Original Paper

An 11-Item Measure of User- and Human-Centered Design for Personal Health Tools (UCD-11): Development and Validation

Holly O Witteman^{1,2,3}, PhD; Gratianne Vaisson^{1,3}, MSc; Thierry Provencher¹, BA; Selma Chipenda Dansokho¹, PhD; Heather Colquhoun⁴, PhD, OT Reg (Ont); Michele Dugas^{1,2}, MSc; Angela Fagerlin^{5,6}, PhD; Anik MC Giguere^{1,2}, PhD; Lynne Haslett⁷, BSCN, MEd, NP-PHC; Aubri Hoffman⁸, PhD; Noah M Ivers^{4,9}, MD, PhD; France Légaré^{1,2}, MD, PhD; Marie-Eve Trottier^{1,3}, MSc; Dawn Stacey^{10,11}, RN, PhD; Robert J Volk⁸, PhD; Jean-Sébastien Renaud^{1,2}, PhD

¹Université Laval, Quebec City, QC, Canada

²VITAM Research Centre for Sustainable Health, Quebec City, QC, Canada

³CHU de Québec-Université Laval, Quebec City, QC, Canada

⁴University of Toronto, Toronto, ON, Canada

⁷East End Community Health Centre, Toronto, ON, Canada

⁸The University of Texas MD Anderson Cancer Center, Houston, TX, United States

⁹Women's College Hospital, Toronto, ON, Canada

¹⁰University of Ottawa, Ottawa, ON, Canada

¹¹Ottawa Hospital Research Institute, Ottawa, ON, Canada

Corresponding Author:

Holly O Witteman, PhD Université Laval 1050, avenue de la Médecine Quebec City, QC Canada Phone: 1 418 656 2131 Email: holly.witteman@fmed.ulaval.ca

Abstract

Background: Researchers developing personal health tools employ a range of approaches to involve prospective users in design and development.

Objective: The aim of this paper was to develop a validated measure of the human- or user-centeredness of design and development processes for personal health tools.

Methods: We conducted a psychometric analysis of data from a previous systematic review of the design and development processes of 348 personal health tools. Using a conceptual framework of user-centered design, our team of patients, caregivers, health professionals, tool developers, and researchers analyzed how specific practices in tool design and development might be combined and used as a measure. We prioritized variables according to their importance within the conceptual framework and validated the resultant measure using principal component analysis with Varimax rotation, classical item analysis, and confirmatory factor analysis.

Results: We retained 11 items in a 3-factor structure explaining 68% of the variance in the data. The Cronbach alpha was .72. Confirmatory factor analysis supported our hypothesis of a latent construct of user-centeredness. Items were whether or not: (1) patient, family, caregiver, or surrogate users were involved in the steps that help tool developers understand users or (2) develop a prototype, (3) asked their opinions, (4) observed using the tool or (5) involved in steps intended to evaluate the tool, (6) the process had 3 or more iterative cycles, (7) changes between cycles were explicitly reported, (8) health professionals were asked their opinion and (9) consulted before the first prototype was developed or (10) between initial and final prototypes, and (11) a panel of other experts was involved.

⁵University of Utah, Salt Lake City, UT, United States

⁶Salt Lake City VA Center for Informatics Decision Enhancement and Surveillance, Salt Lake City, UT, United States

Conclusions: The User-Centered Design 11-item measure (UCD-11) may be used to quantitatively document the user/human-centeredness of design and development processes of patient-centered tools. By building an evidence base about such processes, we can help ensure that tools are adapted to people who will use them, rather than requiring people to adapt to tools.

(J Med Internet Res 2021;23(3):e15032) doi: 10.2196/15032

KEYWORDS

patient-centered care; patient participation; health services research; validation studies as topic; surveys and questionnaires; humans; user-centred design, human-centred design; user-centered design; human-centered design; co-design; instrument; scale; index; patient and public involvement

Introduction

Many products and applications aim to support people in managing their health and living their lives. These include physical tools like wheelchairs [1] or eating utensils [2], medical devices like insulin pumps [3] or home dialysis equipment [4], assistive devices like screen readers [5] or voice aids [6], digital applications like eHealth tools [7] or mHealth (mobile health) tools [8,9], tools for collecting patient-reported outcome or experience measures [10,11], patient decision aids [12], and a variety of other personal health tools.

None of these tools can achieve their intended impact if they are not usable by and useful to their intended users. Accordingly, designers and developers frequently seek to involve users in design and development processes to ensure such usability and utility. In a previous systematic review of the design and development processes of a range of personal health tools, we documented that the extent and type of user involvement varies widely [13]. Structured ways to describe this variation could help capture data across projects and may serve to build an evidence base about the potential effects of design and development processes.

The systematic review was grounded in a framework of user-centered design [14], shown in Figure 1, that we had synthesized from foundational literature. In this framework, a user is any person who interacts with (in other words, uses) a system, service, or product for some purpose. User-centered design is a long-standing approach [15], sometimes referred to as human-centered design [16], that is both conceptually and methodologically related to terms like design thinking and co-design [17]. It is intended to optimize the user experience of a system, service, or product [18-21]. While user-centered design is not the only approach that may facilitate such optimization, it served as a useful overall framework for structuring the data reported in the papers included in our systematic review. In our work, we define user-centered design as a fully or semistructured approach in which people who currently use or who could in the future use a system, service, or product are involved in an iterative process of optimizing its user experience. This iterative process includes one or more steps to understand prospective users, including their needs, goals, strengths, limitations, contexts (eg, the situations or environments in which they will use a tool), and intuitive processes (eg, the ways in which they currently address the issue at hand or use similar systems, services, or products). The iterative process also includes one or more steps to develop or

```
https://www.jmir.org/2021/3/e15032
```

refine prototypes, and one or more steps to observe prospective users' interactions with versions of the tool.

