sMUX: A Short-form Mobile User Experience Instrument

Journal:	*23rd Americas Conference on Information Systems*
Manuscript ID	AMCIS-0034-2017
Track:	12. Human-Computer Interaction (SIGHCI)
Mini Track:	SIGHCI: Interface Design, Evaluation and Impact < 12. Human-Computer Interaction (SIGHCI)
Abstract:	Users' subjective evaluation of digital products and services is a major focus in human-computer interaction (HCI) research. Despite the increasing availability of services that can be accessed via mobile computing, little work has been done to capture specific dimensions of experience that are unique to this domain. The Mobile User Experience (MUX) instrument was developed to evaluate joint usability of mobile software and devices, which the developers refer to as holistic assessment of the mobile user experience. It features three scales (Nuisance, Access, and Mobility) that were derived through a multi-stage instrument development and validation process. In this paper, we develop a short-form version of MUX (sMUX) that is especially suitable for use in practical settings. In addition to evaluating convergent and discriminant validity of instrument constructs, we further assess known groups validity and measurement sensitivity of the instrument by comparing two learning management systems (LMS) software applications across three distinct computing device form factors: smartphones/tablets, laptop computers, and desktop computers. As a benchmark, we contrast these results from sMUX vs. the well-known System Usability Scale (SUS).

sMUX: A Short-form Mobile User Experience Instrument

Completed Research Full Paper

Introduction

User experience (UX) is an important driver of market success for mobile devices as well as software applications designed to run on them. Thus, conducting evaluations to promote positive user experiences is essential to successful product development in the mobile computing domain. Several survey instruments are available to researchers and practitioners who conduct evaluations of mobile user experience, however, each of these instruments has important limitations.

One prominent example is the well-known System Usability Scale (SUS), which has been used to assess user experience with software and other products over the past three decades (Bangor et al., 2009; Brooke, 1996). Practitioners like the 10-item SUS because it is speedy to administer, it produces a score that is easily-interpreted, and a large inventory of prior scores is available to use as a benchmark for analyzing new results. Despite these qualities, SUS is a generic, single-score instrument that does not address key issues of mobile computing, such as compromises in size and capabilities that frequently are required in pursuing the objective of any-place, any-time device use.

A second example is the recently-developed Mobile Application Usability (MAU) instrument (Hoehle & Venkatesh, 2015). MAU appears to be an important advance for scholarly research, yet it is not an ideal solution for UX practitioners. First, given its approach of measuring six second-order constructs, MAU is neither speedy to administer nor easy for practitioners to interpret. Second, although it captures important usability aspects of mobile applications, it does not address the device on which the software is running. This approach ignores device-related factors that can strongly influence user experience, including speed of use, convenience, connectivity, mobility, portability, and differences in device feature sets and user interaction models.

In recognition of these limitations in existing instruments, Djamasbi and Wilson (2017) undertook to develop a set of scales intended to measure joint usability of mobile software and mobile device, which they refer to as *holistic* assessment of the mobile user experience. The resulting Mobile User Experience (MUX) instrument comprises three scales that were derived through a multi-stage instrument development and validation process. The scales are:

- *Nuisance,* encompassing perceptions that the device/software combination slows use, is inconvenient, and evokes feelings of isolation and disconnectedness;
- *Access,* encompassing perceptions that the device/software combination provides easy viewability of screen images, entry of text, and access to links and buttons; and
- *Mobility,* encompassing perceptions that the device/software combination promotes personal mobility and is easily portable.

The developers of MUX originally intended it to augment the SUS instrument with factors focused toward mobile computing. However, initial testing and validation results suggested that MUX could serve as an effective replacement for SUS.

In this paper, we document our development of a short-form version of MUX. Although lengthy surveys are common in academic settings, it can be difficult to convince working professionals to respond to a questionnaire that takes more than a few minutes to complete (Hayes, 2005), and response rates typically drop as length increases both in printed questionnaires (Heberlein & Baumgartner, 1978) and web surveys (Galesic & Bosnjak, 2009).

Development was guided by multiple objectives of 1) meeting survey length constraints of the UX practitioner community, 2) maintaining sufficient validity and robustness to meet standards of scholarly research, and 3) demonstrating that a shortened MUX can provide new insights into user experience with mobile computing beyond what is available using other instruments. To meet the first objective, we developed and evaluated a short-form version of MUX. To meet the second objective, we conducted new validation testing on the shortened scales. To meet the third objective we examined the capabilities of this reduced instrument in distinguishing differences between two software applications across three computing device form factors, thereby providing and evaluation of known groups validity and measurement sensitivity in the reduced instrument.

Research Method and Results

MUX was developed as a 15-item instrument comprising three five-item scales (Djamasbi & Wilson, 2017). Although this is not lengthy by academic standards, practitioners prefer shorter instruments. For example, SUS has been successful among practitioners in large part because its short length (10 items) allows for quick administration. One objective in creating a short-form version of MUX was to reduce the total length of this instrument to 10 items or fewer without excessive loss of discriminant or predictive abilities. (For brevity, the short-form MUX is referred to as *sMUX* and the full 15-item instrument as *MUX* through the remainder of this paper.)

Each MUX scale is comprised of measurement items which are considered to “reflect” the latent factor. Such reflective measurement items may be interchanged or removed without conceptually altering the overall scale (Mackenzie et al., 2005; Podsakoff et al., 2006), assuming that sufficient items remain for effective analysis, preferably at least three items per factor (Bagozzi, 2011). Accordingly, sMUX was formed by drawing three items each to represent Nuisance, Access, and Mobility scales. Selected items were those that produced the highest loadings during confirmatory factor analysis (CFA) validation testing by Djamasbi and Wilson (2017).

Administration Process

An online survey containing the MUX items, SUS items, and demographic items was administered to 214 students at two universities, 106 at a private Eastern university and 108 at a public Midwest university. Students were recruited from undergraduate and graduate business courses and offered extra course credit for completing the online survey or a comparable alternative assignment.

Item / Factor (% of total variance explained by this factor)	1 (25%)	2 (24%)	3 (12%)
Nuisance-A: I felt using a [device] to access [software] would slow me down.	**0.810**	-0.243	-0.037
Nuisance-B: Using [device] to access [software] was inconvenient.	**0.736**	-0.163	-0.018
Nuisance-C: Using a [device] to access [software] made me feel disconnected.	**0.694**	-0.122	-0.009
Access-A: I had no trouble viewing text when using a [device] to access [software].	-0.115	**0.649**	0.140
Access-B: Clicking on links or buttons was easy to accomplish using a [device] to access [software].	-0.140	**0.787**	0.093
Access-C: I have no problem entering text when using a [device] to access [software].	-0.049	**0.680**	-0.025
Mobility-A: Using a [device] to access [software] would improve my ability to be mobile.	0.140	0.092	**0.869**
Mobility-B: I would be able to use a [device] to access [software] on the go.	0.019	0.196	**0.805**
Mobility-C: I feel a [device] used to access [software] would be very portable.	0.063	0.178	**0.863**
SUS01: I think that I would like to use a [device] to access [software] frequently.	-0.174	**0.741**	0.214
SUS02: I found using a [device] to access [software] unnecessarily complex.	**0.724**	-0.228	0.083
SUS03: I thought a [device] was easy to use to access [software].	-0.164	**0.809**	0.030
SUS04: I think that I would need the support of a technical person to be able to use a [device] to access [software].	**0.775**	-0.133	0.089
SUS05: I found the various functions in using a [device] to access [software] were well integrated.	-0.158	**0.670**	0.215
SUS06: I thought there was too much inconsistency in using a [device] to access [software].	**0.774**	-0.269	0.003
SUS07: I would imagine that most people would learn to use a [device] to access [software] very quickly.	-0.150	**0.750**	0.058
SUS08: I found using a [device] to access [software] very cumbersome to use.	**0.701**	0.028	0.026
SUS09: I felt very confident using a [device] to access [software].	-0.310	**0.761**	0.088
SUS10: I needed to learn a lot of things before I could get going with using a [device] to access [software].	**0.747**	-0.132	0.146
* Extraction was conducted via Principal Components Analysis limited to factors above 1.0 eigenvalue; rotation was conducted via Varimax with Kaiser Normalization; KMO measure = 0.890; bolding indicates loading values above 0.5			

