

Revisiting the Factor Structure of the System Usability Scale

James (Jim) R. Lewis

Senior HF Engineer
IBM Corp.
5901 Broken Sound Parkway
Suite 514C
Boca Raton, FL 33487
USA
jimlewis@us.ibm.com

Jeff Sauro

MeasuringU
Principal
jeff@measuringu.com

Abstract

In 2009, we published a paper in which we showed how three independent sources of data indicated that, rather than being a unidimensional measure of perceived usability, the System Usability Scale apparently had two factors: Usability (all items except 4 and 10) and Learnability (Items 4 and 10). In that paper, we called for other researchers to report attempts to replicate that finding. The published research since 2009 has consistently failed to replicate that factor structure. In this paper, we report an analysis of over 9,000 completed SUS questionnaires that shows that the SUS is indeed bidimensional, but not in any interesting or useful way. A comparison of the fit of three confirmatory factor analyses showed that a model in which the SUS's positive-tone (odd-numbered) and negative-tone (even-numbered) were aligned with two factors had a better fit than a unidimensional model (all items on one factor) or the Usability/Learnability model we published in 2009. Because a distinction based on item tone is of little practical or theoretical interest, we recommend that user experience practitioners and researchers treat the SUS as a unidimensional measure of perceived usability, and no longer routinely compute Usability and Learnability subscales.

Keywords

System Usability Scale, SUS, factor structure, perceived usability, perceived learnability, confirmatory factor analysis



Introduction

In this section, we discuss our reasoning as to why we revisited the factor structure of SUS, what is the SUS, the psychometric properties of SUS, and our objectives for this study.

Why Revisit the Factor Structure of the System Usability Scale (SUS)?

There are still lessons to be learned in the domain of standardized usability testing—still work to do. For example, what is the real factor structure of the SUS? (Lewis, 2014, p. 675).

The SUS (Brooke, 1996) is a very popular (if not the most popular) standardized questionnaire for the assessment of perceived usability. Sauro and Lewis (2009), in a study of unpublished industrial usability studies, found that the SUS accounted for 43% of post-test questionnaire usage. It has been cited in over 1,200 publications (Brooke, 2013).

The SUS was designed to be a unidimensional (one factor) measurement of perceived usability (Brooke, 1996). Once researchers began to publish data sets (or correlation matrices) from sample sizes large enough to support factor analysis, it began to appear that SUS might be bidimensional (having a structure with two factors). Factor analyses of data from three independent studies (Borsci, Federici, & Lauriola, 2009; Lewis & Sauro, 2009, which included a reanalysis of the SUS item correlation matrix published by Bangor, Kortum, & Miller, 2008) indicated a consistent two-factor structure (with Items 4 and 10 aligning on a factor separate from the remaining items). Lewis and Sauro named the two factors Usability (all items except 4 and 10) and Learnability (Items 4 and 10).

This was an exciting finding, with support from three independent sources. These new scales had good psychometric properties (e.g., coefficient alpha greater than 0.70). A sensitivity analysis using data from 19 tests provided evidence of the differential utility of the new scales. The promise of this research was that practitioners could continue to use the standard SUS—but, at no extra cost, could also take advantage of the new scales to extract additional information from their SUS data. Google Scholar metrics (visited 9/17/2016) indicate the paper that reported this finding (Lewis & Sauro, 2009) has been cited over 350 times.

Unfortunately, analyses conducted since 2009 (Kortum & Sorber, 2015; Lewis, Brown, & Mayes, 2015; Lewis, Utesch, & Mayer, 2013, 2015; Sauro & Lewis, 2011) have typically resulted in a two-factor structure but have not consistently replicated the item-factor alignment that seemed apparent in 2009 (a separation of Items 4 and 10). Research by Borsci, Federici, Bacci, Gnaldi, and Bartolucci (2015) suggested the possibility that one- versus the two-factor structure (Usability/Learnability) might depend on the level of user experience, but Lewis, Utesch, and Maher (2015) were not able to replicate this finding. Otherwise, the more recent analyses have been somewhat consistent with a general alignment of positive- and negative-tone items on separate factors—the type of unintentional structure that can occur with sets of mixed-tone items (Barnette, 2000; Davis, 1989; Pilotte & Gable, 1990; Schmitt & Stuits, 1985; Schriesheim & Hill, 1981; Stewart & Frye, 2004; Wong, Rindfleisch, & Burroughs, 2003). Specific reported structures have included the following (and note that in every case the second factor has included Items 4 and 10, but not in isolation):

- Factor 1: Items 1, 3, 5, 7, 9; Factor 2: Items 2, 4, 6, 8, 10 (Kortum & Sorber, 2015; Lewis, Brown, & Mayes, 2015)
- Factor 1: Items 1, 3, 5, 6, 7, 8, 9; Factor 2: Items 2, 4, 10 (Kortum & Sorber, 2015)
- Factor 1: Items 1, 2, 3, 5, 7, 9; Factor 2: Items 4, 6, 8, 10 (Sauro & Lewis, 2011)
- Factor 1: Items 1, 9; Factor 2: Items 2, 3, 4, 5, 6, 7, 8, 10 (Borsci et al., 2015; Lewis, Utesch, & Maher, 2015)

When we published our 2009 paper, we were following the data. Our paper has been influential, with over 350 recorded citations. Unfortunately, as clear as the factor structure appeared to be in 2009, analyses since then have failed to replicate the reported Usability/Learnability structure with alarming consistency. We believe it is time to reassess the factor structure of the SUS, and have brought together the largest collection of completed SUS questionnaires of which we are

aware ($N > 9,000$) to, as definitively as possible, compare the fit of various models of the factor structure of the SUS.

What Is the SUS?

As shown in Figure 1, the standard version of the SUS has 10 items, each with five steps anchored with "Strongly Disagree" and "Strongly Agree." It is a mixed-tone questionnaire in which the odd-numbered items have a positive tone and the even-numbered items have a negative tone. The first step in scoring a SUS is to determine each item's score contribution, which will range from 0 (a poor experience) to 4 (a good experience). For positively-worded items (odd numbers), the score contribution is the scale position minus 1. For negatively-worded items (even numbers), the score contribution is 5 minus the scale position. To get the overall SUS score, multiply the sum of the item score contributions by 2.5, which produces a score that can range from 0 (very poor perceived usability) to 100 (excellent perceived usability) in 2.5-point increments.

The System Usability Scale Standard Version		Strongly Disagree					Strongly Agree				
		1	2	3	4	5	1	2	3	4	5
1	I think that I would like to use this system frequently.	0	0	0	0	0					
2	I found the system unnecessarily complex.	0	0	0	0	0					
3	I thought the system was easy to use.	0	0	0	0	0					
4	I think that I would need the support of a technical person to be able to use this system.	0	0	0	0	0					
5	I found the various functions in this system were well integrated.	0	0	0	0	0					
6	I thought there was too much inconsistency in this system.	0	0	0	0	0					
7	I would imagine that most people would learn to use this system very quickly.	0	0	0	0	0					
8	I found the system very awkward to use.	0	0	0	0	0					
9	I felt very confident using the system.	0	0	0	0	0					
10	I needed to learn a lot of things before I could get going with this system.	0	0	0	0	0					

Figure 1. The standard System Usability Scale. Note: Item 8 shows "awkward" in place of the original "cumbersome" (Finstad, 2006; Sauro & Lewis, 2009).

Psychometric Properties of the SUS

The SUS has excellent psychometric properties. Research has consistently shown the SUS to have reliabilities at or just over 0.90 (Bangor et al., 2008; Lewis, Brown, & Mayes, 2015; Lewis & Sauro, 2009; Lewis, Utesch, & Maher, 2015), far above the minimum criterion of 0.70 for measurements of sentiments (Nunnally, 1978). The SUS has also been shown to have acceptable levels of concurrent validity (Bangor, Joseph, Sweeney-Dillon, Stettler, & Pratt, 2013; Bangor et al., 2008; Kortum & Peres, 2014; Lewis, Brown, & Mayes, 2015; Peres, Pham, Philips, 2013) and sensitivity (Kortum & Bangor, 2013; Kortum & Sorber, 2015; Lewis & Sauro, 2009; Tullis & Stetson, 2004). Norms are available to guide the interpretation of the SUS (Bangor, Kortum, & Miller, 2008, 2009; Sauro, 2011; Sauro & Lewis, 2016).

Objective of the Current Study

The objective of this current study is to revisit the factor structure of the SUS. The strategy is to use a very large sample of completed SUS questionnaires to (a) use exploratory factor analysis to reveal the apparent alignment of items, then (b) use confirmatory factor analysis to assess the goodness of fit for three models of item-factor alignment: the Unidimensional model (all 10

SUS items on one factor), the Usability/Learnability model (Items 4 and 10 on one factor, all other items on a second factor), and a Tone model (based on the tone of the SUS items, with positive tone items on one factor, negative tone items on a second factor).

Method

For this study, we assembled a data set of 9,156 completed SUS questionnaires from 112 unpublished industrial usability studies and surveys from a range of software products and websites. Most of the datasets did not have a sufficient sample size for factor analysis, but combined, this is the largest collection of completed SUS questionnaires of which we are aware and provides considerable power for statistical analysis (MacCallum, Browne, & Sugawara, 1996). All analyses were conducted using standard SUS item contribution scores rather than raw scores so score directions were consistent (0-4 point scales; low = poor experience; high = good experience).

Results

In the following sections, we discuss the results as they relate to the exploratory analyses and the confirmatory factor analyses.

Exploratory Analyses

Investigators have used a variety of methods to explore the structure of the SUS. To address the variety of techniques in the literature, we used three popular methods available in IBM SPSS Statistics Version 23: principal components analysis (PCA—strictly speaking, not a factor analytic method, but commonly used), unweighted least squares factor analysis (ULSFA—minimizes the sum of the squared differences between the observed and reproduced correlation matrices), and maximum likelihood factor analysis (MLFA—produces parameter estimates that are most likely to have produced the observed correlation matrix if the sample is from a multivariate normal distribution). The use of these three methods allows the determination that the observed factor structure is robust across the different analytical approaches.

The eigenvalues from the exploratory analyses were 5.637, 1.467, 0.547, 0.491, 0.379, 0.344, 0.317, 0.309, 0.257, and 0.251. Parallel analysis of the eigenvalues (Ledesma & Valero-Mora, 2007; Patil, Singh, Mishra, & Donovan, 2007) indicated a two-factor solution. As shown in Table 1, all three methods (with Varimax-rotated two-factor structures) were consistent with the Tone model (positive and negative tone items loading more strongly on different components/factors).

Table 1. Component/Factor Loadings for Three Exploratory Structural Analyses

Item	PCA 1	PCA 2	ULSFA 1	ULSFA 2	MLFA 1	MLFA 2
1	0.048	0.771	0.638	0.115	0.638	0.116
2	0.739	0.372	0.388	0.686	0.391	0.689
3	0.361	0.798	0.790	0.348	0.793	0.347
4	0.852	0.061	0.108	0.777	0.108	0.772
5	0.211	0.819	0.770	0.219	0.767	0.223
6	0.771	0.339	0.354	0.725	0.348	0.732
7	0.321	0.753	0.706	0.320	0.712	0.316
8	0.767	0.422	0.431	0.742	0.428	0.745
9	0.364	0.778	0.756	0.356	0.751	0.356
10	0.833	0.180	0.213	0.773	0.216	0.766

Confirmatory Factor Analyses

Confirmatory factor analysis (CFA) differs from exploratory factor analysis (EFA) in that an EFA produces unconstrained results that the researcher examines for structural clues, but a CFA is constrained to a precisely defined model (Cliff, 1987). Researchers can conduct CFAs on multiple proposed models and compare their indices of goodness-of-fit to assess which model has the best fit to the given data. Jackson, Gillaspay, Jr., and Purc-Stephenson (2009) have recommended reporting fit statistics that have different measurement properties such as the comparative fit index (CFI—a score of 0.90 or higher indicates good fit), the root-mean-square error of approximation (RMSEA—values less than 0.08 indicate acceptable fit), and the Bayesian information criterion (BIC—lower values are preferred). It is common to also report chi-square tests of absolute model fit, but when sample sizes are very large, such tests almost always lead to rejection of the hypothesis of adequate fit (Kline, 2011), making them uninformative. Instead, we have focused on comparative fit metrics.

We used the lavaan package in the statistical program *R* (Rosseel, 2012) to conduct CFA on the three models of the SUS described in the introduction. Figures 2, 3, and 4 illustrate the three models (created using SPSS AMOS 24). Model 1 (Figure 2) represents the unidimensional model of SUS, which was over-identified with 55 sample moments and 20 parameters ($df = 35$). Model 2, the two-factor Usability and Learnability model shown in Figure 3, was also over-identified with 55 sample moments and 21 parameters ($df = 34$). Model 3, the two-factor positive-negative model shown in Figure 4, was also over-identified with 55 sample moments and 21 parameters ($df = 34$). Table 2 shows the results of the comparative fit analyses of the three models (with 90% confidence intervals for RMSEA produced in lavaan by default).

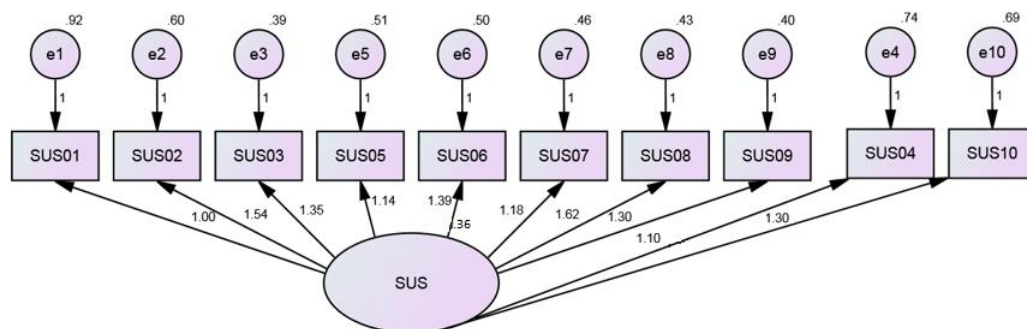


Figure 2. Model 1, the unidimensional SUS.

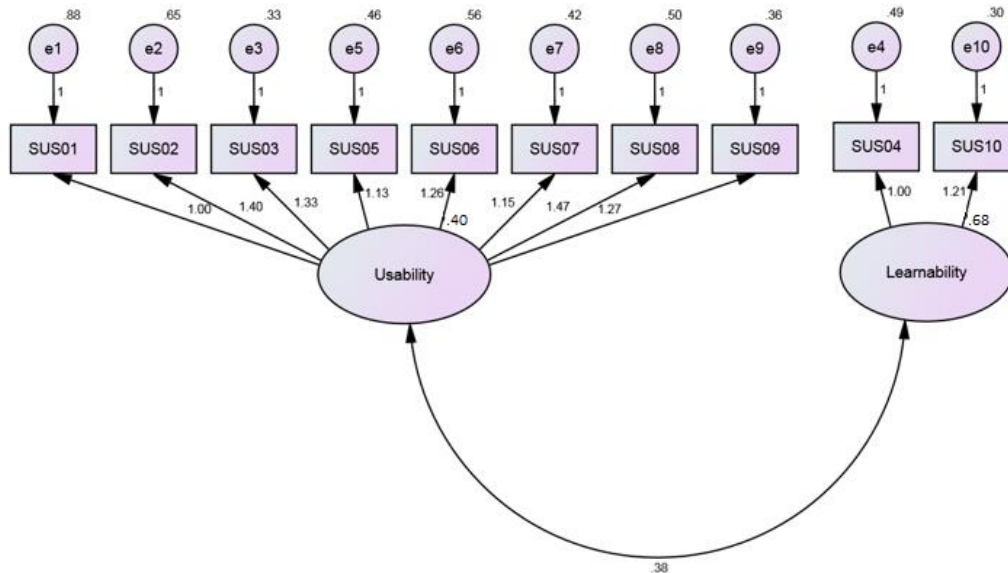


Figure 3. Model 2, the bidimensional SUS (Usability/Learnability).

