

Further Human-Computer Interaction

University of Cambridge

Ashwin Ahuja

Computer Science Tripos Part IB
Paper 7

April 2019

Contents

1	Theory Driven HCI	2
1.1	Using Theory in Design Process	2
1.1.1	Three Waves of HCI	2
1.1.2	First Wave (1980s)	2
1.1.3	Second Wave (1990s)	3
1.1.4	Third Wave (2000s)	3
1.1.5	Further Waves	3
2	Design of Visual Displays	4
2.1	Typography and Text	4
2.2	Maps and Graphs	4
2.3	Schematic Drawings	4
2.4	Pictures	4
2.5	Node-and-link Diagrams	4
2.6	Icons and Symbols	4
2.7	Visual Metaphor	5
2.8	Unified Theories of Visual Representation	5
3	Goal-oriented Interaction	6
3.1	Models of Human Decision Making	6
3.2	Behavioural Economics in HCI	7
3.3	Limitations of Goal-based HCI	7
3.4	Wicked Problems (Rittel and Webber)	7
4	Designing Smart Systems	8
4.1	Smarter Error Correction	8
4.2	Combining typing, speech and gesturing	9
5	Designing Efficient Systems	9
5.1	Constraints on Speed and Accuracy	9
5.2	Keystroke Level Model	9
5.3	Hypothesis-Testing User Studies	9
6	Designing Meaningful Systems	10
7	Evaluating Interactive System Designs	11
7.1	Randomised Controlled Trials	11
7.2	Analysing Qualitative Data	12
7.3	Summary of Analytic Options (Analysing Design)	12
7.4	Summary of Empirical Options (Collecting Data)	12
7.5	Evaluating Non-HCI Projects	13
8	Designing Complex Systems	13
8.1	Cognitive Dimensions of Notations	13
8.2	Representative Cognitive Dimensions	13
8.3	Notational Activities	14
8.4	Beyond Cognition and Interaction	14

1 Theory Driven HCI

Goal of HCI: to make interactions with computers faster, more productive, more creative, more social, more fun etc. Should ask the following questions about a design:

1. How to go about designing user interface?
2. Is it good or bad?
3. How do we know if it is good or bad? Can this be related to scientific evidence or established theories of human behaviour?
4. How could this be improved?
5. Who designed it and why did they do it this way?

Theories can be used to give a critical perspective, which gives us a way of systematically comparing cases to each other.

1.1 Using Theory in Design Process

Gestalt Theory of Perceptual Organisation Application

1. Figure-ground relationship: group elements as either figure or ground - this affects legibility
2. Proximity: group by distance or location
3. Similarity: group by type
4. Symmetry: group by meaning
5. Continuity: group by alignment
6. Closure: perceive shapes that are not there

This can be used to predict how some design choices may be superior to others. The principles can be applied within a design principles which is iterative with phases of

1. **divergent:** creative exploration: Theories can provide a basis for invention, recognising new alternatives or gaps in the range of options that have been considered
Generate new designs from existing one, using creative techniques or exploring metaphors
2. **convergent:** selection and evaluation: Theories give us a basis for critique to compare the relative benefits of possible designs.
Merge features from two candidate designs to produce a better one and discard ideas that poorly fit the desired outcome

1.1.1 Three Waves of HCI

1.1.2 First Wave (1980s)

- Took in theories from (1) Human Factors Engineering, (2) Ergonomics and (3) Cognitive Science
- At this time, the user interface (man-machine interface) was often a separate module, designed independently of the main system
- Main design goal is efficiency (speed and accuracy) for a human operator to achieve **well-defined** tasks
- Researchers use methods to model users' perception, decision and action processes and predict usability
- *Control panel with fixed switches, dials and lamps - group related information together and using the inherent structure of operator sequences of operations can be facilitated*

1.1.3 Second Wave (1990s)

- Theories from social science fields - Anthropology, Sociology and Work Psychology - understanding that computer is one part of social system and design of systems is socio-technical experiment.
- Design takes into account other sources and channels of information including conversations, paper and physical environment
- Focus on studying context where people work and using Ethnography and Contextual Inquiry to understand the other ways of seeing the world