Table 1. EFA* of sMUX Scales and SUS Items

University	Age	Sex	Population n: Device Breakout
Midwest U.	25.4 yrs. (10.1 s.d.)	61 males, 47 females	108: 14 phone/tablet, 81 laptop, 13 desktop accessing Desire2Learn LMS software
Eastern U.	25.3 yrs. (3.7 s.d.)	55 males, 50 females, 1 unreported	106: 11 phone/tablet, 85 laptop, 10 desktop accessing Blackboard LMS software

Table 2. Demographic and Device Breakout Data

On entry to the survey, students were asked to identify the computing device (smartphone/tablet, laptop computer, or desktop computer) they had most recently used to access their university's learning management system (LMS). They were then presented with MUX and SUS items asking their perceptions regarding use of the identified computing device to access the LMS software. These aspects are denoted respectively by the terms "[device]" and "[software]" in Table 1, which shows results of exploratory factor analysis (EFA) across the combined subject population[1]. Item-ordering of MUX and SUS items was individually randomized for administration to each subject as recommended by Wilson and Lankton (2012). Demographic items were administered at the end of the survey (see Table 2).

sMUX Construct Validation[2]

Convergent Validity

We assessed convergent validity by calculating Cronbach's Alpha and composite reliability (CR) of the three sMUX scales, including comparable MUX scales as benchmarks (see Table 3). For all sMUX and MUX scales, both statistics exceeded the .70 criterion proposed by Hair et al. (2009). In addition, we assessed that average variance extracted (AVE) statistics are above .50 in all cases. Both sMUX and MUX scales demonstrated satisfactory convergent validity across these analyses.

We do note that Cronbach's alpha statistics average approximately .10 lower for sMUX scales than for MUX scales. This is not surprising, as alpha is known to be correlated with the number of items under analysis (Hair et al., 2009). Differences in reliability between sMUX and MUX scales are substantially reduced when measured by composite reliability, a statistic which is generally considered to provide a more accurate assessment than is offered by alpha (Peterson & Kim, 2013).

Discriminant Validity

We assessed discriminant validity by analyzing the square root of AVE. We found this statistic to be substantially higher than any correlation of that scale with any other scale, thereby exceeding criteria proposed by Fornell and Larcker (1981). These results correspond closely with those reported by Djamasbi and Wilson (2017) in their initial MUX validation. Both sMUX and MUX scales demonstrate satisfactory discriminant validity, and we did not find any important differences between the instruments.

---

[1] No differences in EFA loading patterns (i.e., loading values above 0.50) were observed between separate university populations and the combined population.

[2] The five-item MUX scales were previously validated by Djamasbi and Wilson (2017) following the scale development process recommended by Hinken (1998).

sMUX Scales						
Factor	AVE	Alpha	CR	Nuisance	Access	Mobility
Nuisance	.705	.790	.878	**.840**		
Access	.631	.707	.837	-.306	**.794**	
Mobility	.729	.814	.890	.025	.255	**.854**
MUX Scales						
Factor	AVE	Alpha	CR	Nuisance	Access	Mobility
Nuisance	.654	.867	.904	**.809**		
Access	.605	.835	.884	-.362	**.778**	
Mobility	.733	.908	.932	.020	.303	**.856**
* Square roots of average variances extracted (AVEs) are shown in bold on the diagonals						

Table 3. AVE, Reliability, and Correlations Related to sMUX and MUX Factors*

Known Groups Validity and Measurement Sensitivity Assessment

Known groups validity (Robinson et al, 1991) is the essential capability of scientific instruments to distinguish as expected between conditions where differences are known to exist. Portney and Watkins (2000, p. 89) write,

> "The most general type of evidence in support of construct validity is provided when a test can discriminate between individuals who are known to have the trait and those who do not. Using the known groups method, a criterion is chosen that can identify the presence or absence of a particular characteristic, and the theoretical context behind the construct is used to predict how different groups are expected to behave. Therefore, the validity of a particular test is supported if the test's results document these known differences."

