**EMCS2430: Human Factors: People and Software**

Assignment: Evaluate a Study

Brian Russel Davis, [brian\_davis@brown.edu](mailto:brian_davis@brown.edu)

*Based on the study report (above), identify at least 5 flaws that jeopardize the validity of the conclusions. For each flaw identified, (1) concretely and succinctly explain why it is a problem, and (2) describe how the flaw could be prevented through different methodological decisions.*

I must admit that I spent most of the time squinting my eyes as I read through this brief, the number of holes in the approach of GraphicPass’ approach is pretty amazing. I know “graphical passwords” seem to be popular on phones ( a very personal device that is most likely always in our possession ), but I very little confidence that this approach translates well to a corporate desktop computer or a laptop. Furthermore, the approach used to analyze the results seem pretty narrow. Was the team really trying to measure the amount of security that was produced or just the likeability of the product? The method of evaluation and the evaluation criteria seemed to be focused on whether the users liked the product ( and could remember their password), not really on whether the product made the organization more secure. Even with the most liberal of assumptions, I don’t find anything about the product redeeming and I could think of stacks of security violations. Here are the most glaring examples.

**Visual passwords generated by users probably have a very small range of variation** and are easy to guess. It’s a well-known fact that people are not very creative when asked to make a password, hence the fact many people use “password” as a password. So what makes the researchers think that visual passwords will be any different? What’s worse is that the evaluation doesn’t even attempt to measure this. To fix the application should teach the user how to create and remember a complex pattern that is not easily guessed. The application should also not allow users to pick obviously easy patterns.

**The fact that the focus group participants are all engineers limits our ability to understand how people with different job functions interact differently** with the product. One would assume that engineers are slightly more aware of security risks which will skew the results. With a little effort, I am sure they can find more people to participate then the other engineering team. Opening the study to the entire company seems like a better idea, so the product can get a true cross-section of all types of users. There is more security when we embrace diversity.

**The fact that researches are watching the participants during the testing ( looking over their shoulders ) doesn’t make the test seem very objective or reflective of real usage**. The fact is that no-one should be looking at someone swipe their password. The fact that the password can be easily seen from far away is actually a security risk, not to mention the fact that the swipe marks can be easily seen on whatever screen the user is swiping on. ( this is the part where I shake my head ) There are many other ways to track users other than physically watching them. Screen recording software, cameras, one-way mirrors … anything but standing there and watching them.

**The failure to test them on their use of traditional passwords removes another important baseline and demographics.** Understanding how this product performs with people who understand how to create a strong password and those that don’t is important. If a person who doesn’t know how to make s strong text password is able to perform better and create a complex visual pattern, then the product works. To fix this all we need to do it ask users to create a new text password, then ask them to create a “Graphic Password” and compare the strength.

**Last but not least, where is the pen-testing?**  The evaluation seems like security theater without a real penetration test. How secure is Graphic Pass? Well, let's see! After the user creates their password, a hacker should be allowed to come in right behind and try and crack the system. Without this type of testing the evaluation is not a security evaluation it's just a measurement of how much people like the product, which as I said, in the beginning, doesn’t seem to matter as much as whether the product increases security.