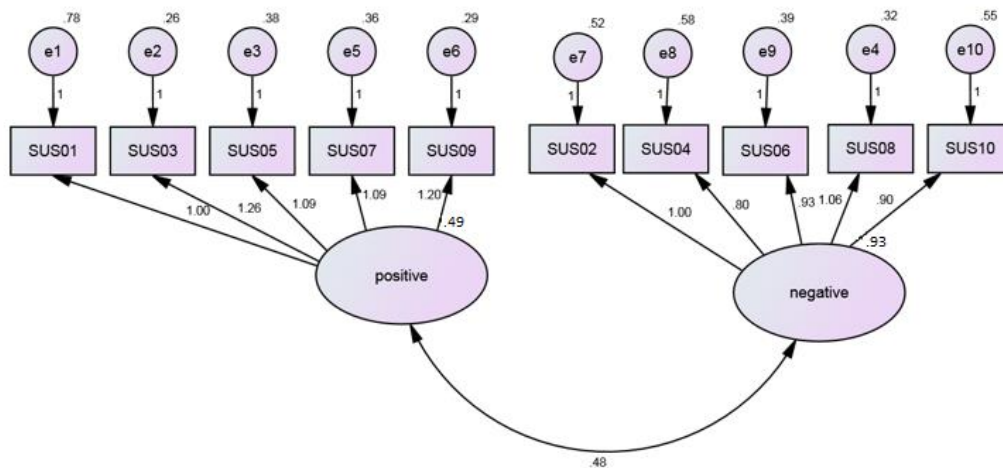


Figure 4. Model 3, the bidimensional SUS (Positive/Negative Tone).

Table 2. Results of CFAs of Three Structural Models of the SUS

Model	Description	CFI	90% lower	RMSEA	90% upper	BIC
1	Unidimensional	0.799	0.187	0.190	0.193	11801
2	Usability/Learnability	0.838	0.170	0.173	0.176	9543
3	Positive/Negative Tone	0.958	0.085	0.088	0.091	2449

Consistent with the results from the EFA, the multiple fit statistics indicated that the best-fitting model was the Positive/Negative Tone model. That was the only one of the three models that had a CFI greater than 0.90, and its RMSEA, despite not quite achieving the criterion of being less than 0.08 for acceptable fit, was about half of that for the other two models. Notably, there

was no overlap among the RMSEA confidence intervals, which is evidence of statistically significant differences. The Bayesian information criterion (BIC) was also lowest (best) for the Positive/Negative Tone model.

Conclusion

One of the strengths of the scientific method is its self-correction when the accumulation of evidence indicates a need to do so. It can be disappointing when an interesting finding fails to survive continuing scrutiny, but this is how our knowledge advances—by keeping a distant reaction to results rather than rooting for a particular outcome.

In 2009, we published a paper (Lewis & Sauro, 2009) in which we showed how three independent sources of data indicated that, rather than being a unidimensional measure of perceived usability, the System Usability Scale apparently had two factors: Usability (all items except 4 and 10) and Learnability (Items 4 and 10). In that paper, we called for other researchers to report attempts to replicate that finding, and we also continued this investigation in our own research. That paper has been cited over 350 times.

The published research since 2009 has consistently failed to replicate that Usability/Learnability factor structure. In this paper, we reported an analysis of over 9,000 completed SUS questionnaires that shows that the SUS is indeed bidimensional, but not in any interesting or useful way. A comparison of the fit of three confirmatory factor analyses showed that a model in which the SUS's positive-tone (odd-numbered) and negative-tone (even-numbered) were aligned with two factors had a better fit than a unidimensional model (all items on one factor) or the Usability/Learnability model we published in 2009.

Thus, the factor structure of the SUS appears to be bidimensional, but apparently not in any interesting way. It is well known that mixed tone questionnaires like the SUS often exhibit this type of nuisance structure when factor analyzed (Barnette, 2000; Davis, 1989; Pilotte & Gable, 1990; Schmitt & Stuitts, 1985; Schriesheim & Hill, 1981; Stewart & Frye, 2004; Wong et al., 2003). The same pattern has been reported for the Usability Metric for User Experience (UMUX) (Lewis, Utesch, & Maher, 2013), another metric of perceived usability that has items with mixed tone. Davis (1989), in his development of the Technology Acceptance Model, started with a pool of mixed tone items, but found that the mixed tone was causing problems in his attempt to get clear factors for Perceived Usefulness and Perceived Ease-of-Use. He consequently eliminated the negative-tone items from consideration.

