hci+d lab.

Week 07

Scientific Foundations
CSCW / Social Computing

HCI 연구방법론 2019 Fall

Human-Computer Interaction+Design Lab _ Joonhwan Lee

오늘 다룰 내용

- Definition of Research
- Characteristic of Research
- Research Methods
- Measures
- Research Questions
- Research Topics

- CSCW / Social Computing Overview

Definition of Research

What is Research?

- “Research” means different things to different people
- Often just a word adding weight to an assertion (“Our research shows that…”)
  - from an AD.

    “Independent research proves our Internet service is the fastest and most reliable – period.”
- *Research* has at least three definitions

Research – Definition #1

- Research is…
  - Careful or diligent **search**

- Examples
  - Searching one's garden for weeds
  - Searching a computer to find all files modified on a certain date

Research – Definition #2

- Research is…
  - **Collecting information** about a particular subject

- Examples
  - Survey voters to collect information on political opinions in advance of an election
  - Observe people using computers and collect information, such as the number of times they
    - Consulted the manual
    - Clicked the wrong button
    - Retried an operation
    - Uttered an expletive

Research – Definition #3

- Research is…
  - **Investigation or experimentation** aimed at the discovery and interpretation of facts, the revision of accepted theories or laws in light of new facts.

- Example
  - Design and conduct a user study to test whether a new interaction technique improves on an existing interaction technique

Experimentation

- A central activity in HCI research
- An experiment is sometimes called a *user study*
- Formal, standardized methodology preferred
  - Brings consistency to a body of work
  - Facilitates reviews and comparisons between different user studies

Facts, Theories, Laws

- Facts
  - Building blocks of evidence
  - Evidence is tested to confirm hypotheses (more later)
- Theory
  - An hypothesis assumed for the sake of argument
  - A scientifically accepted body of principles that explain phenomena
- Law
  - More constraining, more formal, more binding
  - A relationship that is invariable under given conditions
  - HCI involves humans, so laws are of questionable value

Fitts' Law

- HCI's best known "law"
- Fitts proposed a model, not a law
- Fitts' law is a behavioral, predictive, and descriptive model of human motor behavior
- It is a "law" only in that other researchers took up the label as a celebration of the robustness and importance of Fitts' work

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47, 381-391

Fitts' Law

- 한 지점에서 다른 지점으로 마우스 포인터를 이동하려고 할 때 걸리는 시간
  - 이동하려는 거리에 비례
  - 이동하려는 목표의 크기에 반비례

$$T = a + b \log_2 \left(\frac{D}{W} + 1 \right)$$

T: 이동에 걸린 시간 (msec)

a: 디바이스의 시작/멈춤에 소요되는 시간

(실험값. 보통 50 정도 사용)

b: 디바이스 본래의 속도

(실험값. 보통 150 정도 사용)

D: 타겟 중심까지의 거리

W: 대상물의 넓이

Fitts' Law

- Fitts' Law 가 적용된 예

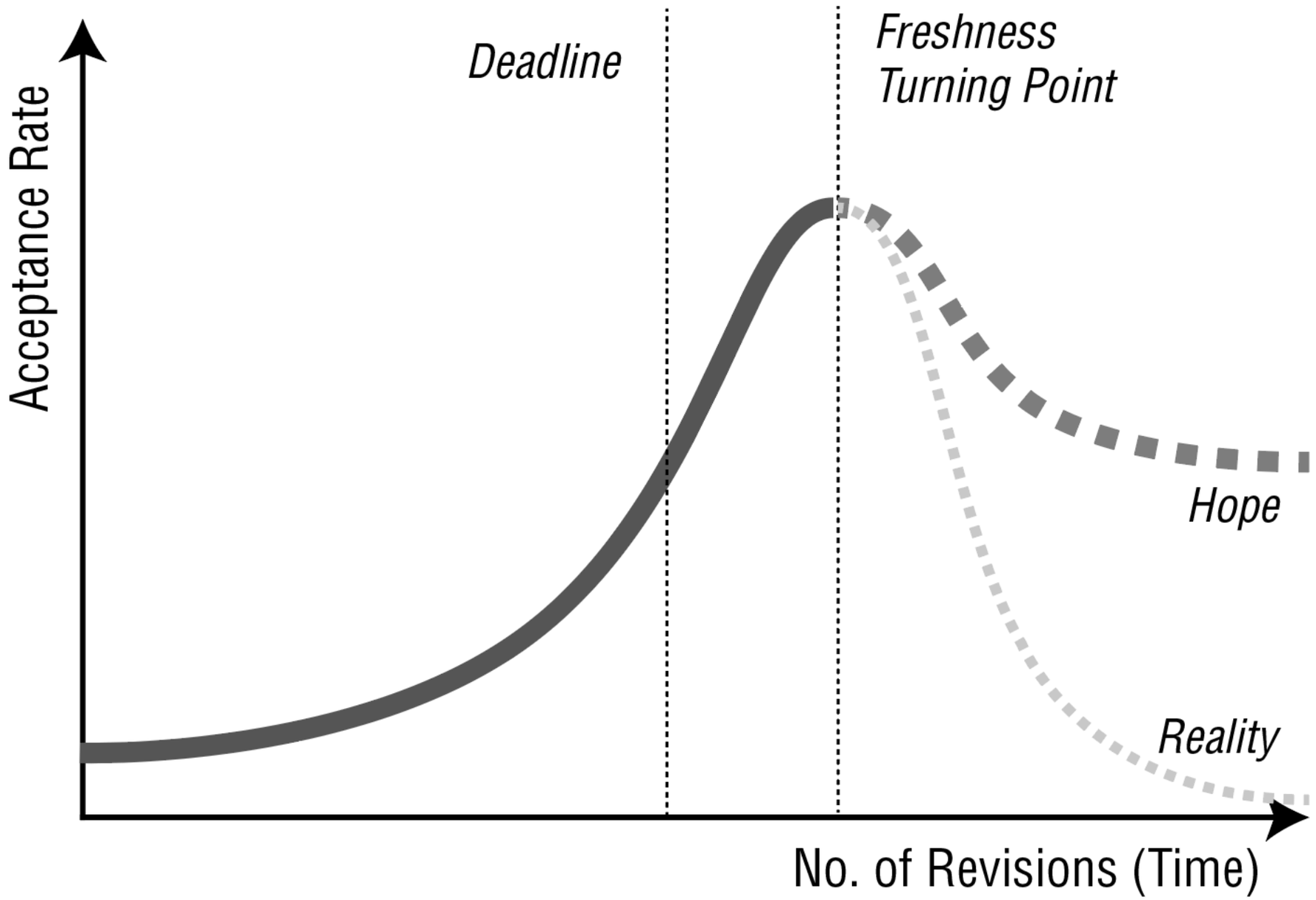
Windows XP

Windows 7

Characteristic of Research

Research Must Be Published

- Publication is the final step
- Also an essential step
- *Publish or perish!*
  - Edict for researchers in all fields, and particularly in academia
- Until it is published, research cannot achieve its critical goal:
  - Extend, refine, or revise the existing body of knowledge in the field



A Fate of Academic Research Paper

Peer Review

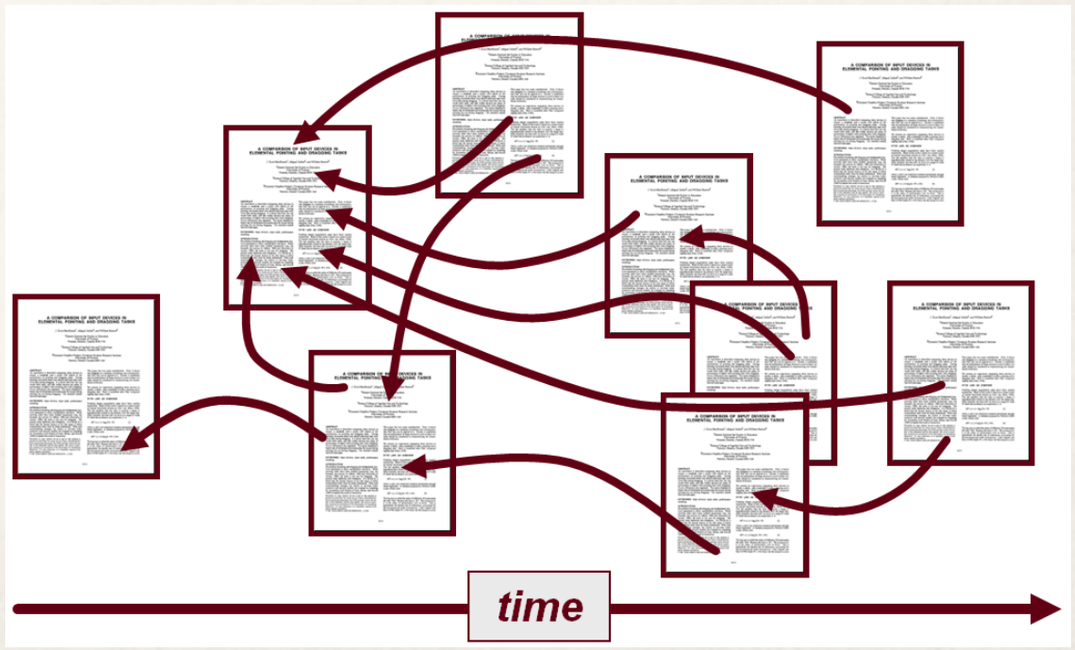
- Research submitted for publication is reviewed by peers – other researchers doing similar research
- Only research meeting a high standard of scrutiny is accepted for publication
  - Are the results novel and useful?
  - Does the evidence support the conclusions?
  - Does the methodology meet the expected standards for the field?
- Accepted research is published and archived
- The final step is complete

