Using the Population Specific User Mastery (PSUM) Scale to determine training needs

B. D. Sawyer[1,2], P. Oppold[1], K. MacArthur, P. A. Hancock[1]

[1]Psychology Department, University of Central Florida, Orlando, USA

[2]Oak Ridge Institute for Science and Education, Wright-Patterson Air Force Base, USA

Abstract:

The MIT2 lab at The University of Central Florida has been testing an in-house tool for quickly and efficiently teasing apart these differences. The Population Specific User Mastery (PSUM) Scale (Sawyer et al., *in press*) was created in order to contrast tools within a given population. Given a tool, a series of tasks within that tool, and expected completion times for each task, PSUM scales provides a snapshot of the measured population's ability, or alternately contrasts multiple populations' abilities.

Summary:

In Human Factors research, human interaction with similar tools is often compared. For example, a researcher might evaluate the distraction potential of two in-vehicle communication devices, or the workload demands of two displays. In order to pilot successfully, researchers need a way to quickly evaluate subtasks of interest within the device use scenario they are studying. The PSUM (Population Specific User Mastery) scale seeks to provide just such a tool (Sawyer et al., *in press;* Greenstein et al, 2013; MacArthur, Greenstein, Sawyer & Hancock, 2013, 2014). A constructed PSUM scale provides a meaningful interval scale of population success at a task within a device, highlighting training needs.

In experiment 1, sixteen undergraduate students at the University of Central Florida aged 18-20 were given class credit for participation. An iPhone and Android smartphone were compared across 10 different tasks, including turning the device on and off, using the phone's navigation system to find a given address and sending a SMS message. Participants completed tasks were timed on a stopwatch by the investigator, and after were asked to complete a demographics questionnaire.

When considering the criterion times for each task, stark differences emerge. Highlights follow: In the allotted 40 seconds for the navigation task, only 12.5% of Android users successfully completed the task, compared to 93.7% of iPhone users. The SMS messaging tasks was allotted 45 seconds, as was completed by 31.3% of Android users and 68.8% of iPhone users. The Facebook login process was allotted 30 seconds, and was completed by 50% of Android users and 87.5% of Android users. Raw time results indicated that participants took longer to complete the navigation task on the Android ($t(11) = -2.05$, $p = 0.05$; $M = 83.81$s; $SD = 85.11$s) than on the iPhone ($M = 28.50$s; $SD = 10.46$s). SMS messaging took longer on the Android ($t(11) = -2.11$, $p = 0.05$; $M = 48.01$s; $SD = 21.64$s) than on the iPhone ($M = 32.46$s; $SD = 10.82$s). Also, logging in to the Facebook application took longer on the Android ($t(11) = -2.66$, $p = 0.02$; $M = 67.87$s; $SD = 47.13$s) than on the iPhone ($M = 35.28$s; $SD = 15.74$s). All above comparisons showed significant differences in time.

These differences showed this young, smartphone savvy population overall more comfortable and efficient using the iPhone. Using the PSUM Scale helped MIT2's researchers choose iPhones for an upcoming experiment, as less training was needed to make participants capable of using the phone for the required tasks in a timely manner.

In this paper two examples of such an inquiry will be reported. In the first, above, two smartphone OSes are compared. There is not space in this abbreviated summary to report the results of Experiment 2, in which an Android™ smartphone is compared to Google Glass™ after 5, 10 and 15 minutes of training (as mentioned in MacArthur, Greenstein, Sawyer & Hancock, 2013 and Sawyer, Finomore, Calvo and

Hancock, 2014). This effort will be covered in detail in the final presentation. Methods for comparing devices, comparing populations, and integrating PSUM into existing research protocols will be discussed.

References:

Greenstein, S., Sawyer, B. D., Niederman, E., Oppold, P., & Hancock, P. A. (2013). Piloting with the PSUM scale: Establishing usability first. *Presented at International Academy, Research, and Industry Association: Vehicular 2013*. Nice, France.

MacArthur, K. R., Greenstein, S., Sawyer, B. D., & Hancock, P. A. (2014). PSUM: Training in Google Glass and Android. *Presented at the American Psychological Association 2014 annual convention*. Washington D.C.

MacArthur, K. R., Greenstein, S., Sawyer, B. D., & Hancock, P. A. (2014). Mastering Google Glass. *Presented at the Human Factors & Applied Psychology Student Conference at Embry-Riddle Aeronautical University*. Daytona Beach, FL.

Sawyer, B. D., Finomore, V. S., Calvo, A. A., & Hancock, P. A. (2014). Google Glass A Driver Distraction Cause or Cure? *Human Factors*, *56*(7), 1307–1321.

Sawyer, B.D., Oppold, P., & Hancock, P.A. (*In Review*). PSUM: A unified metric linking designers and trainers to user goals. Submitted.

Tables and Figures: