MeasuringU (https://measuringu.com/)

(https://measuringu.com/my-account/)

MEASURING USABILITY WITH THE SYSTEM USABILITY SCALE (SUS)

by Jeff Sauro, PhD | February 2, 2011

(http://www.flickr.com/photos/ajmexico/3281139507/sizes/m/in/photostrea is the 25th anniversary of the creation of the most used questionnaire for measuring perceptions of usability.

The System Usability Scale (SUS) was released into this world by John Brooke in 1986.

It was originally created as a "quick and dirty" scale for administering after usability tests on systems like VT100 Terminal ("Green-Screen") applications.

SUS is technology independent and has since been tested on hardware, consumer software, websites, cell-phones, IVRs and even the yellow-pages.

It has become an industry standard with references in over 600 publications.

The System Usability Scale

The SUS is a 10 item questionnaire with 5 response options.

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

The SUS uses the following response format:

Strongly Disagree 1	2	3	4	Strongly Agree 5
○	○	○	○	○

Scoring SUS

MeasuringU (https://measuringu.com/)
(https://measuringu.com/my-account/)

- For odd items: subtract one from the user response.
- For even-numbered items: subtract the user responses from 5
- This scales all values from 0 to 4 (with four being the most positive response).
- Add up the converted responses for each user and multiply that total by 2.5. This converts the range of possible values from 0 to 100 instead of from 0 to 40.

Interpreting SUS Scores

Despite the wide usage of SUS, there has been little guidance on interpreting SUS scores, acceptable modifications to the items and information on reliability and validity.

Over the years I've used SUS a lot in my own research and during usability evaluations. During this time I've reviewed the existing research on SUS and analyzed data from over 5000 users across 500 different evaluations.

This data shows that SUS is a reliable and valid measure of perceived usability. It performs as well or better than commercial questionnaires and home-grown internal questionnaires.

I've put these findings in a 150 page detailed report (http://www.measuringusability.com/products/SUSpack) which contains valuable insights on background, benchmarks and best practices for anyone using the SUS. Here are a few highlights.

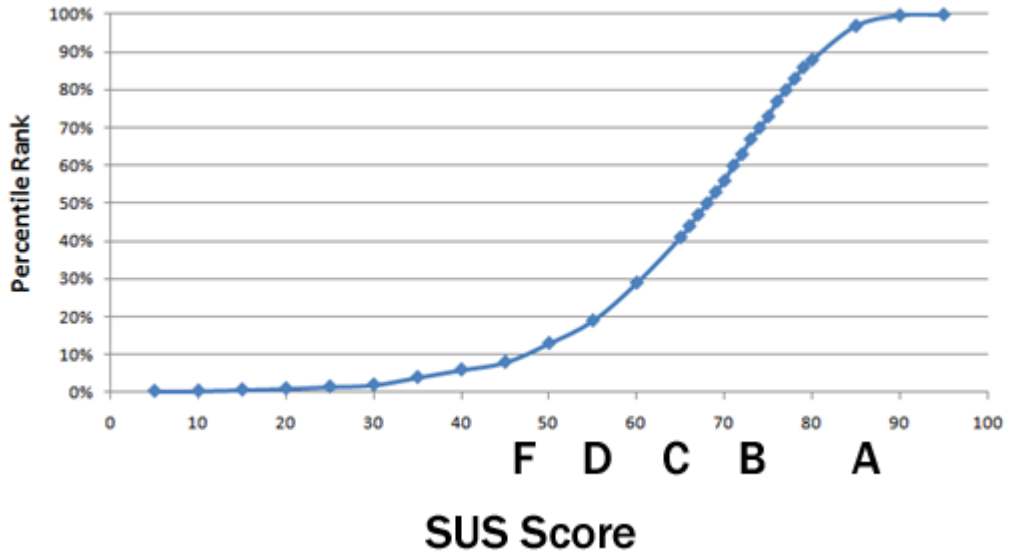
What is a Good SUS Score?

