IOWA STATE UNIVERSITY
Digital Repository

2016

The WEAR Scale: Developing a Measure of the Social Acceptability of a Wearable Device

Norene Kelly
Iowa State University

Stephen B. Gilbert
Iowa State University, gilbert@iastate.edu

Follow this and additional works at: http://lib.dr.iastate.edu/imse_conf

Part of the Applied Mechanics Commons, Ergonomics Commons, and the Industrial and Product Design Commons

Recommended Citation

The WEAR Scale: Developing a Measure of the Social Acceptability of a Wearable Device

Norene Kelly

Iowa State University

Ames, IA, 50011 USA

nbkelly@iastate.edu

Stephen Gilbert

Iowa State University

Ames, IA, 50011 USA

gilbert@iastate.edu

Abstract

The factors affecting the social acceptability of wearable devices are not well understood, yet they have a strong influence on whether a new wearable succeeds or fails. Factors uniquely affecting wearable acceptability as compared to other technology include manners, moral codes, the symbolic communication of dress, habits of

dress, fashion, context of use, form, and aesthetics. This paper describes the development of the WEarable Acceptability Range (WEAR Scale), designed to predict acceptance of a particular wearable. First, the construct "social acceptability of a wearable" was defined using literature and an interview study. Second, the WEAR Scale's item pool was composed, and reviewed by experts. Third, the resulting scale will be administered to sample respondents along with validation measures. The data will be evaluated for reliability and validity, and the scale's length will be adjusted, culminating in a validated WEAR Scale useful to both industry and academia.

Author Keywords
Wearable device; social acceptance; scale development; scale validation.

ACM Classification Keywords
H.1.2. Human/machine Systems: human factors; J.4. Social and behavioral sciences: Sociology; K.8.m Personal Computing: Miscellaneous.

Introduction
A new world of wearable devices (or "wearables") is on the horizon, but it will only arrive if people consent to wearing them [1]. Prior to wide adoption, a wearable must first find social acceptance. Such social acceptance is a crucial factor for wearables (as compared to other technologies) because they are worn on the body, typically in public view. The problem in developing commercial applications for wearables is that these technologies intersect with manners, moral codes, the symbolic communication of dress, habits of dress, fashion, context of use, form, and aesthetics. Social issues in wearable development cannot be

ignored [2]. Dunne et al. [3] termed this increasingly important issue "social wearability," and noted that this aspect of wearability has been less explored than the physical and psychological aspects.

Thus, a predictive instrument for measuring wearable social acceptability is needed to understand the factors affecting not only whether people will consider wearing a device themselves, but also whether they will consider it acceptable for others' use. The overall objective of this ongoing line of research is to develop the WEarable Acceptability Range (WEAR Scale). Because there is no existing theory, starting with data (literature and interviews) using grounded theory, rather than an a priori theory, is a suitable methodological choice. The first sidebar shows the steps implemented in the development of the WEAR Scale, which are typical for the development of a validated survey instrument [4]. In the case of WEAR, first the literature and interview results were used to define the construct and generate scale items (Study 1), thus supporting an argument for content validity [4], [5]. Experts in Study 2 then reviewed this initial item pool (v.1), resulting in revision or deletion of problematic items, leading to v.2. Related measures were then selected to assess construct validity. Following a pilot review, WEAR v.2.1 and the validation items will be administered to a sample of respondents in Study 3. Finally, the results will be evaluated using exploratory factor analysis, the scale length will be adjusted, and the construct validity will be evaluated, ideally leading to the final validated WEAR Scale.

For wearable devices to reach their potential—for users to adopt a particular wearable—people must first deem the device acceptable for themselves and others to

Scale Development Methodology per DeVellis [4]

1. Determine what is being measured (**Study 1**).

2. Compose item pool (v.1).

3. Determine scale format.

4. Expert review of initial item pool to create v.2 (**Study 2**).

5. Determine comparison items or scales for testing construct validity.

6. Administer items v.2.1 to sample of respondents (**Study 3**).*

wear. Acceptability research is the investigation of the perceived attributes of an ideal innovation, and is used to guide research and development so as to create such an innovation [6]. The WEAR Scale is intended to aid developers in directing their efforts, as well as help social scientists test hypotheses and build theories pertaining to the construct "social acceptability of a wearable."

