[WK1]-USABILITY ENGINEERING

Process

Analyze: Understand users and their needs (goals and tasks)

Design: Explore possible approaches to address user's needs and match user preferences, knowledge, context… **Implement:** Many, many tools to prototype and build interfaces

Evaluate: Study users

Usability

- **Learnability** - accomplish basic tasks the first time
- **Efficiency** - how quickly can they perform tasks they have learned the design
- **Memorability** - re-establishing proficiency after some time
- **Less Errors** - number and severity of errors made and how easily can they recover
- **High Satisfaction** - the pleasance of the design

Utility (features you need) + Usability (easy/pleasant) = Useful

User Experience (UX) - More than usability which focuses on performance

- Emotional and visceral response / - Fun (hedonic response), exciting /
- Artistic appeal /- gestalt / - Match to fashions and user expectations

Desirable Aspects	Undesirable aspects
satisfying, helpful, fun, enjoyable, motivating, provocative, engaging, challenging, surprising, pleasurable, enhancing sociability, rewarding, exciting, supporting creativity, emotionally fulfilling, entertaining, cognitively stimulating	boring, unpleasant, frustrating, patronizing, making one feel guilty, making one feel stupid, annoying, cutesy, childish, gimmicky

2018 ISO Usability: The extent to which a system, product or service can be used by **specified users** to **achieve specified goals** with **effectiveness**, **efficiency** and **satisfaction** in a **specified context** of use. **Usability** is relevant when designing or evaluation interactions with a system, product or service for the purposes of **development / procurement / review or comparison / marketing and market research**

Usability is relevant to: — regular ongoing use, to enable users to achieve their goals effectively, efficiently and with satisfaction;

— learning, to enable new users to be become effective, efficient and satisfied when starting to use a system, product or service;

— infrequent use, to enable users to be effective, efficient and satisfied, with the system on each reuse; — use by people with the widest range of capabilities;

— minimizing the risk and the undesirable consequences of use errors;

— maintenance, in that it enables maintenance tasks to be completed effectively, efficiently and with satisfaction.

Usability engineering (UE) is building upon science and practitioner research and experience, **using systematic ways to tackle the task of creating usable interfaces.** It considers user experience as additional to usability

- methods, theory, application, processes, checklists - practitioner research and experience

Challenge of UE:

- Hard to think like the users: understanding the domain / context of use / what the user knows / users experiences and users' interpretations (different background and experience)
- Creating specifications / some interfaces need to be complex / Tension between utility and beauty, fun, marketing / different purposes, contexts, cultures with different devices / Evaluation of purposes and software / legal issues in copying other designs

Trade-offs of UE

- Standards – style guides, related product
- Graphic Design
- Technical Writing – documentation
- Internationalisation
- Performance
- Multiple Platforms – hardware, browsers
- High-level and low-level details
- External factors – social issues
- Legal issues
- Time

Concept Mapping: A way to show your understanding, showing the key ideas relevant to a focus question

Approach to UE

Steps: iterate

- **Design** – create design concepts
- **Implement** – realise alternatives, prototype
- **Evaluate** – verify and refine design
- **Analyse** – understand user work & needs

Project Parameters – risks, goals, resources, system type, organisational culture, dev. stage

MAPPED TO

Project choices – development activities, level of rigor, methods & techniques, iterations

System complexity	=	Work Domain	x	Interaction

[WK2]-Contextual Enquiry and Hierarchical task analysis

Landscape of methods to gain understanding of users and their activities

Contextual Inquiry - observations and interviews, combined with other methods e.g. TA

Interviews, Surveys, Focus Groups	Contextual Inquiry
Ask people what they want	**Observe** what people actually do
- Human memory	-Can you be there at the right time
- Politeness	- Observer effect
- Self-awareness	- Intrusiveness
+ User involvement, commitment, buy-in	+Insights into actual behavior
+ Access to what users know, think, perspectives, contexts that matter	+ Contexts + Stakeholder perspectives

Contextual Analysis / Design – Summarised in diagrams for model / define and refine: system concept statement

User-needs & requirement – develop tasks the users need to do & the interface must support

Properties of good concrete tasks

They are concrete -- a user understands what they need to do after reading them: the instructions need to be clear, easily understood provide all the information the user needs, different people should interpret them in the same way

You can judge success -- There should be a clear start and clear end at which point you can assess whether the user was able to do the task successfully, needed help, or could not do it. There should be a clear outcome.

They do not lead the user -- Tell WHAT to do, not HOW to do the task. Avoid any words that appear at the interface.

They are relevant --Users of your real system would be expected to do them

The set of tasks gives good coverage

Frugal: each task tests different things (unless you explicitly want to study repeat use)

Effective: each task covers an important aspect

They are at the right level of difficulty -- Start with easy tasks / Ensure you time the test and tell the user ahead of time an estimate of the duration

They are respectful and avoid offending the user – No humorous, frivolous or offensive

System Concept Statement

A SCS is a concise descriptive summary of the envisioned system or product stating an initial system vision or mandate; in short, it is a mission statement for the project.

- Mission statement for the system / - Includes – system name, target users, what the system is intended to do, the problems the system should solve & aspects of the user experience

[WK3]-EMPIRICAL STUDIES, THINK-ALOUD

Think Aloud – Users think aloud as they use the interface & interpret actions through qualitative data. It gives insight into the user's mental model because you will see the details of their using process. It is quite inexpensive to do and useful with early prototypes. Even inexperience evaluators can get good results. - **Qualitative** vs Quantitative

- **Empirical** vs Predictive - **Summative** vs Formative - Naturalistic vs **Laboratory**

Design Cycle – User Centered Design, Prototype Construction, Usability Study, Repeat

Participants – similarity to intended user population / age / gender / experience in area / interest or motivation / computer literacy

- If your interface is to be used by very different groups of people, you need to do the think-aloud with each group / **Iterate:** Use same budget of effort to test

Number of Users - 5 enough to uncover 80% of problems to stay within budget restrictions, but need to do with each user group, iteration will improve results – other views is not enough or difference in problems

Pros and Cons of TA

++ finds many usability problems / finds why they occur (process data)
+ small number of test users-relatively inexpensive / generates colourful quotes
+ usable early in development process / requires little facilitator expertise
+ show actions to avoid rationalisations / uncover causes of problems
-- slows users down by about 17% giving access to mental model
- depending on the instructions given to the user, having to think aloud can change the user's problem-solving behaviour / not quantitative directly
/ cannot provide performance data (bottom-line data)
- Unnatural context and situation - People filter, want to please, do not want to look foolish or inept - Hawthorn effect - Experimenter can bias results

User's "theories" must be interpreted with care

Naturalised think-aloud - Multi-user interaction: Two (or more) users work on task / conversation is natural / observer collects dialogue

Stages of running an Evaluation

Preparation: materials for consent, set up machine, room, environment, check user instruction, do a mental run-through **Introduction:** welcome, explain purpose, system tested, confidentiality, anonymity, invite questions

The test: record, support, help if stuck

What are you thinking now?
What do you think that message means?

How do you think you can do it?
Is that what you expected to happen?

Questionnaire/ interview: open and closed

[Pitfalls] Right concrete tasks (Test all key aspects/Multiple tasks for same aspects)

Instructions (no leading/don't use words that are identical on the interface)

Debriefing: thank user, remind of usefulness, make sure all data collected & notes written, user confirms

Analysis, reflection, summarizing, reporting, conclusions for action (After)

Empirical Studies: These are studies that involve people. You observe them and ask them questions and ask them to do questionnaires. You collect data as "personal data". You need to be careful in designing and conducting empirical studies and in managing the data you collect

Formative Evaluation

Conducted at various stages through the development of an interface
The results inform the next stage of development
Better to spread your evaluation budget over several stages
Can stop if you have strong evidence on problems that really need to be addressed

[WK4]-HEURISTIC EVALUATION

HE – record all the problems you find. Careful thinking and creativity in trying things that you think users may do, including slips and errors.

- Describe each problem identified / List the heuristics violated in this system / rate the assessment of the severity
- **Step**: Select set of design heuristics – Recruit experts in HE – Each Expert (take through tasks / sb records violations of the heuristics / expert review all the problems found and rate them) – summarise & rate problems - Combine all expert reports – compare with TA

What is one key disadvantage of HE compared with TA evaluations.

Compared with TA, participants always are real users, HE based on the experts, who have many experiences to check and give feedback for an interface. In particular, the experts commonly find "problems" that do not pose difficulties for actual users and commonly misses some problems that real users have.

What is one key advantage of HE compared with TA evaluations.

The advantage of HE is it doesn't need for user studies in all forms. They can find some minor problems easily which the real users may ignore, Additionally, since they have experiences, so they don't need too much time than the real users

Heuristic Evaluation (Usability Inspection) is an expert-base method for usability evaluation, on contrast to TA which is user-based method.

Neilsen Heuristics Guideline

Visibility of system status	The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
Match between system and real world / Metaphors	The system should **speak the users' language**, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
User control and freedom	Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. **Support undo and redo.**
Consistency and standards	Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
Error prevention	Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
Recognition rather than recall	Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
Flexibility and efficiency of use	Accelerators – unseen by the novice user – may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
Aesthetic and minimalist design	Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
Recognize, diagnose, recover from errors	Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
Help and documentation	Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Tips: Video often really help / Tooltips / Overlay tutorial

Main stages of heuristic evaluation

Preliminaries - Agreed set of heuristics to use; Programming team member overviews system; Team member is available throughout; **Set of tasks**

Evaluation - Each expert works independently through the UI; Team member records problems, so expert can simply state them (single person works well, so they bring all details together); Also answers any questions (eg expert gets stuck, cannot find how to do a task, may need help with domain expertise aspects); Multiple passes (overview, then detailed); Results in a set of identified failures to match the heuristics (part of interface, violated heuristic)

Concluding summary - Summarise all the flaws; Rate these in terms of severity

Severity ratings – (0 to 4) = frequency x impact x persistence

The Severity of flaw: The severity of a usability problem is a combination of 3 factors
The **frequency** with which the problem occurs: Is it common or rare?
The **impact** of the problem if it occurs: Will it be easy or difficult for the users to overcome?
The **persistence** of the problem: Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?
0: I don't agree this is a usability problem 1: Cosmetic problem
2: Minor usability problem 3: Major usability problem 4: Usability catastrophe

HE is referred to as discount evaluation when 5 evaluators are used.

When 5 evaluators are used, can discover 75-80% of usability problems, the ration of benefits to costs is at the optimum. **Selecting Heuristics = context + validity + usability**
Results – Description of problem, Heuristic Violated, Severity rating

Why is Heuristic Evaluation described as a predictive method?

Compare with TA, which is Empirical studies, total tested by the real users, HE is an expert-based method for usability evaluation. Designers can change their interface according to the feedback anytime. The whole process is like an expert predicting the reaction of real users, so it is a predictive method.

Benefits and Problems of HE

Benefits	Problems
- fewer ethical and practical considerations - experts have knowledge of domain and users - important and minor problems discovered - experts are not designers - no need for user studies	- difficult and expensive to find good experts (domain / users in that domain / interface and design heuristics) - may miss trivial problems that real users have - expert bias and problems do not pose difficulties for actual user

[WK5]-USABILITY EVALUATION

The risk: - the participants – are they target users / - the tasks – how well designed they were

Formative Evaluation: a diagnostic evaluation to improve the user interface on the next iteration on the cycle - Typically qualitative / usually "informal"

Summative evaluation: used to make a judgement – Typically combination of quantitative and qualitative / often "formal"

Science: Discover new knowledge that is not actionable for usability engineer **Engineering:** Provides knowledge that can be used by people who create and evaluate interface.

Lab study: artificial environment created for usability evaluation (Benefit of controlling the environment / ensure focus on aspects you want to evaluate / avoid unknown factors that may affect result)

Field study: people use the interface in an "authentic setting" as they would in real use (ecological validity / potentially many unknown "noise variables" affect results)

Quantitative (number): can be from objective measures such as how many participants could do task1 without help in 3 minutes / can be from subjective measures
Qualitative: comments, free text, facial expression made by users

Rigorous: high cost and high accuracy usability evaluation (maximise effectiveness / minimize risk / gold standard)

Rapid: low cost and lower accuracy

Empirical: based on "observations" of people using the interface. -- TA

Analytic: no users needed -- HE

Critical incidents: Arguably the single most important source of qualitative data in formative evaluation / a critical incidence is identified in your observations in the TA. Judgement that this problem is very important indicator the interface is good (or bad)

Objective data: based on "hard facts"

Subjective data: observations involving interpretation or judgement, includes 情感和感觉

How to select appropriate heuristics: context / validity / usability
Guidelines can conflict and overlap and in practice, one must make trade-offs

Formative, rapid, Think-Aloud is far more valuable if you have limited resources
Good guidelines should be **empirically** based.

How to write User-Friendly Content

- Use the words your users use
- Chunk your content
- Front-load the important information
- Use pronouns
- Use active voice
- Use short sentences and paragraphs
- Use bullets and numbered lists
- Use clear headlines and subheads
- Use images, diagrams, or multimedia
- Use white space
- Make action sequences clear
- Avoid jargon and use familiar words
- Define acronyms and abbrev
- Use abbrev sparingly
- Use mixed case with prose
- Limit the number of words and sentences
- Limit prose text on navigation pages
- Use active voice
- Write instructions in the affirmative
- Make first sentence descriptive

Golden rules for dialogue

Consistency	Closure in dialogues	Support internal locus of control
Short Cuts	Simple error handing	Reduce short term memory load
Informative feedback	Easy reversal of actions	

Bruce Togazzini

Anticipation	Efficiency of the User	Metaphors
Autonomy	Explorable Interfaces	Readability
Color	Human Interface Objects	Track State
Blindness	Latency Reduction	Visible Navigation
Consistency	Learnability	Fitts' Law
Defaults		

Others: Company specific / Device / OS specific / ISO, ANSI Standards / National Standards / Military Standards / Accessibility Standards

OSX Mental Model: Familiarity, Simplicity, Availability (functionality available), Discoverability

[WK7]-Prototyping and Personas

What is a prototype in HCI/UX/UE

A prototype is a **draft version** of a product that **allows you to explore your ideas** and show the **intention** behind a feature or the overall design concept to users before investing time and money into development. A prototype can be anything from **paper drawings (low-fidelity)** to something that allows click-through of a few pieces of content to a **fully functioning site (high-fidelity).**

Prototyping Techniques (Lo-Fi---Hi-Fi)

Paper-based sketches – Paper-based storyboard – Computer-aided sketches/ storyboard – Wizard of OZ /Video-prototyping – Computer-based scenario simulation – Computer-based Horizontal simulation – Computer-based Vertical simulation – Computer-based full functionality simulation

Low-Fidelity: Quick to construct / Easy to explore many design alternatives / Have limited or no functionality / Aim to show the general look and feel / Help communication and exchange of ideas with users / People may be more willing to criticize / But: limits testing possible

High-Fidelity: Actual working system from user perspective (eg can interact but no real backend database Partial to complete functionality)

Enables users to explore with the system (lo-fi enables users to follow a small set of paths)

But, **more costly** to build / **Reluctance to change** the design / Users involved in evaluation often focus on superficial finish Management may think they have a full system / A single bug can lead to a complete halt in evaluation

	Lo-fi	Hi-fi
Time to produce	Cheap and quick	Time consuming
Look	Rough & Sketchy	Polished & Final
Focus	Proposition, concept & Structure	Interaction Design, Appearance & Usability
Usability testing	Requires a guided narrative	Users can complete tasks by themselves

Design

Ideation Iteration: Lightning-fast / Loosely structured iteration / For purpose of exploring design ideas / Role of prototype played by sketches

Conceptual Design Iteration: Critique and compare multiple design concepts / sort out best one / weigh concept feasibility / prototypes: low-fidelity paper, storyboards

Intermediate Design Iteration: To arrive at one intermediate design for layout and navigation / Prototypes might evolve / Fully interactive high-fidelity mockups as vehicle for demonstrations and design reviews

Detailed Design Iteration: To decide screen design and layout details / Includes □visual comps of skin / For look and feel appearance / Design will be fully specified with complete descriptions of – Look and feel – Behavior – How all workflows, exception cases, and settings will be handled

Design Refinement Iteration: Prototype usually medium to high fidelity / Evaluation: Rapid method / Full rigorous process

Doing Intermediate Design: Screen layout and navigational structure / Represent keywork flows with sequences of wireframes, click-through prototypes

Benefits of sketching interfaces: Very low cost / quickly try many ideas / look immature so people are more likely to critique them especially when they compare

Sketching: It involves sets of hand drawn sketches of design ideas for the interface / One series of these is a storyboard that explores a path through the interface /Sketching involves rapid drawings that have the essential elements only / Is a form of brainstorming ... creativity

Persona: create reliable and realistic representations of key audience segments for reference – Represent a **major user group** for your system– Express and focus on the **major needs** and **expectations** of **the most important** user groups– **how they're likely to use the system**– Aid in uncovering **universal features** and **functionality**– Describe **real people** with **backgrounds**, **goals**, and **values**"

Personas are a set of fictional, representative user archetypes based on the behaviors, attitudes, and goals of the people we interview in our research phase.

Personas have names, personalities, pictures, personal backgrounds, families, and, most importantly, goals; **they are not "average" users but specific characters.** A persona is a stand-in for a unique **group of people who share common goals;**

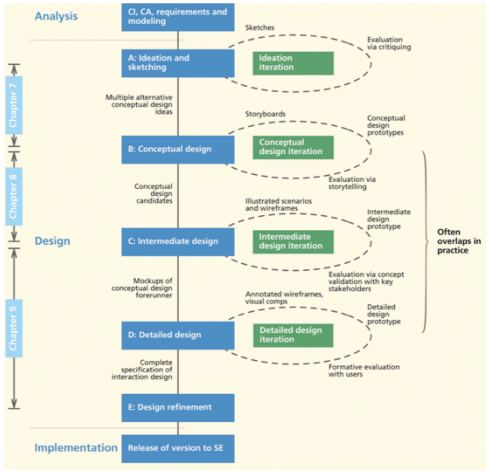
Why is a persona useful: Builds on our ability to think about people, when we "know" them / Valuable as a **communication tool** for the **design team** and other stakeholders / Helps in **discussions in design group** and making **design trade-offs** / Designs can be constantly evaluated against the personas

Perfecting your personas: Add life to the personas, but remember they're design tools first. Use the right goals / Consider experience goals describe how the persona wants to feel when using a product; Most persona goals should be **end goals** that focus on what the persona could get out of using a well-designed product or service. **Personas must be specific to the design problem**

Reconciling market segments and personas: Market **segmentation** and **personas** are two different ... complementary tools that organizations can use to **design** and **sell** successful products

ASS: Capturing Content For Use Outside of The e-Book Reader
Get a hard copy of a page of the e-textbook 1: Suppose you wanted to get the hard copy of pages 170-173, how would you go about doing this? **2:** Suppose you wanted to get the hard copy of chapter 10, how would you go about doing this?
Performance data for all users / qualitative observations over all users

[WK8]-Cognitive Walkthrough



Focus on learnability / Appropriate for novice or casual users / Does not focus on speed / Sequence is not known prior to inspection / Assesses user success and recovery from errors /Conducted by **experts**

Gulf of execution: distance between the users goals and the means of achieving them through the system e.g. what do I do now?, how do I?

Gulf of evaluation: the amount of effort required to determine the system state / When problems occur in the evaluation phase "the gulf of execution" widens

Overcoming Gulfs – how easily determined 1. Function of device 2. Actions possible 3. Mapping from intention to movement 4. Perform the action 5. Tell what state the system is in 6. Tell if the system is in the desired state 7. Mapping from system state to interpretation

Mental Models: what people **believe** for users of a system and system designers. It strongly impacts all aspects of interaction. Mental models drive predictions, planning actions. Relevant things that a person believes

Designer – how the system works / about the user's mental model / – About the user's mental model ; **User** – Beliefs about the system

Cognitive Walkthrough: a systematic way to think like the user by drawing on Norman's gulfs and a model of Exploratory Learning. Conducted by experts who assess the users success and recovery from errors. Focus on learnability not on speed. Sequence not know.

Uses of CW: Relatively inexpensive in our very, very lightweight approach
Desk check – no users / better with expert evaluators
Generally applicable / Novice, casual, intermittent users / Focus on learnability

Usefulness: Really useful technique, even for designer / Better if done by outsider / expert / But students and non-expert evaluators still can gain from using it / Part of early usability evaluation because low cost and no user needed.

Extended Cognitive Walkthrough – adding users mental model
- What does user believe? How do you find this out? What did we assume about the user's mental model?- What differences are there in the mental model for: novice or expert user
+: - relatively inexpensive / - desk check, no users, better with evaluators / - generally applicable and focus on learnability **—:** - need a prototype but its an early technique
- need to define, users, MMs, tasks, steps - repeated over tasks and each class

[WK9]- Questionnaires

Landscape of methods to understand users

Interview, Surveys, Focus Group (ask people what they believe)		Observations of what people actually do (watch)	
- Human memory - Politeness - Self-awareness	+ User involvement, commitment, buy-in + Access to what users know, think, perspectives, contexts that matter	- Stay at the right time - Observer effect Intrusiveness	+ Insights into actual behavior + Contexts + Stakeholder perspectives

Questionnaires (Self - reporting)

A data collection technique based on participant's answers to questions and also widely used collect data about user satisfaction. This can include emotional aspects like how much fun it is use the interface and can explore the broader user experience.
These are an essential way to collect quantitative subjective data. They supplement objective data from the think-aloud observations and you need to consider them together
To use a questionnaire that has a small number of questions is quick and easy.
It gives quantitative data. But require skill to produce so that data are valid and reliable.

Closed questionnaires: These are closed in the sense that the participant has a fixed number of responses and must select one of them.

Likert Scales: Semantic differentials and Closed response questions
Indicate your agreement with the statement. / Whether need a neutral
Strongly agree – Neutral – Strongly disagree – Prefer not to answer
A 7-point scale gives people the level of resolution in usability testing workforce management application.

Open questionnaires: These allow the participant to give free responses.

Simple questions: What was the best thing about this interface? If you could change one thing in this interface what would it be? Any other comments?

Standard questionnaires:

Reliability: How consistent responses are to the questions?
Validity: correlates with other evidence

Sensitivity: discriminates between good interfaces and bad interfaces. The ability to detect differences at even small sample sizes (<10) was one of the major reasons why the System Usability Scale (SUS) and the Single Ease Question (SEQ) are recommended.
Objectivity: Give independent instrument
Quantification: Fine grain of reporting and statistical analysis
Economy: cheap to design
Comparability: + scientific generalization
Modifying a questionnaire can damage its validity and compromises comparability, however, you often can choose a subset of questions and change the word in some questions. Addition, you can add questions for things you need to learn about and using different scale value.

SUS - Goals: effectiveness, efficiency, satisfaction (usability)

50 potential questionnaire statements	The average SUS score is a 68
SUS measures usability & learnability	Reversing the items causes more harm than good
Familiarity breeds content	Usability predicts customer loyalty
Raw SUS scores aren't normally distributed	Can use SUS on small sample sizes
SUS scores were not meant to be diagnostic	SUS might not always be the best questionnaire

SUS is used to measure "usability"- especially learnability, ease-of-use, it is very widely used, free, reliable, valid and useful even with small numbers of participants, particularly useful because the scores have been calibrated. It is as reliable as commercial questionnaires. Many usability professionals understand SUS scores – 68 average. It is not diagnostic (TA is)
Critique - The alternating positive and negative questions to reduce extreme response and following recommended practice at the time. / The conversion to 100 is tedious and people tends to think of the score as a percentage. / The wording could be simpler. /The 5-point scale rather than 7-point scale

PSSUQ- Post-Study Usability Questionnaire
- 16 items survey of satisfaction…sub scales: System Quality / Information Quality / Interface Quality. - The PSSUQ is highly reliable

SEQ – Single ease question
10 questions, highly reliable, generic / only assesses usability

SUPR-Q (usability/credibility/loyalty/appearance)
Score: sum of responses for the first 12 questions + 1/2 the score for the likelihood to recommend question (how likely are you to recommend this website to a friend?). And then compare to the industry benchmarks. / Psychometric validation
Score is a percentile (percentile rank)e.g. Score of 75%: >75% of all websites tested in the db
Compare: The SUPR-Q usability correlates with SUS score / both are post-study
SUS is free and comparative scoring is available freely, SUPR-Q costs
SUS is old and easy to see ways to improve, but it is still very widely used and calibrated.

UMUX (UMUX-lite only positive)
- **question**: This system's capabilities meet my requirements. / Using this system is a frustrating experience. / This system is easy to use. /I have to spend too much time correcting things with this system.
- UMUX-lite is probably good enough. It probably gives similar scores to SUS and so you probably can interpret them as such, do consider it in workplace.

NPS (Net Promotor Score) – very widely used
How likely are you to recommend this product to a friend or colleague?
11 point scale (0:not at all likely to / 10:extremely likely to recommend).
People with: 9, 10 -promoters / 7, 8 – passives / 0-6 detractors
Score = %age of promoters - %age detractors
100promoters, 30passive, 80detractors – 20/210*100 = 9.5%
score > 0 is good, >50 excellent, > 70 exceptional
has been widely used but it ignores the "passives" so: - people are poor predictors of their future behavior / - Better to ask "why" for detractors

hedonism: the unrestrained pursuit of pleasure
(in usability) the aspects of a user interface that appeal to a person's desire of pleasure and avoidance of boredom and discomfort. The aspects that are fun, original, interesting, engaging, and cool. A positive subjective experience.

AttrackDiff: The model separates the 4 essential aspects: – The product quality intended by the designer. / – The subjective perception of quality and subjective evaluation of quality [by the user]. / – The independent pragmatic and hedonic qualities. / – Behavioural and emotional consequences.
Participants are asked to rate the interface on a likert scale 1..7 (pragmatic/hedonic/appeal)

p-values – P < 0.05 means there is less than 5% chance that these results were observed were due to chance
– usual threshold is .05 in much usability work (but in other fields and contexts .01 or even .001)

[WK10]- HCI Laws

Fitts' Law: The time to move quickly to a target / Further takes longer / smaller targets take longer/ This is really important in the fine-grained design of interfaces / The theory provides detailed formulae / The constants differ for different forms of interaction.
Movement divided into 2 phases: high velocity and deceleration

$$MT = a + b\log_2(\frac{D}{W} + 1)$$

MT: time to reach target
D: distance W: is width of target
a,b: constants
It's also important to account for high risk interactive elements that you don't want the user to accidentally activate. **Those should often be kept further away from heavily used interactive items.**

Fitts' Rule: Create Larger Targets / Minimize Cursor Movement but at the same time, group things by function and class / Avoid Muscular Tension Take account of movements that are ackward. **Engineers put more time** into simplifying an application versus making **millions of users spend an extra minute** using the more complex system.

Fitts': Empirical foundations / Heuristics for designers – Grouping items for flow of action / Special locations (prime pixel / magic pixels)

In HCI, smaller end is often best

Hick's Law: The time to choose between equally likely alternatives increases as the log of the choices. Time=b.Log2 (n+1). This is a matter of time for processing choices. Response time for unanticipated choice
Applicability of Hick's Law is limited. It does NOT apply when: alternatives are not equally likely or there is more happening than just the choice

Power of Practice Law: Time to do many mechanical ("sensory-motor") and cognitive skills, gets faster with repetition

$$T_n = T_1 * n^{-a} \text{ (with a ~ 0.2-0.6)}$$

Stages in building competence: Unconscious incompetence / Conscious incompetence: Working on learning / **Conscious competence:** Have learnt task but still need to concentrate / **Unconscious competence:** Task has been automated

[WK11]- GOMS: Evaluating speed / efficiency

GOMS: method for expert users assessing the speed of use, done through tasks
- Goals. - Operations – keystrokes, clicks. - Methods – sets of operations
- Selection rules – decide between methods

Who – experts evaluators in using the application to use method and select tasks - trained in domain and in the tools - use method well and selection of appropriate tasks to consider
What – focus on performance speed (of task completion), tools that people use extensively, automating interactions. **Character**: predictive / no users needed / low cost
When: early design / potentially also when revisiting the design in Agile approaches / when considering what interaction options to implement
Role and Time: designing the interaction methods or deciding what interaction methods to implement

Keystroke analysis

K - keypress	**P – point with mouse**	**C – click with mouse**
H – home hands on new device	**M – mentally prepare**	**R(t) – System response time**

- M before K/C or P - MPMC becomes MPC because C anticipated
- MKMKMK…MKKK for cognitive unit
GOMS analysis helps a designer think about alternative designs for the interactions for tasks.

0.08-1.20	0.8-1.5	0.2
0.4	1.35	?

[WK12]- Accessibility and links to visual perception

Accessibility: enabling web access to people with disabilities through browsers and media players, meshing with assistive technologies, authoring tools to produce accessible content
- important for disabilities. - Overall better design
- Critical information and services on the web. - Legal requirements and financial merits
- marketplace issues – widespread, aging population, legal costs
* Good design helps everyone. Thinking about accessibility for those with handicaps. Makes for sites that are more accessible to others too. **Clarity over cuteness.**
Nature of disabilities – vision, hearing, physical limitations, cognitive and neurological
Web content accessibility guidelines 2.0: perceivable / operable / understandable / robust
Approaches - Text is king – for assistive and search technologies
- Alternative modalities – image plus text, audio plus text caption - Use of style sheets

[WK13]- A/B testing, UX versus XP

Different Types of data: Quantitative or Categorical
If p < **0.05**, reject null hypothesis; there is a "significant" difference between A and B.
Never draw a conclusion merely based on a p value.
Always include "**statistically**" as a prefix when using the word "**significant**" to describe a p value-based finding.

A/B testing: split testing, bucket testing, and multivariant testing

	A/B Testing	Usability	Radical Innovation
Cost	Low	Low–medium	High (unless lucky)
Benefits	1–10%	10–100%	100–1,000%
Risk	None	Low	High
Who can do?	Everybody	Everybody	Geniuses
How often?	Weekly	Monthly	Every 10 years
GDP impact	Medium	High	Medium

An example of an Empirical form of Field study beyond the scope of this unit.
Controlled experiments with website and authentic users
Control and test variant, coolies used to ensure version consistency and metric used to measure

Reading

An experience Sampling Study of User Reactions to Browser Warnings

Two challenges: increasing both comprehension and adherence rates.
When encountering warning messages in situ, participants have a wide variety of reasons for choosing to adhere to or proceed past a given warning.
Habituation plays a smaller role in user decision making.
ESM (Experience Sampling Method): ①how chrome & Firefox users make decisions about contemporary warning; ②let them take a range of factors into account when deciding whether to adhere to a warning – focus questions: ①why do users still not adhere to or comprehend warnings? ②when they adhere to warnings, why do they do?

Informal Cognitive Walkthroughs (ICW):

Abstract: Agile software teams' frequent releases and fast iterations present a growing need for rigorous user experience research methods that are faster, lighter-weight, and more flexible. To this end, we developed the Informal Cognitive Walkthrough (ICW). This agile research methodology grew organically, over the course of three years, while working with a very large agile software development team. ICWs involve conducting one or more Simplified 'Streamlined Cognitive Walkthroughs' (SSCW), followed by one or more Simplified 'Pluralistic Walkthroughs' (SPW). In this paper, we present the ICW and provide a real-world example of its application. Preliminary experiences with the method revealed potential advantages over traditional lab studies, ranging from more quickly uncovering and fixing usability issues, to a stronger collaboration between the disciplines, and to acting as a forcing function in aligning diverse engineers to deliver on a common user goal.

Cognitive Walkthrough (CW): The Cognitive Walkthrough (CW) is a usability inspection method which involves "simulating a user's problem- solving process at each step in the human-computer dialog, checking to see if the user's goals and memory for actions can be assumed to lead to the next correct action"
-: criticized for its overhead (time consuming, tedious form-filling)

	SPW	PW	Think-aloud
Process	Users were provided with blank generic comment forms, as opposed to printouts with the step-by-step screens for PWs	The users make their points first at each step	1.Recruit representative users 2.Give them representative tasks to perform. 3.Shut up and let the users do the talking.
Benefits	Agility Flexible	Individually write down their answers before sharing	Cheap Flexible Convincing Easy to learn
Notes	Simplification	Complex	No standard

Similarities	Both involve users, so both are a user-method (not an expert-method) Both require a well-designed set of tasks

Quantifying visual preferences around the world

Computational model: combines a user's demographic information with computational image metrics for assessing a website's colorfulness and complexity in order to predict a user's subjective perception of visual appeal.
This report supposed that **complexity** is more important as a predictor of appeal than colorfulness, and demographic background significantly affects preferences for colorfulness and complexity.
Visual complexity –clutter, strong predictor of appeal, more important than colorfulness, lower less severe consequences, low-to-medium appeal to most
Colorfulness – affects perceived trust worthiness, significant effect, inverted U shape, not as dramatic response, medium to high colorfulness appeal to most
Influence of demographics – AGE -medium age lower colorful, older more complex, GENDER – more colorful females, visual complexity similar, COUNTRY, EDUCATION – lower more colorful & more complex
- limitations: causes of results, interactions between demographic variables, highly subjective nature of visual appeal, cultural differences difficult to understand at the moment.

Getting the Right Design / the Design Right: testing many is better than one.

When presented with a single design, users give significantly higher ratings and were more reluctant to criticize than when presented with the same design in a group of three. The author thinks that multiple designs can bring a more comprehensive comment which include negative and positive, and which aspects could be better according to those alternatives.
- Aesthetics (matches 8: aesthetic & minimalist design)
- Consistency (matches 4: consistency & standards, real-world metaphors)
- explorability (matches 3: user control & freedom)

Utilizing Employees as Usability Participants:

The main goal is exploring shall we do some usability testing with the company's employees. Or whether employee participation is significantly different from public participation. For the company's products, there doesn't have too much difference between the employees and non-employees. For competitor products, compared with non-employees, the company's employees may spend less time but give a more critical response. **Different devices has different view.**

System Concept Statement (SCS)

A SCS is a concise descriptive summary of the envisioned system or product stating an initial system vision or mandate; in short, it is a mission statement for the project.
Mission statement for the system
Includes – system name, target users, what the system is intended to do, the problems the system should solve & aspects of the user experience

Review questions

TA when users find the task difficult – help them as soon as it was clear they were struggling, ended the tasks and thanked the user – should not allow user to struggle violated need for participants not to have frustrations, record as failure but make notes about the help given and observations

Show me how you would get started on that. / Show me how you would do that. / Suppose your assignment is to learn about UX goals. Find the section about that and read the title to me. / Please do that now.

Abstract Task 1.1: Highlight text on the current page

Rationale: This makes it easier for a reader to locate aspects they consider important
* Concrete -- They are ready for use in a think-aloud study – a user understands what they need to do after reading them
* Relevant -- This is a valid case to include in a user study for assessing abstract task because it is a concrete instance of that abstract task.
A UX measure is a usage attribute to be assessed in evaluating a UX goal.
A measuring instrument is the means for providing values for a particular UX measure.
Benchmark task descriptions we create as measuring instruments
A benchmark task is a representative task that you will … take qualitative data and quantitative data.

Mental Models
TA – gives access to mental model
HE – takes into account mental model, makes assumptions
CW – based around invidual mental model
GOM – relationship between mental models and time

HE
Hard to select a particular paragraph sometimes. - User control and freedom.
Couldn't choose page range to download (only page by page)
Couldn't print directly (only print to PDF)
Some buttons have similar function, such as "Chapter Down" and "Print to PDF" - Consistency and standards

	CW	HE	COMS	TA	OSM
Learnable?	Y	Y	N	Y	Y
Efficient?	N	N	Y	~N	Y
Memorable?	N	~Y	N	Y	Y
Errors?	~Y	~N	N	Y	Y
Satisfaction	N	N	N	~Y	~Y
Useful	N	N	N	~N	Y

(extended) Cognitive Walkthrough / Observational study, monitoring
1. text is low contrast: hard for low-vision users to read.
2. labels on the left and fields on the right: hard for many people to associate them, and almost impossible for someone who needs to zoom in to use the page.
3. "Remember details?" label isn't associated with the checkbox: so you have to tap or click only on the tiny square rather than just clicking the label; also, someone using a screen reader would have trouble figuring out the association.

Important things for **TA**: preparation / introduction / test / questionnaire or interview / debrief
Important things for **HE**: preliminaries / evaluation / concluding summary

Why screen elements should not blink - hard to ignore, reduces legibility, can be annoying
Relationship between Fitts Law and GOMS - assumption that P is varied according to the pixels, magic pixels don't have the deceleration phase
Fitt's law & hierarchical menus – Rule 2: reduces travel distance, minimises cursor movement Improving the design – vertical drop down, colour consideration