COMP5048 VISUAL ANALYTICS

HCI Evaluation Methods

Tony Huang
University of Tasmania

October 5, 2017

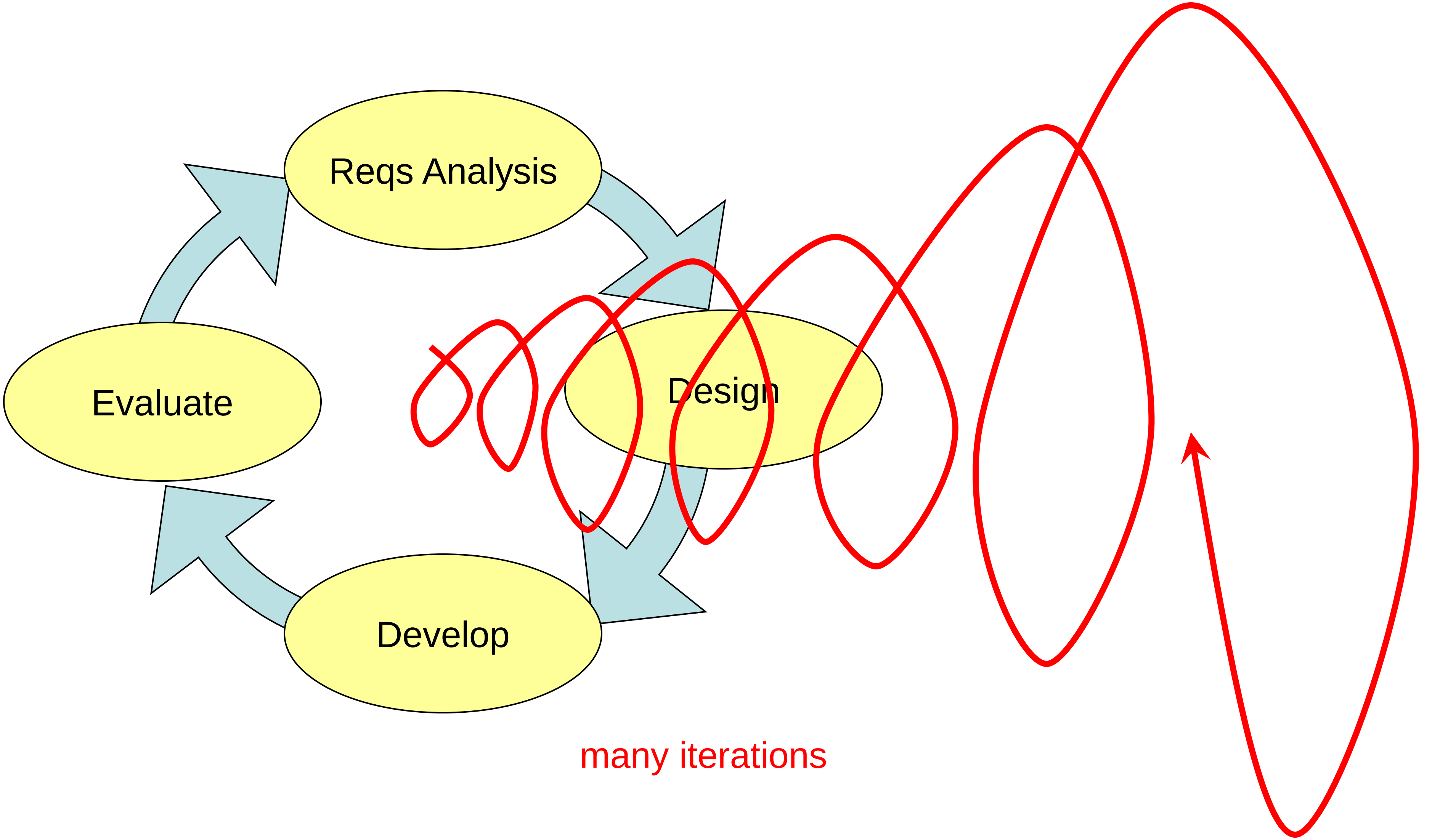
Outline

- **What is HCI**
- Evaluation framework
- Evaluation methods
- Case studies

HCI – Human-Computer Interaction

- Human-computer interaction is the study of interaction between people and computers.
- Interaction between humans and computers occurs at the user interface.
- User interface provides means of:
  - **Input**: allowing users to manipulate a system
  - **Output**: allowing the system to indicate the effects of the users' manipulation.

System development cycle



System development cycle

- Designers need to check whether the their ideas are really what users need/want; or whether the final product works as expected.
- To do that, we need some form of methods, or more specifically, empirical methods for HCI.

Usability

- Usability refers to the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of user
- Usability measures the quality of a user's experience when interacting with a product or system
  - Ease of learning
  - Efficiency of use
  - Memorability
  - Error frequency and severity
  - Subjective satisfaction

Why usability evaluation?

- Identify any usability problems that the product has.
- Collect empirical data on users' performance.
- Determine users' satisfaction with the product.

- Test early;
- Test often.

Outline

- What is HCI
- **Evaluation framework**
- Evaluation methods
- Case studies

DECIDE: a framework to guide evaluation

- Determine the goals.
- Explore the questions.
- Choose the evaluation approach and methods.
- Identify the practical issues.
- Decide how to deal with the ethical issues.
- Evaluate, analyze, interpret and present the data.

Determine the goals

- What are the high-level goals of the evaluation?
- Who wants it and why?
- The goals influence the approach used for the study.
- Some examples of goals:
  - Check to ensure that the final interface is consistent.
  - Investigate how technology affects working practices.
  - Improve the usability of an existing product .

Explore the questions

- All evaluations need goals & questions to guide them.
- E.g., the goal of finding out why many customers prefer to purchase paper airline tickets rather than e-tickets can be broken down into sub-questions:
  - What are customers’ attitudes to these new tickets?
  - Are they concerned about security?
  - Is the interface for obtaining them poor?

Choose the evaluation approach & methods

- The evaluation approach influences the methods used, and in turn, how data is collected, analyzed and presented.

- E.g., field studies typically:
  - Involve observation and interviews.
  - Do not involve controlled tests in a laboratory.
  - Produce qualitative data.

Identify practical issues

- For example, how to:
  - Select users
  - Stay on budget
  - Stay on schedule
  - Find evaluators
  - Select equipment

Decide about ethical issues

- Develop an informed consent form

- Users have a right to:
  - Know the goals of the study;
  - Know what will happen to the findings;
  - Privacy of personal information;
  - Leave when they wish;
  - Be treated politely.

Evaluate, interpret & present data

- The approach and methods used influence how data is evaluated, interpreted and presented.
- The following need to be considered:
  - Reliability: can the study be replicated?
  - Validity: is it measuring what you expected?
  - Biases: is the process creating biases?
  - Scope: can the findings be generalized?
  - Ecological validity: is the environment influencing the findings?

Key points

- There are many issues to consider before conducting an evaluation study.
- These include the goals of the study, the approaches and methods to use, practical issues, ethical issues, and how the data will be collected, analyzed and presented.
- The DECIDE framework provides a useful checklist for planning an evaluation study.

Outline

- What is HCI
- Evaluation framework
- **Evaluation methods**
- Case studies

Usability evaluation

- Survey:
  - Interview
  - Questionnaire
  - Focused group
- Analytic inspection:
  - Heuristic Evaluation
    - Principles, Guidelines
  - Cognitive walkthroughs
    - Based on task scenarios
- Empirical evaluation:
  - Observational experiment
    - Observation, problem identification
  - Controlled Experiment
    - Formal controlled scientific experiment
    - Comparisons, statistical analysis

Evaluation techniques

- Which technique(s) to use?
  - Depends on testing purposes
  - Depends on the stage in the development cycle
  - Depends on resources available
    - E.g. time, availability or expertise & equipment, access to users
  - Can be used in combination

Usability evaluation

- **Survey:**
  - Interview
  - Questionnaire
  - Focused group
- Analytic inspection:
  - Heuristic Evaluation
    - Principles, Guidelines
  - Cognitive walkthroughs
    - Based on task scenarios
- Empirical evaluation:
  - Observational experiment
    - Observation, problem identification
  - Controlled Experiment
    - Formal controlled scientific experiment
    - Comparisons, statistical analysis

Survey

- A technique of gathering data by asking questions to people who are thought to have desired information.
  - Surveys may be conducted once; at repeated intervals, or concurrently with multiple samples
  - They may be used to collect information from a few or many
- The purpose is to make statistical inferences about the population being studied
- A single survey is made of at least:
  - a sample (or full population in the case of a census),
  - **a method of data collection** (e.g., a questionnaire) and
  - individual questions or items that become data that can be analysed statistically.

Interview

- Qualitative technique
  - Gathering information about users by talking directly to them
  - A method for discovering facts and opinions of the users.
- Format:
  - It is usually done by one interviewer speaking to one user at a time.
  - Structured interviews: a pre-defined set of questions and users
  - Open-ended interviews: allows for an exploratory approach to uncover unexpected information.
- Problems:
  - The unstructured nature of the resulting data can be easily misinterpreted.

Questionnaire

- Qualitative technique
  - But results can be quantified
- Preparation
  - Keep questions simple, be clear and concise
  - Group questions appropriately & give explanation
- Pilot questionnaire before distributing it
  - It is still unreasonable to think that any one person can anticipate all the potential problems
- Problems
  - It is only as good as the questions it contains

Question types

- Open-ended:
  - What are the features you think helpful, if any?
  - What are the features you think can be improved, if any?
- Closed:
  - Which of the following have you used? (tick all that apply)

    1) word processor 2) database 3) spreadsheet
  - How easy was it to understand the drawing?

    1) Very easy 2) Easy 3) Average 4) Difficult 5) Very Difficult

Question types

- Scale: are generally used to measure opinions, attitudes, feelings and habits.
  - Please indicted how much effort you devoted for this task based on a scale from 0-6?

    0 (extremely easy) 1 2 3 4 5 6 (extremely difficult)
- Ranking: are used to determine degrees of importance.
  - Indicate the importance you attach to the following 'job motivators' by writing the appropriate number (1-5) (1 = most important, 5 = least important)

    _____ Recognition for good work
    _____ A say in decision making
    _____ A good chance for promotion
    _____ Good pay and benefits
    _____ Fair treatment from the company

Questionnaire

- Established questionnaires will give more reliable and repeatable results than ad hoc questionnaires.
- Three questionnaire for assessing the perceived usability of an interactive system:
  - Questionnaire for User Interface Satisfaction (QUIS) (1988)
  - Computer System Usability Questionnaire (CSUQ) (1995)
  - System Usability Scale (SUS) (1996)

Focused group

- A focus group is a small-group discussion guided by a trained leader (moderator) in an interactive setting.
- It is used to learn more about opinions on a designated topic, or a product, and then to guide future action.
- A focus group:
  - is an interview,
  - conducted by a trained moderator among a small group of respondents,
  - conducted in an unstructured and natural way where respondents are free to give views from any aspect.

