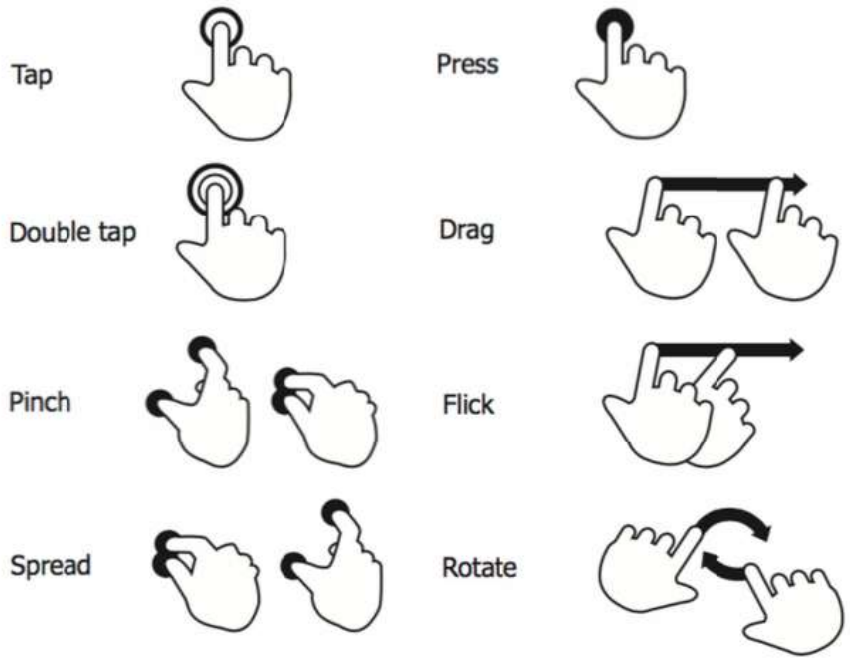
Zero Contact Research

Survey on Expert-Level Gesture Use and Adoption on Multi-touch Tablets

- Small component of Jeff's Ph.D. thesis
  - Point study
- Defines a set of basic gestures and notes existence of enhanced gestures
  - Are enhanced gestures used?



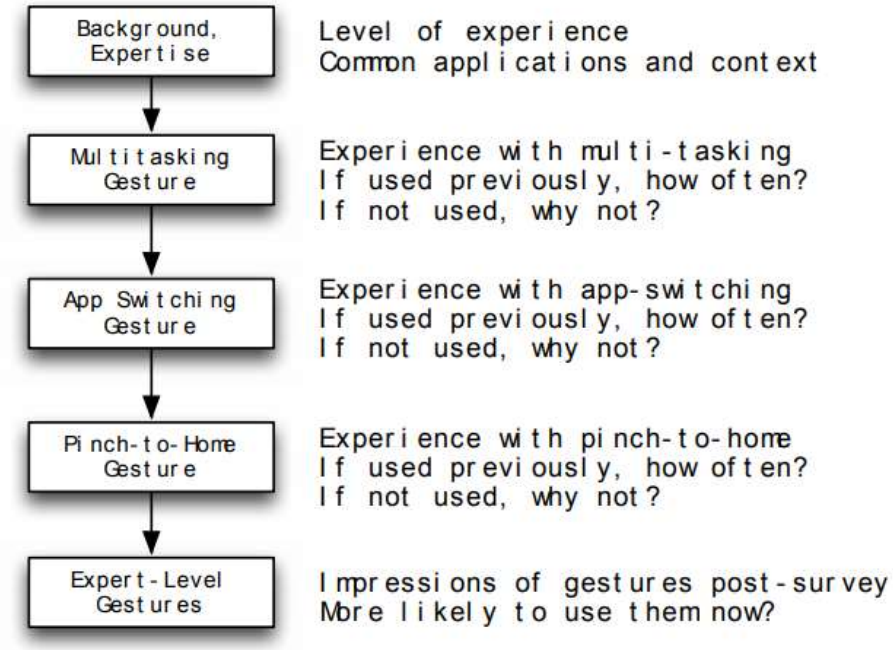
Method: On-line Survey

- Defensive writing
  - Any method of data collection has advantages and disadvantages
  - Surveys?

