Sandelowski M. Rigor or rigor mortis: the problem of rigor in qualitative research revisited. ANS Adv Nurs Sci 1993;16(2):1–8.

[48] Angen MJ. Evaluating interpretive inquiry: reviewing the validity debate and opening the dialogue. Qual Health Res 2000;10(3):378–95.