Patents

- Some research develops into bona fide inventions
- A researcher/company may wish to maintain ownership of (profit from) the invention
- Patenting is an option
  - The patent application describes
    - Previous related work
    - How the invention addresses a need
    - The best mode of implementation
  - If the application is granted, the patent is issued
  - Note: A patent is a publication; thus patenting meets the must-publish criterion for research

Citations, References, Impact

- Citations, like hyperlinks, connect research to other research
- Through citations, a body of research takes shape
- The number of citations to a research paper is an indication of the paper's impact
- Can you spot the high-impact paper below? (arrows are citations)



Research Must Be Reproducible

- Research that cannot be replicated is useless
- A high standard or reproducibility is essential
- The research write-up must be sufficiently detailed to allow a skilled researcher to replicate the research if he/she desired
- The easiest way to ensure reproducibility is to follow a standardized methodology
- Many great advances in science pertain to methodology (e.g., Louis Pasteur's detailed disclosure of the methodology used in his research in microbiology)

Research vs. Engineering vs. Design

- Researchers often work closely with engineers and designers, but the skills each brings are different
- Engineers and designers are in the business of building things, bringing together the best in *form* (design emphasis) and *function* (engineering emphasis)
- One can image that there is a certain tension, even trade-off, between form and function
- Sometimes, things don’t go quite as planned

Form Trumpeting Function

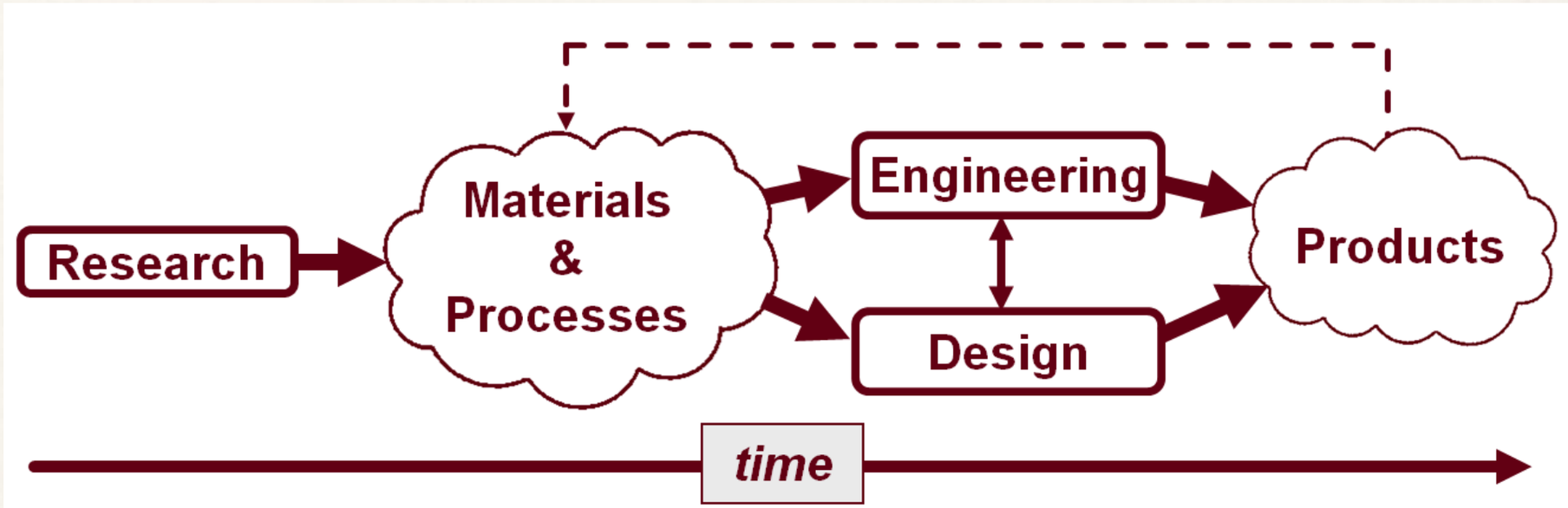
- The photo below shows part of a laptop computer
- The form is elegant – smooth, shiny, metallic
- The touchpad design (or is in engineering?) has a problem
- No tactile sense at the sides of the touchpad

Form Trumpeting Function

Research Milieu

- Engineering and design are about products
- Research is not about products
- Research is narrowly focused
- Research questions are small in scope
- Research is incremental, not monumental
- Research ideas build on previous research ideas
- Good ideas are refined, advanced (into new ideas)
- Bad ideas are discarded, modified
- Products come later, much later

Schematic


Research
Materials
&
Processes
Engineering
Design
Products
time

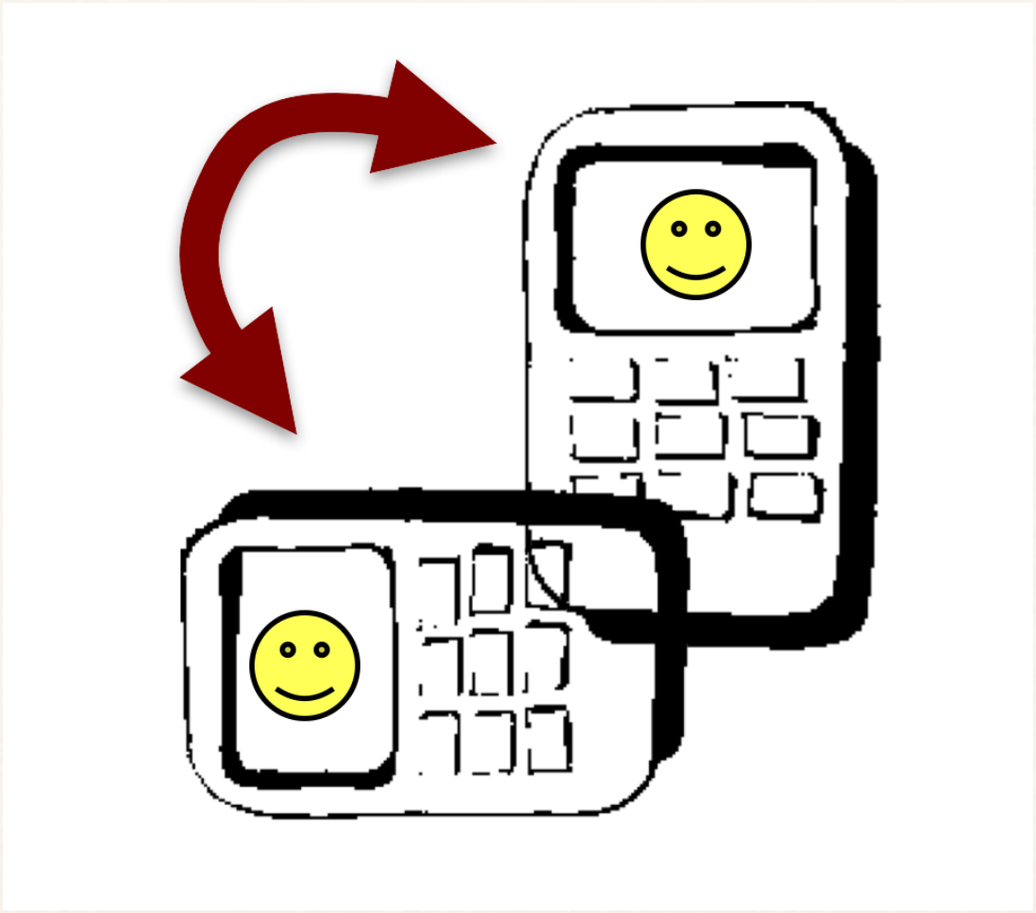
Example: Apple iPhone (2007)

iPhone Gestures:

- Tilt
- Multitouch
- Flick

Tilt

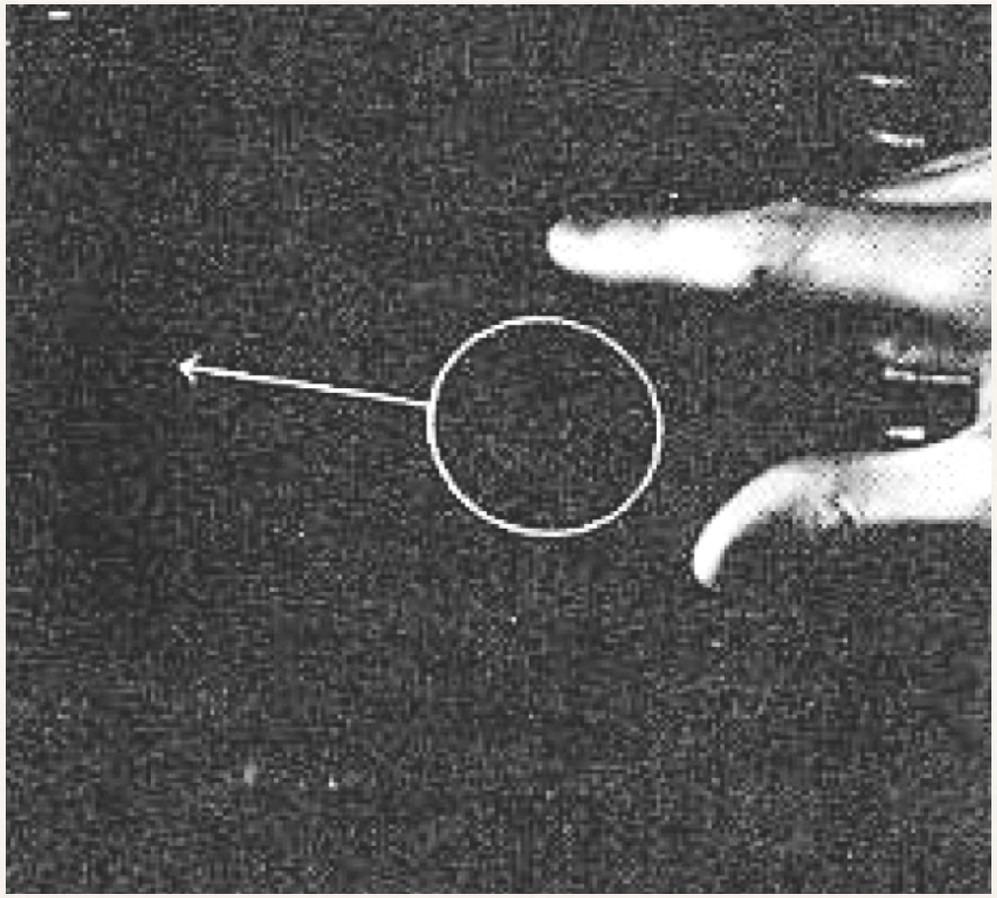
- Research on tilt as an interaction primitive dates at least to 1998



Harrison, B., Fishkin, K. P., Gujar, A., Mochon, C., & Want, R. (1998). Squeeze me, hold me, tilt me! An exploration of manipulative user interfaces. Proc CHI '98, 17-24, New York: ACM.

Multitouch

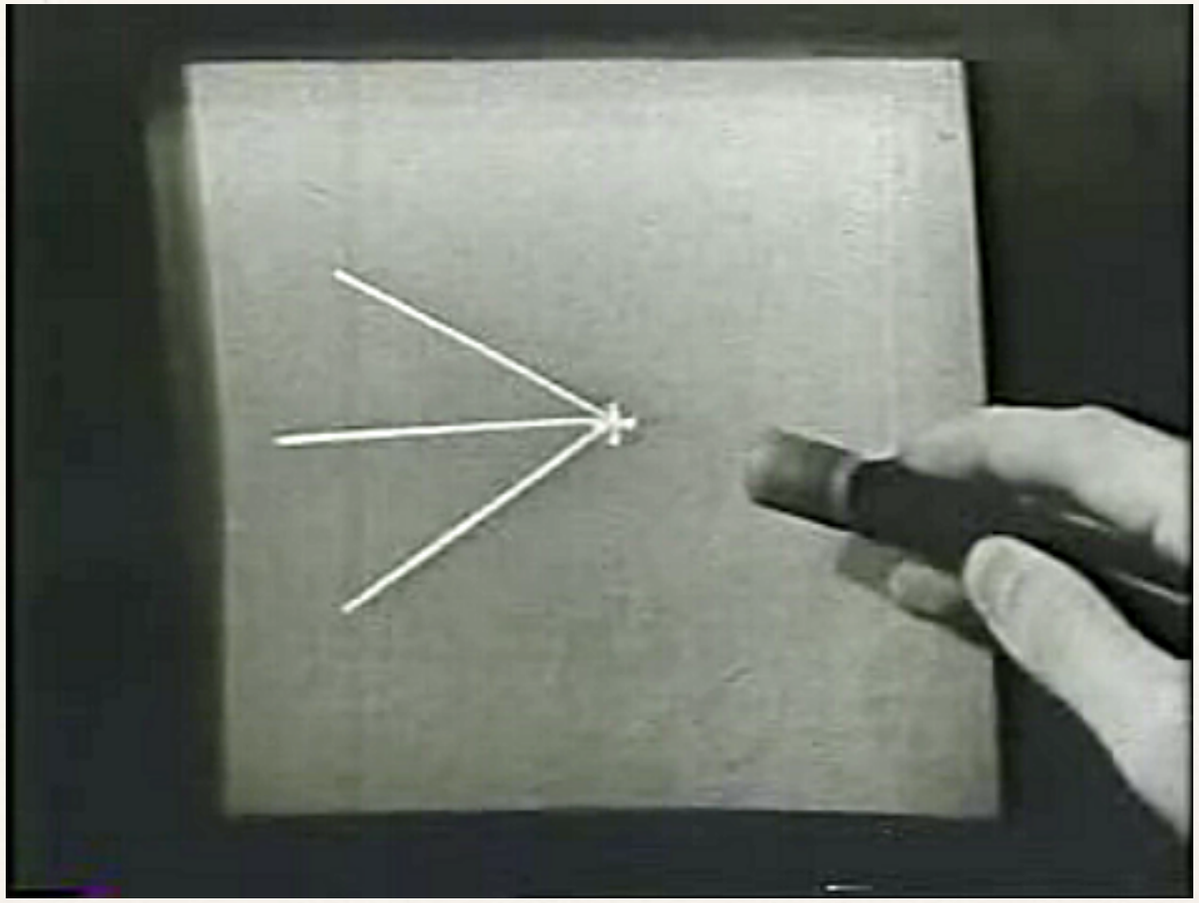
- Research on multitouch as an interaction primitive dates at least to 1978



Herot, C. F., & Weinzapfel, G. (1978). One-point touch input of vector information for computer displays. Proceedings of SIGGRAPH 1978, 210-216, New York: ACM.

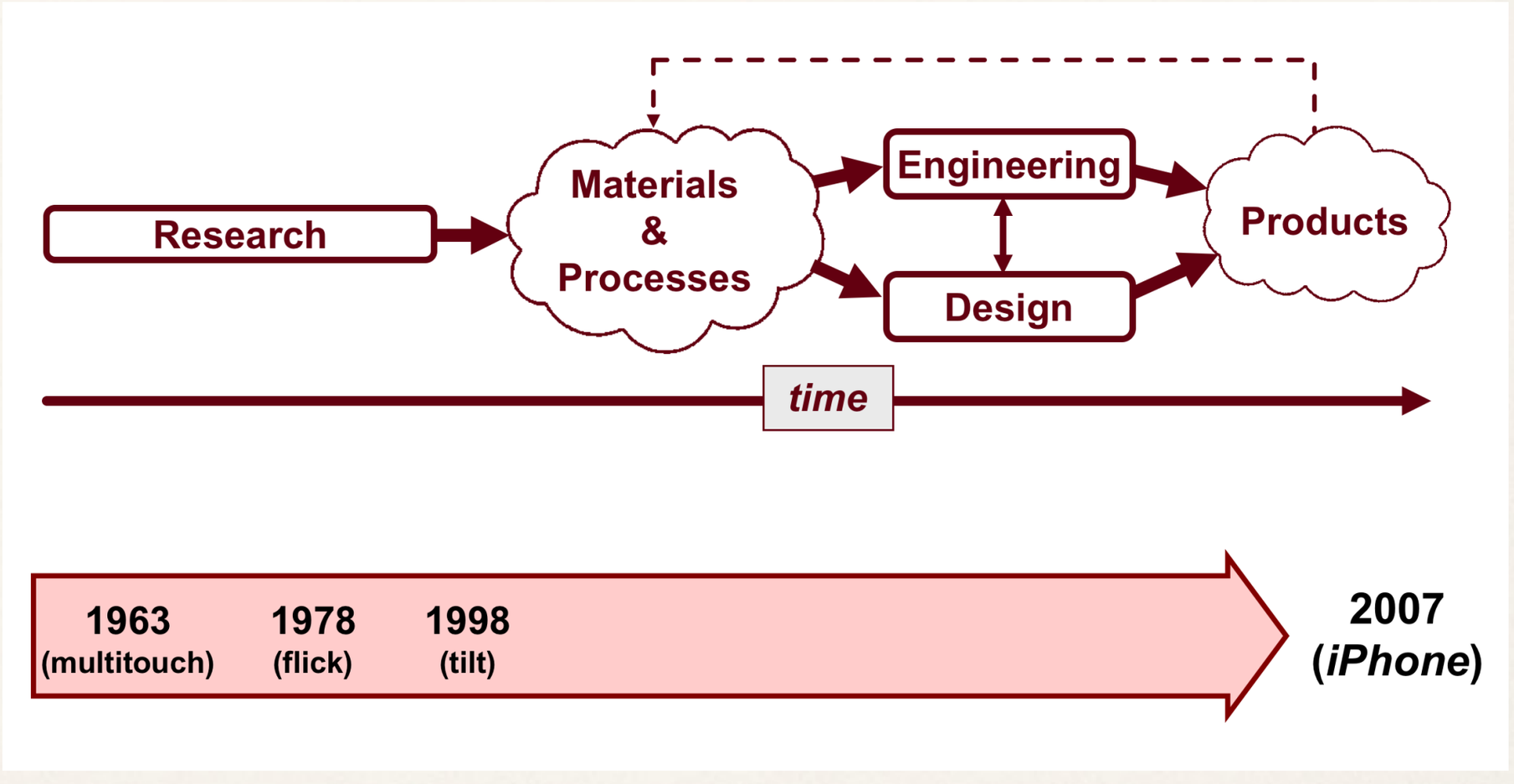
Flick

- Research on flick as an interaction primitive dates at least to 1963



Sutherland, I. E. (1963). Sketchpad: A man-machine graphical communication system. Proceedings of the AFIPS Spring Joint Computer Conference, 329-346, New York: ACM.

Schematic (updated)



Design as Research

- Gaver and Bowers opine on the struggle for designers to also be researchers:
  - Do we need to add research questions or methodological rigour to design practice for it to count as research?
  - Do we have to change design practices to make our contributions to HCI look more like research?
  - Is the result still design, or have we lost something in the process?
  - These questions have been vexing the HCI design community – and us – for some time. The problem is that novel products alone do not seem sufficient to count as research.

Gaver, B., & Bowers, J. (2012, July/August). Annotated portfolios. interactions, 40-49.

Empirical Research

- Empirical:
  - Originating in or based on observation or experience
  - Relying on experience or observation alone without due regard for system or theory (i.e., don't be blinded by pre-conceptions)
- Example: Nicolas Copernicus (1473-1543)
  - Prevailing system or theory: celestial bodies revolved around the earth
  - Copernicus made astronomical observations that cut against this view
  - Result: heliocentric cosmology (the earth and planets revolve around the sun)

Empirical Research

- Empirical: (by another definition)
  - Capable of being verified or disproved by observation or experiment
- HCI research
  - Framed by hypotheses
  - Methodology to test hypotheses
  - Experiments (aka user studies) are the vehicle
  - Hypotheses must be sufficiently narrow and clear to allow for verification or disproval

Research Methods

Research Methods

- Observational method
- Experimental method
- Correlational method

Observational Method

- Example methods:
  - Interviews, field investigations, contextual inquiries, case studies, field studies, focus groups, think aloud protocols, story telling, walkthroughs, cultural probes, etc.
- Focus on qualitative assessments (cf. quantitative)
- Relevance vs. precision
  - High in relevance (behaviours studied in a natural setting)
  - Low in precision (lacks control available in a laboratory)
- Goal: discover and explain reasons underlying human behaviour (*why* or *how*, as opposed to what, where, or when)

Experimental Method

- Aka *scientific method*
- Controlled experiments conducted in lab setting
- Relevance vs. precision
  - Low in relevance (artificial environment)
  - High in precision (extraneous behaviours easy to control)
- At least two variables:
  - *Manipulated variable* (aka independent variable)
  - *Response variable* (aka dependent variable)
- Cause-and-effect conclusions possible (changes in the manipulated variable caused changes in the response variable)

Correlational Method

- Look for relationships between variables
- Observations made, data collected
  - Example: are user's privacy settings while social networking related to their age, gender, level of education, employment status, income, etc.
- Non-experimental
  - Interviews, on-line surveys, questionnaires, etc.
- Balance between relevance and precision (some quantification, observations not in lab)
- Cause-and-effect conclusions not possible

Observe and Measure

- Foundation of empirical research
- Observation is the starting point; observations are made…
  - By the apparatus
  - By a human observer
- Manual observation
  - Log sheet, notebooks
  - Screen capture, photographs, videos, etc.
- Measurement
  - With measurement, anecdotes (*April showers bring May flowers*) turn to empirical evidence
  - *“When you cannot measure, your knowledge is of a meager and unsatisfactory kind”* (Kelvin)

Measures

Scales of Measurement

- **Ratio**
- **Interval**
- **Ordinal**
- **Nominal**

Sophisticated

Crude

Nominal Data

- Nominal data (aka *categorical data*) are arbitrary codes assigned to attributes; e.g.,
  - 1 = male, 2 = female
  - 1 = mouse, 2 = touchpad, 3 = pointing stick
- The code needn't be a number; i.e.,
  - M = male, F = female
- Obviously, the statistical mean cannot be computed on nominal data
  - Usually it is the count that is important
  - "Are females or males more likely to…"
  - "Do left handers or right handers have more difficulty with…"
  - Note: The count itself is a ratio-scale measurement

Nominal Data – HCI Example

- Task: Observe students “on the move” on university campus
- Code and count students by…
  - Gender (male, female)
  - Mobile phone usage (not using, using)

Gender	Mobile Phone Usage		Total	%
	Not Using	Using		
Male	683	98	781	51.1%
Female	644	102	746	48.9%
Total	1327	200	1527	
%	86.9%	13.1%		

Ordinal Data

- Ordinal data associate an order or rank to an attribute
- The attribute is any characteristic or circumstance of interest; e.g.,
  - Users try three GPS systems for a period of time, then rank them: 1st, 2nd, 3rd choice
- More sophisticated than nominal data
  - Comparisons of “greater than” or “less than” possible

Ordinal Data – HCI Example

How many email message do you receive each day?

1. None (I don’t use email)
2. 1-5 per day
3. 6-30 per day
4. 31-100 per day
5. More than 100 per day

Interval Data

- Equal distances between adjacent values
- Classic example: temperature (°F, °C)
- Statistical mean possible
  - E.g., the mean midday temperature during July
- Ratios not possible
  - Cannot say 10 °C is twice 5 °C

Interval Data – HCI Example

- Questionnaires often solicit a level of agreement to a statement
- Responses on a Likert scale
- Likert scale characteristics:
  1. Statement soliciting level of agreement
  2. Responses are symmetric about a neutral middle value
  3. Gradations between responses are equal (more-or-less)
- Assuming “equal gradations”, the statistical mean is valid (and related statistical tests are possible)

Interval Data – HCI Example

Please indicate your level of agreement with the following statements.

	Strongly disagree	Mildly disagree	Neutral	Mildly agree	Strongly agree
It is safe to talk on a mobile phone while driving.	1	2	3	4	5
It is safe to read a text message on a mobile phone while driving.	1	2	3	4	5
It is safe to compose a text message on a mobile phone while driving.	1	2	3	4	5