Iivari and Iivari [22] noted that the different ways in which user-centeredness is described in the literature imply four distinct meanings or dimensions: (1) user focus, meaning that the system is designed and developed around users' needs and capabilities; (2) work-centeredness, meaning that the system is designed and developed around users' workflow and tasks; (3) user involvement or participation, meaning that the design and development process involves users or users participate in the process; and (4) system personalization, meaning the system is individualized by or for individual users. Our definition of user-centeredness and framework of user-centered design draw most strongly upon the third of these (user involvement or participation) as a means to achieve the first (user focus) and fourth (system personalization). The second meaning (work-centeredness) is less relevant here as it refers to paid work in the original definition. However, it may be worth noting the considerable work that people may need to undertake to make health decisions or to live with illness or disability [23-26].

In our previous systematic review, we used the above framework of user-centered design to extract and organize the extracted data from 623 articles describing the design, development, or evaluation processes of 390 personal health tools, predominantly patient decision aids, which are tools intended to support personal health decisions [13]. We documented a wide range of practices, leading us to question whether it might be possible to use this structured data set to develop a measure to capture aspects of the user-centeredness of design and development processes, similar to how other measures capture complex concepts or processes that are not directly observable; for example, social capital [27], learning processes [28], health-related quality of life [29], and health care quality [30]. We posited that although a high-level summary of design and development processes would not be able to capture nuances within each project, it may nonetheless be valuable to be able to capture information that would otherwise be difficult to synthesize across diverse projects. Of the 390 included personal health tools in our previous systematic review, 348 met our prespecified criterion regarding sufficient information related to the design and development processes, while the other 42 reported information only about their evaluation. Therefore, in this study, using an existing structured data set describing the design and development of 348 personal health tools, we aimed to derive a measure of the user- or human-centeredness of the design and development of personal health tools.

Figure 1. User-centered design framework.



Methods

Validity Framework and Overall Approach

Guided by an established validity framework [31], we developed and validated a measure using classical test theory. Classical test theory is a set of concepts and methods developed over decades [32-35] based on the earlier work of Spearman [36,37]. It posits that it is possible to develop items that each assess part of a construct that we wish to measure but is not directly observable; for example, patient-reported outcomes [38,39], responsibility and cooperation in a group learning task [40], or, in our case, the user-centeredness of a design and development process. Classical test theory further posits that each item captures part of what we wish to measure, plus error, and assumes that the error is random. This means that as the number of items increases, the overall error drops toward zero. Classical test theory is simpler than other methods (eg, item response theory, generalizability theory) and therefore satisfied the criterion of parsimony, which refers to choosing the simplest approach that meets one's measurement and evaluation needs [41].

The validity framework reflects consensus in the field of measurement and evaluation about what indicates the validity of a measure, particularly in domains such as education that focus on assessment. Specifically, validity refers to the extent to which evidence and theory support interpretations of the score for its proposed use [31]. The validity framework therefore proposes five ways in which a measure may or may not demonstrate validity: its content validity, its response process, its internal structure, its relationship to other variables, and the consequences of the measure [31,42]. Because our aim was to develop a new measure in an area with few metrics, our study directly addresses the first three of these five. We discuss how

```
https://www.jmir.org/2021/3/e15032
```

related and future research might inform the fourth and fifth ways of assessing validity.

Content Validity

Content validity (point 1 in the validity framework [31]) refers to how well items match the definition of a construct. To ensure content validity of items, in our original systematic review, we had used foundational literature [15,16,18,43-45]; held monthly or bimonthly consultations in person and by teleconference over the course of 2 years within our interdisciplinary group of experts, including patients, caregivers, health professionals, academic researchers, and other stakeholders; and consulted with 15 additional experts outside the research team [13]. Discussions over the years of the project centered on the items themselves as well as prioritization of items according to their relevance within our conceptual framework.

Response Process

Response process (point 2 in the validity framework [31]) refers to quality control when using a measure [42]. In our case, it is the extent to which analysts are able to accurately and consistently assign a value to each item in the measure. We had refined the response process for each item through an iterative process of data extraction and data validation. This included consultation with 15 external experts and four rounds of pilot data extraction and refinement of response processes across randomly selected sets of five articles each time (total: 20 articles). We had also confirmed the accuracy of the extracted data with the authors of the original articles included in the systematic review and found very low rates of error [13].

Internal Structure

Internal structure (point 3 in the validity framework [31]) addresses to what extent items in a measure are coherent among

XSL•FO RenderX

themselves and conform to the construct on which the proposed score interpretations are based. In our case, good internal structure would indicate that although the items are distinct, they are all measuring the same overall construct. We would therefore be able to detect patterns reflecting this construct. Specifically, processes that are more user-centered would score higher, and processes that are less user-centered would score lower. To assess this, we first identified which prioritized items formed a positive definite matrix of tetrachoric correlations. Tetrachoric correlations are similar to correlations between continuous variables (eg, Pearson correlations) but instead calculate correlations between dichotomous (ie, yes/no, true/false) variables. A matrix can be thought of as something like a table of numbers. A matrix of correlations is a square matrix, meaning it has the same number of rows as columns, in which any given row or column of the matrix represents a vector made up of an item's correlations with each of the other items in the set. The diagonal of the matrix will contain values of 1 because those cells represent each item's correlation with itself. Positive definite matrices are matrices that are able to be inverted. For readers unfamiliar with matrix algebra, a useful analogy may be that inversion is to matrices as division is to numbers. Inversion is possible when the vectors (in our case, vectors of tetrachoric correlations between potential items in the measure) that make up the matrix are sufficiently independent of each other. Matrix inversion is required to conduct principal component analysis.