Pros and Cons

- Pros
  - It is flexible.
  - It generates quick results.
  - It costs little to conduct.
  - Group dynamics often bring out aspects of the topic or reveal information about the subject that may not have been anticipated by the researcher or emerged from individual interviews.
- Cons
  - The researcher has less control over the session than he or she does in individual interviews.
  - Data are often difficult to analyse.
  - Moderators require certain skills.

Usability evaluation

- Survey:
  - Interview
  - Questionnaire
  - Focused group
- **Analytic inspection:**
  - Heuristic Evaluation
    - Principles, Guidelines
  - Cognitive walkthroughs
    - Based on task scenarios
- Empirical evaluation:
  - Observational experiment
    - Observation, problem identification
  - Controlled Experiment
    - Formal controlled scientific experiment
    - Comparisons, statistical analysis

Analytic inspection

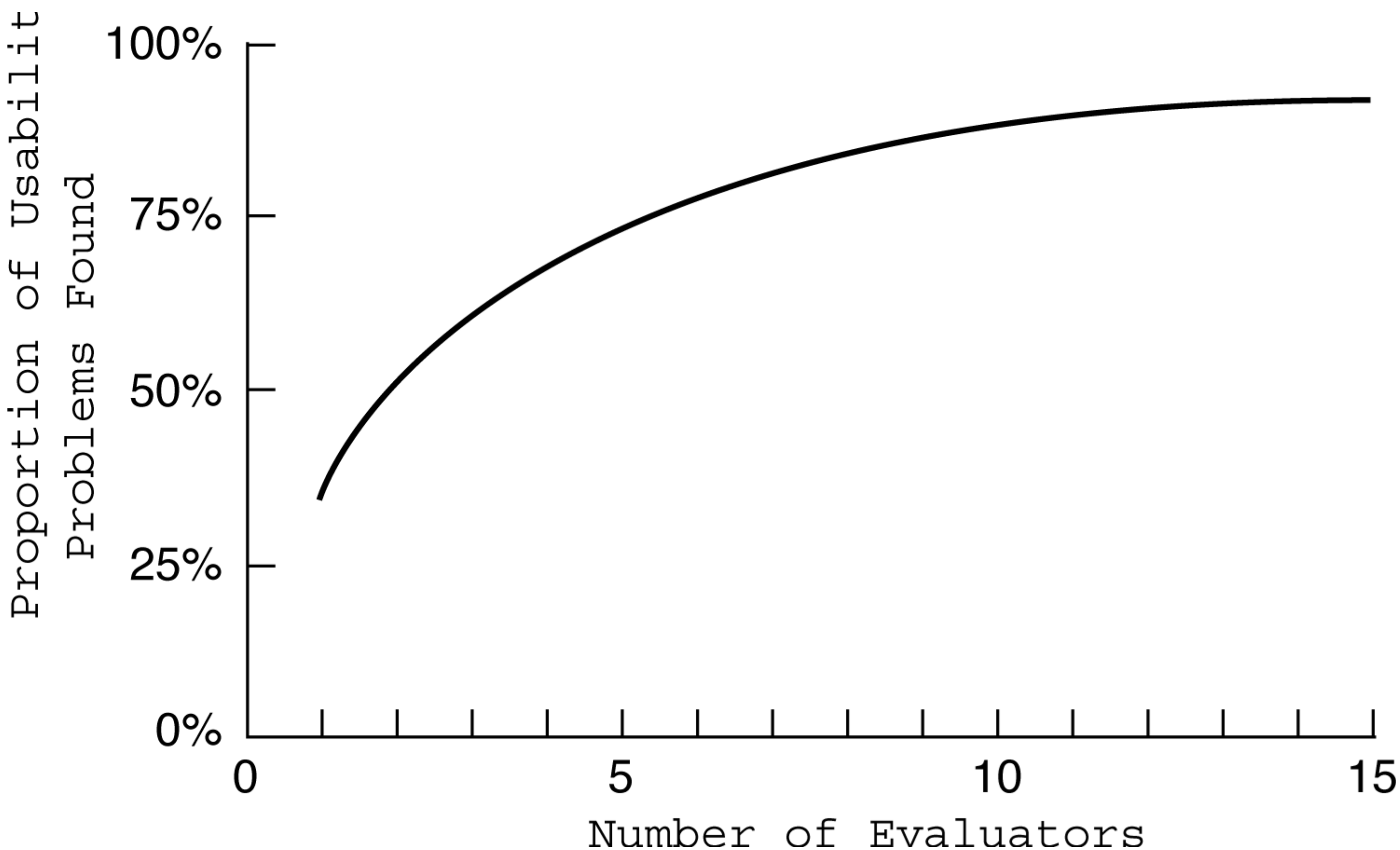
- Experts use their knowledge of users & technology to review system usability.
- Expert critiques can be formal or informal reports.
- Benefits:
  - Generate results quickly with low cost.
  - Can be used early in the design phases

- **Heuristic evaluation is a review guided by a set of heuristics.**
- **Walkthroughs involve stepping through a pre-planned scenario and noting potential problems.**

Nielsen’s heuristics

- Visibility of system status.
- Match between system and real world.
- User control and freedom.
- Consistency and standards.
- Error prevention.
- Recognition rather than recall.
- Flexibility and efficiency of use.
- Aesthetic and minimalist design.
- Help users recognize, diagnose, recover from errors.
- Help and documentation.

No. of evaluators & problems



3 stages for doing heuristic evaluation

- Briefing session to tell experts what to do.
- Evaluation period of 1-2 hours in which:
  - Each expert works separately;
  - Some time is used to get a feel for the product;
  - Other time is used to focus on specific features.
- Debriefing session in which experts work together to prioritize problems.

Advantages and problems

- Few ethical & practical issues to consider because users not involved.
- Can also be difficult & expensive to find experts.
- Best experts have knowledge of application domain & users.
- Biggest problems:
  - Important problems may get missed;
  - Many trivial problems are often identified;
  - Experts have biases; they are not real users.

Cognitive walkthroughs

- Focus on ease of learning.
- Designer presents an aspect of the design & usage scenarios.
- Expert is told the assumptions about user population, context of use, task details.
- One or more experts walk through the design prototype with the scenario.
- Experts are guided by 3 questions.

The 3 questions

- Will the user try to achieve the effect that the subtask has?
- Will the user notice that the correct action is available?
- Will the user associate and interpret the response from the action correctly?

As the experts work through the scenario they note problems.

Advantages

- May be done without first hand access to users.
- Unlike some usability inspection methods, takes explicit account of the user's task.
- Provides suggestions on how to improves learnability of the system
- Can be applied during any phase of development.
- Is quick and inexpensive to apply if done in a streamlined form.

Disadvantages

- The value of the data is limited by the skills of the evaluators.
- Tends to yield a relatively superficial and narrow analysis that focuses on the words and graphics used on the screen.
- The method does not provide an estimate on the frequency or severity of identified problems.
- Following the method exactly as outlined in the research is labour intensive.

Heuristic Evaluation vs Cognitive Walkthrough

- Heuristic Evaluation
  - How: guide by existing design principles and heuristics
  - What: examines if the system in question abides by recognized usability principles
  - Why: to see if the system can comfortably be used based on prior experience in similar systems
- Cognitive Walkthrough
  - How: guide by asking questions from the perspective of the user.
  - What: performs list of specific tasks
  - Why: to see if the tasks can be performed in the correct sequence of actions they were designed in

Usability evaluation

- Survey:
  - Interview
  - Questionnaire
  - Focused group
- Analytic inspection:
  - Heuristic Evaluation
    - Principles, Guidelines
  - Cognitive walkthroughs
    - Based on task scenarios
- **Empirical evaluation:**
  - Observational experiment
    - Observation, problem identification
  - Controlled Experiment
    - Formal controlled scientific experiment
    - Comparisons, statistical analysis

Empirical evaluation

- **Observational experiment**
  - Observation, problem identification
- Controlled Experiment
  - Formal controlled scientific experiment
  - Comparisons, statistical analysis

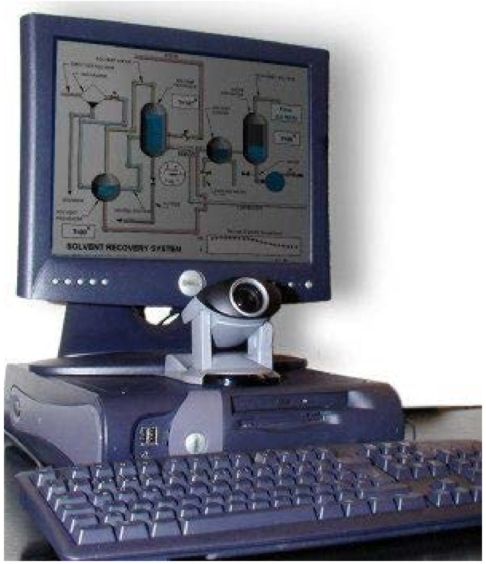
Observational testing

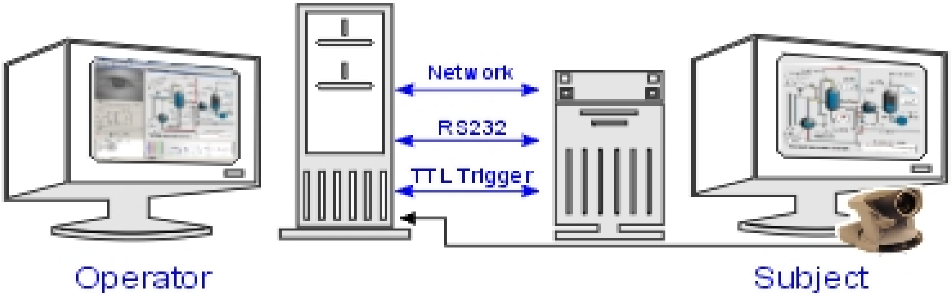
- In an observational usability test, representative users try to do typical tasks with the product, while observers, including the development staff, watch, listen, and take notes.
- Examples:
  - Think aloud
  - Eye tracking

Think aloud

- Users are asked to say whatever they are looking at, thinking, doing, and feeling, as they go about their task.
- Observers are asked to objectively take notes of everything that users say, without attempting to interpret their actions and words.
- Test sessions are often audio and video taped.

