**Background**

- Nielsen's mathematical model suggesting the use of 5 participants to find eighty-five percent of critical errors when conducting usability tests is a staple within usability testing (1993).

- Faulkner (2003) suggests that twenty participants would be more accurate in reducing the possibility of missing critical errors.

- However, Nielsen (1993), Faulkner (2003), and others only assessed software usability. Concerns exist whether such methodology is also acceptable in hardware assessments where a larger variety of interactions exist.

**Hypothesis**

- The purpose of this work was to investigate the amount of participants required to consistently discover 85% of the errors within a hardware usability study.

- Additionally, this work assessed whether technology specific familiarity has a reductive effect on required sample size in hardware usability studies.

**Methods**

**Participants**

- 20 Participants from the University of Houston-Clearlake

**Materials**

- Wireless mouse & keyboard (Familiar)
- All-in-one desktop
- Manual Typewriter (Unfamiliar)
- Observer Media recorder

**Procedure**

- **Task one**
  Talk Aloud procedure where participants detailed issues with a wireless mouse and keyboard when pairing to a computer.

- **Task Two**
  Talk Aloud procedure where participants detailed issues while replicating a paragraph on a manual typewriter.

**Analysis**

- Usability groups created from sample sizes of 5, 10, 15
- Sampling for each group was repeated 25 times to create an effectiveness distribution.
- An additional analysis was completed examining the distribution as determinant by criticality of errors.

**Results**

- **5 participants**
  - Keyboard/Mouse
  - Typewriter

- **10 participants**
  - Keyboard/Mouse
  - Typewriter

- **15 participants**
  - Keyboard/Mouse
  - Typewriter

**Discussion**

- While the data supported the use of five participants to consistently identify 85% of the errors when assessing more familiar technology, less familiar technology required a larger sample size, approaching fifteen.

- However, although a clear positive shift is evident when increasing sample sizes in unfamiliar hardware assessments, no significant benefit was found when increasing sample sizes for familiar hardware assessments.

- In familiar technology hardware usability assessments, sample sizes of 5 demonstrated very high capture rates of errors with high criticality. However, because of the limited number of high and moderate-high severities a severe reductions in range was demonstrated.

**Future Work**

- Future research will investigate whether criticality of errors shows a similar sampling pattern.

- This experimental technique will also be used to assess this usability practice in software.