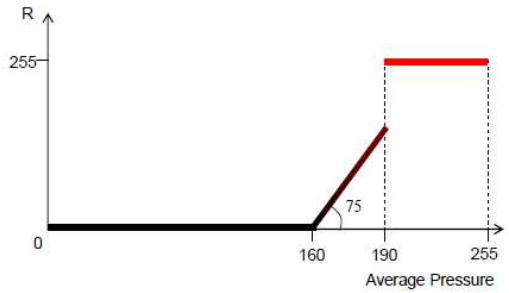
Experimental Analysis of Mode Switching Techniques, Yang Li et al.

- Premise: Mode switching is an important and common task in pen-tablet interfaces
  - Based on the need to overload pen behavior
  - So evaluate different techniques to see which one is better in speed and errors
- Five mode switching techniques
  - Button in toolbar, press-and-hold, non-preferred hand, pressure, flipping pen

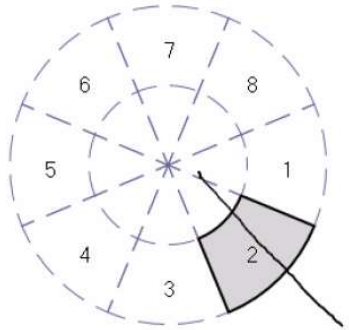
Methodology

- First a pilot
  - Tests pressure levels
- Experimental task
  - Pie cutting task
    - Baseline and compound (control and experimental)



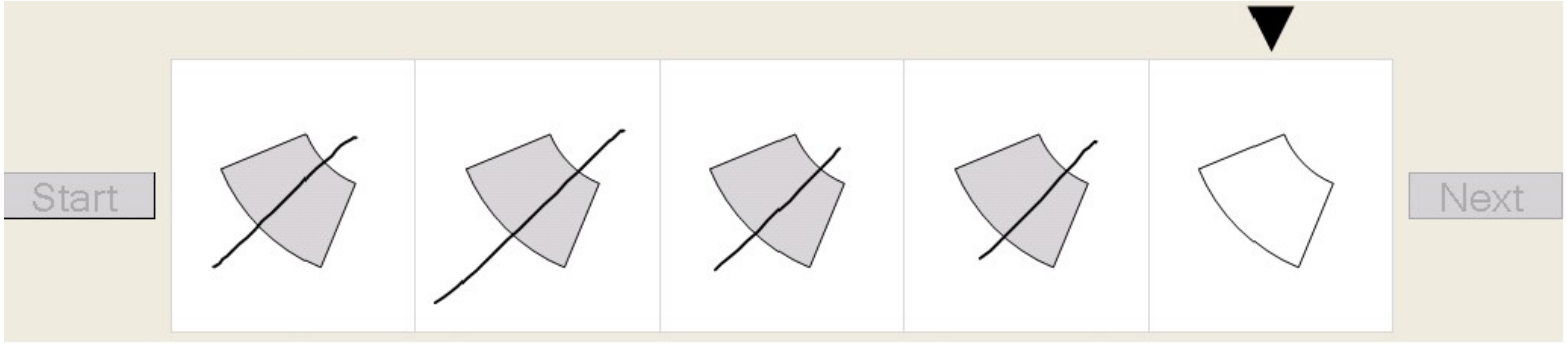
To capture the nature of sketch-based interactions, which are normally informal and fluid, we designed a pie-crossing task as an abstraction of the action of gesturing and inking (see Figure 2). A pie slice is shown with one of eight orientations corresponding to the eight major geographical directions. A participant was required to quickly cross a slice from its inner edge towards its outer edge according to a target's orientation. This design examines the drawing of various directions without requiring precise positioning and careful alignment by participants. This design also captures a realistic use scenario of gestures, i.e., marking menus [8], where users can cross a series of objects with marks to perform different commands.