The average SUS score from all 500 studies is a 68. A SUS score above a 68 would be considered above average and anything below 68 is below average.

MeasuringU (https://measuringu.com/) (https://measuringu.com/my-account/)

The best way to interpret your score is to convert it to a percentile rank through a process called normalizing. I've created a calculator and guide (http://www.measuringusability.com/products/SUSpack) which takes raw SUS scores and generates percentile ranks and letter-grades (from A+ to F) for eight different application types.

The graph below shows how the percentile ranks associate with SUS scores and letter grades.



This process is similar to "grading on a curve (http://en.wikipedia.org/wiki/Grading_on_a_curve)" based on the distribution of all scores. For example, a raw SUS score of a 74 converts to a percentile rank of 70%. A SUS score of 74 has higher perceived usability than 70% of all products tested. It can be interpreted as a grade of a B-.

You'd need to score above an 80.3 to get an A (the top 10% of scores). This is also the point where users are more likely to be recommending the product to a friend (http://www.measuringusability.com/usability-loyalty.php). Scoring at the mean score of 68 gets you a C and anything below a 51 is an F (putting you in the bottom 15%).

SUS Scores are not Percentages

Even though a SUS score can range from 0 to 100, it isn't a percentage. While it is technically correct that a SUS score of 70 out of 100 represents 70% of the possible maximum score, it suggests the score is at the 70th percentile. A score at this level would mean the application tested is above average. In fact, a score of 70 is closer to the average SUS score of 68. It is actually more appropriate to call it 50%.

When communicating SUS scores to stakeholders, and especially those who are unfamiliar with SUS, it's best to convert the original SUS score into a percentile so a 70% really means above average.

SUS Measures Usability & Learnability

While SUS was only intended to measure perceived ease-of-use (a single dimension), recent research (https://measuringu.com/wp-content/uploads/2017/07/Lewis_Sauro_HCII2009.pdf)[pdf] shows that it provides a global measure of system satisfaction and sub-scales of usability and learnability. Items 4 and 10 provide the learnability dimension and the other 8 items provide the usability dimension. This means you can track and report on both subscales and the global SUS score.

SUS is Reliable

Reliability refers to how consistently users respond to the items (the repeatability of the responses). SUS has been shown to be more reliable and detect differences at smaller sample sizes than home-grown questionnaires and other commercially available ones.

Sample size and reliability are unrelated, so SUS can be used on very small sample sizes (as few as two users) and still generate reliable results. However, small sample sizes generate imprecise estimates of the

unknown user-population SUS score. You should compute a confidence interval around your sample SUS score to understand the variability in your estimate.

SUS is Valid

Validity refers to how well something can measure what it is intended to measure. In this case that's perceived usability. SUS has been shown to effectively distinguish between unusable and usable systems as well as or better than proprietary questionnaires. SUS also correlates highly with other questionnaire-based measurements of usability (called concurrent validity).

SUS is not Diagnostic

SUS was not intended to diagnose usability problems. In its original use, SUS was administered after a usability test where all user-sessions were recorded on videotape (VHS and Betamax). Low SUS scores indicated to the researchers that they needed to review the tape and identify problems encountered with the interface. SUS can be used outside of a usability test for benchmarking, however, the results won't shed much light on why users are responding the way they are.

Modest Correlation between SUS and Task-Performance

Users may encounter problems (even severe problems) with an application and provide SUS scores which seem high. Post-test SUS scores do correlate with task performance, although the correlation is modest (around r= .24 for completion rates and time), which means that only around 6% of the SUS scores are explained by what happens in the

usability test. This is the same level of correlation found (http://www.measuringusability.com/papers/Sauro_Lewis_CHI2009.pdf) [pdf] with other post-test questionnaires.

https://measuringu.com/ (https://measuringu.com/my-account/)

Quick and Not So Dirty

At only 10 items, SUS may be quick to administer and score, but data from over 5000 users and almost 500 different studies suggests that SUS is far from dirty. Its versatility, brevity and wide-usage means that despite inevitable changes in technology, we can probably count on SUS being around for at least another 25 years.