Eye tracking

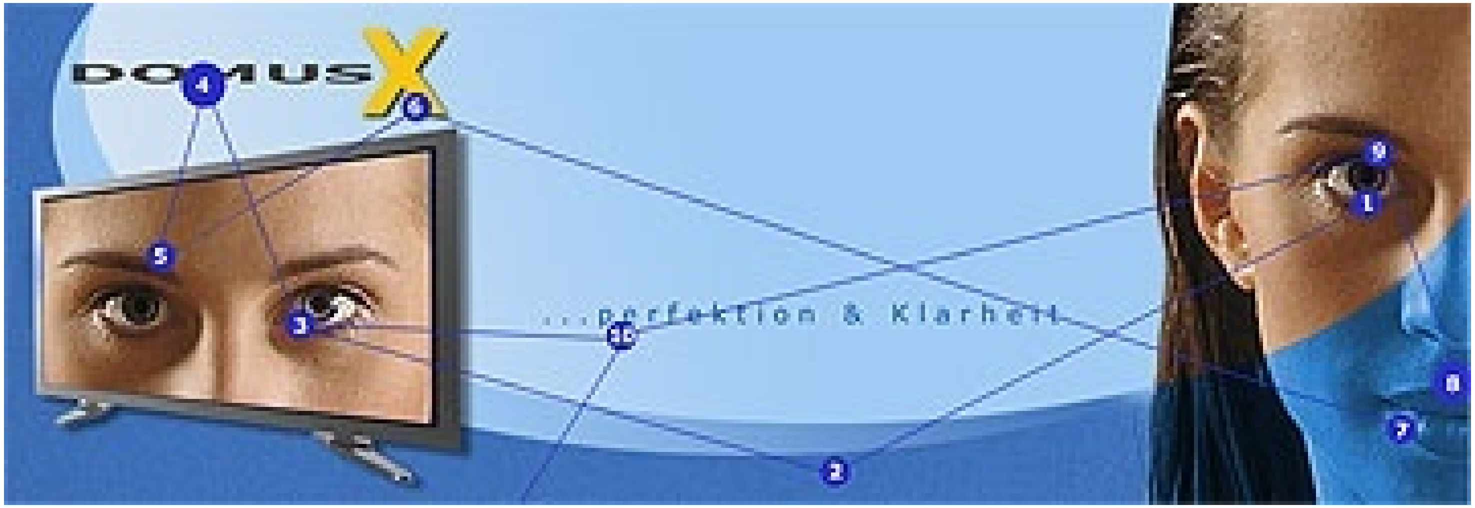




Pictures from SensoMotoric Instruments (SMI)

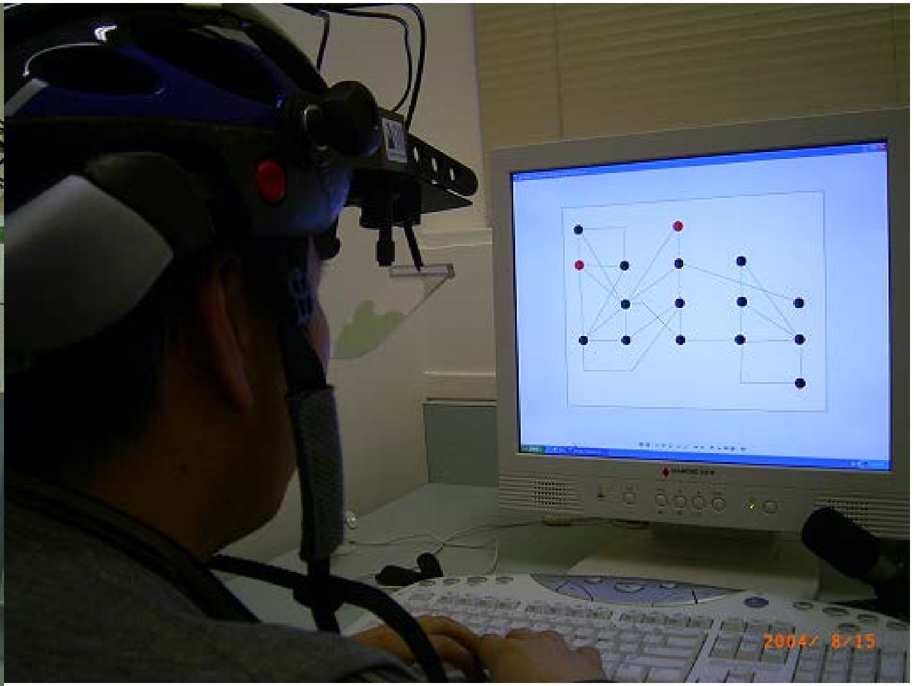
Eye tracking

- Fixations and saccades
- Scanpath: the resulting series of fixations and saccades.



Picture from Internet

Eye tracker at the CHAI lab



Empirical evaluation

- Observational experiment
  - Observation, problem identification
- **Controlled Experiment**
  - Formal controlled scientific experiment
  - Comparisons, statistical analysis

Observational test vs. controlled experiment

- Observational test
  - Formative: helps guide design
  - Single UI, early in design process
  - Few subjects
  - Identify usability problems
  - Qualitative feedback from users
- **Controlled experiment**
  - Summative: measure final result
  - Compare multiple UIs
  - Many subjects, strict protocol
  - Quantitative results, statistical significance

Controlled experiment

- It is a test of the effect of a single variable by changing it while keeping all other variables the same.
- A controlled experiment generally compares the results obtained from an experimental sample against a control sample.
- General terms
  - Participant (subject)
  - Independent variables (test conditions), dependent variable,
  - Control variable, random variable
  - Confounding variable
  - Within subjects vs. between subjects
  - Counterbalancing, Latin square

Independent Variable

- Independent Variable (IV, what you very)
  - Independent of participant behavior
  - Examples: interface, visual layout, gender, age
- Test conditions: levels, or value of an IV
- Provide a name for both IV and its levels (test conditions)

IV	Conditions (levels)
Device	Mouse, trackball, joystick
visualization	2d, 3d, animated
Search engine	Google, Bing
Feedback mood	Audio, video, tactile, none

Dependent variable

- Dependent Variable (DV, what you measure)
  - User performance time
  - Accuracy, errors
  - Subjective satisfaction

Confounding variable

- A confounding variable is one that provides an alternative explanation for the thing we are trying to explain with our IVs.
- Example: we want to compare two systems (windows 7 vs. windows 8)
  - All participants have prior experience with windows 7, but no experience with windows 8
  - “Prior experience” is a confounding variable
- A major issue in observation studies is that we often don't always know what the potential confounding factors may be.

General process

1. Determine research questions
2. Start with a testable hypothesis
   – Interface X is faster than interface Y
3. Manipulate independent variables
   – different interfaces, tasks
4. Measure dependent variables
   – times, errors, satisfaction
5. Use statistical tests to accept or reject the hypothesis

Hypothesis Testing

- How to “prove” a hypotheses in science?
    - In most cases, it is impossible to prove the hypothesis directly.
- This is done by disproving the null hypothesis.
    - Easier to disprove things, by counter-example
    - First we suppose the null hypothesis true: Null hypothesis=opposite of hypothesis
    - Then a conflicting result is found
    - Disprove the null hypothesis
    - Hence, the hypothesis is proved

Population and sample

- Statistical tests calculate the probability that the pattern of results we have observed in our data could have arisen by chance alone.

population

IT students are better than psychology students in memorizing numbers?

sample

100 IT students and 100 psychology students were tested

This is true for the sample. What about population?

Mean for IT = 80;
Mean for psychology =60

Two possibilities for the population:

1) There really is a difference in the population, or

2) There is no difference in the population.

Statistical tests allow us to decide which of the two possible explanations is the most likely by calculating a p value.

P value

- Probability refers to the likelihood of any given event occurring out of all possible events.
- It tells us how likely is an event to occur by chance.
- How likely is “very likely”?
  - By convention, we consider “significant” those differences that occur less than .05 (5%) by chance alone.

# $P < 0.05$

- Woohoo!
- E.g. $p=0.001$
- $P<0.05$
- Found a “statistically significant” difference
- Effect is likely to be resulted from IV
- Null hypothesis rejected
- Hypothesis confirmed

$P > 0.05$

- e.g. $p=0.25$ in an experiment
- $P>0.05$ significance level
- So we conclude effect could have happened by chance
- Cannot say that IV effected DV

P > 0.05

- Hence, no difference?
- Not sure
  - Did not detect a difference, but could still be different
  - Potential real effect did not overcome random variation
  - Boring, basically found nothing
- How?
  - Not enough users
  - Measuring of DVs are not accurate enough
  - Need better tasks, data, …

Experimental design

- Main types:
  - Between-subjects (B-S)
  - Within-subjects (W-S)
- Between-subjects
  - One subject is exposed to only one condition, and contributes one entry to the whole data.
  - Subjects are randomly allocated to one of the conditions

Condition 1	Condition 2
subject 1	subject 13
subject 2	subject 14
….	….
subject 12	subject 24

Experimental design

- Within-subjects
  - Each subject is exposed to each of the conditions
  - Each subject contributes one entry for each of the conditions
- Usually, WS design is used in HCI evaluations, and requires 10-30 subjects

Condition 1	Condition 2
subject 1	subject 1
subject 2	subject 2
….	….
subject 12	subject 12

Within-subjects design

- Advantages
  - fewer subjects
  - Less time
  - Less expensive

  - Increased control of subjects variability: comparisons between conditions happen within each subject.
  - More power to detect significant difference

Within-subjects design

- Disadvantages
  - Learning effect
  - Carryover effect
  - Fatigue, boredom
- Solutions:
  - More practice before testing; randomization; counterbalancing (Latin square)

A	B	C
B	C	A
C	A	B

  - Have rest between tasks
  - Limit the testing time

Parametric vs. non-parametric

- There are two types of statistical methods.
- Parametric
  - Based on highly restrictive assumptions about type of data. For example: normal distribution
  - Compare means
- Non parametric
  - More general purpose
  - Compare medians
- Parametric methods are more powerful in detecting significant differences, but have more strict assumptions on data.

Statistical methods

- How do you choose which method to use?
- Depends on
  - Experimental design
  - Distribution of DVs
  - Levels of IVs (conditions)
- It is important to use the right method, or the results will be little meaningful.

	1 condition	**2 conditions**		**3 or more conditions**	
Parametric	t test	t test	paired t test	B-S ANOVA	W-S ANOVA
Non-parametric	Wilcoxon rank sum	Mann-Whitney	Wilcoxon	Krusakl-Wallis	Friedman

Validity

- Validity is concerned with the study's success at measuring what the researchers set out to measure.
  - Internal Validity: the extent to which a causal conclusion based on a study is warranted
  - External Validity: the extent to which the results of a study can be generalized to other situations and to other people
  - Ecological validity: the extent to which the conditions simulated in the laboratory reflect real life conditions.
  - Construct Validity: the extent to which a test measures what it claims, or purports, to be measuring
  - Statistical validity: The extent to which an observed result, such as a difference between two measurements, can be relied upon and not attributed to random error in sampling or in measurement.

Points to note

- Pilot study is essential and critical for success.
- Make sure each experimental run is as similar as possible to all of the subjects.

- Put subjects at their ease.
- Test experimental variables, not subjects.
- Subjects may need to be motivated in some way.

- No experiments are perfect! Results obtained should be interpreted within the limitations.

Ethics

- Should apply for ethics approval from research office of USYD before starting your experiment.
- Information sheet
  - Voluntary
  - Withdraw at any time without penalty
  - Contact details: complaints, questions
- Participant consent form: Named and signed
- You must not cause your subjects distress: Invading privacy, physical abuse, unpleasant emotions, etc.
- Original material produced by subjects must be kept confidential.
- In reports, subjects should not be identifiable in any way.

Outline

- What is HCI
- Evaluation framework
- Evaluation methods
- **Case studies**

Graph Visualization

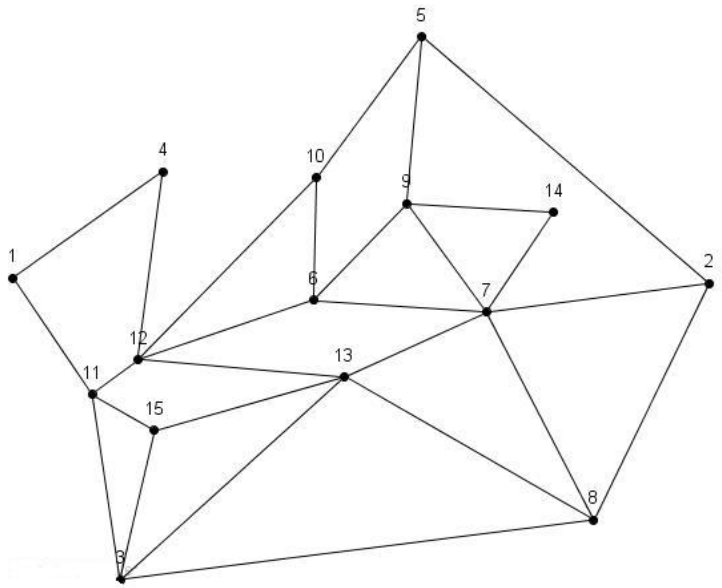
- Efficiency: The running time of algorithms should be reasonably fast.
- Elegance: algorithms should be easy to understand and easy to code; final drawings should be beautiful.
- **Effectiveness:** graph viewers should understand the underlying data quickly and correctly.

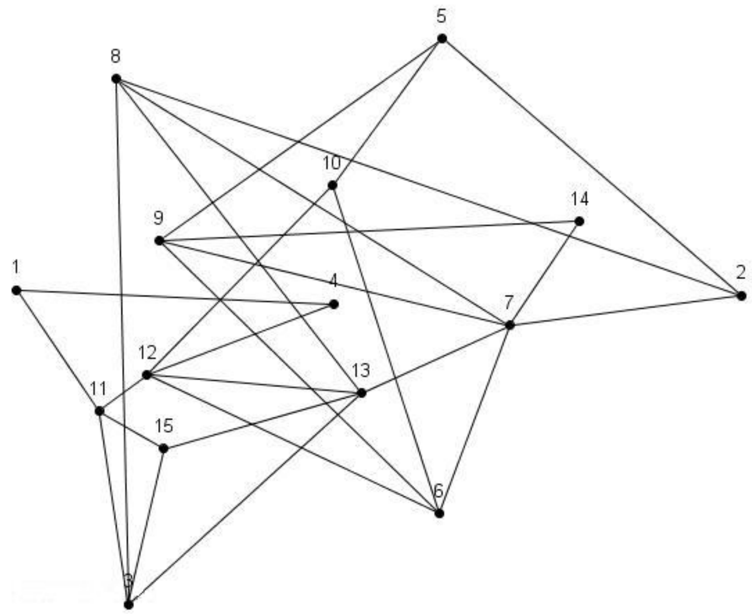
Effectiveness

- Visualization designers are often satisfied with the “coolness” of the technologies they introduced.
  - technology that looks “cool” to the designer might be too complex or superfluous for real users
- It is assumed that graphs should be effective when drawn conforming to some predefined criteria.
  - Maximise symmetry
  - Minimise edge crosses
  - Maximise angular resolution
- However, common senses and intuitions are not reliable.
  - Users, data sets, tasks are different

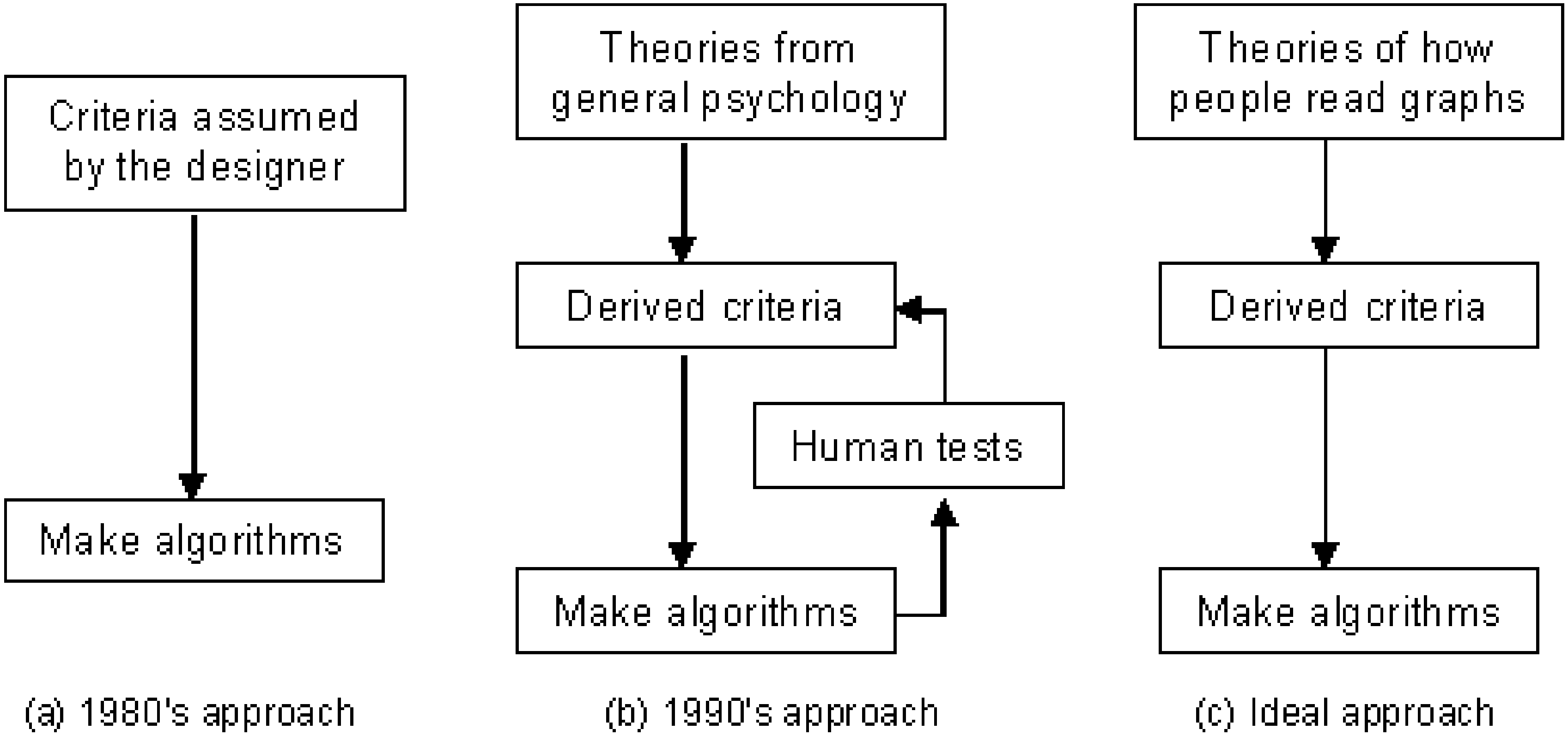
Which one is better?

- We do not know until we actually evaluate them.



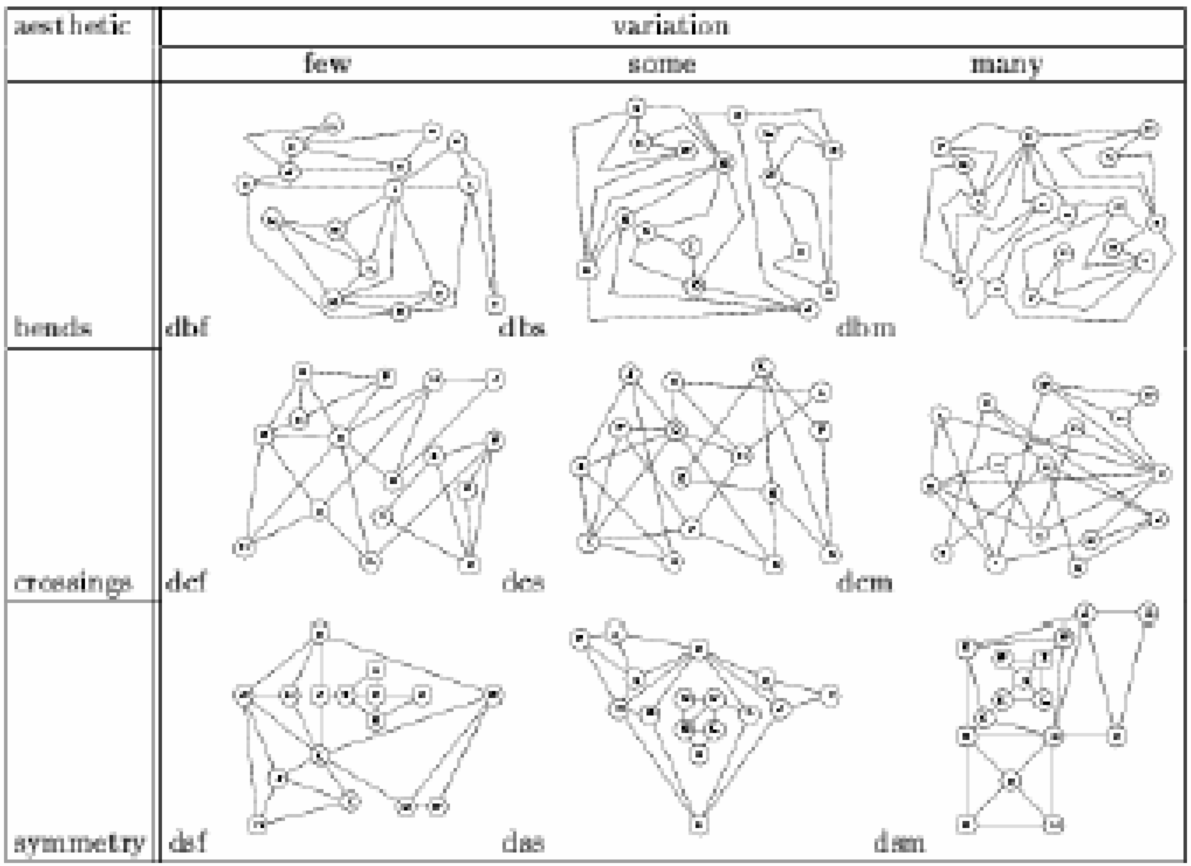


History of graph visualization technology



Controlled experiment on performance

- Purchase et al. (1995) Validating graph drawing aesthetics. GD95, 435-446.



Design and analysis

- One dense graph
- IV1: bends, crosses, symmetry
- IV2: high/medium/low presence
- DV: accuracy in fixed time (45 seconds)
- Task: find the shortest path between two nodes

- Paper based, within-subjects, random order, a “filler” task

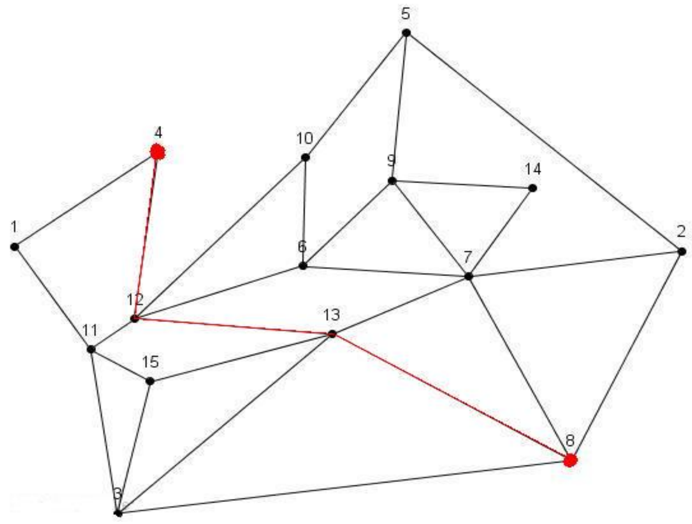
- Friedman ANOVA.

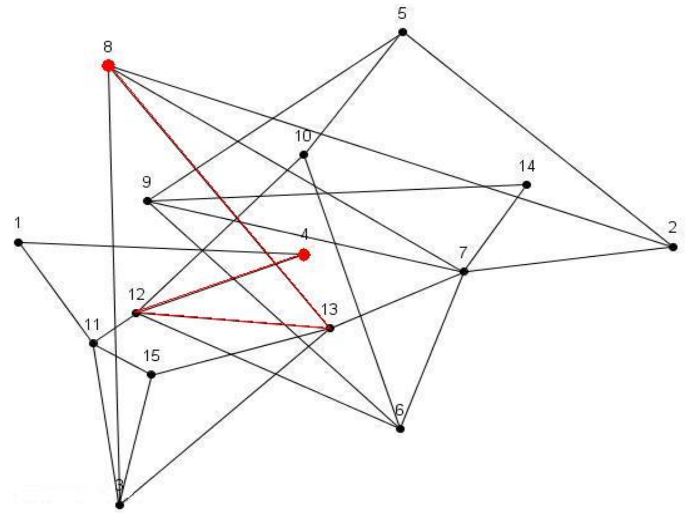
Results and discussion

- Bends and crossings are important. Symmetry needs further examination.
- A pioneering work that provides empirical evidence for intuition-based aesthetic criteria.

Limitations

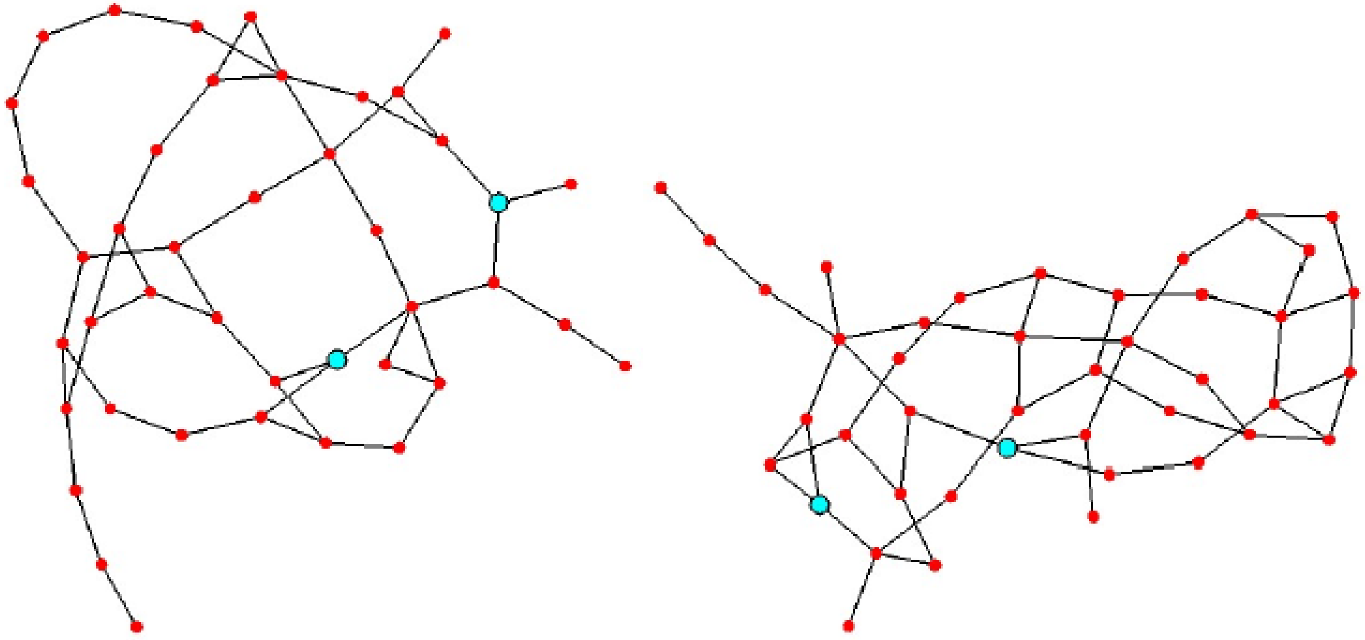
- Graphs are difficult to control (confounding factors)
  - Elements are interconnected to each other.
- Change in on criteria can lead to change in another.
- For example, making more crossings can also make more sharp turns on the paths.
  - Gestalt principle of continuity





An experiment without manipulating IVs

- Ware et al. (2003) Cognitive measurements of graph aesthetics. Information visualization. 1(2), 103-110.



Design and analysis

- Create a set of random graphs, drawn with a spring algorithm
- Highlight two nodes, record the measurements of the predictor variables
  - **Continuity, number of crossings, crossing angles, number of branches, shortest path length….**
- Record the response variable
  - Response time
- Regress response variable on predictor variables to detect their relationships.
- Task: shortest path between two highlighted nodes
- Analysis: **Correlation and multiple regression**

Results and discussion

	spl	con	cr	aca	tcr	br	all	tll	rt
spl	1	0.484	0.191	0.134	0.059	0.379	0.064	0.930	0.736
con		1	0.019	0.082	0.125	0.119	-0.294	0.331	0.633
cr			1	0.141	0.347	0.267	0.428	0.332	0.449
aca				1	0.064	0.208	0.099	0.167	0.148
tcr					1	0.116	0.011	0.066	0.216
br						1	0.353	0.475	0.462
all							1	0.419	0.050
tll								1	0.623
rt									1

$$rt = 0.414spl + 0.406con + 0.317cr + 0.172br$$

- Path continuity (con) and number of crossings (cr) on the shortest path are important factors.

Beyond time and error

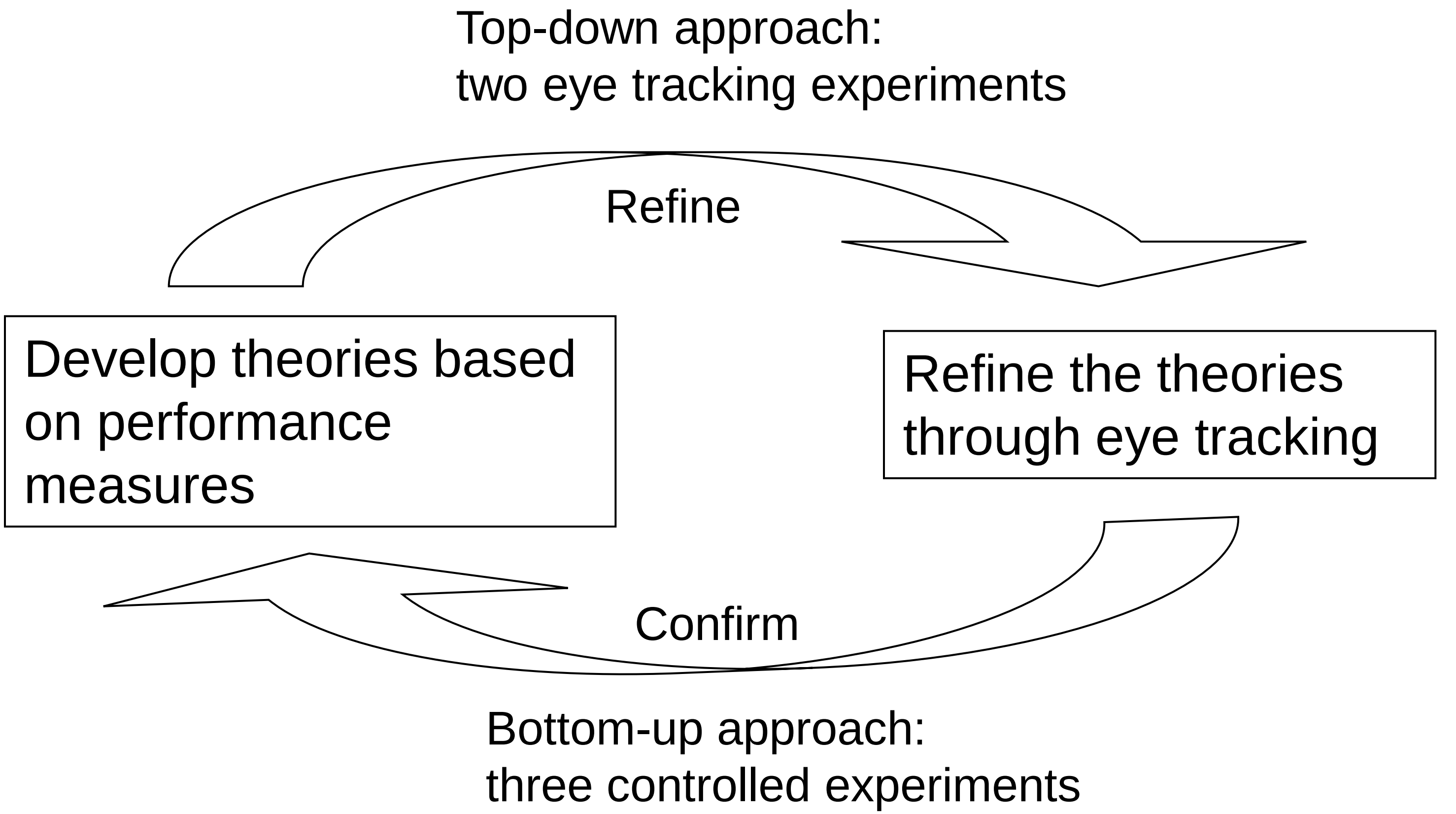
- Where the time is spent and how the performance is affected?
- What is the mechanism of crossings affecting performance?

- Time and error performance logging
  - treat the human as a “black box”, which tell us ***what***, but not ***how*** and ***why***

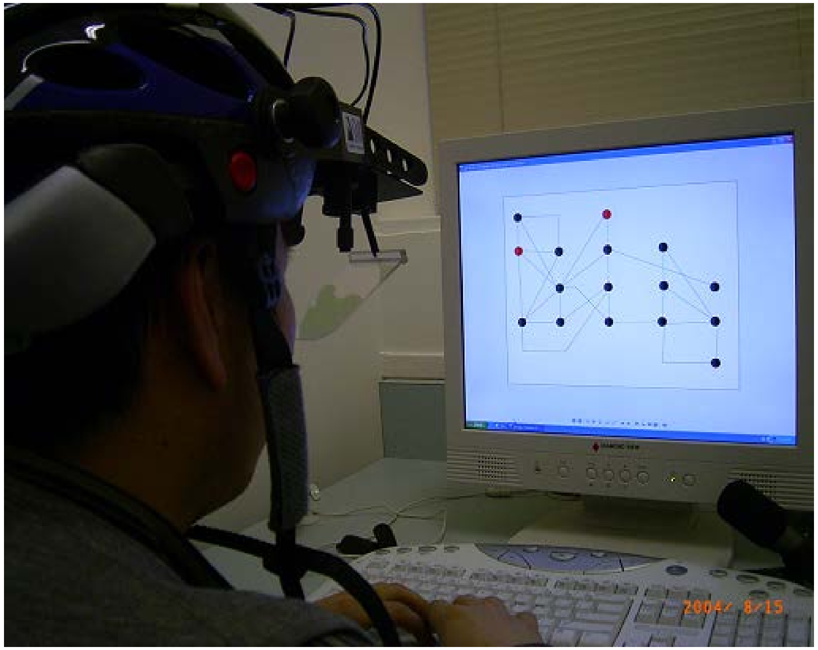
- Eye tracking may give insight as to ***how***
- Post-interview and questionnaire tell us ***why***

- Two exploratory eye tracking experiments
  - Ex1: small and sparse graphs
  - Ex2: larger and denser graphs
- Three confirmatory controlled experiments
  - Ex3a: existence of geodesic-path tendency
  - Ex3b: effects of geodesic-path tendency
  - Ex4: effects of crossing angles

Research Methodology



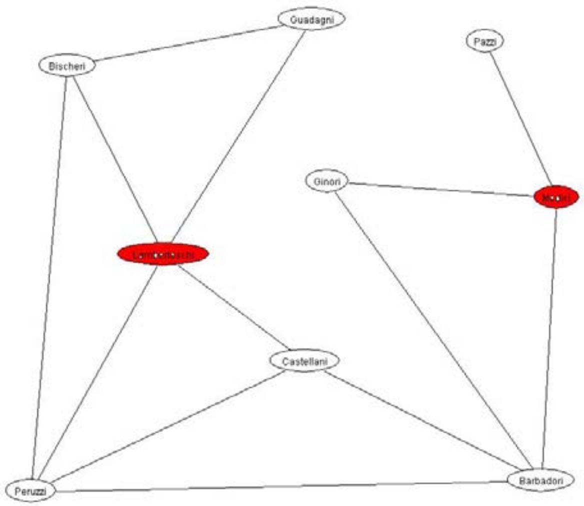
Eye Tracker

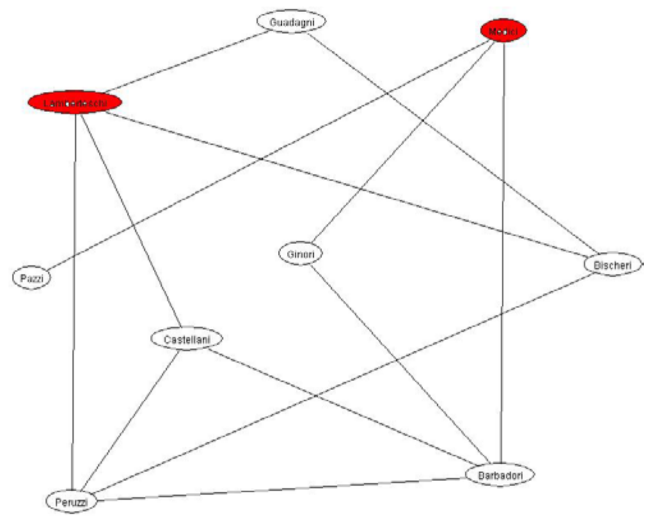


Experiment 1

- Task: find the shortest path between two highlighted nodes.
- Time, error and eye movements were recorded.
- Questionnaires and interviews.

Examples of Stimuli



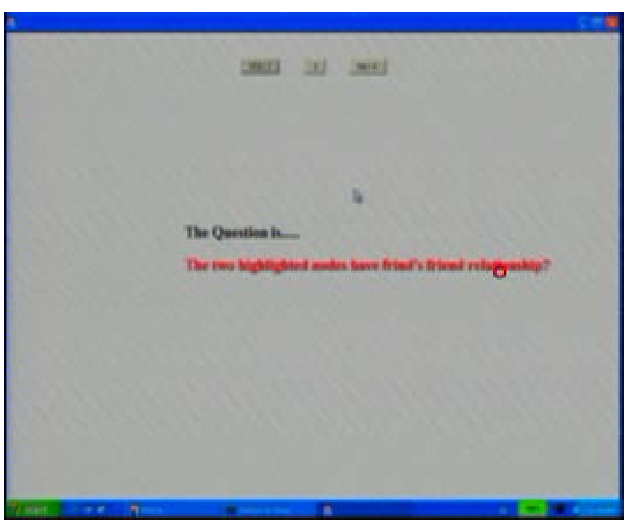


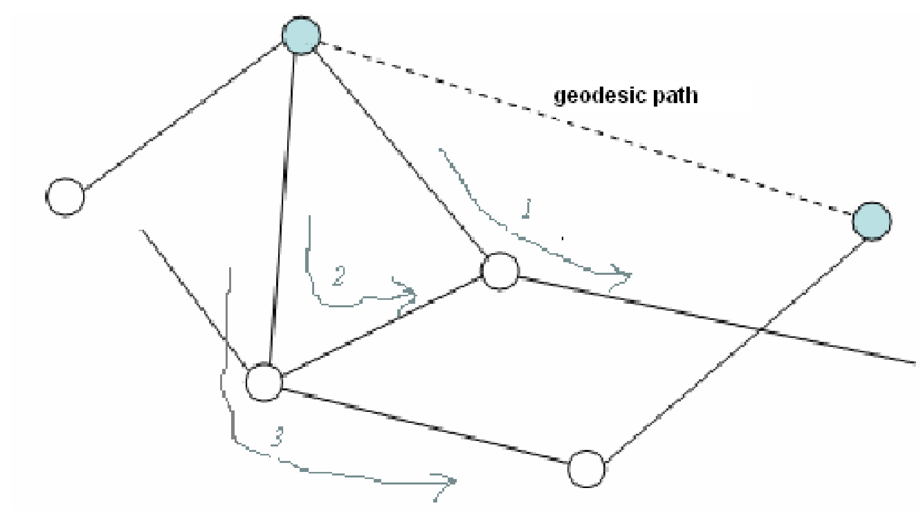
Results: Time and Error

- Overall, subjects spent significantly more time with crossing drawings than with non-crossings
- However, on some specific instances, this was not the case

Results: Eye Tracking Video Data

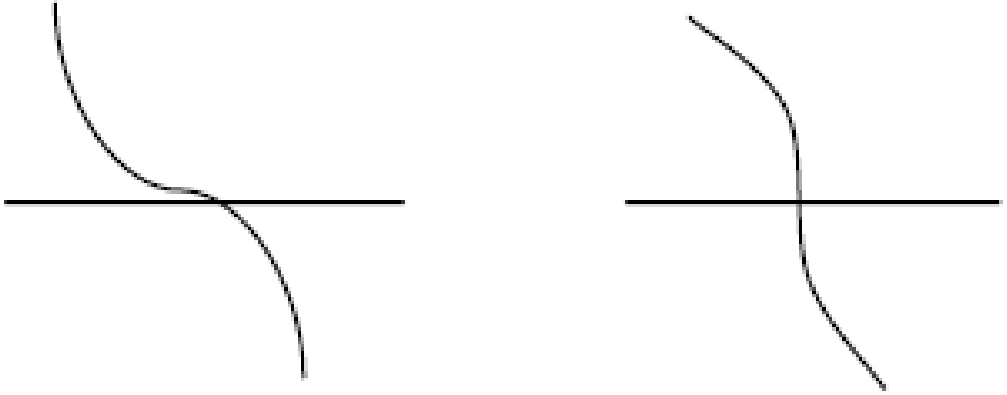
- Crossings had little impact on eye movements.
- **Geodesic-path tendency**: subjects seemed to follow the geodesic path between the current node and target node.





Possible Reasons for the Lack of Crossing Effects

- ***Crossing angles*** may inhibit readability [Ware et al. 2003].