Task

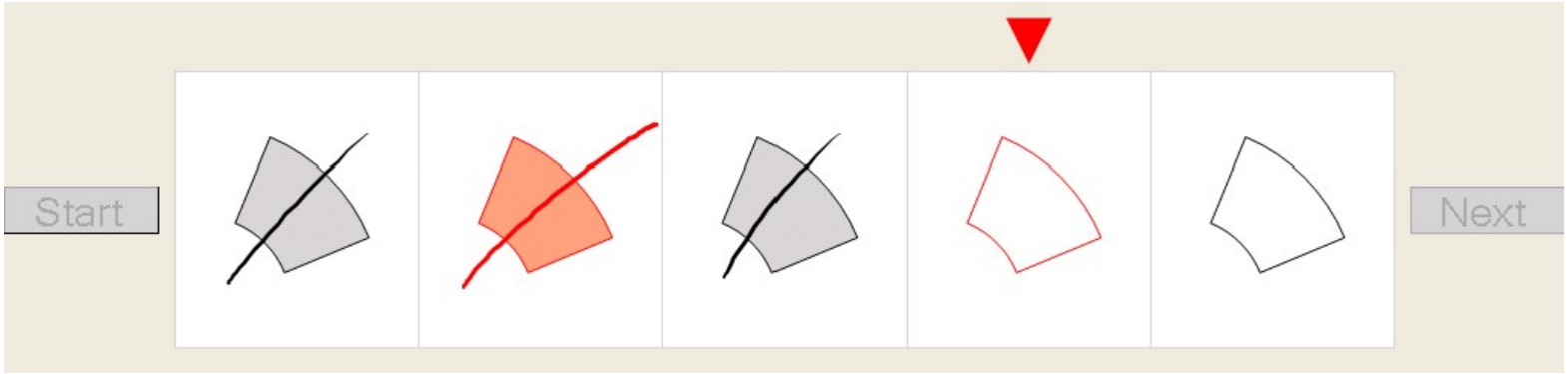


- Pie cutting

Control



Experimental



Errors

- One set of measures is accuracy of mode switching
- What are errors?
- Mode errors
  - Mode-in and Mode-out
- Crossing errors
- Out of target errors

Procedure

- Training phase + test phase
  - In this case training data ignored
- 5X5 Latin square counter-balanced techniques
- 9 Blocks
  - First baseline/control, then experimental, alternating for 5 control, 4 experimental

The experiment included a training phase for the baseline tasks, five experimental sessions with one session for each technique, and a post-study questionnaire. The experiment took about 80 minutes in total. A 5x5 Latin Square was used to counterbalance the order of the techniques. Each session was divided into two parts. The first part involved learning to use a mode switching technique and extensive practice. The second part was the experimental phase in which a participant was given 9 blocks of trials. The first block was a *baseline task* and then a *compound task*, alternating until the ninth block ended with a *baseline task*. A participant could take a break between blocks. In total, the experiment consisted of:
15 participants x
5 mode switching techniques x
9 block of trials x
8 screens (8 orientations) x
5 pie-crossing tasks
= 27,000 pie-crossing tasks

Counter-balanced: Latin Square Design

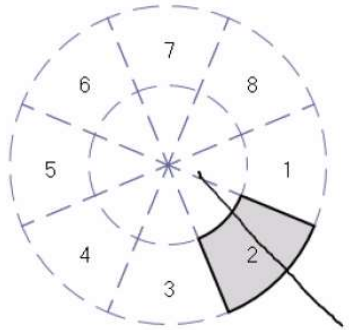
A	*B*	*C*	*D*	*E*
B	*A*	*D*	*E*	*C*
C	*E*	*A*	*B*	*D*
D	*C*	*E*	*A*	*B*
E	*D*	*B*	*C*	*A*

Measures

- Dependent variables are time, errors, preference
- In 9 experimental blocks
  - Two used as warm up, seven analysed
- Duration divided into 3 cycles
  - Break after first, third, last pie
  - Last two cycles have a mode-switch
- Mode switch time = average cycle duration for last two cycles with mode-switch in compound – average cycle duration for last two cycles in control (see slide 4)