We identified the items to compose the set whose correlations would make up the matrix by first rank ordering possible items in the data set according to their priority in our conceptual framework, using the expertise of our interdisciplinary team (see the Patient Partnership section). We then built the matrix in a stepwise fashion, adding items until the matrix of correlations was no longer invertible. Then, based on classical item analysis in which we required discrimination indices >0.2 [46-48], we formed a group of items with an acceptable value of Kaiser's measure of sampling adequacy (>0.6 [49]), meaning that they share enough common variance to allow principal component analysis. We then conducted this analysis with Varimax rotation. Using the resultant scree plot and content expertise based on our conceptual framework, we identified components that explained sufficient variance in the data, retaining items with loadings over 0.4 on at least one factor. We also performed classical item analysis to assess the resultant psychometric properties of the items in the measure. Finally, we used confirmatory factor analysis with unweighted least squares estimation to test our hypothesis of the existence of a latent construct of user-centeredness explaining the variance in

the three components. In other words, we tested whether or not our data suggested that the components we found in our analysis shared a common root.

Applying the Measure Within the Data Set

We applied the resulting measure within the data set to examine and compare scores for the two groups of projects within the original study: patient decision aids, which could have been developed in any way, and other personal health tools that specifically described their design and development method as user- or human-centered design. To explore potential changes in design and development methods over time, we plotted scores within the two groups according to the year of publication of the first paper published about each project. To provide further information about the distribution of scores within the data set used to develop the measure, we calculated percentile ranks of the scores within the data set, applying the definition of a percentile rank that, for example, being in the 97th percentile indicates that the score was higher than 96% of those tested [50].

We conducted analyses in SAS, version 9.4 (SAS Institute Inc) and in R, version 3.3.2 (The R Foundation).

Patient Partnership

Patients and other stakeholders participated in every aspect of the research for this project overall as members of the research team. For the development of the measure, patient and caregiver partners were most involved in the prioritization of items for analysis.

Availability of Data and Materials

Data used in this study are available via Scholars Portal Dataverse [51].

Results

Items Retained in the User-Centered Design 11-Item Measure (UCD-11)

Out of 19 identified potential variables, we retained 11 items in a three-factor structure explaining 68% of the variance in the data, which refers to the variance within the 19 variables. The Kaiser's measure of sampling accuracy was 0.68, which is considered acceptable [49]. Each item is binary and is scored as either present or absent. Table 1 and Figure 2 show the 11 retained items and factor structure. The Cronbach alpha for all 11 items was .72, indicating acceptable internal consistency [52].



Witteman et al

 Table 1. Final measure with factor loadings.

Items ^a	Explanations and examples	Factors		
		Preprototype involvement	Iterative respon- siveness	Other expert involvement
1. Were potential end users (eg, pa- tients, caregivers, family and friends, surrogates) involved in any steps to help understand users (eg, who they are, in what context might they use the tool) and their needs?	Such steps could include various forms of user research, including formal or informal needs assessment, focus groups, surveys, contextual inquiry, ethnographic obser- vation of existing practices, literature review in which users were involved in appraising and interpreting existing literature, development of user groups, personas, user profiles, tasks, or scenarios, or other activities	0.82	b	_
2. Were potential end users involved in any steps of designing, developing, and/or refining a prototype?	Such steps could include storyboarding, reviewing the draft design or content before starting to develop the tool, and designing, developing, or refining a prototype	0.83	_	_
3. Were potential end users involved in any steps intended to evaluate proto- types or a final version of the tool?	Such steps could include feasibility testing, usability testing with iterative prototypes, pilot testing, a random- ized controlled trial of a final version of the tool, or other activities	_	0.78	_
4. Were potential end users asked their opinions of the tool in any way?	For example, they might be asked to voice their opinions in a focus group, interview, survey, or through other methods	—	0.80	_
5. Were potential end users observed using the tool in any way?	For example, they might be observed in a think-aloud study, cognitive interviews, through passive observation, logfiles, or other methods	_	0.71	_
6. Did the development process have3 or more iterative cycles?	The definition of a cycle is that the team developed something and showed it to at least one person outside the team before making changes; each new cycle leads to a version of the tool that has been revised in some small or large way	_	0.64	_
7. Were changes between iterative cy- cles explicitly reported in any way?	For example, the team might have explicitly reported them in a peer-reviewed paper or in a technical report. In the case of rapid prototyping, such reporting could be, for example, a list of design decisions made and the ratio- nale for the decisions	_	0.87	_
8. Were health professionals asked their opinion of the tool at any point?	Health professionals could be any relevant professionals, including physicians, nurses, allied health providers, etc. These professionals are not members of the research team. They provide care to people who are likely users of the tool. Asking for their opinion means simply asking for feedback, in contrast to, for example, observing their in- teraction with the tool or assessing the impact of the tool on health professionals' behavior	_	_	0.80
9. Were health professionals consulted before the first prototype was developed?	Consulting before the first prototype means consulting prior to developing anything. This may include a variety of consultation methods	0.49	_	0.75
10. Were health professionals consulted between initial and final prototypes?	Consulting between initial and final prototypes means some initial design of the tool was already created when consulting with health professionals	_	_	0.91
11. Was an expert panel involved?	An expert panel is typically an advisory panel composed of experts in areas relevant to the tool if such experts are not already present on the research team (eg, plain lan- guage experts, accessibility experts, designers, engineers, industrial designers, digital security experts, etc). These experts may be health professionals but not health profes- sionals who would provide direct care to end users	_	_	0.56

^aAll items are scored as yes=1 and no=0. When assigning scores from written reports of projects, if an item is not reported as having been done, it is scored as not having been done. The total score on the User-Centered Design 11-item scale (UCD-11) is the number of yes answers and therefore ranges from 0 to 11.

^bFactor loadings <0.40 are not shown. This is because loadings <0.40 indicate that the item does not contribute substantially to that factor.