We chose an online survey as a data collection method for the same reasons proposed by Kjeldskov et al. [5]: external validity. Our goal is to collect responses related to software features used "in-the-wild" during day-to-day interactions, which is difficult to accomplish in a controlled lab environment. A similar approach was taken by Snowdon et al. [13] when assessing a mobile map navigation technique.

Method: On-line Survey

- Defensive writing
  - Any method of data collection has advantages and disadvantages
  - Surveys?
- Design
  - Tree structure (yes/no options with follow-up including teaching gesture.



Results: Descriptive Statistics

- Different categorical variables (i.e. counts)
- Results are descriptive statistics plus correlations across demographic factors
  - Numerical -> numerical correlation = correlation coefficient (r)
  - Categorical -> categorical, i.e. tabular data = chi-square statistic

Our users are relatively experienced: 60.8% have owned an iPad for at least 2-3 years (62/102), while 15.7% have owned it for a "longer period" (16/102). Daily usage was high: 72.6% use it for at least 30 minutes (77/106), and 41.5% use it for more than 60 minutes each day (44/106). As expected, most users use their iPads for media consumption and social media (Figure 3). Although there was no correlation between age and experience-level ($X2 = 17.1$, NS), there were differences in usage patterns: for instance, younger participants watched more videos ($X2 = 15.9$, $p < 0.01$).

Amazon Mechanical Turk Studies

- Implement studies (e.g. in Javascript, html/CSS, etc.)
- Deploy on mechanical turk
- Pay workers and get your data
- Benefits
  - Large, heterogeneous group of users.
  - If you pay well, lots of data fast.
- Risks
  - Data quality

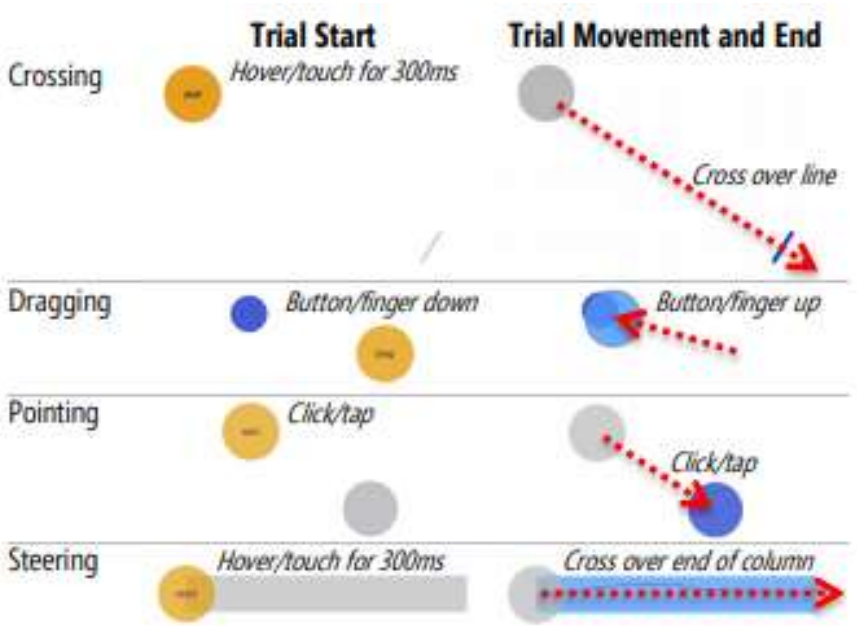
Turkers and Quality Control

Silberman et al. [30] recently noted that demographics have shifted in the past five years and that “professional Turkers” now complete most tasks in the system and have a stronger incentive than other workers to seek out high paying tasks and perform them well

[A] number of quality control mechanisms have become popular, such as redundancy, reputation systems, ground truth seeding, statistical filtering, and expert review. Providing feedback through “shepherding” can also lead to higher quality work. At a high level, these approaches can be grouped into up-front task design approaches versus posthoc result analysis approaches.