The ability to distinguish between different types of software and devices is especially pertinent in the context of mobile computing, where usability can be impacted by seemingly small differences in features, such as the size of display (Raptis et al., 2013) or the method of page navigation (Kim et al., 2016). The first step in establishing relevance to the domain of mobile computing is to ensure that instruments accurately detect differences between research conditions where these differences are known to exist in advance of the analysis.

Our student subjects were drawn from two universities that use distinct LMS software (Blackboard at Eastern U. vs. Desire2Learn at Midwest U.). These students were asked to respond to questions about the LMS in the context of the computing device they had used most recently to access their university's LMS. Thus, for a subject at Midwest U. who reported most recently using a laptop to access the university LMS, the item denoted as *Nuisance-A* in Table 1 would have been presented to the subject as "I felt using a laptop to access Desire2Learn would slow me down".

We propose certain differences will exist between software instances and device form factors. First, Chawdhry et al. (2011) studied university students' perceptions of Blackboard and Desire2Learn software, reporting that nearly two-thirds of students preferred Desire2Learn. This research suggests similar results will emerge in our usability-focused comparison of LMS software. Second, structural differences exist in mobility characteristics between desktop computers and the other device form factors we are studying (smartphone/tablet and laptop computers). Our expectation that desktop computers will be perceived as less mobile is based on logical analysis related to these structural differences.

Measurement sensitivity, describing the ability to detect small differences between conditions in a research domain, is especially important when developing short-form versions of previously-validated survey instruments (Garratt et al., 1994; Katz et al., 1992). The most relevant conditions we observe in the research domain of mobile computing are those representing specific combinations of mobile device and software application. If sMUX is to be a valuable addition to mobile UX research, we propose that it should exhibit sufficient measurement sensitivity to distinguish between distinct conditions.

In order to assess known groups validity and measurement sensitivity of sMUX, we performed a series of SPSS ANOVAs on the data. We entered *Software* (Desire2Learn coded as 1 and Blackboard coded as 2) and *Device Form Factor* (smartphone/tablet coded as 1, laptop computer coded as 2, and desktop computer coded as 3) as fixed factors, and summated data (calculated as the average value of items making up each factor) for each of the sMUX scales was entered as a dependent variable in a series of three separate analyses. We conducted one further ANOVA in which the calculated SUS score (Brooke, 1996) was entered as the dependent variable. Results of these four ANOVAs are shown in Table 4.

Significant differences were found in sMUX scales on both Software and Device Form Factor dimensions. sMUX Nuisance scores were higher for Blackboard software than for Desire2Learn ($p = .042$), and Device Form Factor values varied significantly on each sMUX scale. SUS Scores were higher for Desire2Learn than Blackboard ($p = .016$).

In order to identify which device form factors varied significantly from one another, homogeneous subsets were calculated in SPSS using the Scheffé test (see Table 5). The smartphone/tablet form factor scored significantly higher in sMUX-Nuisance and lower in sMUX-Access and SUS than the other form factors. Not surprisingly, desktop computer scored lower than other form factors in sMUX-Mobility.

Our findings demonstrate known groups validity in that the sMUX-Nuisance scale correctly identified differences between LMS software conditions and the sMUX-Mobility scale identified differences between desktop computers and the other device form factors, two relationships in our analysis that were anticipated in advance based on prior research or logic.

In addition, sMUX scales demonstrated measurement sensitivity by providing a fine-grained explanation of the nature of those differences. For example, where SUS shows the smartphone/tablet form factor provided a lower-quality overall user experience than laptop or desktop computers, sMUX reveals that these differences centered on perceptions of higher Nuisance and lower Access characteristics in smartphones and tablets and that they were counterbalanced to some degree by higher perceptions of Mobility for these devices. We note in our analysis that SUS scores did not distinguish differences between desktop computers and other device form factors or provide a similarly fine-grained explanation of differences among the conditions we assessed.