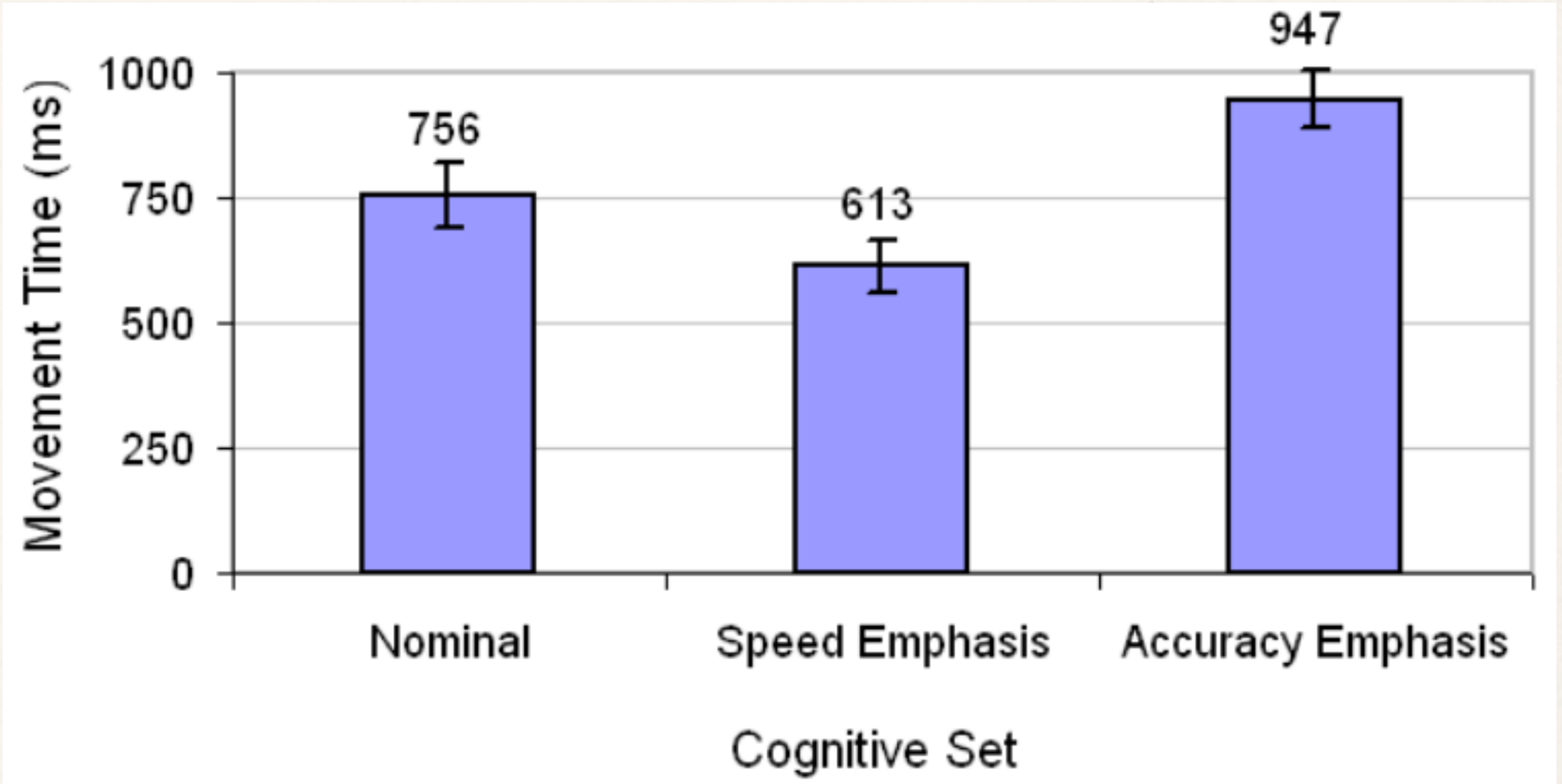
Ratio Data

- Most sophisticated of the four scales of measurement
- Preferred scale of measurement
- Absolute zero, therefore many calculations possible
- Summaries and comparisons are strengthened

Ratio Data

- A “count” is a ratio-scale measurement
  - E.g., “time” (the number of seconds to complete a task)
- Enhance counts by adding further ratios where possible
  - Facilitates comparisons
  - Example – a 10-word phrase was entered in 30 seconds
    - Bad: *t* = 30 seconds
    - Good: Entry rate = 10 / 0.5 = 20 wpm

Ratio Data – HCI Example



$F(2,34) = 372.7, p < .0001$

MacKenzie, I. S., & Isokoski, P. (2008). Fitts' throughput and the speed-accuracy tradeoff. Proc CHI 2008, 1633-1636, New York: ACM.

Four Measurement Comparison

Provides:	Nominal	Ordinal	Interval	Ratio
The "order" of values is known		✔	✔	✔
"Counts," aka "Frequency of Distribution"	✔	✔	✔	✔
Mode	✔	✔	✔	✔
Median		✔	✔	✔
Mean			✔	✔
Can quantify the difference between each value			✔	✔
Can add or subtract values			✔	✔
Can multiple and divide values				✔
Has "true zero"				✔

Research Questions

Research Questions

- We conduct empirical research to answer questions about UI designs or interaction techniques
- Consider the following questions:
  - Is it viable?
  - Is it better than current practice?
  - Which design alternative is best?
  - What are the performance limits?
  - What are the weaknesses?
  - Does it work well for novices?
  - How much practice is required?

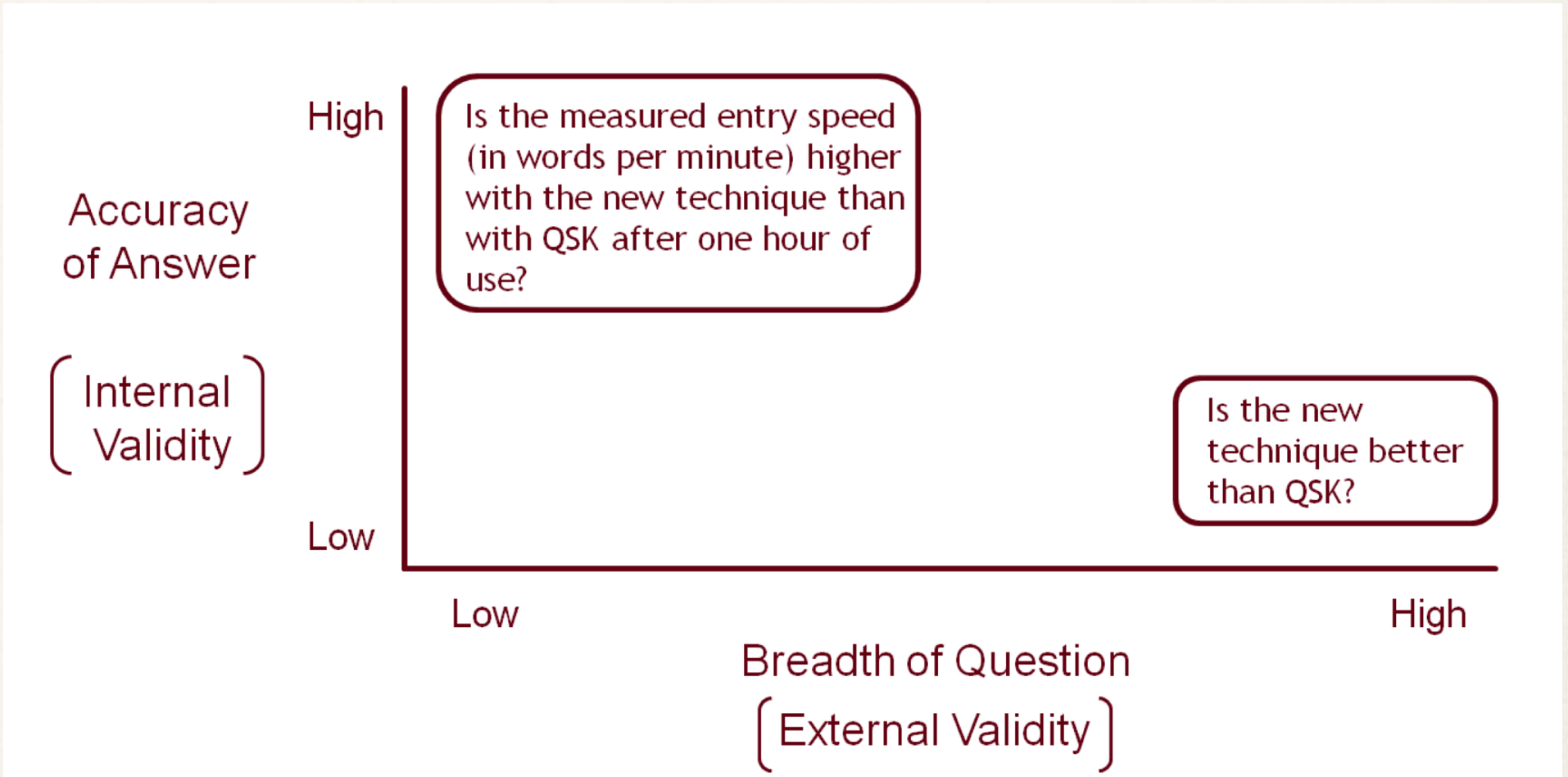
Testable Research Questions

- Try to re-cast as testable questions (even though the new question may appear less important)
- Scenario…
  - You have **invented a new text entry technique for touchscreen mobile phones**, and you think it's pretty good. In fact, you think it is **better than the Qwerty** soft keyboard (QSK). You decide to undertake a program of empirical enquiry to evaluate your invention. What are your research questions?

Research Questions Example

- Very weak
  - Is the new technique any good?
- Weak
  - Is the new technique better than QSK?
- Better
  - Is the new technique faster than QSK?
- Better still
  - Is the measured entry speed (in words per minute) higher for the new technique than for QSK after one hour of use?