It is possible that the SUS might have internal structure that is obscured by the effect of having mixed tone items, but we found no significant evidence supporting that hypothesis. It is interesting to note in Table 1 that the magnitude of the factor loadings for Items 4 and 10 in all three exploratory analyses were greater than those for Items 2, 6, and 8 on the negative tone factor, suggesting (but not proving) that there might be some research contexts in which they would emerge as an independent factor.

Because a distinction based on item tone is of little practical or theoretical interest when measuring with the SUS, it is, with some regret but based on accumulating evidence, that we recommend that user experience practitioners and researchers treat the SUS as a unidimensional measure of perceived usability, and no longer routinely compute or report Usability and Learnability subscales.

Recommendations for Researchers

Researchers should be cautious in their use of the Usability/Learnability factor structure reported by Lewis and Sauro (2009). As shown in Table 1, Items 4 and 10 loaded more strongly on the negative tone factor than the other three items. It might be the case that the Usability/Learnability structure appears in certain special circumstances (e.g., as reported by Borsci et al., 2015 in their investigation of the amount of experience users have with a product), but such findings require replication. Although the evidence strongly suggests that the SUS is bidimensional as a function of item tone, these dimensions are of little theoretical or practical interest. Unless there is compelling evidence in a specific domain of research to support interpretation of an alternative structure, the best research policy is to interpret the SUS as a unidimensional measure of perceived usability.

Tips for Usability Practitioners

The following are some guidelines for practitioners:

- Do not routinely compute Usability and Learnability subscales from SUS data.
- Instead, routinely compute the standard overall SUS and interpret it as a unidimensional measure of perceived usability.
- Only if you are working in a context in which the Usability and Learnability subscales have been shown to reliably occur, should you compute and report them.

References

- Bangor, A., Joseph, K., Sweeney-Dillon, M., Stettler, G., & Pratt, J. (2013). Using the SUS to help demonstrate usability's value to business goals. In *Proceedings of the Human Factors Society and Ergonomics Society Annual Meeting* (pp. 202–205). Santa Monica, CA: HFES.
- Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An empirical evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction, 24*, 574–594.
- Bangor, A., Kortum, P. T., & Miller, J. T. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of Usability Studies, 4*(3), 114–123.
- Barnette, J. J. (2000). Effects of stem and Likert response option reversals on survey internal consistency: If you feel the need, there is a better alternative to using those negatively worded stems. *Educational and Psychological Measurement, 60*, 361–370.
- Borsci, S., Federici, S., Bacci, S., Gnaldi, M., & Bartolucci, F. (2015). Assessing user satisfaction in the era of user experience: Comparison of the SUS, UMUX and UMUX-LITE as a function of product experience. *International Journal of Human-Computer Interaction, 31*(8), 484–495.
- Borsci, S., Federici, S., & Lauriola, M. (2009). On the dimensionality of the System Usability Scale: A test of alternative measurement models. *Cognitive Processes, 10*, 193–197.
- Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In P. Jordan, B. Thomas, & B. Weerdmeester (Eds.), *Usability Evaluation in Industry* (pp. 189–194). London, UK: Taylor & Francis.
- Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies, 8*(2), 29–40.
- Cliff, N. (1987). *Analyzing multivariate data*. Orlando, FL: Harcourt, Brace, Jovanovich.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*, 319–339.
- Finstad, K. (2006). The System Usability Scale and non-native English speakers. *Journal of Usability Studies, 1*(4), 185–188.
- Jackson, D. L., Gillaspay, Jr., J. A., & Purc-Stephenson. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14*, 6–23.
- Kline, R. B. (2011). *Principles and practices of structural equation modeling* (3rd ed.). New York, NY: The Guilford Press.
- Kortum, P., & Bangor, A. (2013). Usability ratings for everyday products measured with the System Usability Scale. *International Journal of Human-Computer Interaction, 29*, 67–76.
- Kortum, P., & Peres, S. C. (2014). The relationship between system effectiveness and subjective usability scores using the System Usability Scale. *International Journal of Human-Computer Interaction, 30*, 575–584.
- Kortum, P., & Sorber, M. (2015). Measuring the usability of mobile applications for phones and tablets. *International Journal of Human-Computer Interaction, 31*, 518–529.
- Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1–11.