Literature Review

The main research question of this work—*What are the factors affecting social acceptability of a wearable device?*—is sparsely addressed in the current literature. There are existing measures pertaining to the acceptance and adoption of technology in general, e.g., Technology Acceptance Model (TAM) [7], [8] and Unified Theory of Acceptance and Use of Technology (UTAUT) [9]. However, placing technology on one's body is a significantly different matter than using it as a stand-alone device. This is because a wearable is more like a piece of clothing than a PC or an appliance, and clothing has been shown to help define identity and supply clues to categorize oneself and others in the culture [11]. People embody not just the clothes they wear but also the clothing's symbolic meaning [11], and presumably the same is true for wearables. But many of the existing measures of technology acceptance address only particular contexts and functions, such as the use of software in a work situation, which deviate quite critically from wearing a device for personal use. While one would expect a positive correlation between measures such as affinity for technology [12] and the acceptability of a wearable, the two are likely separate constructs.

How individuals perceive each other and themselves is affected by society's increasing use of and reliance on technology [13]. Even though we are a technology-driven society, persons wearing technology may be perceived as less human-like, and there has been and continues to be a negative stigma attached to the excessive use of technology [13], [14]. The technology industry's lack of understanding of the rapidly evolving landscape of new technology and shifting social norms leads to (at the least) bad press and (at the worst) investigations and lawsuits [15]. Navigating public expectations and regulatory requirements demands a better understanding of the factors affecting the social acceptability of wearable devices.

Study 1 and Related Development

The first step in scale development is to determine what is being measured, which was accomplished via a qualitative analysis of the literature and an interview study of the intended population (Study 1). Item generation then proceeded directly from the literature and study findings, thus establishing an argument for content validity [4].

For the WEAR Scale, the target population was defined as persons aged 18 to 30 because technology is typically developed by younger people for the use of younger people and marketed at younger target groups [16]. Based on this previous result, and Spector's [17] finding that the college population can represent the general population well, depending on the scale, the authors sampled from a population of college students for Study 1. Johnson et al. [18], in their comprehensive review of published research on dress and human behavior, found that college students were used in 19.1 percent of the studies.

Definition of Social Acceptability of a Wearable Device (result from Study 1)

A wearable, for the purposes of developing the WEAR Scale, is a small computational device or accessory that is worn on the body in public. A wearable is personal, and personally-owned, and is at least minimally visible and comfortable. Donning a wearable requires action, which is preceded by decision-making about the social acceptability of the action and the anticipated reaction of others.

A socially acceptable wearable is most notably marked by an absence of negative reactions or judgments from others.

Nine interviews were conducted with the target population at a Midwestern university to gain information on this population's conceptualization of the construct "social acceptability of a wearable" and the language they use to talk about it. The objective of such interviewing in scale development is to use the resulting key phrases and ideas gathered from the target population in defining the construct and in writing the initial item pool [4], [19].

An interview protocol was formulated based on a comprehensive review of the literature (models and theories of technology acceptance and adoption, and literature and theories pertaining to dress, the body/self, and fashion). Each participant was interviewed about the term "wearable" and the concept of "social acceptability." After each interview, author Kelly analyzed her notes and entered key words and phrases in a spreadsheet, to collect frequencies on responses that could be quantified and to distill comments into common categories. Combining the interview data with definitions found in the literature ([3], [13], [20], [21], [22], [23]) resulted in *Definition of Social Acceptability of a Wearable Device* (see sidebar on previous page).

A number of questions in the interview study resulted in concepts that were then categorized. These concepts and categories were a useful starting point in organizing the writing of scale items. The question *What makes a wearable socially acceptable or unacceptable?* most directly addresses the construct and was therefore the starting point for item generation. Using this as an example, four categories were identified in the responses to this question: available/ordinary, consequences, aesthetics, and

functionality. For further example, in response to this question, interviewees stated that a socially acceptable device is accessible, affordable, and not in limited release. These data resulted in the item: *This device seems to be accessible, that is, affordable and not in limited release*. This process was continued to compose items for the initial pool (v.1), with responses being a 1 (strongly agree) to 6 (strongly disagree) Likert scale.