Discussion

Our objectives in designing this study were to develop a reduced instrument that satisfies brevity needs of UX practitioners, maintains sufficient robustness and validity to meet standards of scholarly research, and provides new insights into user experience with mobile computing beyond what is available using other instruments. We interpret that sMUX has met each of these objectives.

The initial assessment we conducted suggests that sMUX has We found the three sMUX factors that emerged through EFA explained 62% of total variance vs. 60% reported by Djamasbi and Wilson (2017), and all items loaded consistently with other items in their theorized construct in both studies. We also found that all SUS items in this study loaded primarily with items comprising the sMUX Nuisance or Access constructs. This finding reinforces the assessment by Djamasbi and Wilson (2017) that SUS can be effectively replaced by MUX or, as we found here, by sMUX in circumstances where it is necessary to limit survey length.

Scale	Factor	SS	df	Mean Square	F	Sig	Partial Eta²
sMUX Nuisance	Software	2.825	1	2.825	4.170	.042	.020
	Device Form Factor	7.037	2	3.519	5.193	.006	.048
	SW. x Device	0.869	2	.434	0.641	.528	.006
sMUX Access	Software	0.704	1	.704	1.545	.215	.007
	Device Form Factor	6.559	2	3.279	7.201	.001	.065
	SW. x Device	0.303	2	.151	0.332	.718	.003
sMUX Mobility	Software	0.399	1	.399	0.642	.424	.003
	Device Form Factor	47.027	2	23.513	37.854	.000	.267
	SW. x Device	2.384	2	1.192	1.919	.149	.018
SUS	Software	1304.9	1	1304.9	5.89	.016	.028
	Device Form Factor	2900.8	2	1450.4	6.547	.002	.059
	SW. x Device	379.1	2	189.5	.856	.427	.008

Table 4. ANOVA Between-subjects Effects for sMUX and SUS Scales

Measure	Device Form Factor	N	Subset*	
			1	2
sMUX Nuisance	Smartphone/Tablet	23		2.67
	Laptop Computer	166	2.14	
	Desktop Computer	25	2.09	
sMUX Access	Smartphone/Tablet	23	3.55	
	Laptop Computer	166		4.07
	Desktop Computer	25		4.12
sMUX Mobility	Smartphone/Tablet	23		3.97
	Laptop Computer	166		3.77
	Desktop Computer	25	2.24	
SUS	Smartphone/Tablet	23	62.8	
	Laptop Computer	166		72.7
	Desktop Computer	25		76.6

* Responses measured on a five-position Likert scale where 1 = "Strongly Disagree", 2 = "Disagree", 3 = "Neither Agree Nor Disagree", 4 = "Agree", and 5 = "Strongly Agree"

Table 5. Homogeneous Subsets for sMUX Scales and SUS

In addition, our assessment of convergent and discriminant validity in sMUX compares favorably with prior validation of the MUX instrument. As shown in Table 3, all sMUX dimensions surpass reliability thresholds and demonstrate low levels of intercorrelation with each other relative to average variance explained (AVE), producing overall statistics that are similar to those reported for MUX (Djamasbi & Wilson, 2017). These findings are especially supportive of the robustness of sMUX given that our research design in this study expanded to test across two different software applications rather than the single software application that was studied previously.

Development of sMUX provides new insights in several ways. sMUX provides finer-grained detail regarding the sources of usability than is possible to achieve with general instruments, such as SUS. For example, data from Table 4 shows usability difference between software is driven by higher sMUX Nuisance value of Blackboard vs. Desire2Learn.

In addition, these details are not limited in focus to the mobile software domain, as is assessed by the MAU instrument (Hoehle & Venkatesh, 2015). As data from Table 5 demonstrates, prominent usability differences arose between smartphone/tablet and other form factors *using the same software*. These differences are driven by the higher sMUX Nuisance values and lower sMUX Access values users perceived during actual experiences with these devices.

Recommendations for Applying sMUX and MUX

Both MUX and sMUX instruments now have been successfully validated, and each has proved to be capable of distinguishing between conditions that are representative of mobile computing contexts. The combined findings of this study and Djamasbi and Wilson (2017) suggest that either instrument could be applied effectively to study usability of mobile software, mobile devices, or combinations of mobile software and devices. Nonetheless, we recommend several factors for consideration when selecting which instrument to apply.

We expect sMUX to be the obvious choice in circumstances where survey length is highly constrained, for example, where measurement is applied as one of several components in formative usability evaluations (Redish et al, 2002). Where survey length is not constrained, we recommend that researchers implement the complete MUX instrument. This conservative approach offers the most flexibility in case there is some emergent interaction between research context and measurement item that was not encountered in initial validation testing. In addition, we recognize that few results of MUX/sMUX research exist currently, and it will be some time before the existing inventory of results will match that of the SUS instrument. At some point, we anticipate that benchmarks will be developed for interpreting MUX/sMUX results in the context of SUS. Until such benchmarks are developed, however, we recommend that researchers incorporate both SUS and MUX/sMUX measures into their research designs where it is possible to do so.

Future Research Directions

The successful applications to date of MUX and sMUX instruments encourage future research extending in several directions. Usability researchers will be benefited greatly by development of an inventory of MUX/sMUX research outcomes that can be used for benchmarking. This effort may be guided by Bangor et al. (2008), who evaluated results from 2,324 SUS surveys gleaned from 206 usability studies. Based on this evaluation, they were able to map SUS scores onto categories that greatly assist interpreting results in practical settings, for example, denoting “Good” SUS scores as those in the range of 55-75 and “Excellent” scores in the range of 75-87.5.

Where our initial testing used relatively broad mobile device form factors, we propose MUX/sMUX should be applied in research designs where more detailed differences are compared, for example, between smartphone and tablet form factors or between different brands of smartphones. We anticipate fine-grained distinctions will be supported through the approach of separately evaluating users’ perceptions of Nuisance, Access, and Mobility as opposed to simply assessing overall usability.

A further opportunity for expanded research will address a wider array of software than has been studied to date. It will be interesting in these cases to contrast MUX/sMUX results to those of other measures, such as MAU instrument (Hoehle & Venkatesh, 2015) or to use MUX/sMUX results to augment other measures.

Finally, we propose the MUX/sMUX instruments can play a central role in understanding user adoption and continuance decisions in use of mobile devices and mobile software. Hoehle and Venkatesh (2015, p. 465) write that usability of mobile software should be considered "in combination with other theories, such as IS continuance theory (Bhattacherjee, 2001), IS success model (DeLone and McLean, 1992) unified theory of acceptance and use of technology (Venkatesh et al., 2003), and task-technology fit (Goodhue and Thompson, 1995) to study why individuals use mobile applications." We concur with this assessment and add that further benefits may be gained by also considering usability of mobile devices and combinations of mobile device and software within these theories. The MUX and sMUX instruments make such extended considerations possible.

Conclusion

Use of mobile computing is becoming more of a norm than exception. For this reason, evaluating the holistic experience of users related to mobile software and mobile devices is both relevant and important to HCI researchers. In this paper we presented the continuing refinement of a mobile user experience (MUX) instrument to produce a short-form version titled sMUX, which we anticipate will be especially valuable to practitioners. We successfully evaluated convergent and discriminant validity of sMUX and further demonstrated the known groups validity and measurement sensitivity of this instrument across multiple combinations of mobile software and mobile device. Both sMUX and the larger MUX instrument offer potential to increase understanding of usability issues in ways that could not be possible with prior usability instruments.

References

Bagozzi, R. P. (2011). Measurement and meaning in information systems and organizational research: Methodological and philosophical foundations. MIS Quarterly, 35(2), 261-292.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS Scores mean: Adding an adjective rating scale. Journal of Usability Studies, 4(3), 114–123.

Bhattacherjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. MIS Quarterly, 25(3), 351-370.

Brooke, J. (1996). SUS: A quick and dirty usability scale. Usability Evaluation in Industry, 189(194), 4-7.

Chawdhry, A., Paullet, K., & Benjamin, D. (2011). Comparatively assessing the use of Blackboard versus Desire2Learn: student perceptions of the online tools. Issues in Information Systems, 12(2), 273-280.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. Information Systems Research, 3(1), 60-95.

Djamasbi, S., & Wilson, E. V. (2017). MUX: Development of a Holistic Mobile User Experience Instrument. In Proceedings of the 50th Hawaii International Conference on System Sciences, Waikaloa Village, HI.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. Journal of marketing research, 382-388.

Galesic, M., & Bosnjak, M. (2009). Effects of questionnaire length on participation and indicators of response quality in a web survey. Public Opinion Quarterly, 73, 2, 349-360.

Garratt, A. M., Ruta, D. A., Abdalla, M. I., & Russell, I. T. (1994). SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. Quality in Health Care, 3(4), 186-192.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. MIS Quarterly, 19(2) 213-236.

Hair, J., Black, W., Babin, B., & Anderson, R. (2009). Multivariate data analysis (7th edition). Upper Saddle River, NJ: Pearson Prentice Hall.

Hayes, A. F. (2005). A computational tool for survey shortening applicable to composite attitude, opinion, and personality measurement scales. Paper presented at the meeting of the Midwestern Association for Public Opinion Research, Chicago, IL. http://www.afhayes.com/public/alphamax.pdf

Heberlein, T. A., & Baumgartner, R. (1978). Factors affecting response rates to mailed questionnaires: A quantitative analysis of the published literature. American

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. Organizational Research Methods, 1(1), 104-121.

Hoehle, H., & Venkatesh, V. (2015). Mobile application usability: conceptualization and instrument development. MIS Quarterly, 39(2), 435-472.

Katz, J. N., Larson, M. G., Phillips, C. B., Fossel, A. H., & Liang, M. H. (1992). Comparative measurement sensitivity of short and longer health status instruments. Medical Care, 30(10), 917-925.

Kim, J., Thomas, P., Sankaranarayana, R., Gedeon, T., & Yoon, H. J. (2016). Pagination versus scrolling in mobile web search. In Proceedings of the 25th ACM International on Conference on Information and Knowledge Management (pp. 751-760). ACM.

MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The problem of measurement model misspecification in behavioral and organizational research and some recommended solutions. Journal of Applied Psychology, 90(4), 710-730.

Peterson, R. A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. Journal of Applied Psychology, 98(1), 194-189.

Podsakoff, N. P., Shen, W., & Podsakoff, P. M. (2006). The role of formative measurement models in strategic management research: review, critique, and implications for future research. In D. J. Ketchen and D. D. Bergh (Eds.), Research methodology in strategy and management (pp. 197-252). Emerald Group Publishing Limited.

Portney, L. G., & Watkins, M. P. (2000). Foundations of clinical research: applications to practice (Vol. 2). Upper Saddle River, NJ: Prentice Hall.

Raptis, D., Tselios, N., Kjeldskov, J., & Skov, M. B. (2013, August). Does size matter? Investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. In Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services (pp. 127-136). ACM.

Redish, J. G., Bias, R. G., Bailey, R., Molich, R., Dumas, J., & Spool, J. M. (2002, April). Usability in practice: Formative usability evaluations-evolution and revolution. In CHI'02 extended abstracts on Human Factors in Computing Systems (pp. 885-890). ACM.

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). Criteria for scale selection and evaluation. Measures of personality and social psychological attitudes, 1(3), 1-16.

Sociological Review, 43, 447-462.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425-478.

Wilson, E. V. & Lankton, N. K. (2012). Some unfortunate consequences of non-randomized, grouped-item survey administration in IS research. Proceedings of the 2012 International Conference on Information Systems (ICIS), Orlando, FL.

Wilson, E. V., & Lankton, N. K. (2013). Effects of prior use, intention, and habit on IT continuance across sporadic use and frequent use conditions. Communications of the Association for Information Systems, 33 (Article 3).

Wilson, E. V., Mao, E., & Lankton, N. K. (2010). The distinct roles of prior IT use and habit strength in predicting continued sporadic use of IT. Communications of the Association for Information Systems, 27(1), 185-206.