A Tradeoff


Accuracy of Answer
High
Low
Is the measured entry speed (in words per minute) higher with the new technique than with QSK after one hour of use?
Is the new technique better than QSK?
Internal Validity
Low
High
Breadth of Question
External Validity

Internal Validity

- Definition:
  - The extent to which the effects observed are due to the test conditions (e.g., multitap vs. new)
- Statistically, this means…
  - Differences (in the means) are due to inherent properties of the test conditions
  - Variances are due to participant differences (“pre-dispositions”)
  - Other potential sources of variance are controlled or exist equally or randomly across the test conditions

External Validity

- Definition:
  - The extent to which results are generalizable to other people and other situations
- People
  - The participants are representative of the broader intended population of users
- Situations
  - The test environment and experimental procedures are representative of real world situations where the interface or technique will be used

Test Environment Example

- Scenario…
  - You wish to compare two input devices for remote pointing (e.g., at a projection screen)
- External validity is improved if the test environment mimics expected usage
  - Test environment should probably…
    - Use a large display or projection screen (not a desktop monitor)
    - Position participants at a significant distance from screen (rather than close up)
    - Have participants stand (rather than sit)
    - Include an audience!
- But… is internal validity compromised?

Experimental Procedure Example

- Scenario…
  - You wish to compare two text entry techniques for mobile devices
- External validity is improved if the experimental procedure mimics expected usage
- Test procedure should probably have participants…
  - Enter personalized paragraphs of text (e.g., a paragraph about a favorite movie)
  - Edit and correct mistakes as they normally would
- But… is internal validity compromised?

Research Topics

Research Topics

- Finding a research topic is a challenge (for students… and for seasoned researchers too!)
- Four tips:
  1. Think small
  2. Replicate
  3. Know the literature
  4. Think inside the box

Tip #1 - Think Small

- Looking for that big idea?
- Advice: Forget it (besides, it isn't necessary)
- Research questions are small, narrowly focused
- Pursue several small, related research topics and before you know it, a dissertation topic is formed

Tip #2 - Replicate

- Seems odd: where's the research in simply replicating what was done before?
- Of course, there is no research in replication, but the trick is in the path to replicating
- Replicating prior research is a lot of work
- Along the way, you will undoubtedly discover small and novel improvements – things to try
- A little tweak here, a small modification there
- BTW, you might not find a novel idea until well into the process (perhaps afterward!)

Tip #3 – Know The Literature

- Whatever topic interests you, read the literature
  - E.g., social networking, gaming
- If too broad, narrow (e.g., privacy settings in social networking, avatars in gaming)
- Read papers, open a spreadsheet, tabulate variables in the methodology and the findings
- Chaotic at first, order and shape will emerge (eventually)
- With some luck (and further study) a research topics will emerge

Tip #3 – Know The Literature (Example)

Study (1st author)	Number of Keys[a]	Direct/ Indirect	Scanning	Number of Participants	Speed[b] (wpm)	Notes
Bellman [2]	5	Indirect	No	11	11	4 cursors keys + SELECT key. Error rates not reported. No error correction method.
Dunlop [4]	4	Direct	No	12	8.90	4 letter keys + SPACE key. Error rates reported as "very low."
Dunlop [5]	4	Direct	No	20	12	4 letter keys + 1 key for SPACE/NEXT. Error rates not reported. No error correction method.
Tanaka-Ishii [25	3	Direct	No	8	12+	4 letters keys + 4 keys for editing, and selecting. 5 hours training. Error rates not reported. Errors corrected using CLEAR key.
Gong [7]	3	Direct	No	32	8.01	3 letter keys + two additional keys. Error rate = 2.1%. Errors corrected using DELETE key.
MacKenzie [16]	3	Indirect	No	10	9.61	2 cursor keys + SELECT key. Error rate = 2.2%. No error correction method.
Baljko [1]	2	Indirect	Yes	12	3.08	1 SELECT key + BACKSPACE key. 43 virtual keys. RC scanning. Same phrase entered 4 times. Error rate = 18.5%. Scanning interval = 750 ms.
Simpson [24]	1	Indirect	Yes	4	4.48	1 SELECT key. 26 virtual keys. RC scanning. Excluded trials with selection errors or missed selections. No error correction. Scanning interval = 525 ms at end of study.
Koester [10]	1	Indirect	Yes	3	7.2	1 SELECT key. 33 virtual keys. RC scanning with word prediction. Dictionary size not given. Virtual BACKSPACE key. 10 blocks of trials. Error rates not reported. Included trials with selection errors or missed selections. Fastest participant: 8.4 wpm.

[a] For "direct" entry, the value is the number of letter keys. For "indirect" entry, the value is the total number of keys.
[b] The entry speed cited is the highest of the values reported in each source, taken from the last block if multiple blocks.

MacKenzie, I. S. (2009). The one-key challenge: Searching for an efficient one-key text entry method. Proc ASSETS 2009, 91-98, New York: ACM.

Tip #4 – Think Inside The Box

- ***Think outside the box*** → dispense with accepted beliefs and assumptions (in the box), and think in a way that assumes nothing and challenges everything
- ***Think inside the box***: just get on with your day; but at every juncture, every interaction, think and question
- Our everyday foibles are fertile ground for research topics

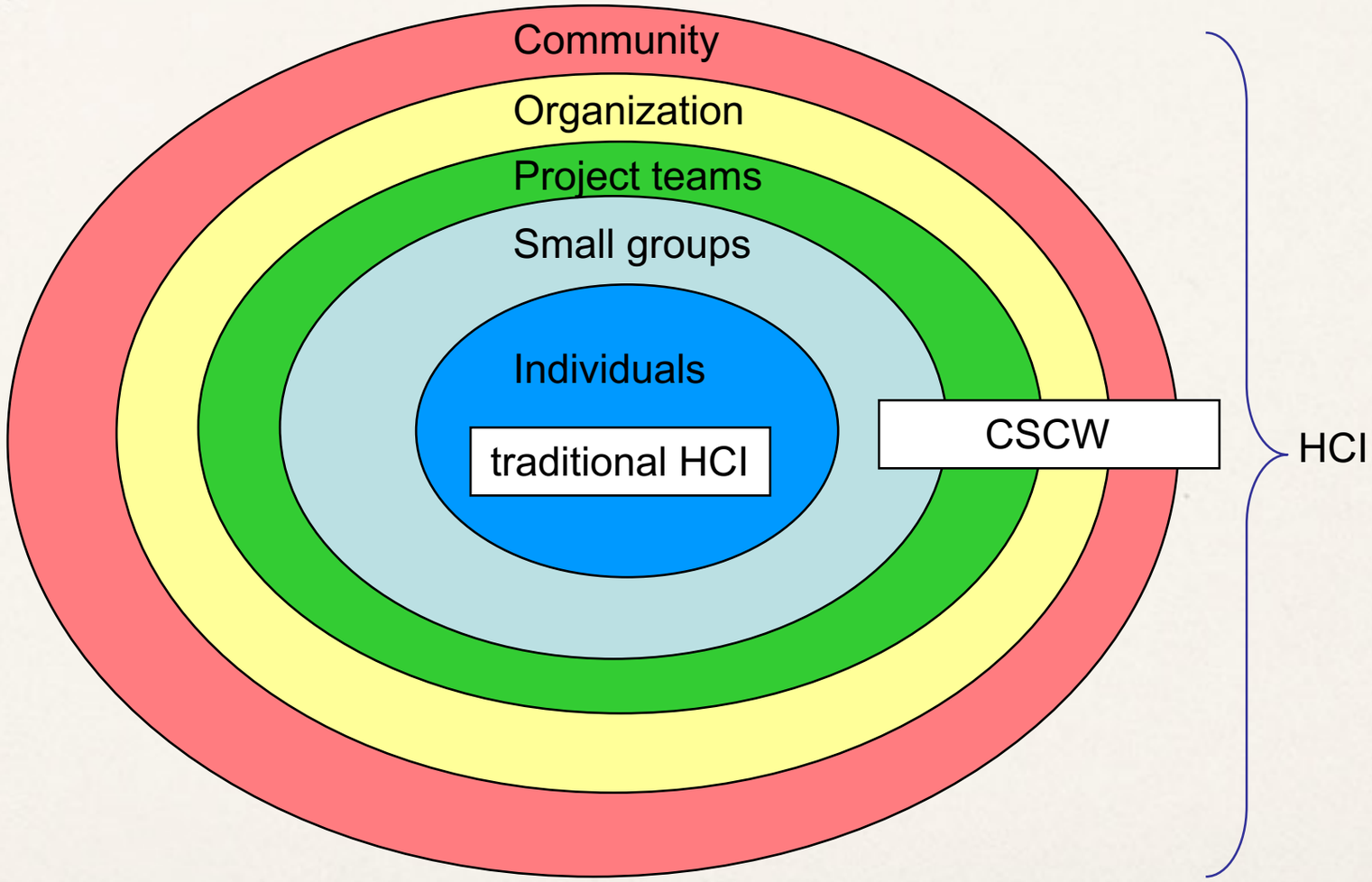
CSCW / Social Computing Overview

Computer Supported Cooperative Workspace (CSCW)