Ethnography: systematic study of peoples and their cultures

Contextual Inquiry: semi-structured interview method to obtain information about the context of use, where users are first asked a set of standard questions and then observed and questioned while they work in their own environments

- Diverse stakeholders are often integrated using prototyping and participator design workshops in order to observe and acknowledge other value systems

Value Systems: A hierarchy of values that all moral beings have, reflected in their choices

Participatory Design: Participatory design is an approach to design attempting to actively involve all stakeholders in the design process to help ensure the result meets their needs and is usable.

- Social science can be used to observe, analyse and understand the design process itself

Large Open Office space: many devices that have been used to facilitate and coordinate collaborative work

1.1.4 Third Wave (2000s)

- Theories from Art, Philosophy and Design (emotional and theories of aesthetics) - HCI as a culture and experience
- Idea is that outside of the workplace, efficiency is not a priority, instead the User Experience is important - the usage of the product is discretionary.

Speculative Design: Created to provoke and question how interesting or thoughtful the interaction will be

1.1.5 Further Waves

- New places for new idea - alt.chi venue as a place specifically created for controversial research that has been rejected from the CHI conference
- Ideas include focuses on:
 1. Wellbeing
 2. Flow
 3. Empathy
 4. Mindfulness
 5. Altruism (*Calvo and Peters*)
 6. Inclusion and Accessibility
 7. Ageing
 8. Low income
 9. Human rights
 10. Bardzell's Feminist Utopianism - design critique directly attacks the mechanisms of institutional privilege using practises that are designed to amplify marginalised voices when thinking about the future.

2 Design of Visual Displays

Fundamental Principle: necessary to work out a correspondence between the information structure (Consists of some number and variety of individual elements and relationships between them) and the visual marks that the user can see. Each element or relationship may correspond to a visible mark or arrangement of marks on a display surface. Display design involves choosing the correspondences, marks and arrangements of marks in ways such that the overall result makes sense to the user.

Many possible design choices are influenced primarily by the historical conventions of how other visual displays have been created in the past.

2.1 Typography and Text

- Originally, computer displays represented paper documents - structured using tabulated columns, alignment, indentation and emphasis, borders and shading - restricted to operations of the typewriter (and now the keyboard interaction with text is limited and frustrating in comparison with writing on paper)

Locations constrained to a specific set of places on the screen - grid system

- Human readers therefore benefit from centuries of refinement in text document design - multiple refined writing systems (for example for maths)
- Therefore, can follow conventions of magazine design, poster advertising, form design, textbooks, etc, when combining media and text

2.2 Maps and Graphs

Cathode Ray Tube displays came before actual displays and could be interpreted easily as being similar to previous paper conventions. Diagrams used quantitative correspondence between direction on surface and continuous quantity (time / distance)

Radar screen shows direction and distance of objects according to a single central reference point, just as Hereford Mappa Mundi of 1300 organised places according to approximate distance from Jerusalem

2.3 Schematic Drawings

Ivan Sutherland: Introduced sophisticated alternative input systems - pen - therefore allows you to properly emulate engineering drawings (as on pen). White space indicated an interpretive break, so independent representations can share the same divided surface

2.4 Pictures

As with the examples above, pictorial representations, including paintings, etc rely upon being similar to 'real' physical artwork. They can be used for bringing out emotions in the person who is looking at the image. However, people need to be aware that the way the item is framed (in this case, being on a computer screen), can change the meaning of the art.

2.5 Node-and-link Diagrams

Node-and-link diagram appears to be effectively considered like a graph - in that we consider the connectivity. These connectivity diagrams appear to come out of circuit schematics - where the designer has complete freedom to place items wherever they want in the diagram itself as the only relevant factor is the connections. Additionally, addition of text (**secondary notation**) can be used to carry more information.

2.6 Icons and Symbols

Types of symbols:

1. Pictorial
2. Arbitrary - for example the sign for a train station - these are generally just conventions and are the norm over time. As computer resolutions increase, can reduce the number of these as we have more specific icons which can be made out. Need to just ensure the meaning is clear - **semiotics** offers a sophisticated way of analysing the basis on which marks correspond to meanings.

3. Icons - small pictures symbolising different kinds of system object - these are generally pictorial
4. Brand Logos - are generally combination of an icon and might have some text. In practise the icons generally simply add decoration to text labels and are designed to be self-explanatory without the text
5. Symbolic - make use of symbols, whether they be letters (foreign or otherwise), or other characters

2.7 Visual Metaphor

Desktop: First visual interfaces of computers were designed to look like a desktop - even though they did not act in anyway like a desk top. Therefore, it's important to note that the strict correspondence to physical objects is more obstructive than instructive.

Desktop is particularly effective as it allows **secondary notation** - allowing the user to create their own meaning by arranging items as desired. The icons use a large number of conventions to indicate symbolic correspondence to software operations or company brands

2.8 Unified Theories of Visual Representation

Important to note that most of these items of screen design have derived heavily from fields of professional study where the designs have been worked out over time. However, it is sometimes necessary to invent new visual representations.

An approach for doing this is using a holistic perspective on visual language, information design, notations or diagrams.

- Addresses relevant factors from low level visual perception to critique of visual culture - including that is necessary to ignore technical and marketing claims and remember all visual representations are marks on a surface intended to correspond to things understood by the reader
- Two dimensions of the surface correspond to: (1) Physical space, (2) dimensions of an object, (3) pictorial perspective, (4) continuous abstract scales.
- Surface can also be partitioned into regions that should be interpreted differently - elements can be aligned, grouped, connected or contained in order to express their relationships. This relationship should be understood by convention or explained
- Individual elements assigned meaning according to semiotic principles of correspondence

	Graphic Resources	Correspondence	Design Uses
Marks	Shape Orientation Size Texture Saturation Colour Line	Literal (visual imitation of physical features) Mapping (quantity, relative scale) Conventional (arbitrary)	Mark position, identify category (shape, texture colour) Indicate direction (orientation, line) Express magnitude (saturation, size, length) Simple symbols and colour codes
Symbols	Geometric elements Letter forms Logos and icons Picture elements Connective elements	Topological (linking) Depictive (pictorial conventions) Figurative (metonym, visual puns) Connotative (professional and cultural association) Acquired (specialist literacies)	Texts and symbolic calculi Diagram elements Branding Visual rhetoric Definition of regions
Regions	Alignment grids Borders and frames Area fills White space Gestalt integration	Containment Separation Framing (composition, photography) Layering	Identifying shared membership Segregating or nesting multiple surface conventions in panels Accommodating labels, captions or legends
Surfaces	The plane Material object on which the marks are imposed (paper, stone) Mounting, orientation and display context Display medium	Literal (map) Euclidean (scale and angle) Metrical (quantitative axes) Juxtaposed or ordered (regions, catalogues) Image-schematic Embodied/situated	Typographic layouts Graphs and charts Relational diagrams Visual interfaces Secondary notations Signs and displays

3 Goal-oriented Interaction

Idea is to find out how to use cognitive theories to understand user behaviour and what they find hard - the idea is that the user interaction can be modelled as a search. General purpose search algorithms are familiar in computer science, where an objective function can be combined with a state space and dependency graph to recursively search for optimal solutions use BFS or DFS.

Can model interaction with a user interface as a search process if we clearly define the **goal**. This is effectively a Cognitive Walkthrough:

1. Analyse the user interface by identifying the next user goal
2. Determine whether the necessary actions are available
3. Ensuring that they are labelled in the way that the user will recognise them
4. Confirm that the system will give appropriate feedback of progression towards the goal

3.1 Models of Human Decision Making

- Challenge is where it is tough for the designer to correctly anticipate the user's goal in more complex situations **Bounded Rationality**: Even if there is an optimal solution, amount of time necessary to find it may be too long. Therefore, bounded rationality is where a model takes into account the cost of computation - the user can follow a plan that is satisfactory, within constraints
- **Prospect Theory (Kahneman and Tversky)**: Describes human behaviour in terms of a utility model that considers the outcome of possible actions, with weighting of estimated benefits by likelihood - does not assume complete knowledge of the state space
- **Heuristics and Biases**: External considerations that account that for observed patterns in human decision making that cannot be explained by optimising
 1. Availability Heuristic: reasoning based on examples close at hand

2. Affect Heuristic: decisions based on emotion rather than calculating cost to benefit
3. Representative Heuristic: Probability is judged based based on resemblance to a class of similar situations.

Biases are used to ensure that the consequences of estimation error are within tolerable bounds

1. Loss aversion: losses hurt more than gains feel good
2. Expectation Bias: People observe results that they expect
3. Bandwagon Effect: Prefer actions taken by other people

Can use these theories to affect people's behaviours, for example by the UK government in nudge theory.

3.2 Behavioural Economics in HCI

Attention Investment theory of abstraction use is a model of end-user programming which explains why users without prior experiences of programming take decisions that favour repeated manual actions rather than automated shortcuts - this is some sort of loss aversion where bugs and other issues might mean there might be more time taken to fix the mess. In this case bounded rationality will then take place.

Important to consider that much of the routine computer usage is done through memorised patterns of interaction with no clear mental model or goal.

3.3 Limitations of Goal-based HCI

- Rational models assume user doesn't make mistakes - completely unrealistic
- We need to find where errors are happening and why and then create a **cognitive model of this**: decision process that is not consistent with the identified goals, constraints and search space, including information loss due to (1) cognitive limitations, (2) incorrect mental models and (3) misleading designs
- Anticipating all the factors needs a description of a user journey that accounts for (1) problem identification, (2) diagnosis, (3) debugging, (3) testing, (4) iteration and procedures that *characterise the user's own activities as a kind of design process*
- **Persuasive Design**: field of HCI that considers how users choose alternative goals or otherwise modify their goals - useful where trying to change user behaviour - applies nudge methods to modify biases.
Need to ensure the design isn't patronising to the user

3.4 Wicked Problems (Rittel and Webber)

Class of problems that cannot be addresses with classical goal based problem solving methods - have the following **characteristics**:

1. No definitive formulation
2. No stopping rule
3. Solutions aren't true or false, but good or bad
4. No immediate and no ultimate test of a solution
5. Do not have an enumerable set of potential solutions, nor a well-defined set of permissible operations
6. Unique
7. Can be considered to be a symptom of another problem
8. Existence of a discrepancy representing a problem can be explained in numerous ways - this choice determines the nature of the problem's resolution
9. Planner has no right to be wrong

4 Designing Smart Systems

- Probabilistic model of the user's model can be used to create more efficient user interfaces
- An example of this is through the use of predictive text entry which is particularly important for things like smartphones and smartwatches, etc where you can't have a full size keyboard
- Issue is that people are reluctant to adopt a radically different text-entry method that takes more than few minutes of training
- However, improvement of sensors (in quality and reduced cost) has led to probabilistic text-entry methods - eg touch-screen error-correction, calculates the posterior word probabilities by combining likelihoods from a touch model with prior probabilities from a statistical language model
 - However, hypothesis space is absolutely enormous
 - Exploit the hypothesis space by converting it into a word-confusion network - time-ordered series of connected word confusion clusters. Each cluster contains a set of word hypotheses and the probabilities of these hypotheses sum to one
 - Can use this as a user interface to reduce error correction by letting users access the next-best word candidates in the hypothesis space
 - Also provides flexibility in the use of multiple modalities
- Speech recognition: assigns posterior probabilities to word sequences by combining likelihoods from an acoustic model with prior probabilities from a statistical language model
- In general, this probabilistic model methodology has been used to provide high performance, single-modality inputs, but can be used to provide multiple text-entry methods with lots of flexibility offering higher-performance without forcing user to learn new text entry methods.