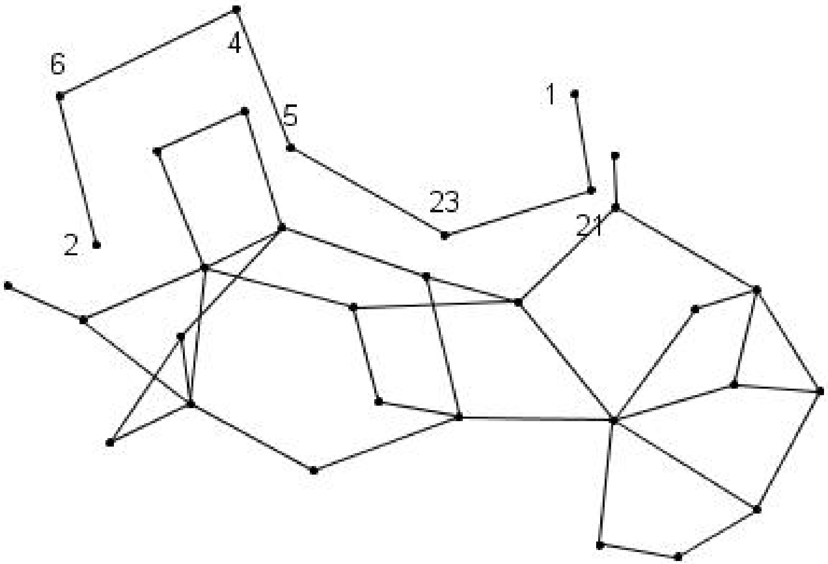


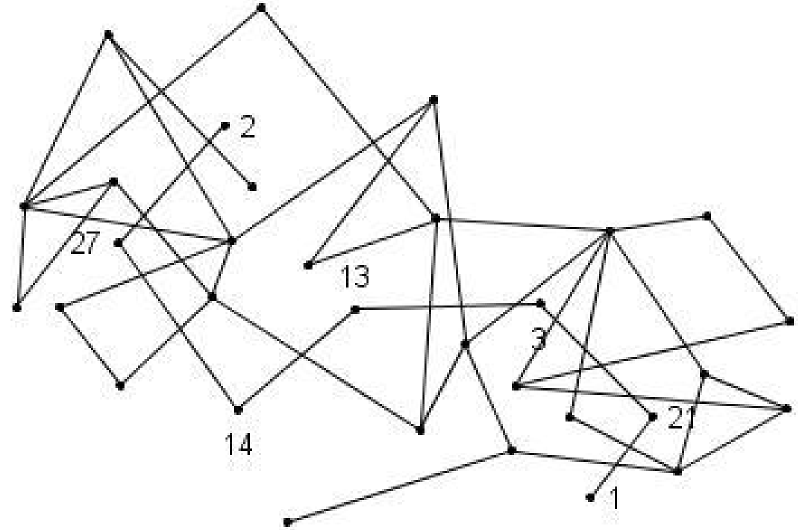
Experiment 2

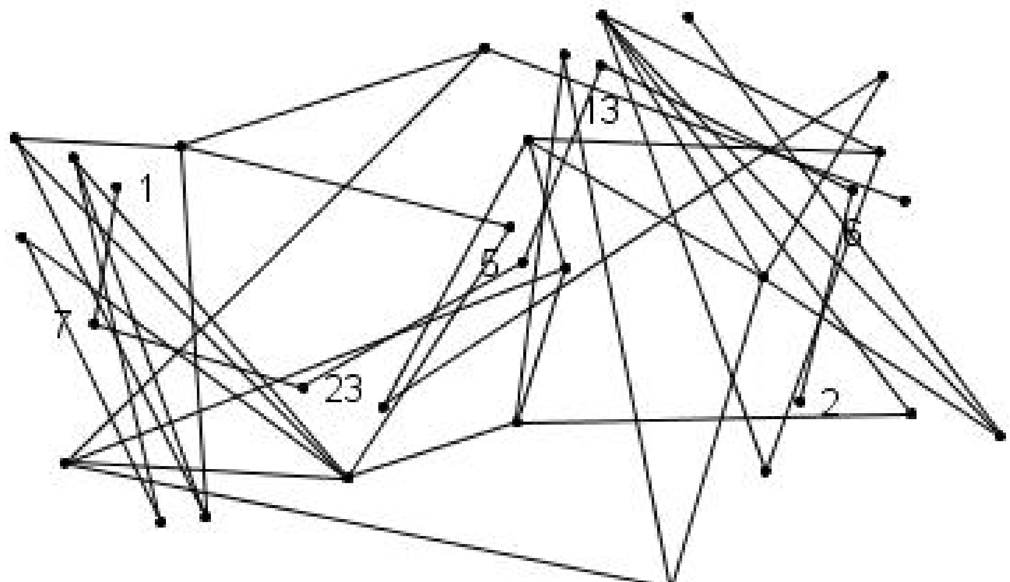
Crossing angle: graphs were drawn with three conditions:

- No crossings on the path
- Small-angle crossings
- Large-angle crossings

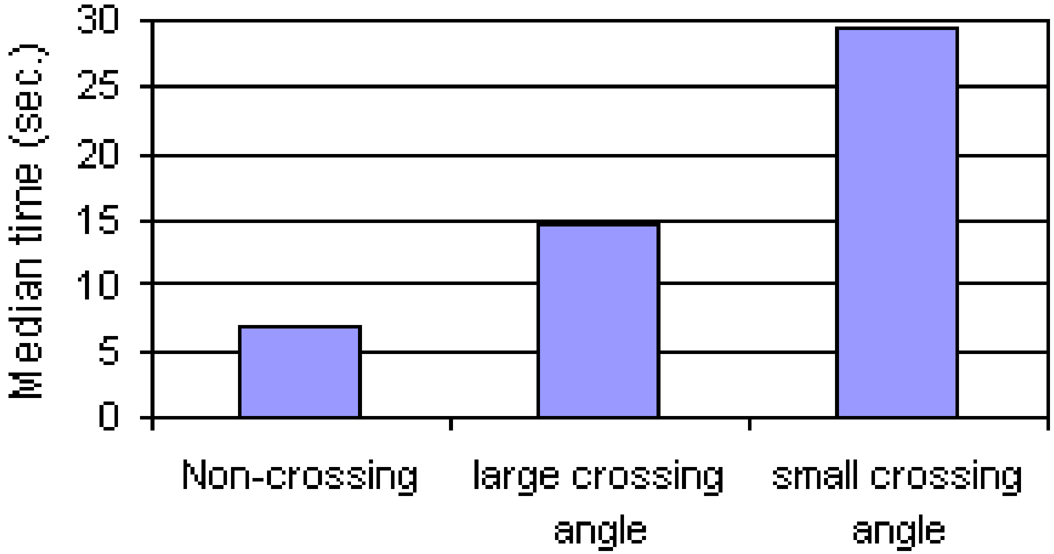
Stimuli







Results



- Effects of crossing angles were significant on time

Results: Eye Tracking Video Data

No crossings: eye movements were smooth and fast.

Large crossing angle: eye movements were still smooth, but slower.

Small crossing angle: eye movements were very slow and no longer smooth (back-forth moves at crossing points).

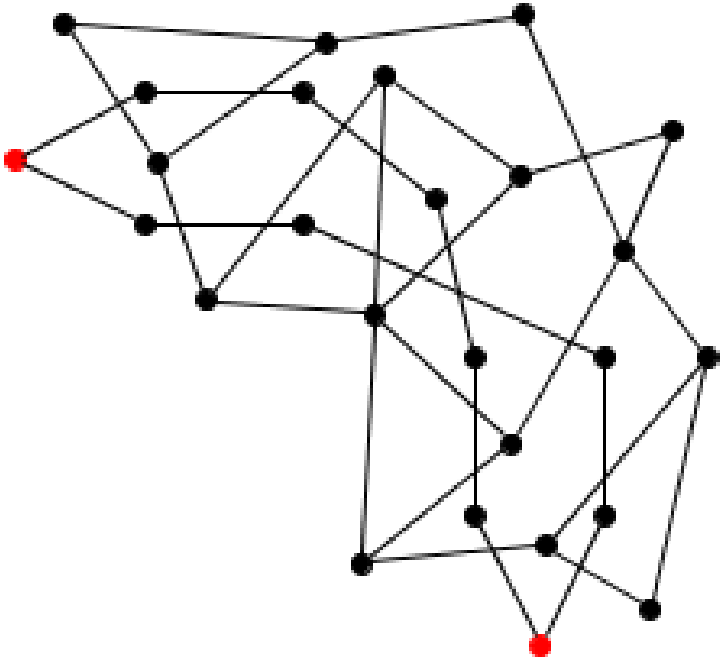
An Example


The Question is....
what is the shortest path between A and B ?
Nodes A and B are already highlighted for you.

- Observations of eye-tracking videos need ***<u>confirmation</u>***:
  - Existence of geodesic-path tendency (Ex3a).
  - Effects of geodesic-path tendency (Ex3b).
  - Effects of crossing angles (Ex4).

Experiment 3a: Existence of geodesic-path tendency

- Two separate paths between the two highlighted nodes



Results

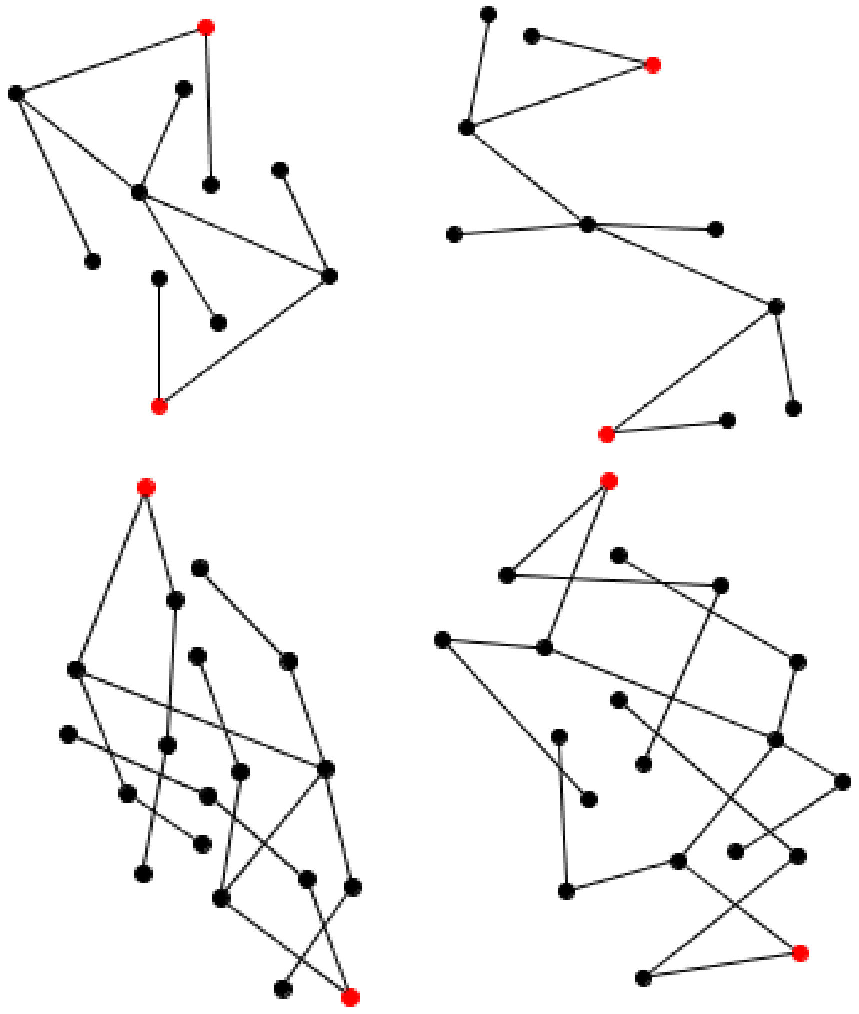
- Subjects followed geodesic-closest path 75% of the time
- People have a “***geodesic-path tendency***”