- CSCW grew from discontent with **single user HCI methods** applied to **multi-user technologies and settings**
- Focus on
  - Workplace activity
  - Understanding nature of collaborative tasks
  - Co-evolution of technologies and communities
- Early apps
  - CAD, computer integrated manufacturing, computer aided software engineering, office automation

CSCW

- CSCW focuses on people working with others



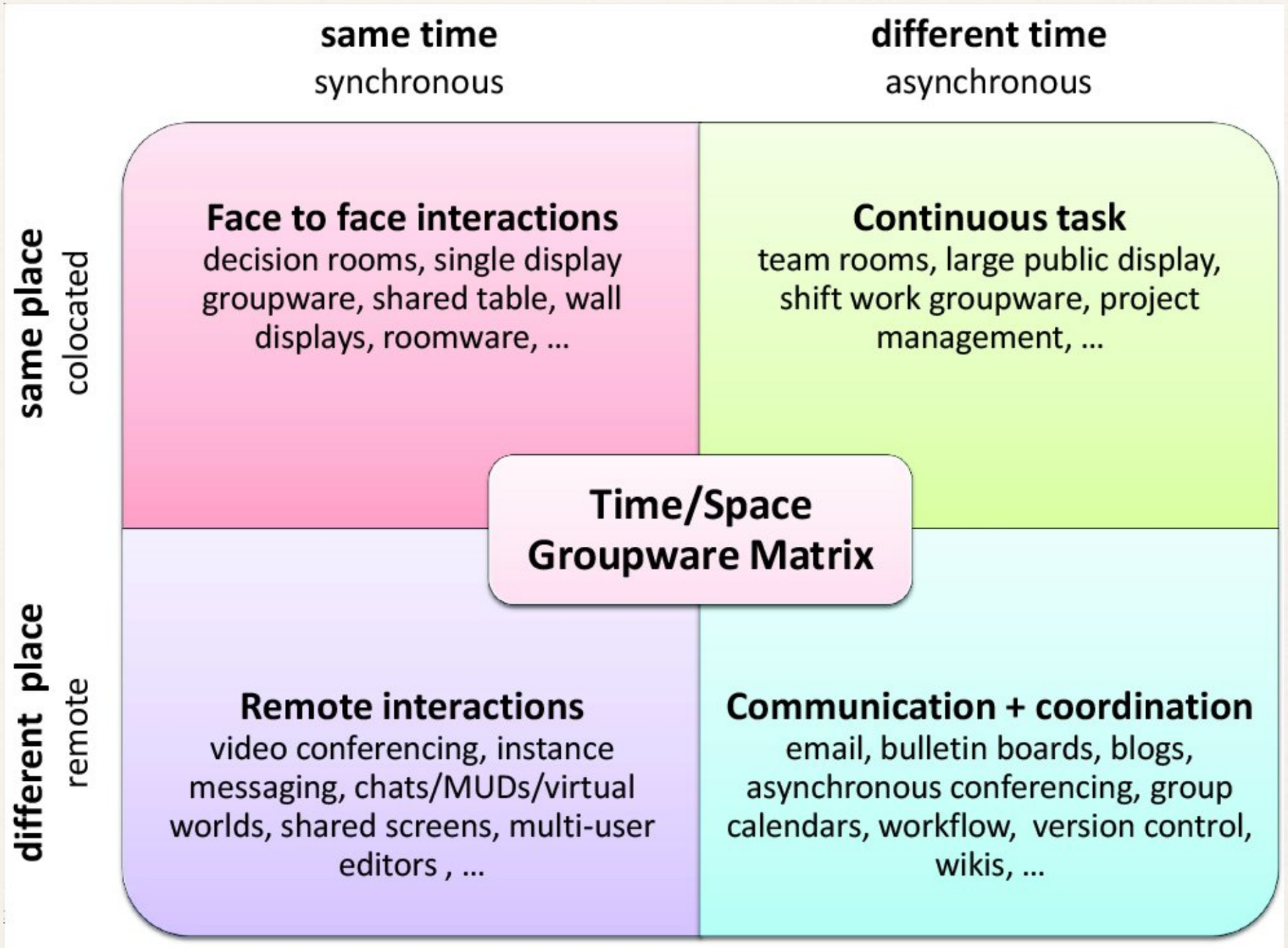
Designing CSCW is hard, why?

- Multiple users
- “Virtual” (not physical) presence
- The Network!!

- Virtual presence could be “Beyond Being There”
- Some distinguishing features of CSCW:
  - asynchronous communication
  - anonymous communication
  - automatically archive of communication

Hollan, Jim and Stornetta, Scott. “Beyond Being There.” CHI 1992.

CSCW Matrix



Wikipedia. Johansen, 1988 in Baecker, R.M.; Others, (1995). Readings in human-computer interaction: toward the year 2000. Morgan Kaufmann Publishers.

Social Computing

- Group of University Researchers to Make web Science a Field of Study

- Steve Lohr, New York Times, Nov. 2, 2006
- http://www.nytimes.com/2006/11/02/technology/02compute.html?_r=3&adxnnl=1&oref=slogin&adxnnlx=1212113936-DbHS7WsdpYrJCC4d1pZXmw&

Social Computing

- “Web science, the researchers say, has **social and engineering dimensions**. It extends well **beyond traditional computer science**, they say, to include the **emerging research in social networks and the social sciences** that is being used to study **how people behave on the Web**. And Web science, they add, shifts the center of gravity in engineering research from how a single computer works to how huge decentralized Web systems work.”

Social Computing

- “Computer science is at a turning point, and it has to go **beyond algorithms and understand the social dynamics of issues** like trust, responsibility, empathy and privacy in this vast networked space.”

- Ben Shneiderman, a Professor at the University of Maryland

Social Computing

- Social computing is an area of computer science that is concerned with the **intersection of social behavior and computational systems**.
- Social computing is the collaborative and interactive aspect of online behavior.

- Wikipedia, http://en.wikipedia.org/wiki/Social_computing

Social Computing

ACM DIGITAL LIBRARY

Search ACM DL

Advanced Search

See also: Digital Library Home All ACM Journals

ACM Transactions on
Social Computing (TSC)

seeks to publish work that covers the full spectrum of social computing including theoretical, empirical, systems, and design research contributions.

Association for Computing Machinery

Home Authors Editorial Board Reviewers Archive About TSC Announcements Contact Subscribe

Search TSC in the ACM DL

Latest Articles

Introduction to *ACM Transactions on Social Computing*

Kevin Crowston

Current Issue: Volume 1 Issue 1, February 2018

Table of Contents
SIGN UP FOR TOC SERVICES :

Why Social Computing?

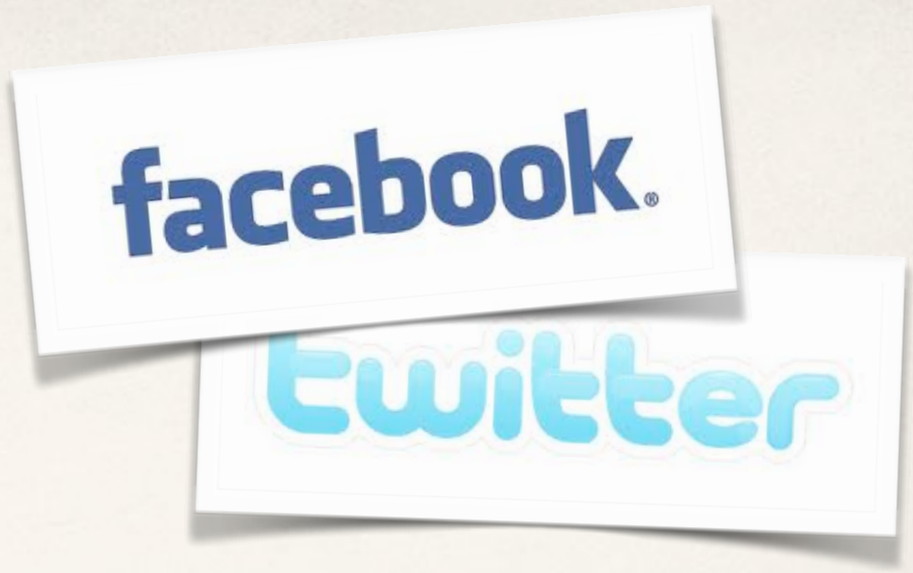
- We have new data that need computational methods.
  - Worldwide, there are over **2.23 billion monthly active Facebook users** for Q2 2018 (Facebook MAUs) which is an 11 percent increase year over year.
  - There are **1.15 billion mobile daily active users** (Mobile DAU) for December 2016, an increase of 23 percent year-over-year.

https://zephoria.com/top-15-valuable-facebook-statistics/

Why Social Computing?