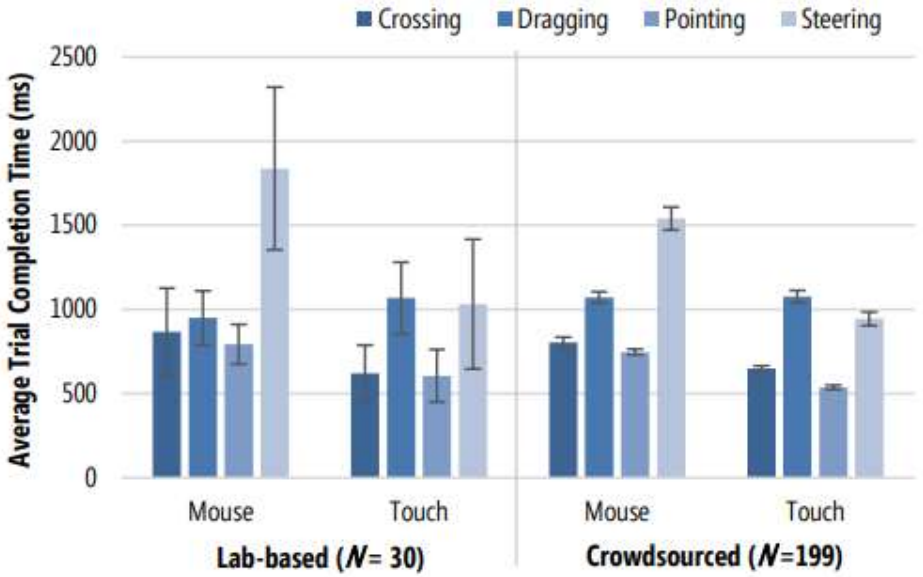
Crowdsourced vs. Lab-based Performance Data

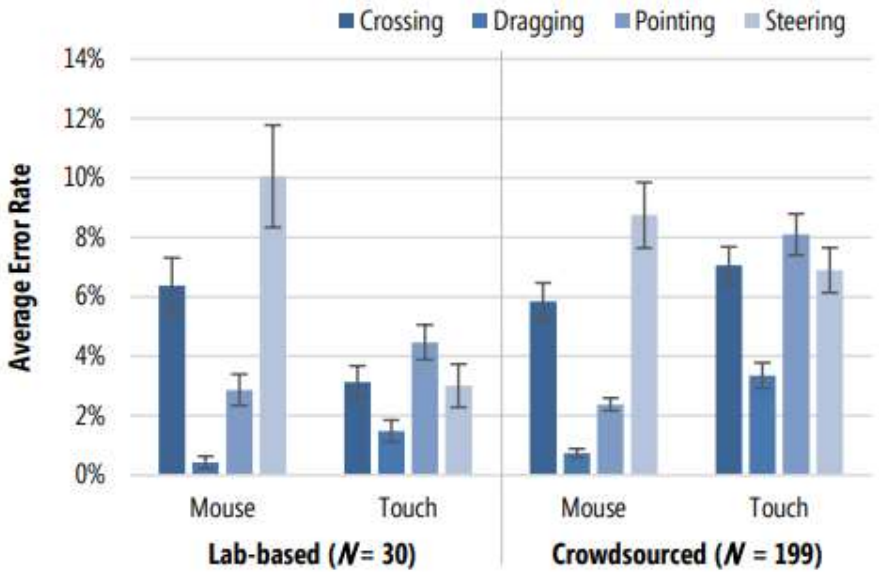
- Some work existed on desktop (mouse) performance data
- Wanted to compare with touch-based performance data
- Tasks described to right



Experiment

- N = 30 in lab
- N = 202 mechanical turkers
  - Half each for mouse/touch
- Results show lower accuracy for touch, not much difference for mouse





Experiment

- N = 30 in lab
- N = 202 mechanical turkers
  - Half each for mouse/touch
- Results show lower accuracy for touch, not much difference for mouse