An Example

Ex3b: Effects of Geodesic-path Tendency

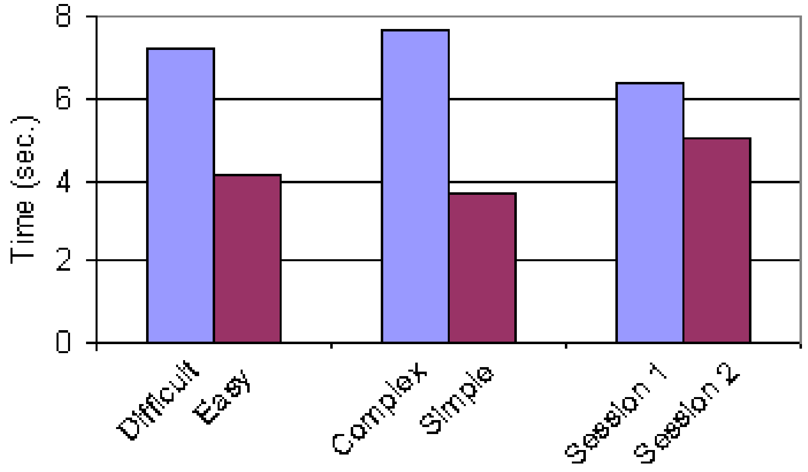
Each graph was drawn in two ways:

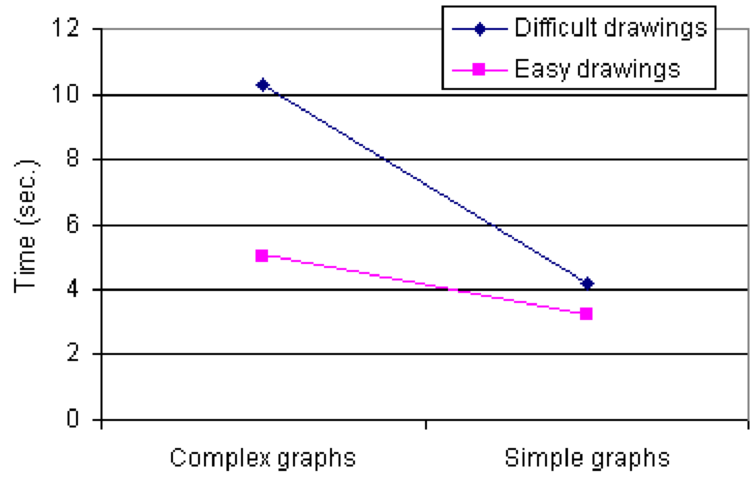
- Difficult: dead-end branches going toward target
- Easy: dead-end branches away from the target



Results

- Difficulty significantly affected response time and errors

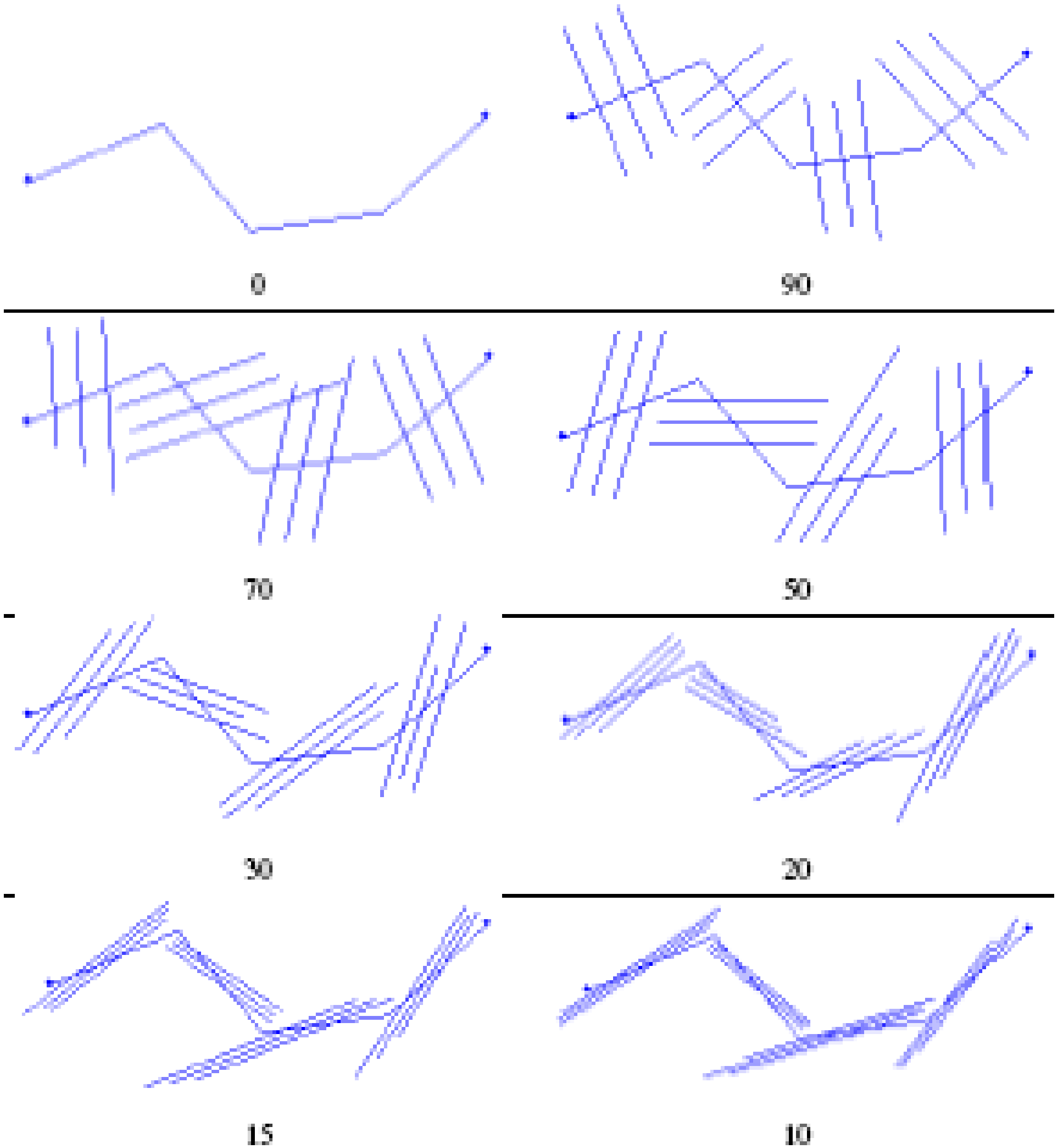




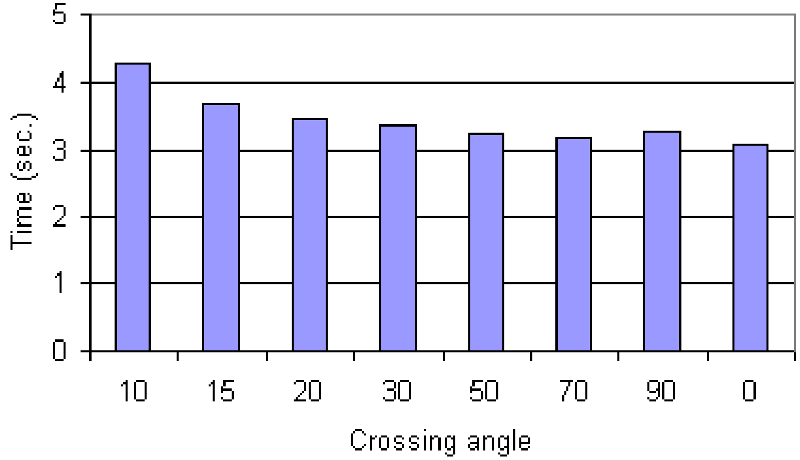
Discussion and an Example of Video

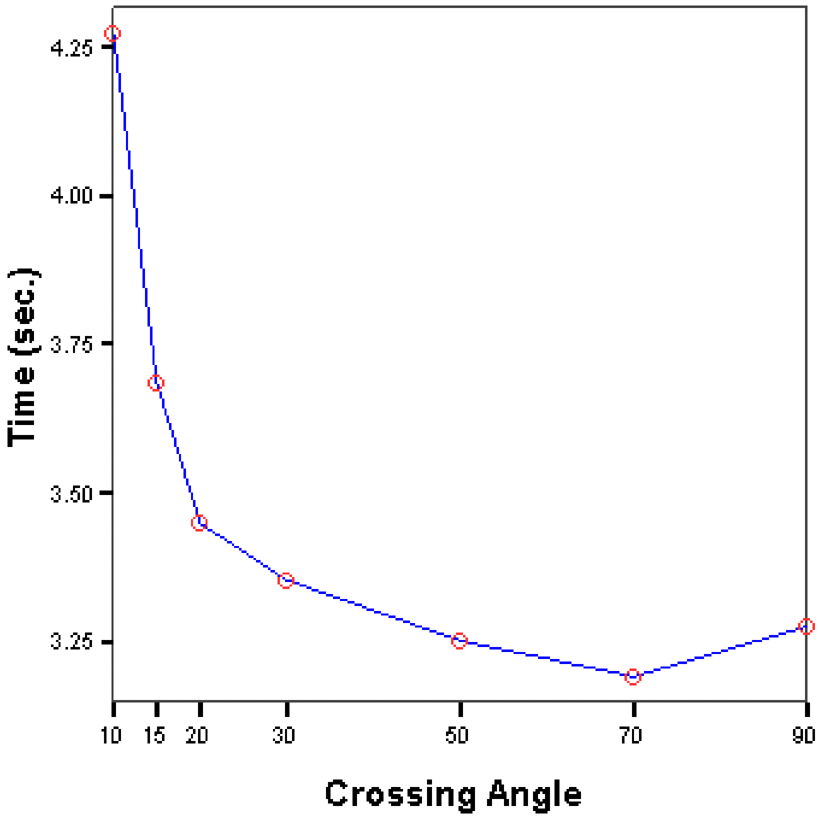
- Geodesic-path tendency affected performance of shortest path tasks significantly.

Ex4: Effects of Crossing Angles



Results


Time (sec.)
5
4
3
2
1
0
10
15
20
30
50
70
90
0
Crossing angle


Time (sec.)
4.25
4.00
3.75
3.50
3.25
10
15
20
30
50
70
90
Crossing Angle

Results

- Linear component and a quadratic component in the relationship between time and angle.
- 70^+ degree angle was equivalent to that with no-crossings.

Conclusion

- Eye movements tell us *how:*
  - How crossings affect eye movements and performance
  - Impact of crossings differs with crossing angle and size of graphs
  - People have geodesic-path tendency in searching shortest paths
- This gives guidance to designers of graph visualization technology
- To obtain insights on *why*, post-task interviews were used.

Acknowledgement

- Book: Interaction Design: beyond human-computer interaction
  - By Jenny Preece et al.
  - companion website: http://www.id-book.com/index.php
- Lecture Notes for CS 5764: Information Visualization
  - By Chris North
  - http://infovis.cs.vt.edu/cs5764/Fall2009/
- Some slides were compiled based on materials from Internet.