XSL•FO RenderX

Figure 2. Items and scoring of the User-Centered Design 11-item measure (UCD-11).



The preprototype involvement factor included 2 items: (1) whether prospective users (ie, patient, family, caregiver, or surrogate users) were involved in steps that help tool developers understand users, and (2) whether prospective users were involved in the steps of prototype development. The iterative responsiveness factor included 5 items: (3) whether prospective users were observed using the tool; (5) whether they were involved in steps intended to evaluate the tool; (6) whether the development process had 3 or more iterative cycles; and (7) whether changes between iterative cycles were explicitly reported. The other expert involvement factor included 4 items: (8) whether health professionals were consulted before the first prototype was

https://www.jmir.org/2021/3/e15032

RenderX

developed; (10) whether health professionals were consulted between initial and final prototypes; and (11) whether an expert panel of nonusers was involved. As shown in Table 1, each of the 11 items is formulated as a question that can be answered by "yes" or "no," and is assumed to be "no" if the item is not reported. The score is the number of "yes" answers and therefore ranges from 0 to 11.

Items Not Retained in UCD-11

The 8 items not retained due to a lack of sufficient explanation of variance were whether or not: (1) the users involved were currently dealing with the health situation, (2) a formal patient organization was involved, (3) an advisory panel of users was involved, (4) there were users who were formal members of the

research team, (5) users were offered incentives or compensation of any kind for their involvement (eg, cash, gift cards, payment for parking), (6) people who were members of any vulnerable population were explicitly involved [53], (7) users were recruited using convenience sampling, and (8) users were recruited using methods that one might use to recruit from populations that may be harder to reach (eg, community centers, purposive sampling, snowball sampling).

Classical Test Theory and Confirmatory Factor Analysis Results

Classical item difficulty parameters ranged from 0.28 to 0.85 on a scale ranging from 0 to 1 and discrimination indices from 0.29 to 0.46, indicating good discriminating power [46-48]. This means that the items discriminate well between higher and lower overall scores on the measure. Confirmatory factor analysis demonstrated that a second-order model provided an acceptable to good fit [54] (standardized root mean residual=0.09; goodness of fit index=0.96; adjusted goodness of fit index=0.94; normed fit index=0.93), supporting our hypothesis of a latent construct of user-centeredness that explains the three factors. This means that UCD-11 provides a

single score or a single number rather than multiple numbers, and may therefore be used as a unidimensional measure. Had we not observed a single latent construct, the measure would have always needed to be reported with scores for each factor.

Scores Within the Data Set

As expected when applying a measure to the data set used to develop it, scores within the data set were distributed across the full range of possible scores (ie, 0 to 11). The median score was 6 out of a possible 11 (IQR 3-8) across all 348 projects. Median scores were 5 out of a possible 11 (IQR 3-8) for the design and development of patient decision aids, and 7 out of a possible 11 (IQR 5-8) for other personal health tools in which the authors specifically described their design and development method as user- or human-centered design. The 95% CI of the difference in mean scores for patient decision aid projects compared to projects that described their approach as user- or human-centered design was (-1.5 to -0.3). Figure 3 shows scores over time within the two groups. There were no discernable time trends in UCD-11 scores.

Table 2 provides percentiles for each possible UCD-11 scorewithin the data set of 348 projects.

Figure 3. User-Centered Design 11-Item scale (UCD-11) scores by year of publication of the first paper describing a project. UCD refers to other personal health tools explicitly naming user- or human-centered design as the guiding process. PtDA: patient decision aids.





Table 2.	Percentile ranks o	f the User	-Centered Design	n 11-item scale	(UCD-11)	scores
----------	--------------------	------------	------------------	-----------------	----------	--------

UCD-11 score	Percentile rank	Interpretation
0	Oth	The score is not higher than any other scores in the data set.
1	4th	The score is higher than 3% of scores in the data set.
2	8th	The score is higher than 7% of scores in the data set.
3	17th	The score is higher than 16% of scores in the data set.
4	27th	The score is higher than 26% of scores in the data set.
5	36th	The score is higher than 35% of scores in the data set.
6	49th	The score is higher than 48% of scores in the data set.
7	61st	The score is higher than 60% of scores in the data set.
8	74th	The score is higher than 73% of scores in the data set.
9	87th	The score is higher than 86% of scores in the data set.
10	95th	The score is higher than 94% of scores in the data set.
11	99th	The score is higher than 98% of scores in the data set.

Discussion

Principal Results and Comparisons With Prior Work

Our study aimed to derive a measure of user-centeredness of the design and development processes for personal health tools. Applying a conceptual framework of user-centered design allowed us to identify indicators of this construct and develop an internally valid measure. This measure includes items that address the involvement of users and health professionals at every stage of a framework of user-centered design [14] as well as the importance of designing and developing tools in iterative cycles. Given the creative nature of design and development and a wide range of possible tools, the items are high-level assessments of whether or not particular aspects of involvement were present or absent, not assessments of the quality of each aspect.

To the best of our knowledge, ours is the first such validated measure for health applications. Other broadly applicable measures exist that assess the usability or ease of use of tools (eg, the System Usability Scale [55,56]). However, this measure assesses the quality of the resulting tool or system, not the process of arriving at the end product. Process measures do exist, for example, in software, consumer product development, and information systems [57-59].

Barki and Hartwick [54] developed measures centered around the design and development of information systems in professional contexts, with items reported by users. The items in their measures included "I was able to make changes to the formalized agreement of work to be done during system definition" and "I formally reviewed work done by Information Systems/Data Processing staff during implementation." Users also indicated, for example, to what extent they felt the system was needed or relevant to them. Our measure has some items similar to the items in their user participation scale; however, in our measure, users themselves do not need to indicate whether or not a step occurred.

Kujala [58] offers a measure intended to assess the quality of system specifications after these have been developed. Items

```
https://www.jmir.org/2021/3/e15032
```

include "Customer or user requirements are completely defined" and "The correctness of the requirements is checked with real users," assessed on a 4-point Likert scale, with responses ranging from "disagree" to "agree." This measure assesses the quality of user research outputs, which should typically be generated early in a project. In contrast, our measure offers a means of measuring user involvement by the aspects of a design and development process that were or were not done during the entire process.

Subramanyam and colleagues [59] assessed user participation in software development using data collected from time sheets and surveys across 117 projects conducted over 4 years at a large manufacturing firm. Projects often consisted of developing manufacturing and supply chain software. They found that users reported higher satisfaction in projects developing new software when the demands on their time were lowest, whereas developers reported higher satisfaction when users' time spent in the project was highest. Users in this case were employees in the firm, who presumably had other work-related tasks to do as well. Our measure differs from this approach in that we assess involvement in a variety of steps as well as other factors (eg, 3 or more iterative cycles) rather than the total time spent by users.

In summary, our measure aligns somewhat with work from other contexts to measure user-centeredness. The key difference between our measure and previous measures is that ours assesses the process of design and development rather than the quality of the end product, is specific to the context of health-related tools rather than that of information systems or more general contexts, and may be reported or assessed by anyone with sufficient knowledge of the design and development process rather than requiring reporting by users. This latter difference offers flexibility of administration and feasibility for assessing the design and development of completed projects. However, this also means that our measure does not capture the quality of involvement, neither from the perspectives of those involved nor in any sort of external way. Future research should compare the relationship-or lack thereof-between whether or not specific steps occurred in a design and development process and users' perspectives on the quality of the design and

Witteman et al

development process. We also suggest that future research focused on the quality of the process might investigate how or whether including experts in design improves the design and development process and resulting tool. Previous research in tools designed for clinicians has shown that including design and human factors engineering experts generally increases the quality of the tools, and also that the extent of improvement varies considerably according to the individual expert [60].

In addition to the strengths of our study, the first external use of our measure, conducted through advance provision of the measure to colleagues, offered some additional promising indications of its validity, specifically with respect to the fourth and fifth items of the validity framework (relationship to other variables and consequences of the measure) that were not possible to assess in our study. Higgins and colleagues [61] conducted a systematic review of 26 electronic tools for managing children's pain. They aimed to investigate the characteristics of tools still available for patients and families to use versus those that were no longer in use. They found that higher UCD-11 scores were associated with the tools still being available for use after the grant and project had ended [61].

Although case reports suggest that involving users in the design and development of health-related tools can lead to more usable, accepted, or effective tools [62,63], and, as mentioned above, emerging evidence suggests that higher scores on our measure are associated with more sustained availability of tools [61], we lack definitive evidence about the extent to which increasing user-centeredness may improve tools. It may be that there is a point beyond which it is either not feasible or not a good use of limited time and resources to increase involvement. For these reasons, UCD-11 should be considered descriptive, not normative.

Limitations

Our study has two main limitations. First, our data came from published reports, not direct capture of design and development processes. Although we have reason to believe the data are of high quality given our rigorous data validation and low rates of error [13], data from a systematic review of this nature may not contain full details of design and development processes. We chose to use these data because we believed they might offer valuable insights across hundreds of projects. Another research team might choose to draft a list of items from scratch, seek to apply them to new design processes, and validate a measure that way, one project at a time. Second, because our largest data source came from reports of the design and development of patient decision aids, our findings may be overly influenced by practices in the field of shared decision making and patient decision aids. We believe that this focus is appropriate for increasing user-centeredness in the context of health care. Shared decision making has been noted as "the pinnacle of patient-centered care" [64] and patient-centered care has been defined as "care that is respectful of and responsive to individual patient preferences, needs, and values," such that, "patient values guide all clinical decisions" [65], a definition that aligns precisely with the goals of shared decision making [66]. However, it is possible that, because patient decision aids are intended to be used to complement consultation with a health professional, this focus in our data may have led to overemphasis on the role of health professionals in developing tools for use by people outside the health system.

Using UCD-11

Our goal in developing UCD-11 was to offer a straightforward, descriptive measure that can be used by teams as part of reporting their own processes or alternatively by researchers who may apply it to written reports of design and development processes. UCD-11 is intended as a complement to—not a replacement for—detailed descriptions of the design and development processes of personal health tools and is intended to be applied at the end of a project. As stated earlier, it is a descriptive, not normative, measure. Although Higgins and colleagues [61] offered evidence that higher UCD-11 scores are associated with positive implementation outcomes of a personal health tool, we do not have evidence that higher scores necessarily indicate higher-quality design and development processes.

Conclusions

Using a framework of user-centered design synthesized from foundational literature, we were able to derive UCD-11, an internally valid descriptive measure of the user-centeredness of the design and development processes of personal health tools. This measure offers a structured way to consider design and development methods (eg, co-design) when creating tools with and for patients and caregivers. Through measurement and reporting, this measure can help collect evidence about user involvement in order that future research might better specify how we can make the best possible use of the time and effort of all people involved in design and development. We hope this measure will help generate structured data toward this goal and help foster more creation of tools that are adapted to the people who will use them, rather than requiring people to adapt to the tools.

Acknowledgments

As with each of our team's papers, all members of the team were offered the opportunity to coauthor this publication. Not all members elected to accept the invitation, depending on their interest and other commitments. The authors thank team members Sholom Glouberman (patient team member), Jean Légaré (patient team member), Carrie A Levin (patient decision aid developer team member), Karli Lopez (caregiver team member), Victor M Montori (academic team member), and Kerri Sparling (patient team member), who all contributed to the broader project that led to the work presented here.