Analysis of the interview data resulted in 68 items, in ten categories: aesthetics, available/ordinary, consequences, ergonomics, functionality, judgment, norms, others' reactions, others' thoughts, and qualities of the device or wearer. Then the concepts from the literature were examined as to whether they were represented in the 68 items generated from the interview data. If they were not represented, a new item number was written for the concept. The additional items composed from the literature resulted in a total of 97 items. A new category emerged from concepts found in the literature: self-identity. This is a concept that did not present in the interview data, but was added as a category due to its importance in the literature.

Next, items in each category from the literature and interviews were compared for redundancies. The 97 items were culled based on redundancy, resulting in 73 items in WEAR v.1 for expert review.

Study 2 and Related Development

The next stage of scale development is to have a few people who are knowledgeable about the content area review the item pool [4]. About 20 potential experts in the field of wearables were emailed an invitation to the

---

WEAR Scale v.2.1 (Likert)

1. I think my peers would find this device acceptable to wear.
2. The size of this device is conveniently small.
3. This device is like the clothing and accessories typically worn in our society.
4. This device looks natural and not out of place on the body.
5. This device could make people uncomfortable. (R)
6. This device is consistent with my self image.
7. This device seems to be useful and easy to use.
8. I like how this device shows membership to a certain social group.
9. I like what this device communicates about its wearer.
10. A wearer of this device would be keeping to the social norms we need to stick to.
11. This device might be considered disfiguring to its wearer. (R)
12. This device seems creepy. (R)
13. This device seems like "too much" technology. (R)
14. This device might restrict movement or physically get in the way. (R)
15. This device is fashionable.
16. This device seems comfortable, not bulky.

study, of which three participated-- a male from academia and two females from industry. It was explained to them that there were three main purposes in having knowledgeable people review the initial pool of items (v.1) of the WEAR Scale: 1) to obtain their ratings of how relevant they think each item is to the construct the scale is meant to measure; 2) to obtain their evaluations of each item's clarity and conciseness; and 3) to allow them to provide feedback on how the item pool thus far may fail to reflect the construct (social acceptability of a wearable) under measurement [4]. The survey thus consisted of the 73 items, each followed by three choices: very relevant to social acceptability of a wearable (scored 1.00), somewhat relevant to social acceptability of a wearable (scored 2.00), and not relevant to social acceptability of a wearable (scored 3.00), as well as a comments box.

Using the experts' ratings, the following items were removed: the one item that received a mean score of 3.00, the one item that received a mean score of 2.67, and the 19 items that received a mean score of 2.33. This pared down the item pool to 52 items. In considering the comments, three additional items were removed. Finally, the authors agreed with the comment that an additional item ("The way this device displays membership to a certain social group is appealing") would make sense as a counter the item to "The way this device displays membership to a certain social group is unappealing." Therefore, the WEAR Scale following expert review consisted of 50 items (v.2).

At this point in the scale development process, the researcher determines which items or scales to use for purpose of testing construct validity [4]. The construct validity of the WEAR Scale may be evidenced by

demonstrating its correlation with related measures. The resulting patterns of relationships will either provide support for validity or provide clues for revisions if the relationships are not as expected [4]. Related constructs are tested as a method of instrument refinement, i.e., improving an instrument's representation of a construct [24]. After examining numerous scales and items with potential for testing construct validity, the Affinity for Technology Scale [12] was selected, as well as items focused on self-reported optimism, age, and likeableness of a person wearing a certain device (see Figure 1).

Study 3 and Related Development

Following expert review, there were 50 items (WEAR Scale v.2) to administer to a sample of respondents. This section describes how these items were subjected to pilot testing and revision, which will be followed by

Figure 1: These photos of the LG Tone+ HBS-730 Wireless Bluetooth Stereo Headset Neckband will be used to test WEAR Scale (v.2.1) in Study 3 [25], [26].

administration to the target population and item evaluation, resulting in WEAR Scale v.3. The resulting data will also be examined to assess its expected relation to the scales and items identified above, to evaluate construct validity.

WEAR Scale v.2.1 (cont'd)

20. This device is cool.
21. This device could allow its wearer to take advantage of people. (R)
22. I can imagine that people would be interested in this device and would not have a problem wearing it.
23. This device is sleek, not clunky.
24. This device seems to offer options for personalization, so that everyone is not wearing the "same thing."
25. This device could help people.
26. This device would be generally accepted by the vast majority of people.
27. This device is not weird.
28. The wearer of this device would not be judged negatively by others.
29. This device is similar to existing acceptable devices or

Two members of the target population were asked to complete the survey as if they were taking it as participants. These pilot reviewers made note of any items that they thought: were not clear; were confusing because they contained multiple ideas; were overly long or wordy; or contained hard to understand words. Five items were revised according to the pilot reviewers' feedback. The final 50 items in the WEAR Scale (v.2.1) are shown in three continuous sidebars.

For Study 3, a sample of 300 participants, aged 18 to 30, will be sought to respond to these 50 items as well as the items/scales for validation. In administering the items in Study 3, a wearable and a description must be presented so that the participants may respond to the scale questions in relation to a particular device. The wearable that was chosen was the LG Tone+ HBS-730 Wireless Bluetooth Stereo Headset Neckband (see Figure 1), because it is an existing product that participants may own themselves, or may have seen on other people, but it is not widely worn and it does not resemble an existing accepted accessory like a wristwatch. Therefore, it should evoke useful variability in participants' responses.

In assessing the performance of the WEAR Scale (v.2), first the correlation matrix will be used to examine correlations between items, to remove items that do not contribute to an internally consistent scale [4], [17]. Second, item variance will be examined; in general, a relatively high variance is desirable [4]. Third, item means will be examined, and in general a mean close to the center of possible scores is advantageous [4]. A low variance or lopsided mean tends to lower an item's correlation, and thus first attention will be given to items that have a high correlation with other items, in terms of retaining the

item. Once the most valuable items are identified via the correlation matrix, variances and means will provide additional guidance [4]. Then, factor analysis will be used to determine which group of items (if any) constitute a unidimensional set. Finally, after discarding the poorer items and retaining the better items per the above methods, coefficient alpha (or Cronbach's alpha) will be used to evaluate how well the weeding out process worked. At this point, it is anticipated that a pool of items that show acceptable reliability (alpha) will exist, and attention will be turned toward balancing brevity and reliability to produce WEAR Scale v.3.

The scale produced from factor analysis and other analyses described above will then be used to test the relationships that were hypothesized, namely: a positive relationship between affinity for technology and acceptability of wearables, a positive relationship between optimism and acceptability of wearables, a negative relationship between age and acceptability of wearables, and a positive relationship between likeableness rating and acceptability of wearables.

Current Limitations and Future Development
A limitation that will exist with WEAR Scale v.3 will be that it was tested with regard to one particular wearable. Although the scale's constructs will have been validated using this one wearable, additional studies using different wearables would be useful in refining the WEAR Scale. Another limitation is that the participants in both Study 1 and Study 3 were recruited from the campus of a Midwestern university. While the intended target population for the WEAR Scale is younger adults in the United States, samples drawn from diverse regions may have produced different results based on people's varying attitudes about fashion in different regions.

WEAR Scale v.2.1 (cont'd)

36. Use of this device raises privacy issues. (R)
37. I could imagine aspiring to be like the wearer of such a device.
38. There is no chance of being ridiculed when wearing this device.
39. Use of this device would be more threatening than exciting. (R)
40. This device is stylish.
41. This device's placement on the body could cause awkwardness or embarrassment. (R)
42. Wearing this device would cause no reaction, or a neutral reaction, from other people.
43. I don't like how this device shows membership to a certain social group. (R)
44. This device is goofy. (R)
45. This device would enhance the wearer's image.

This paper described the ongoing process of developing a scale to measure the social acceptability of a wearable device. While this process is lengthy and validation is an ongoing process [17, [19], it is anticipated that the resulting scale will be highly valuable to both industry (in researching and developing prototypes) and academia (in building models and theories pertaining to wearable technologies). Given the context-dependent nature of the construct *social acceptability*, ecological validity is a concern, and the WEAR Scale should be understood as a tool that is to be used in conjunction with other tools or methods, e.g., Suchman's situated action [27].

