Beyond Magnitude: Dissecting Causes of User Errors for Better Design Recommendations

Michael J. Crites[1] , Noah J. Wheeler[2], Trevor J. McIntyre[3], and Samuel J. Levulis[4]
[1]Resideo Technologies, Golden Valley, MN
[2]Simple Interact, Dallas, TX
[3]Texas Tech University, Lubbock, TX
[4]Microsoft Corporation, Redmond, WA

Introduction

User researchers identify and classify errors during usability tests in order to improve user interface design. Traditionally, the methods employed by researchers focus on the magnitude of errors. For instance, a user researcher might use dimensions such as the number of users impacted as well as the impact of the error on the user's ability to complete the task. However, they might not classify errors according to the underlying cause of the error.

Traditional Methods to Identify Causes of Error

Think-aloud Protocol (Boren & Ramey, 2000)
Widely used method to understand users' motivations for action during testing. Using this method, practitioners have users speak their thoughts during testing.
Issues with Method: Nørgaard & Hornbæk (2006) found that this method may be biased towards problems the evaluator already knows about, which subjectively leads to further exploration of problems they have foreseen or have deemed more severe.

Root Cause Analysis (as detailed by Norman, 2013)
Norman (2013, p. 42) simplifies "root cause analysis" as "asking 'why' until the fundamental cause of the activity is reached."
Issues with Method: This activity is likely to reveal what caused a user to commit an error, but it still may not be prescriptive enough to reveal design activities that would address the error

Affinity Diagramming (Beyer & Holtzblatt, 1998)
Clustering errors and issues together that are reasoned to be similar in nature, creating a hierarchical structure of overall issues.
Issues with Method: This method attempts to achieve a higher-level understanding of the cause of errors, but its subjectivity likely leads to inconsistencies, as Molich et al. (2004).

Output of an Affinity Diagramming activity

Reason's (1990) GEMS Model

Reason (1990) developed a classification system of human error that looked at the causes of errors in reference to the user. We believe this approach can be applied to the classification of user errors.

Error Type	Description	Example
Skill-Based Slip	Incorrect execution of a planned action	Missed key during typing
Skill-Based Lapse	Incorrect omission of a stored, planned action	Going to the kitchen to make tea; instead making coffee
Rule-Based Mistake	Applying an incorrect "if-then" rule	A toddler, with iPad experience, swipes to move the page forward in a paper book
Knowledge-Based Mistake	Incorrect reasoning due to a lack of stored rules for how to deal with the novel situation	Patient presents with collection of symptoms the doctor has seen; diagnosis based on more familiar subset of symptoms

A New Categorization Method

Using Reason's (1990) framework, We propose a new method that categorizes usability errors as "slips" or "mistakes." If the cause of the error appears to be due to an accident in space-time, then this can be categorized as a slip. If the cause of the error appears to be cognitive, then this can be categorized as a mistake.

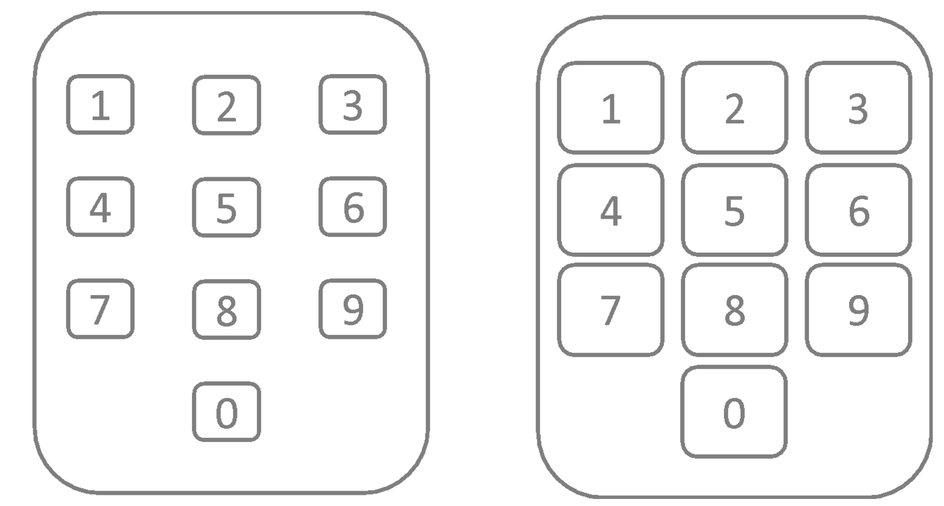
Slips are an incorrect action

Mistakes are an incorrect intention

Practical Example

A user incorrectly types their PIN as 1-2-3-4 instead of the correct 1-2-3-5. First, this might be due to misremembering the number. This would be a mistake. Second, a user may intend to type the correct PIN, but accidentally selected the last digit as 4 instead of 5 (a slip). If it is determined to be a slip, then the researchers might recommend the left keypad instead of some way to improve the user's mental model and keeping the keypad design on the right:

If the usability error was contributed to a slip, then the physical design should be improved



Categorizing errors this way helps researchers provide directed design recommendations. Specifically, the design might be improved by helping users remember their code to prevent a mistake or having input keys with larger ore more spaced out touch targets to prevent a slip.

Conclusion

- This method for categorizing user errors can help researchers provide design recommendations that address the underlying cause of errors.
- The authors have found this method of categorization to aid in providing design recommendations to specific teams and generating more buy-in from stakeholders such as engineering.

References

- Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. IEEE Transactions on Professional Communication, 43, 261-278.
- Hornbaek, K., & Stage, J. (2006). The interplay between usability evaluation and user interaction design. International Journal of Human-Computer Interaction, 21, 117-123.
- Molich, R., Ede, M. R., Kaasgaard, K., & Karyukin, B. (2004). Comparative usability evaluation. Behaviour & Information Technology, 23, 65-74.
- Nielsen, J. (1993). Usability Engineering. San Diego, CA: Morgan and Kaufman.
- Norman, D. A. (2013). The design of everyday things: Revised and expanded edition. New York, NY: Basic Books.
- Reason, J. (1990). Human error. Cambridge university press.