This study was funded by the Patient-Centered Outcomes Research Institute (PCORI): ME-1306-03174 (PI: HOW) and the Canadian Institutes of Health Research (CIHR): FDN-148246 (PI: HOW). PCORI and CIHR had no role in determining the study design, the plans for data collection or analysis, the decision to publish, nor the preparation of this manuscript. HOW is funded

by a Tier 2 Canada Research Chair in Human-Centred Digital Health and received salary support during this work from Research Scholar Junior 1 and 2 Career Development Awards by the Fonds de Recherche du Québec—Santé (FRQS). AMCG received salary support during this work from a Research Scholar Junior 2 Career Development Award by the FRQS. NMI is funded by a Tier 2 Canada Research Chair in Implementation of Evidence-based Practice and received salary support during this work from a New Investigator Award by the CIHR as well as a New Investigator Award from the Department of Family and Community Medicine, University of Toronto. FL is funded by a Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation. DS holds a University of Ottawa Research Chair in Knowledge Translation to Patients.

During the course of this project, Carrie A Levin (patient decision aid developer team member) received salary support as research director for the Informed Medical Decisions Foundation, the research division of Healthwise, Inc, a not-for-profit organization that creates products including patient decision aids.

Authors' Contributions

HOW, GV, and JSR were responsible for study conceptualization and methodology; JSR for validation; HOW, GV, and JSR for formal analysis; HOW, GV, SCD, HC, MD, AF, AMCG, LH, AH, NMI, FL, TP, DS, MET, RJV, and JSR for investigation; GV and TP for data curation; HOW and JSR for writing the original draft; HOW, GV, SCD, HC, MD, AF, AMCG, LH, AH, NMI, FL, TP, DS, MET, RJV, and JSR for writing, review, and editing; HOW and SCD for project administration; and HOW, SCD, HC, AF, AMCG, LH, AH, NMI, FL, DS, RJV, and JSR for funding acquisition.

Conflicts of Interest

No conflicts to declare.

References

- Carrington P, Hurst A, Kane S. Wearables and Chairables: Inclusive Design of Mobile Input and Output Techniques for Power Wheelchair Users. New York, NY, USA: Association for Computing Machinery; 2014 Apr Presented at: Proceedings of the 14th SIGCHI Conference on Human Factors in Computing Systems; 2014; Toronto, ON, Canada p. 3103-3112. [doi: 10.1145/2556288.2557237]
- 2. Renda G, Jackson S, Kuys B, Whitfield TWA. The cutlery effect: do designed products for people with disabilities stigmatise them? Disabil Rehabil Assist Technol 2016 Nov;11(8):661-667. [doi: 10.3109/17483107.2015.1042077] [Medline: 25936430]
- Heller S, White D, Lee E, Lawton J, Pollard D, Waugh N, et al. A cluster randomised trial, cost-effectiveness analysis and psychosocial evaluation of insulin pump therapy compared with multiple injections during flexible intensive insulin therapy for type 1 diabetes: the REPOSE Trial. Health Technol Assess 2017 Apr;21(20):1-278 [FREE Full text] [doi: 10.3310/hta21200] [Medline: 28440211]
- Wallace EL, Lea J, Chaudhary NS, Griffin R, Hammelman E, Cohen J, et al. Home Dialysis Utilization Among Racial and Ethnic Minorities in the United States at the National, Regional, and State Level. Perit Dial Int 2017;37(1):21-29. [doi: 10.3747/pdi.2016.00025] [Medline: 27680759]
- Shilkrot R, Huber J, Liu C, Maes P, Nanayakkara S. FingerReader: A Wearable Device to Support Text Reading on the Go. New York, NY, USA: Association of Computing Machinery; 2014 Apr Presented at: Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems (CHI EA '14); 2014; Toronto, ON, Canada. [doi: 10.1145/2559206.2581220]
- Hawley MS, Cunningham SP, Green PD, Enderby P, Palmer R, Sehgal S, et al. A voice-input voice-output communication aid for people with severe speech impairment. IEEE Trans Neural Syst Rehabil Eng 2013 Jan;21(1):23-31. [doi: 10.1109/TNSRE.2012.2209678] [Medline: 22875259]
- Markham SA, Levi BH, Green MJ, Schubart JR. Use of a Computer Program for Advance Care Planning with African American Participants. Journal of the National Medical Association 2015 Feb;107(1):26-32. [doi: 10.1016/S0027-9684(15)30006-7] [Medline: 27282525]
- Hightow-Weidman LB, Muessig KE, Pike EC, LeGrand S, Baltierra N, Rucker AJ, et al. HealthMpowerment.org: Building Community Through a Mobile-Optimized, Online Health Promotion Intervention. Health Educ Behav 2015 Aug;42(4):493-499 [FREE Full text] [doi: 10.1177/1090198114562043] [Medline: 25588932]
- 9. Sridhar A, Chen A, Forbes E, Glik D. Mobile application for information on reversible contraception: a randomized controlled trial. Am J Obstet Gynecol 2015 Jun;212(6):774.e1-774.e7. [doi: 10.1016/j.ajog.2015.01.011] [Medline: 25582097]
- Hartzler AL, Izard JP, Dalkin BL, Mikles SP, Gore JL. Design and feasibility of integrating personalized PRO dashboards into prostate cancer care. J Am Med Inform Assoc 2016 Jan;23(1):38-47 [FREE Full text] [doi: 10.1093/jamia/ocv101] [Medline: 26260247]
- Sanchez-Morillo D, Fernandez-Granero M, Jiménez AL. Detecting COPD exacerbations early using daily telemonitoring of symptoms and k-means clustering: a pilot study. Med Biol Eng Comput 2015 May;53(5):441-451. [doi: <u>10.1007/s11517-015-1252-4</u>] [Medline: <u>25725628</u>]

- Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
- 13. Vaisson G, Provencher T, Dugas M, Trottier ME, Chipenda Dansokho S, Colquhoun H, et al. User Involvement in the Design and Development of Patient Decision Aids and Other Personal Health Tools: A Systematic Review. Medical Decision Making (forthcoming) 2021. [doi: 10.1177/0272989X20984134]
- Witteman HO, Dansokho SC, Colquhoun H, Coulter A, Dugas M, Fagerlin A, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev 2015 Jan 26;4:11 [FREE Full text] [doi: 10.1186/2046-4053-4-11] [Medline: 25623074]
- 15. Gould JD, Lewis C. Designing for usability: key principles and what designers think. Communications of the ACM 1985;28(3):300-311. [doi: 10.1145/3166.3170]
- 16. Ergonomics of human-system interaction Part 210: Human-centred design for interactive systems. Switzerland: International Standardization Organization (ISO); Jul 2019:1-33.
- 17. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
- 18. Abras C, Maloney-Krichmar D, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications; 2004:445-456.
- 19. Garrett JJ. The elements of user experience, 2nd edition. Berkeley, CA: New Riders Publishing; 2011.
- 20. Tullis T, Albert W. Measuring the user experience. Waltham, MA: Morgan Kaufmann; 2010.
- 21. Goodman E, Kuniavsky M, Moed A. Observing the User Experience, 2nd edition. Waltham, MA: Morgan Kaufmann; 2012.
- 22. Iivari J, Iivari N. Varieties of user-centredness: An analysis of four systems development methods. Information Systems Journal 2011 Mar;21(2):125-153 [FREE Full text] [doi: 10.1111/j.1365-2575.2010.00351.x]
- Strauss AL, Fagerhaugh S, Suczek B, Wiener C. The work of hospitalized patients. Social Science & Medicine 1982 Jan;16(9):977-986. [doi: 10.1016/0277-9536(82)90366-5] [Medline: 7112176]
- 24. Corbin J, Strauss A. Managing chronic illness at home: Three lines of work. Qual Sociol 1985;8(3):224-247. [doi: 10.1007/bf00989485]
- Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. J Am Med Inform Assoc 2015 Jan;22(1):2-10 [FREE Full text] [doi: 10.1136/amiajnl-2014-002826] [Medline: 25125685]
- 26. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. The invisible work of personal health information management among people with multiple chronic conditions: qualitative interview study among patients and providers. J Med Internet Res 2015 Jun 04;17(6):e137 [FREE Full text] [doi: 10.2196/jmir.4381] [Medline: 26043709]
- 27. Lochner K, Kawachi I, Kennedy BP. Social capital: a guide to its measurement. Health Place 1999 Dec;5(4):259-270. [doi: 10.1016/s1353-8292(99)00016-7] [Medline: 10984580]
- 28. Biggs J. What do inventories of students' learning processes really measure? A theoretical review and clarification. Br J Educ Psychol 1993 Feb;63 (Pt 1):3-19. [doi: 10.1111/j.2044-8279.1993.tb01038.x] [Medline: 8466833]
- 29. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec 9;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
- 30. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. Int J Qual Health Care 2001 Dec;13(6):469-474. [doi: <u>10.1093/intqhc/13.6.469</u>] [Medline: <u>11769749</u>]
- 31. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational Psychological Testing. Standards for Educational and Psychological Testing. Washington, DC: American Educational Research Association; 2014.
- 32. Guilford J. Psychometric Methods, 1st ed. New York and London: McGraw-Hill Book Company, Inc; 1936.
- 33. Gulliksen H. Theory of mental tests. New York: Wiley; 1950.
- 34. Lord F, Novick M, Birnbaum A. Statistical theories of mental test scores. Reading, MA: Addison-Wesley; 1968.
- 35. Magnusson D. Test Theory. Reading, MA: Addison-Wesley Pub Co; 1967.
- 36. Spearman C. The Proof and Measurement of Association between Two Things. Am J Psychol 1904 Jan;15(1):72-101. [doi: 10.2307/1412159]
- Spearman C. Demonstration of Formulae for True Measurement of Correlation. Am J Psychol 1907 Apr;18(2):161-169. [doi: <u>10.2307/1412408</u>]
- 38. Sébille V, Hardouin J, Le Néel T, Kubis G, Boyer F, Guillemin F, et al. Methodological issues regarding power of classical test theory (CTT) and item response theory (IRT)-based approaches for the comparison of patient-reported outcomes in two groups of patients--a simulation study. BMC Med Res Methodol 2010 Mar 25;10:24 [FREE Full text] [doi: 10.1186/1471-2288-10-24] [Medline: 20338031]