References

1. Bill Wasik. Jan. 2014. Try it on. *Wired*: 90-99.

2. C. Edwards. 2003. Wearable computing struggles for social acceptance. *IEE Review*, 49, 9: 24-25.

3. Lucy E. Dunne, Halley Profita, Clint Zeagler, James Clawson, Scott Gilliland, Ellen Do, and Jim Budd. 2014, August. The social comfort of wearable technology and gestural interaction. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*: 4159-4162. IEEE.

4. Robert F. DeVellis. 2012. *Scale development: Theory and applications*. Sage Publications.

5. Katherine Regan Sterba, Robert F. DeVellis, Megan A. Lewis, Donald H. Baucom, Joanne M. Jordan, and Brenda DeVellis. 2007. Developing and testing a measure of dyadic efficacy for married women with rheumatoid arthritis and their spouses. *Arthritis Care & Research*, 57, 2: 294-302.

6. Everett M. Rogers. 2003. *Diffusion of innovations.* Free Press.

7. Fred D. Davis. 1986. A technology acceptance model for empirically testing new end-user information systems: Theory and results. Doctoral dissertation, Massachusetts Institute of Technology.

8. Fred D. Davis. 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*: 319-340.

9. Viswanath Venkatesh, Michael G. Morris, Gordon B. Davis, and Fred D. Davis. 2003. User acceptance of information technology: Toward a unified view. *MIS Quarterly*: 425-478.

10. Patricia A. Cunningham and Susan Voso Lab. 1991. Understanding dress and popular culture. In *Dress and popular culture*, Patricia A. Cunningham and Susan Voso Lab (eds.). Bowling Green State University: 5-18.

11. Hajo Adam and Adam D. Galinsky. 2012. Enclothed cognition. *Journal of Experimental Social Psychology*, 48, 4: 918-925.

12. Steve W. Edison and Gary L. Geissler. 2003. Measuring attitudes towards general technology: Antecedents, hypotheses and scale development. *Journal of Targeting, Measurement and Analysis for Marketing*, 12, 2: 137-156.

13. Heather C. Lum, Valerie K. Sims, Matthew G. Chin, Mand Nicholas C Lagattuta. Oct. 2009. Perceptions of humans wearing technology. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53, 13: 864-868. Sage Publications.

14. V.R. Manoj and Jayapaul Azariah. 2001. Cybofree-cyborgs, fantasy, reality, ethics and education. *Eubios Journal of Asian and International Bioethics*, 11: 178-184.

15. Omer Tene and Jules Polonetsky. 2013. A theory of creepy: Technology, privacy and shifting social norms. *Yale Journal of Law & Technology*, 16: 59-102.

16. Susan B. Van Hemel and Richard W. Pew (Eds.). (2004). *Technology for adaptive aging*. National Academies Press.

17. Paul E. Spector. 1992. Summated rating scale construction: An introduction. Sage Publications.

18. Kim K.P. Johnson, Jeong-Ju Yoo, Minjeong Kim, and Sharron J. Lennon. 2008. Dress and human behavior: A review and critique. *Clothing and Textiles Research Journal*, *26*(1), 3-22.

19. Lee Anna Clark and David Watson. 1995. Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 3: 309.

20. Erving Goffman. 1990. The presentation of the self in everyday life. Penguin.

21. Steve Mann. 2014. Wearable computing. In *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Mads Soegaard and Rikke Friis Dam (eds.) The Interaction Design Foundation. Retrieved August 16, 2014 from https://www.interaction-design.org/encyclopedia/wearable_computing.html

22. Julie Rico and Stephen Brewster. 2010. Usable gestures for mobile interfaces: Evaluating social acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 887-896. ACM.

23. Wearable Devices Magazine. 2014. *Introduction to wearable technology*. Retrieved August 16, 2014 from http://www.wearabledevices.com/what-is-a-wearable-device/

24. Gregory T. Smith and Denis M. McCarthy. 1995. Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 300-308.

25. *LG Tone+ HBS-730*. 2015. Retrieved October 18, 2015 from http://vapingunderground.com/threads/lg-tone-hbs-730-wireless-bluetooth-stereo-headset-neckband-style-hbs730.73983/

26. *LG Tone Plus Headphone Review.* Retrieved October 18, 2015 https://www.youtube.com/watch?v=ucFnZS8sOww

27. Suchman, L. (1987). Plans and Situated Actions: The Problem of Human Machine Communication. Cambridge University Press.