Text entry methods: all successful ones share good performance AND high similarity to other mainstream methods

1. Single-stroke letter alphabets
2. Multitap and predictive text entry on keypad
3. Touchscreen keyboard
4. Physical thumb keyboard
5. Gesture keyboard
6. Handwriting recognition
7. Speech recognition

Therefore, when creating a new solution, it must be similar to a familiar one - **path dependency**. Therefore, in learning this method, while the user's initial speed of usage is slow, over time, it will become superior and become better than the familiar method at the *crossover point*.

4.1 Smarter Error Correction

- When correcting a text-entry error, users must tap backspace or select the error's location - however, can use rich hypothesis space to design more **flexible** user interfaces
- Can develop an algorithm to allow users to say a correction without explicit reference to the error location, or can say with more context
 - Works by merging word-confusion networks from different probabilistic text-entry modalities - taking two networks as inputs and using some of the probability mass in each cluster, softens the networks by adding epsilon (no word) and wildcard (can either stay in same cluster and generate any word, or move to next cluster and generate wildcard word that matches any word) transitions
 - Then algorithm searches for the highest joint probability path through both networks using a token passing model

4.2 Combining typing, speech and gesturing

- By merging speech and gesture keyboard text entry, you can reduce the word error rate 53 percent and using only the gesture keyboard to correct spoken words can reduce the word error rates by 44 percent
- Also, allowing users to seamlessly switch between methods of inputs is useful because different methods have different performance envelopes for different sentences

5 Designing Efficient Systems

Idea is that we can optimise systems using probabilistic models to predict human actions, with the models being updated in real time. Additionally, allow efficiency of the user interface to be predicted and also measured at design time.

5.1 Constraints on Speed and Accuracy

There is generally a speed vs accuracy trade-off that can be characterised as an information channel - fast and inaccurate actions result in more channel noise, meaning that the information gain per unit of time does not increase as quickly as the number of movements made - this can be described by **Fitt's Law**: time taken to point at something is proportional to Distance to the target, while inversely proportional to Width of target. The ratio of width to distance is the index of difficulty (k) and can be understood as the potential amount of information gained by the system when the user points

$$Time = k \log\left(\frac{2D}{W}\right) \quad (1)$$

Can use Fitts' Law to design more efficient user interfaces if we have a prior expectation for the actions that a user is likely to take. **Semantic Pointing** modifies the mapping of mouse motion to screen pixels so that the effective width of more likely targets is increased and the effective distance between them is decreased.

5.2 Keystroke Level Model

Mouse movements and key strokes (of an expert) are modelled using KLM:

- Time taken to press key or mouse = 200ms
- Time taken to home that hands on mouse or keyboard = 400ms
- Mental preparation time = 1500ms

Can use more sophisticated model like GOMS (Goal / Operator / Method / Selection) in order to predict mental preparation time.

5.3 Hypothesis-Testing User Studies

A/B Test: Randomised controlled trials giving different users different versions of the interface and seeing which version is more likely to have desired behaviour. (1) Measure the completion times for a task and then (2) Carry out a significance test to see whether the difference is statistically significant.

- **t-test compares the effect size** (difference between sample means) to the variation in experimental data.
- Most straightforward tests for comparing sample distributions rely on data following a normal or Gaussian distribution - if this is not true - should use a sign test to compare sets of matches samples - **within-subject comparison**: get a subject to do task on both old and new interface - note whether sign is positive (improved version faster) or negative (original version faster) and compare proportion of each sign

Successful user studies rely on controlled experiments to minimise the variation in the data from factors unrelated to the effect of the design change.

1. Individual differences between subjects
2. Distractions during the trial
3. Motivation of the participant

4. Accidental intervention by experimenter
5. Other random factors

Difference in means always reports with confidence intervals or error bars - significant is not always interesting - very small effects can be shown to be reliable if the variance is small or the sample size is very large. Instead better to have large effect sizes.