Age-Matched Participants

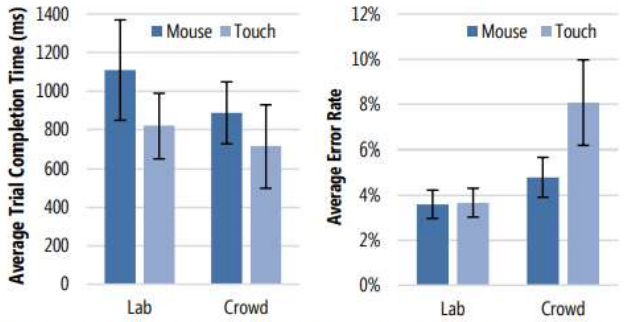
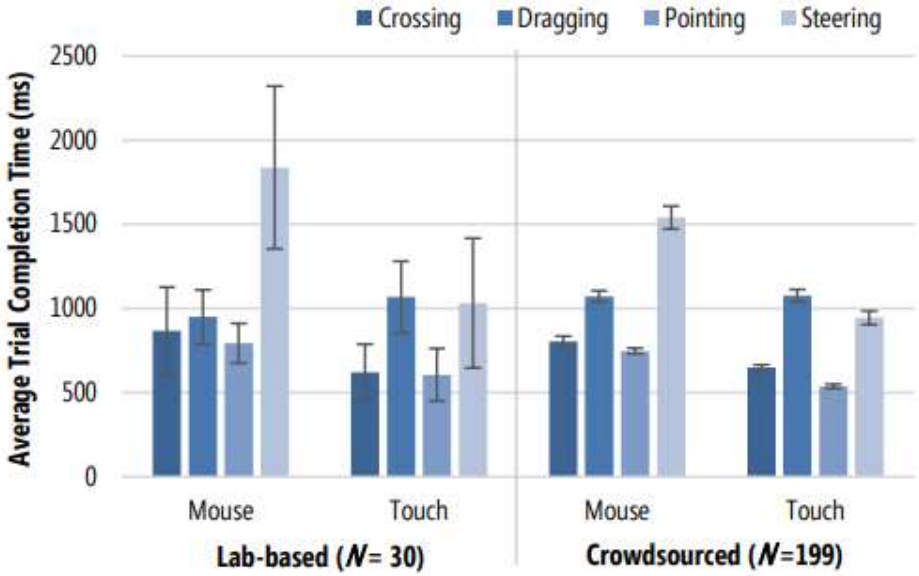
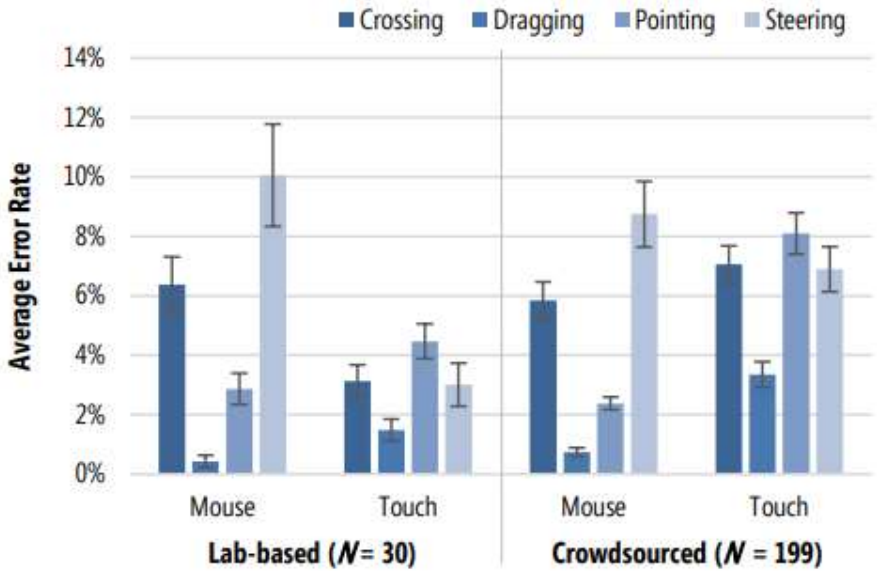


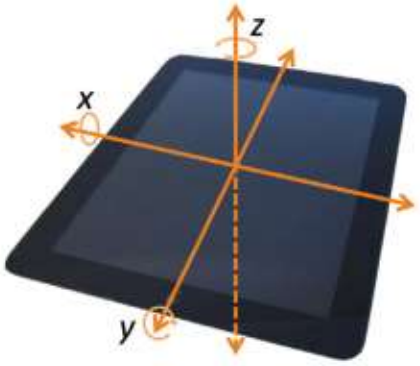
Figure 4. Average trial completion time (left) and error rate (right) for the direct comparison of lab-based participants to a roughly age-matched subset of crowdsourced participants in Experiment 1. (*N*=60 in total; error bars show standard error.)