- 인간의 사회적 행위를 지원
  - 사회적 커뮤니티를 위한 온라인 커뮤니케이션을 지원하는 웹 2.0 서비스와 도구
    - 블로그, 위키, 소셜 네트워크 서비스, 북마킹
- 사람들의 집단적 협력과 지능을 활용
  - 협력 필터링(Collaborative Filtering), 추천 (Recommendation), 예측(Forecasting), 평판 (Reputation) 기술 등

소셜미디어 연구



- 소셜 네트워크의 급속한 성장

  → 트위터, 페이스북 등의 소셜미디어 서비스 사용자 급증

  - 사적 정보의 교류
  - 뉴스와 같은 정보성 메시지의 소비, 공유

- 소셜미디어 사용 목적: 정보 공유와 사회적 친분 관계 유지

  - Java, Song, Finin, & Tseng, 2007; Naaman, Boase, & Lai, 2010; Zhao & Rosson, 2009

- 소셜미디어 사용자의 행동 패턴을 정량적으로 분석하는데에 주로 초점을 맞추고 있음.

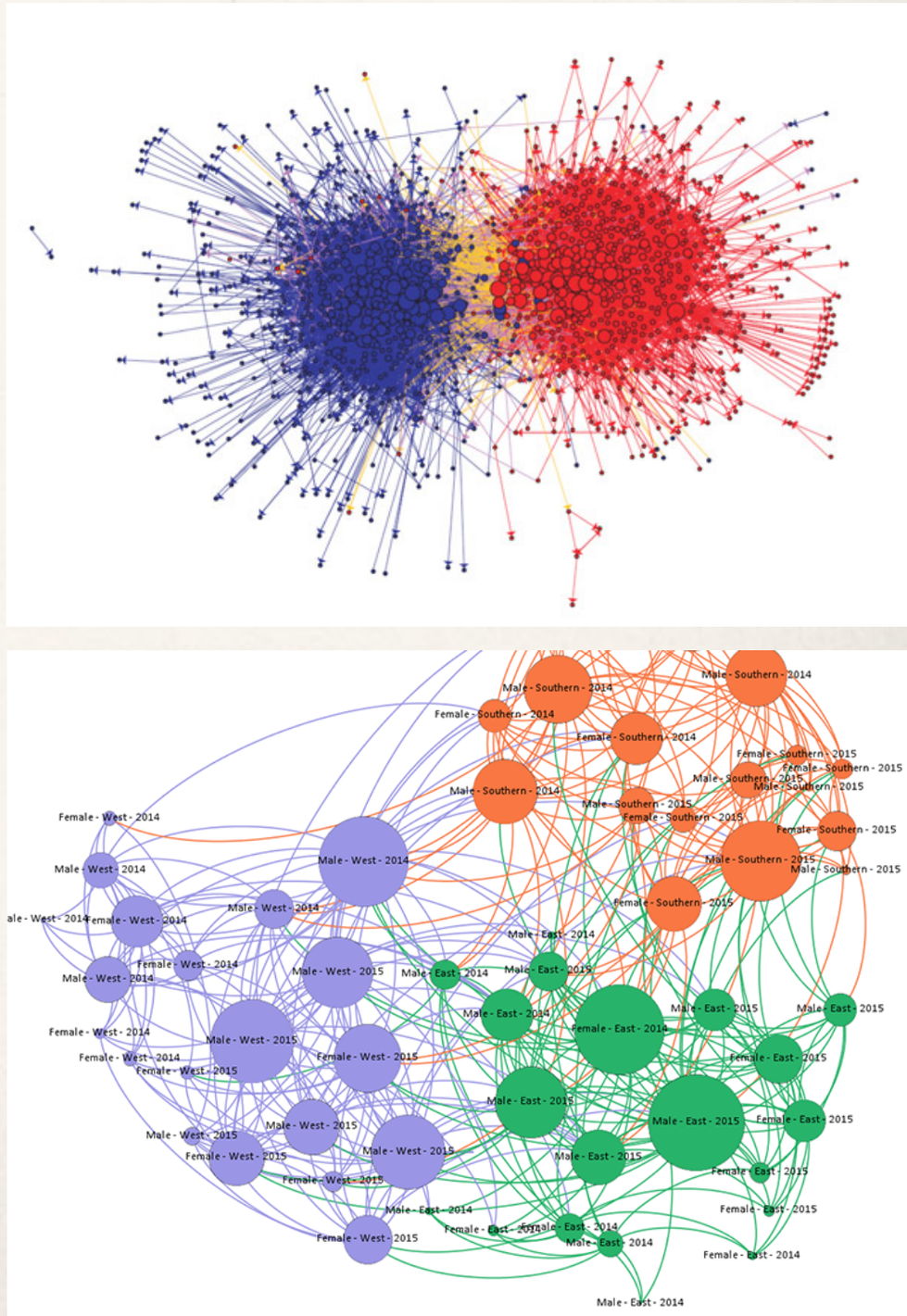
관심 연구 주제

- 소셜미디어에서의 행동 분석
  - **why** - 왜 그런 행동을 하는지…?
    - follower-following study, homophily study, depression study, github study
  - **what** - 무엇을 이야기하고 공유하는지…?
    - text mining, opinion mining, word network analysis
  - **how** - 서로 어떻게 영향을 주고 받는지…?
    - social viewing and political judgements study

→ 이런 결과를 어떻게 활용할 수 있는지…?

연구의 방향

- Analyzing Connections
  - Social Network Analysis
    - Reachability
    - Distance & Number of Paths
    - Degree of Node
    - Centrality
    - Morphology Changes



연구의 방향

- Analyzing Activities
  - 사람들의 행동에는 “어떤” 이유가 있을 것.
    - 예: 우울함의 전조가 보이는 사람들은 소셜미디어에서 어떻게 행동하는가?
      사람들은 왜 소셜미디어에서 친구관계를 맺는가?
  - 행동의 결과를 종속변수로 사용
  - 행동의 내용을 독립변수로 사용
  - 다량의 소셜네트워크 데이터를 API를 통해 수집
  - 사례: Depression Study of Facebook Users
    - 종속변수 - Depression Measure: CES-D
    - 독립변수 - Facebook activities
      - number of likes, number of comments, how often they login to facebook, how often they change their profile picture, how much comments they receive in a day…

연구의 방향

- Analyzing Activities
  - Linear model analysis

$$Y_i = \beta_0 + \beta_1\phi_1(X_{i1}) + \cdots + \beta_p\phi_p(X_{ip}) + \varepsilon_i \qquad i = 1, \ldots, n$$

    - Yi : survey results
    - Xij : independent variables (crawled data)

```
Residuals:
    Min      1Q  Median      3Q     Max
-5.5031 -1.1005  0.0535  1.3613  3.8545

Coefficients:
                          Estimate Std. Error t value Pr(>|t|)
(Intercept)               3.567127   0.073872  48.288  < 2e-16 ***
likes_from_friends       -0.006566   0.006908  -0.950  0.34212
likes_to_friends          0.003180   0.004377   0.727  0.46770
comments_from_friends     0.014643   0.006913   2.118  0.03441 *
comments_to_friends       0.008820   0.005714   1.544  0.12303
photo_tags_from_friends   0.092660   0.090532   1.024  0.30633
photo_tags_to_friends    -0.030312   0.041907  -0.723  0.46966
status_tags_from_friends  0.566128   0.311284   1.819  0.06927 .
status_tags_to_friends    0.936695   0.294069   3.185  0.00149 **
msg_from_friends          0.004335   0.001307   3.315  0.00095 ***
msg_to_friends           -0.004753   0.001354  -3.509  0.00047 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.899 on 956 degrees of freedom
Multiple R-squared:  0.05099, Adjusted R-squared:  0.04106
F-statistic: 5.137 on 10 and 956 DF,  p-value: 2.409e-07
```

연구의 방향

- Analyzing Messages
  - What stories are people talking?
  - Can we predict behavior through their message?
    - “Predicting Postpartum Changes in Emotion and Behavior via Social Media” (de Choudhury et al. 2013)
- Text mining
  - frequency
  - text categorization
  - topic extraction
- Opinion mining (sentiment analysis)
  - computational study of opinions, sentiments

Mr. Clinton 1993

Mr. Obama 2009

Next Week: Reading Assignments

- T2: Human-Computer Interaction
  - Chapter 5. Designing HCI Experiments

Next Week: Reading Assignments

- Weiser, M. (1999). The computer for the 21st century. ACM SIGMOBILE Mobile Computing and Communications Review, 3(3), 3–11.
- Abowd, G. D. and Mynatt, E. D. (2000). Charting past, present, and future research in ubiquitous computing. ACM Trans. Comput.-Hum. Interact. 7, 1, 29-58.
- Kiesler, S., & Hinds, P. (2004). Introduction to This Special Issue on Human-robot Interaction. Human-Computer Interaction, 19(1), 1-8.
- Mutlu, B. & Forlizzi, J. (2008). Robots in organizations: the role of workflow, social, and environmental factors in human-robot interaction. In Proceedings of the 3rd ACM/IEEE international Conference on Human Robot interaction ’08. ACM, 287-294.

Questions…?