Drawbacks

- Statistical comparison not always viewed favourably in commercial applications
- **Hawthorne Effect (1920s)**: measuring effect of factory lighting levels on productivity. Showed that productivity improved if lighting was increased and that productivity improved if lighting was decreased. Turned out that worker motivation, and therefore productivity, improved any time that an experiment was carried out.
- This happens with user interface modifications - an interesting design change may result in apparent efficiency improvements in experimental context but has no long-term benefits
- **Taylorism**: Optimisation of human efficiency in an industrial context. He came up with techniques to measure worker efficiency in comparison to machines to it was possible to measure and manage workers. In second-wave HCI, start to consider social aspects and work with trade unions to understand rights and point of view of the workers
- Also, sometimes efficiency is not sufficient measure of system design - in discretionary use systems, where the person using the system is not an employee but has their own goals, thus can decide whether to be efficient or not. Therefore, brings in third wave HCI to bring in culture.

6 Designing Meaningful Systems

- **Design Ethnography**: Offers a holistic, in-context understanding of how life works so it can be supported, enhanced and changed.
- **Ethnographic Study**: Attempts to observe people in their normal environment, over a substantial period of time. During the observation, the experiences of people are documented in a variety of ways, paying attention to both peoples' activities and the artefacts that they interact with. The outputs of an ethnographic study are diverse, including: (1) task flow diagrams, (2) Journey maps, (3) Concept generation maps, (4) Timeline of people's activities, (5) Written reports and (6) Thick descriptions
- This can be integrated within typical processes of design in implementation - eg combined with design workshops, in order to: (1) Understand the experiences that people were currently having and how those experienced could be supported or changed by a new product and (2) to understand how people relate to the novel scenarios produced by the design team and then to evaluate early product designs in a real-world context.
- Helps teams understand:
 1. What things happen - (1) what the everyday life of the users look like, (2) what kinds of experiences with the technology they have, (3) what are people's strategies for working with products, services and devices to get things done
 2. How things matter - (1) how products help people be social and how they integrate in the social life of people, (2) how products help people make meaning and sense of their life, (3) how products help people be in control of their life

Then understanding included in the product development process to build technology that better fits people's needs.

- **Quantitative Research** and ethnographic research can be combined into **mixed methods** which can result in a better understanding than would be achievable with a single method along

7 Evaluating Interactive System Designs

Formative Evaluation: Used in early stages to compare, assess and refine design ideas. Often involves open research questions, in which researcher is interested in learning further information that can inform the design.

Summative Evaluation: More likely to be used in the later stages of a project and to involve closed research questions, with purposes of testing and evaluating systems according to predefined criteria.

1. **Empirical Methods:** Making observations and measurement of users, in order to learn new information relevant to system design or use - associated with open research questions
 - (a) **Qualitative:** (1) Think-aloud, (2) Interviews, (3) Field observation
 - (b) **Quantitative:** Generally require a working system, so are mostly summative - for example (1) Use of analytics and metrics in A/B experiments and (2) Controlled laboratory trials
2. **Analytical Methods:** Based on applying a theory to analysis and discussion of the design. Generally very useful to formative evaluation, as it may be difficult to observe how a design is used if system design has not yet been completed.
 - (a) **Qualitative:** (1) Cognitive Walkthrough, (2) Cognitive Dimensions of Notations
 - (b) **Quantitative :** Methods including Keystroke Level Model, which can be used to create numerical comparisons of closed research questions

7.1 Randomised Controlled Trials

Requirements:

1. Performance Measure
2. Representative sample of consenting target population
3. Experimental task that can be used to collect performance data

Analysis of results:

- Results measured in terms of effect size, including correlation with other factors and reporting significance measures to check whether the observed effects might have resulted from random variation or other factors.
- **Overcoming natural variation needs large samples**
- **Does not provide understanding of why change occurred - hard to know whether the effect will generalise**
- **If relevant variables that are orthogonal to each other, many separate experiments might be needed to distinguish between their effects and interactions**
- **Proxy Measures:** use measures such as number of days customers actively use the product instead of desired outcome (profit maximisation) as that has a large latency.
- All controlled experiments assessed according to Internal Validity and External Validity
 - **Internal Validity:** was the study done right - (1) reproducibility, (2) scientific integrity, (3) refutability
 - **External Validity:** does the study tell us useful things - focusing on whether the results are generalisable to real world situations such as ensuring (1) representativeness of sample population, (2) experimental task, (3) application context

7.2 Analysing Qualitative Data

1. **Categorical Coding:** qualitative data analysis method that can be used to answer closed questions, for example comparing different groups of people or users of different products.
 - (a) Create coding frame of expected categories of interest
 - (b) Text data then segmented (for example on phrase boundaries)
 - (c) Each segment assigned to one category
 - (d) Frequency and correspondence can then be compared
 - Should incorporate some assessment of inter-rater reliability, where two or more people make the coding decisions independently to avoid systematic bias
 - Then compare how many decisions agree, relative to chance, using a statistic metric (Cohen's Kappa - for two people, Fleiss' Kappa for more people where 0.6 to 0.8 is considered substantial agreement)
 - Can also take into account how many decisions still disagreed after the discussion, involving refining and iterating the coding frame to resolve decision criteria
2. **Grounded Theory:** Can be used to explore open questions where no expectation or theoretical assumption of the insights that the researcher is looking for.
 - (a) First, read data closely, looking for interesting categories (**open coding**)
 - (b) Then collect fragments, writing memos to capture insights as they occur
 - (c) Emerging themes are organised using axial coding across different sources of evidence - and constantly compare memos, themes and findings to the original data in order to ensure that these can be justified.
 - (d) Process ends when the theoretical description has reached saturation in relation to original data, with the main themes complete and accounted for

7.3 Summary of Analytic Options (Analysing Design)

1. **Cognitive Walkthrough:** normally use in formative contexts - generally a time-saving precaution before user studies, to find and fix blatant usability bugs
2. **KLM / GOMS:** These require alternative detailed UI designs, in order to compare them for efficiency
3. **Cognitive Dimensions:** Better suited to less structured tasks than the first two, which rely on having a set goal and task structure

7.4 Summary of Empirical Options (Collecting Data)

1. **Interviews / Ethnography:** Useful in formative / preparation phase - to develop design ideas / to capture user requirements
2. **Think-aloud / Wizard of Oz:** Useful for both paper prototypes and working systems - highly effective at uncovering usability bugs, as long as verbal protocol is analysed rigorously using qualitative methods.
3. **Controlled Experiments:** Can help establish engineering aspects of work - however, important to ensure you can (1) measure the important attributes in a meaningful way with both internal and external validity, (2) test significance, (3) Report confidence interval of observed means and effect sizes
4. **Surveys and Informal Questionnaires:** Important to be clear about what is being measured - use mix of open questions (more qualitative information) and closed questions (easier to aggregate and test hypotheses).

Open questions require a coding frame to structure and compare data, or grounded theory methods. Also, must do a pilot study with a small number of people.

Faster to collect data from large sample size than interviews

Give less insight than interviews now

5. **Field Testing:** If created a working product, possible to make a controlled release and collect data on how it is used.

6. **Standardised Survey Instruments:** Standard psychometric instruments to evaluate mental states such as (1) fatigue, (2) stress, (3) confusion and (4) emotional state. Also approaches to assess individual differences - and then can results compared to existing scientific literature
7. **Purely Affective Reports** appear empirical or quantitative but are incredibly biased
8. **Introspective Reports** made by a single subject - highly biased and subjective - this is known as the **Highest-Paid Person's Opinion**

7.5 Evaluating Non-HCI Projects

Important to say that principles can also be applied to things which do not include human interactive aspects. Can always define goals and hypotheses and understand boundaries and performance limits by exploring them. Can still use analytic and empirical evaluation techniques, combining both formative and summative evaluation.

8 Designing Complex Systems

Issue: may not necessarily have a well-defined goal, instead need to design interaction spaces rather than well defined systems.

In order to do this, need to use a broad brush analysis technique, because attempts to describe individual tasks and specific actions would result in a death by detail, resulting from a combinatorial explosion of possible and potentially relevant details. Necessary to find an analytical frame that structures the description of an interaction, so it can be compared to an ideal characterisation of the application domain in terms of desirable interaction patterns offering a critical perspective.

8.1 Cognitive Dimensions of Notations

Canonical example of a broad brush analysis technique which is used in designing programming languages and APIs is the Cognitive Dimensions of Notations framework.

Physics of Notations - inspired by CDs: tries to establish basic principles of visual perception that might be relevant and **Patterns of User Experience:** which attempts to focus on the subjective experience of the user rather than purely observable behaviour

Cognitive Dimensions are presented as a vocabulary for design discussion - these reflect common usability factors that experienced designers might have noticed, but did not have a name for.

- Based on the observation that there is no perfect user interface
- Any user interface design reflects a set of design trade-offs that the designers have had to make
- Cognitive Dimensions give designers a vocabulary to describe these trade-offs
- Impossible to create a design that has perfect characteristics in every dimensions

8.2 Representative Cognitive Dimensions

1. **Premature Commitment:** constraints on the order of doing things based on dependencies.
2. **Hidden Dependencies:** important links between entities are not visible. If you change parts of the structure of the product, will changes affect other parts.
3. **Secondary Notation:** extra information in means other than formal syntax - possible to make notes to yourself or express information that is not recognised as part of the notation
4. **Viscosity:** resistance to change. How easy would it be to change parts of the product.
5. **Visibility:** ability to view components easily.
 - (a) How easy is it to see or find the various parts of the notation while it is being created or changed
 - (b) What kind of things are difficult to see or find
 - (c) If you need to compare or combine different parts, can you see them at the same time and if not, why not

6. **Closeness of Mapping:** closeness of representation to the domain - how closely related is the notation to the result that is being described
7. **Consistency:** similar semantics are expressed in similar syntactic forms - when to things are similar (in notation), is the similarity clear from the way they appear
8. **Diffuseness:** verbosity of the language
9. **Error-proneness:** the notation invites mistakes
10. **Hard mental operations:** high demand on cognitive resources
11. **Progressive Evaluation:** work-to-date can be checked at any time
12. **Provisionality:** degree of commitment to actions or marks
13. **Role-expressiveness:** purpose of a component is readily inferred
14. **Abstraction:** type and availability of abstraction mechanisms. Can you define new abstractions, do you operate on abstracted elements, etc.

8.3 Notational Activities

When users interact with content, there are limited numbers of activities that they can engage in when considered with the respect to the way the structure of the content might change. A CD's evaluation must consider which classes of activity will be the primary type of interaction for all representative system users. If the needs of different users have different priorities, those activities can be emphasised when design trade-offs are selected as a CD's profile.

Activities include:

1. **Search:** Finding information by navigating through the content structure using facilities provided by the environment. In general notation will not vary though parts of it that the users sees will vary
2. **Incrementation:** Adding further content without altering the structure in any way
3. **Modification:** Changing existing structure, perhaps without adding new content
4. **Transcription:** Copying content from one structure or notation to another notation
5. **Exploratory Design:** Combining incrementation and modification, with added characteristic that the end state is not known in advance. Any kind of viscosity can make this much harder. This is why good languages for hacking may not be strictly types, as maintaining type declarations causes greater viscosity. Loosely types languages are more likely to suffer from hidden dependencies (a trade off with viscosity), but this is not a problem for exploratory design, where the programmer can often hold this information in his head during the short development timescale.
6. **Collaboration:** If the main purpose of the notation is to be shared or discussed with other people, design considerations can be different to those necessary for working by yourself

8.4 Beyond Cognition and Interaction

- Issues like ethics and accountability come up when dealing with automating and/or justifying bias and prejudice
- **Digital Humanities:** treating text and images as meaningful and sophisticated, rather than simple categorical judgements
- Ethical questions about who does the intellectual work of data labelling - who pays for them and where do profits go
- Systems are being designed and shaped by knowledge infrastructure - need to be aware that their biases are being built into the system then