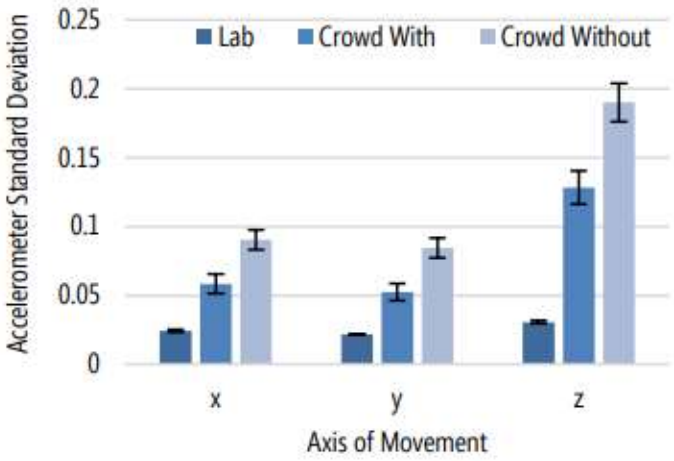


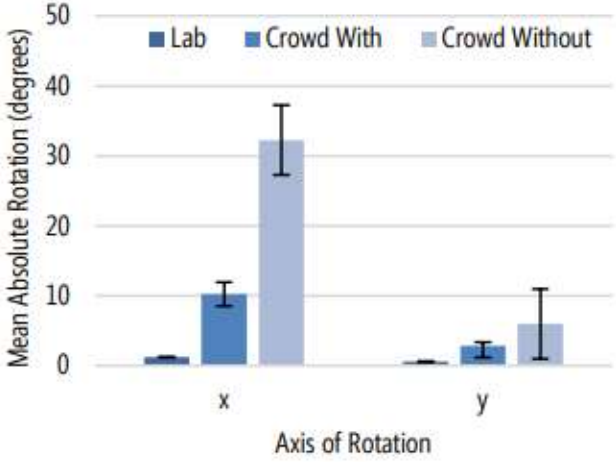


Experiment 2: Some positives

- Crowd will take instruction
- Told crowdworkers to place tablet flat, see data that looks like lab for movement, perhaps a bit more rotation







With pandemic and limited contact

- Mechanical Turk
  - Can be a good source of UI/experimental data if your experiment can be coded to run on personal devices
  - Gain: Can do work with specialized systems if it is (at least somewhat) commercially available
  - Risks: limited participant pool, specialized participants so unsure of generalizability
- On-line surveys
  - Essentially a type of closed interview where participants answer specific questions.
  - Gain: Easy to deploy, flexible, fast, cheap (free in many cases)
  - Risks: Need very careful design, require highly targeted (and preferably closed) questions
- On-line interviews
  - Video-based interviewing to mimic in-person data collection
  - Gain: fast data collection, cheap, more flexibility than questionnaires re format (semi-structured, etc.), easy recording of audio and video.
  - Risks: Somewhat artificial, harder to establish rapport.

With pandemic and limited contact

- Targeted sampling
  - See Caesar's paper, provided as reference
  - Gain: heterogeneous users, fast data collection, relatively inexpensive
  - Risks: need good instruction, aspects of control vanish due to heterogeneous environments, must run on widely available hardware (not good for VR, tracking, etc.)
- Hardware exchange
  - Package specialized hardware and deploy it to participants using limited contact
  - Gain: Full flexibility to use specialized systems
  - Risks: Uncontrolled environment, hardware risk, etc.
- Diary studies
  - Have participants record data of interest
  - Gain: Almost un-impacted by pandemic, can do post-collection interviewing via skype, generalized mobile apps to collect data.
  - Risks: Data that you collect will be impacted by pandemic, may not generalize beyond now.