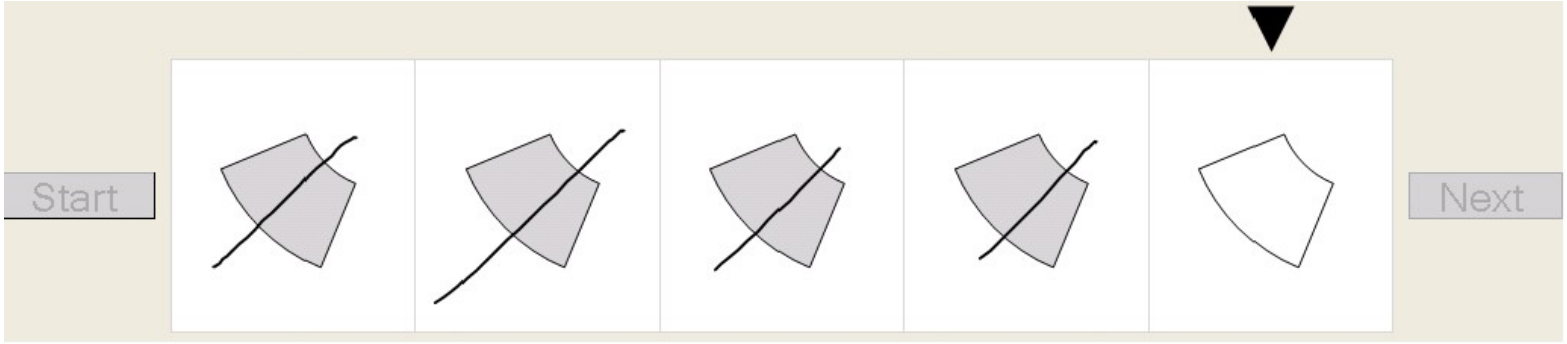
The dependent variables were the *mode switching time*, the total number of errors in a compound task, and the subjective preference of participants. The first two blocks in the experimental phase were for warming up and the data of the seven following blocks were used for analysis.
The timing for each screen is started when the Start button is clicked and automatically ended when the last pie is crossed and the pen is lifted. This duration is divided into three cycles. The first cycle starts when the Start button is clicked and ends when the first pie is crossed. The second cycle starts right after the first cycle and ends after the third pie is crossed. This is followed by the third cycle, which includes crossing the last two slices. Therefore, one target needs to be crossed in the first cycle and two targets need to be crossed in each of the second and the third cycles. We call cycle 2 and 3 *full cycles* and cycle one the *start cycle*.
In a *compound task*, a full cycle contains a complete mode switch process including switching into gesture mode and switching back to ink mode.
The *mode switching time* for each of the three compound blocks was computed by subtracting the mean of the two adjacent baseline tasks' *average cycle durations* from the compound block's *average cycle duration*. Average cycle duration was the mean duration of all correct full cycles in a block.