To help you in your next study with SUS or to interpret your existing SUS data I've assembled a comprehensive guide (http://www.measuringusability.com/products/SUSpack) on how to use benchmarks, compare SUS scores and find the right sample size for your study.

SEARCH BLOG

Search here...

Most Popular Most Recent

User Experience Salaries & Calculator (2018) (https://measuringu.com/salary-survey2018/)

How To Make Personas More Scientific (https://measuringu.com/scientific-personas/)

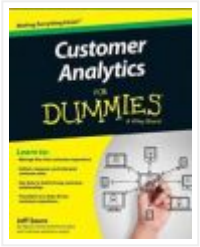
Scientific Thinking For Better Design (https://measuringu.com/scientific-thinking/)

9 Biases That Affect Survey Responses (https://measuringu.com/survey-biases/)
4 Types of Observational Research (https://measuringu.com/observation-role/)

(https://measuringu.com/) (https://measuringu.com/my-account/)

BOOKS

(https://measuringu.com/book/benchmarking-the-user-experience/)
Benchmarking the User Experience
(https://measuringu.com/book/benchmarking-the-user-experience/)

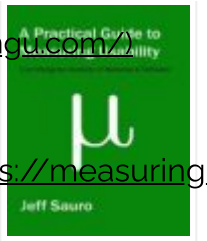


(https://measuringu.com/book/customer-analytics-for-dummies/)
Customer Analytics For Dummies (https://measuringu.com/book/customer-analytics-for-dummies/)

(https://measuringu.com/book/quantifying-the-user-experiencepractical-statistics-for-user-research/)
Quantifying The User Experience: Practical Statistics For User Research
(https://measuringu.com/book/quantifying-the-user-experiencepractical-statistics-for-user-research/)

MeasuringU (https://measuringu.com/)

(https://measuringu.com/my-account/)



(https://measuringu.com/book/a-practical-guide-to-measuring-usability/)
A Practical Guide To Measuring Usability (https://measuringu.com/book/a-practical-guide-to-measuring-usability/)

(https://measuringu.com/book/excel-r-companion-to-quantifying-the-user-experience/)
Excel & R Companion to the 2nd Edition of Quantifying the User Experience (https://measuringu.com/book/excel-r-companion-to-quantifying-the-user-experience/)

3300 E 1st Ave. Suite 370
Denver, Colorado 80206
1 + 303-578-2801 - MST
Contact Us (http://measuringu.com/contact/)
(http://measuringu.com/contact/)

(http://measuringu.com/contact/)
(https://measuringu.com/)
(http://measuringu.com/contact/)

(https://measuringu.com/my-account/)

(http://measuringu.com/contact/)
(http://measuringu.com/contact/)
(http://measuringu.com/contact/)

BLOG

(http://measuringu.com/contact/)
(http://measuringu.com/contact/)
(http://measuringu.com/contact/)

Five Reasons to Use Open-Ended Questions (https://measuringu.com/open-ended-reasons/)

Confidence Intervals for Net Promoter Scores (https://measuringu.com/nps-confidence-intervals/)

48 UX Metrics, Methods, & Measurement Articles from 2020 (https://measuringu.com/ux-2020/)

From Functionality to Features: Making the UMUX-Lite Even Simpler (https://measuringu.com/simpler-umux-lite/)

What a Randomization Test Is and How to Run One in R (https://measuringu.com/randomization-test/)

NEWSLETTER

Sign-up to receive weekly updates.

Your Email*

SUBSCRIBE

MeasuringU

(https://measuringu.com/)

© 2004-2020 MeasuringU

(https://measuringu.com/my-account/)

About (https://measuringu.com/about/) ▪ Blog (https://measuringu.com/blogs/) ▪ Services (https://measuringu.com/services/)

(https://www.facebook.com/