- 39. Cappelleri JC, Jason Lundy J, Hays RD. Overview of classical test theory and item response theory for the quantitative assessment of items in developing patient-reported outcomes measures. Clin Ther 2014 May;36(5):648-662 [FREE Full text] [doi: 10.1016/j.clinthera.2014.04.006] [Medline: 24811753]
- 40. León-Del-Barco B, Mendo-Lázaro S, Felipe-Castaño E, Fajardo-Bullón F, Iglesias-Gallego D. Measuring Responsibility and Cooperation in Learning Teams in the University Setting: Validation of a Questionnaire. Front Psychol 2018;9:326 [FREE Full text] [doi: 10.3389/fpsyg.2018.00326] [Medline: 29593622]
- 41. De Champlain AF. A primer on classical test theory and item response theory for assessments in medical education. Med Educ 2010;44(1):109-117. [doi: 10.1111/j.1365-2923.2009.03425.x]
- 42. Downing SM. Validity: on meaningful interpretation of assessment data. Med Educ 2003 Sep;37(9):830-837. [doi: 10.1046/j.1365-2923.2003.01594.x] [Medline: 14506816]
- 43. Mao J, Vredenburg K, Smith PW, Carey T. The state of user-centered design practice. Commun ACM 2005 Mar;48(3):105-109. [doi: 10.1145/1047671.1047677]
- 44. Nielsen J. The usability engineering life cycle. Computer 1992 Mar;25(3):12-22. [doi: 10.1109/2.121503]
- 45. Norman D. The Design of Everyday Things. New York: Basic Books; 2002.
- 46. Nunnally J, Bernstein L. Psychometric theory. 3rd ed. New York: McGraw Hill; 1994.
- 47. Schmeiser CB, Welch CJ. Test development. Educational measurement 2006;4:307-353.
- 48. Everitt BS, Skrondal A. The Cambridge Dictionary of Statistics. Cambridge: Cambridge University Press; 2010.
- 49. Kaiser HF. An index of factorial simplicity. Psychometrika 1974 Mar;39(1):31-36. [doi: 10.1007/bf02291575]
- 50. Lang TA, Secic M. How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers. Philadelphia: American College of Physicians; 2006:20.
- 51. Data for Development and Validation of UCD-11: An 11-item Measure of User- and Human-Centered Design for Personal Health Tools. Scholars Portal Dataverse. URL: <u>https://doi.org/10.5683/SP2/LJLUWQ</u> [accessed 2021-02-22]
- 52. DeVellis R. Scale Development: Theory and Applications. 3rd ed. Thousand Oaks, CA: Sage Publications; 2011.
- 53. Dugas M, Trottier ME, Chipenda Dansokho S, Vaisson G, Provencher T, Colquhoun H, et al. Involving members of vulnerable populations in the development of patient decision aids: a mixed methods sequential explanatory study. BMC Med Inform Decis Mak 2017 Jan 19;17(1):12 [FREE Full text] [doi: 10.1186/s12911-016-0399-8] [Medline: 28103862]
- 54. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. Methods of Psychological Research Online 2003;8(8):23-74.
- 55. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
- 56. Bangor A, Kortum P, Miller J. Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. Journal of Usability Studies 2009 May;4(3):114-123.
- 57. Barki H, Hartwick J. Measuring User Participation, User Involvement, and User Attitude. MIS Quarterly 1994 Mar;18(1):59-82. [doi: <u>10.2307/249610</u>]
- 58. Kujala S. Effective user involvement in product development by improving the analysis of user needs. Behaviour & Information Technology 2008 Nov;27(6):457-473. [doi: 10.1080/01449290601111051]
- 59. Subramanyam R, Weisstein F, Krishnan MS. User participation in software development projects. Commun. ACM 2010 Mar;53(3):137-141 [FREE Full text] [doi: 10.1145/1666420.1666455]
- 60. Kealey MR. Impact of Design Expertise and Methodologies on the Usability of Printed Education Materials Internet. TSpace. Toronto, ON: University of Toronto; 2015. URL: <u>https://tspace.library.utoronto.ca/handle/1807/70839</u> [accessed 2020-07-31]
- 61. Higgins KS, Tutelman PR, Chambers CT, Witteman HO, Barwick M, Corkum P, et al. Availability of researcher-led eHealth tools for pain assessment and management: barriers, facilitators, costs, and design. PR9 2018 Sep;3(1):e686. [doi: 10.1097/pr9.0000000000686]
- 62. Kilsdonk E, Peute LW, Riezebos RJ, Kremer LC, Jaspers MWM. From an expert-driven paper guideline to a user-centred decision support system: a usability comparison study. Artif Intell Med 2013 Sep;59(1):5-13. [doi: 10.1016/j.artmed.2013.04.004] [Medline: 23684240]
- 63. Wilkinson CR, De Angeli A. Applying user centred and participatory design approaches to commercial product development. Design Studies 2014 Nov;35(6):614-631. [doi: <u>10.1016/j.destud.2014.06.001</u>]
- 64. Barry MJ, Edgman-Levitan S. Shared Decision Making The Pinnacle of Patient-Centered Care. N Engl J Med 2012 Mar;366(9):780-781. [doi: 10.1056/nejmp1109283]
- 65. Committee on Quality of Health Care in America. Crossing the Quality Chasm: A New Health System for the 21st Century. Journal For Healthcare Quality 2002;24(5):52. [doi: <u>10.1111/j.1945-1474.2002.tb00463.x</u>]
- 66. Légaré F, Witteman HO. Shared decision making: examining key elements and barriers to adoption into routine clinical practice. Health Aff (Millwood) 2013 Feb;32(2):276-284. [doi: 10.1377/hlthaff.2012.1078] [Medline: 23381520]

Abbreviations

mHealth: mobile health



UCD-11: User-Centered Design 11-item scale

Edited by G Eysenbach; submitted 13.06.19; peer-reviewed by I Holeman, L van Velsen, D Sakaguchi-Tang; comments to author 03.10.19; revised version received 27.08.20; accepted 03.10.20; published 16.03.21

Please cite as:

Witteman HO, Vaisson G, Provencher T, Chipenda Dansokho S, Colquhoun H, Dugas M, Fagerlin A, Giguere AMC, Haslett L, Hoffman A, Ivers NM, Légaré F, Trottier ME, Stacey D, Volk RJ, Renaud JS
An 11-Item Measure of User- and Human-Centered Design for Personal Health Tools (UCD-11): Development and Validation J Med Internet Res 2021;23(3):e15032
URL: https://www.jmir.org/2021/3/e15032
doi: 10.2196/15032
PMID: 33724194

©Holly O Witteman, Gratianne Vaisson, Thierry Provencher, Selma Chipenda Dansokho, Heather Colquhoun, Michele Dugas, Angela Fagerlin, Anik MC Giguere, Lynne Haslett, Aubri Hoffman, Noah M Ivers, France Légaré, Marie-Eve Trottier, Dawn Stacey, Robert J Volk, Jean-Sébastien Renaud. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 16.03.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