Task

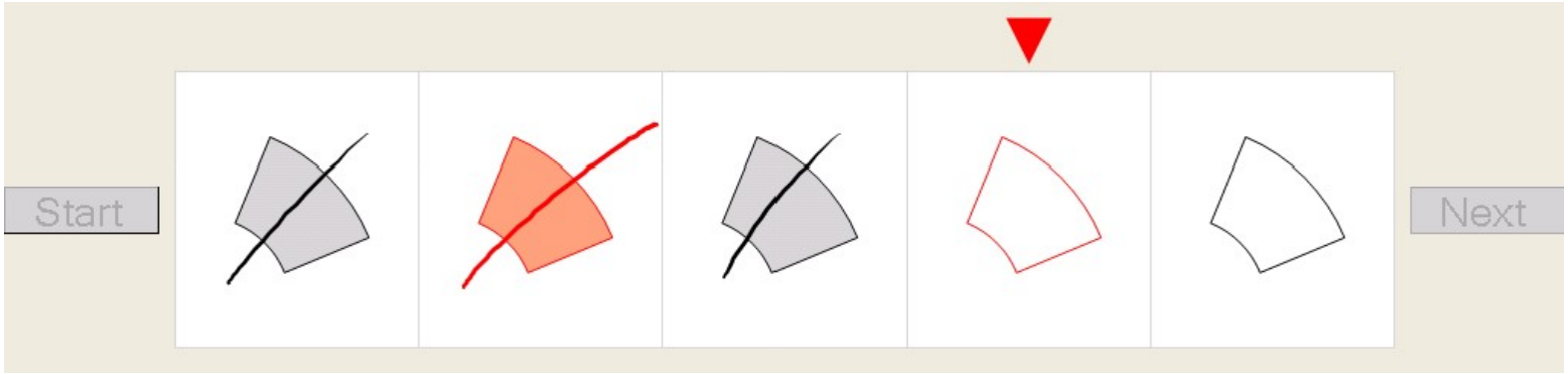


- Pie cutting

Control



Experimental



Results

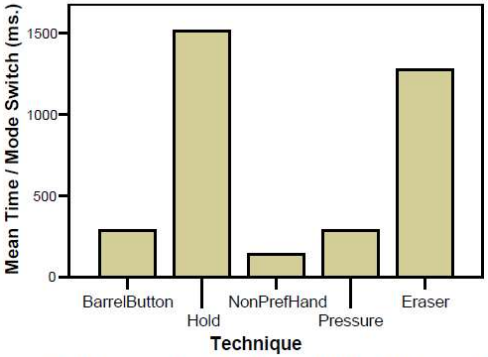


Figure 5. The mean mode switching time of five techniques.

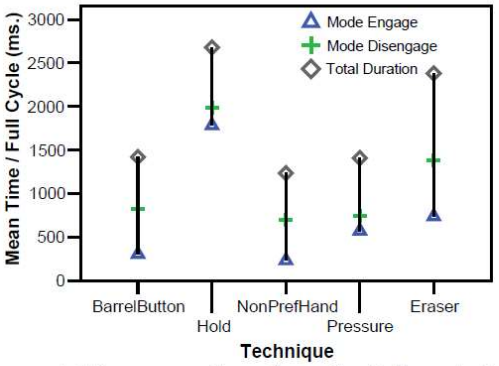


Figure 6. The mean duration of a full cycle in a compound task (before subtraction). The time of mode engagement and disengagement are also shown.

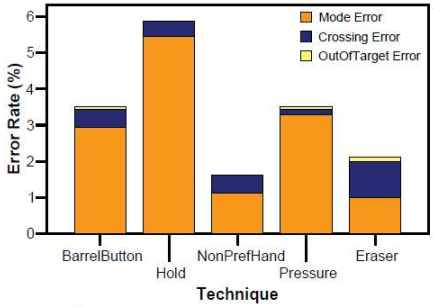


Figure 7. The error rate on each pie-crossing. The mode errors included both Mode-In and Mode-Out errors.

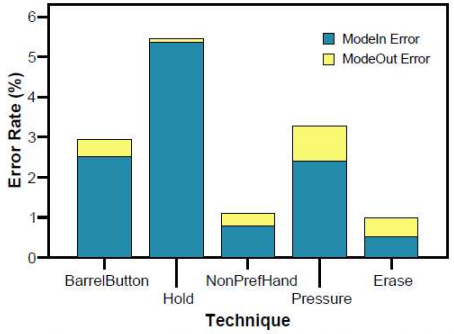


Figure 8. The occurrence rate of Mode-In and Mode-Out errors on each pie-crossing, which is weighted by the frequency of inking and gesturing, respectively. 40% of pie-crossings were for gesturing, while 60% were for inking.

Table 1. The participants' preferences for each technique.

Dimension	Barrel-Button	Hold	NonPref-Hand	Pressure	Eraser
Learning	4.4	3.5	4.7	3.5	4.2
Use	3.7	2.2	4.1	3.4	2.4
Accuracy	3.7	2.9	4.6	3.3	3.6
Speed	4	1.7	4.5	4	1.9
Eye fatigue	4.1	3.3	4.4	3.9	4.2
Hand fatigue	3.5	3.3	4.1	3.3	2.1

Opinions

- Most common goods:
  - Thorough strategy to capture interaction cycle and cost
  - Good figures to describe techniques, good study description.
- Other good points
  - Good use of training sessions
  - Good options for mode switching
  - Elicitation of preference
- Most common bads:
  - University students
  - Ecological validity
  - Number of participants
- Others
  - Male-female ratio
  - Criteria for five techniques
  - Preference (vs something like NASA TLX?)

Bads:

- University students
  - Performance, errors, and age
- Ecological validity
  - Remember questions:
    - Is there a difference
    - Is the difference large or small
    - **Is the difference statistically significant**
    - Does the difference matter
- Number of participants
  - How many would be good?
  - See http://daniellakens.blogspot.com/2016/11/why-within-subject-designs-require-less.html
  - https://statisticsbyjim.com/basics/central-limit-theorem/
  - Normal corrections (Log-normal transforms, Box-Cox transform, aligned rank transforms)
    - http://faculty.washington.edu/wobbrock/pubs/chi-11.06.pdf