- Lewis, J. R. (2014). Usability: Lessons learned . . . and yet to be learned. *International Journal of Human-Computer Interaction*, *30*, 663–684.
- Lewis, J. R., Brown, J., & Mayes, D. K. (2015). Psychometric evaluation of the EMO and the SUS in the context of a large-sample unmoderated usability study. *International Journal of Human-Computer Interaction*, *31*(8), 545–553.
- Lewis, J. R., & Sauro, J. (2009). The factor structure of the System Usability Scale. In Kurosu, M. (Ed.), *Human Centered Design, HCII 2009* (pp. 94–103). Heidelberg, Germany: Springer-Verlag.
- Lewis, J. R., Utesch, B. S., & Maher, D. E. (2013). UMUX-LITE – When there’s no time for the SUS. In *Proceedings of CHI 2013* (pp. 2099–2102). Paris, France: ACM.
- Lewis, J. R., Utesch, B. S., & Maher, D. E. (2015). Measuring perceived usability: The SUS, UMUX-LITE, and AltUsability. *International Journal of Human-Computer Interaction*, *31*, 496–505.
- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149.
- Nunnally, J.C. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.
- Patil, V. H., Singh, S. N., Mishra, S., & D. Donavan, T. (2007). *Parallel Analysis Engine to Aid Determining Number of Factors to Retain* [Computer software]. Available from <http://smishra.faculty.ku.edu/parallelengine.htm>.
- Peres, S. C., Pham, T., & Phillips, R. (2013). Validation of the System Usability Scale (SUS): SUS in the wild. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 192–196). Santa Monica, CA: HFES.
- Pilotte, W. J., & Gable, R. K. (1990). The impact of positive and negative item stems on the validity of a computer anxiety scale. *Educational and Psychological Measurement*, *50*, 603–610.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36.
- Sauro, J. (2011). *A practical guide to the System Usability Scale*. Denver, CO: Measuring Usability.
- Sauro, J., & Lewis, J. R. (2009). Correlations among prototypical usability metrics: Evidence for the construct of usability. In *Proceedings of CHI 2009* (pp. 1609–1618). Boston, MA: ACM.
- Sauro, J., & Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? In *Proceedings of CHI 2011* (pp. 2215–2223). Vancouver, Canada: ACM.
- Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research* (2nd ed.). Cambridge, MA: Morgan-Kaufmann.
- Schmitt, N., & Stuits, D. (1985) Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, *9*, 367–373.
- Schriesheim, C. A., & Hill, K. D. (1981). Controlling acquiescence response bias by item reversals: the effect on questionnaire validity. *Educational and Psychological Measurement*, *41*, 1101–1114.
- Stewart, T. J., & Frye, A. W. (2004). Investigating the use of negatively-phrased survey items in medical education settings: Common wisdom or common mistake? *Academic Medicine*, *79* (Suppl. 10), S1–S3.
- Tullis, T. S., & Stetson, J. N. (2004). *A comparison of questionnaires for assessing website usability*. Paper presented at the Usability Professionals Association Annual Conference, June. Minneapolis, MN, USA: UPA.
- Wong, N., Rindfleisch, A., & Burroughs, J. (2003). Do reverse-worded items confound measures in cross-cultural consumer research? The case of the material values scale. *Journal of Consumer Research*, *30*, 72–91.

About the Authors



James R. (Jim) Lewis, PhD

Dr. Lewis is a senior human factors engineer (at IBM since 1981). He has published influential papers in the areas of usability testing and measurement. His books include *Practical Speech User Interface Design* and (with Jeff Sauro) *Quantifying the User Experience* (now in its second edition).



Jeff Sauro, PhD

Dr. Sauro is a six-sigma trained statistical analyst and founding principal of MeasuringU, a customer experience research firm based in Denver. He has conducted usability tests and statistical analysis for companies such as Google, eBay, Walmart, Autodesk, Lenovo, and Dropbox, and has published over 20 peer-reviewed research articles and 5 books, including *Customer Analytics for Dummies*.