



Ing. Otto Parra González

INTERFAZ HOMBRE- MAQUINA





INTERFAZ HOMBRE-MAQUINA

EVALUATION TECHNIQUES: EVALUATION THROUGH USER PARTICIPATION





INTRODUCTION

- The techniques to evaluate a design or system through analysis by the designer, or an expert evaluator, rather than testing with actual users are useful for filtering and refining the design, they are not a replacement for actual usability testing with the people for whom the system is intended: the users.
- Evaluation through user participation includes empirical or experimental methods, observational methods, query techniques, and methods that use physiological monitoring, such as eye tracking and measures of heart rate and skin conductance.



INTRODUCTION

- User participation in evaluation tends to occur in the later stages of development when there is at least a working prototype of the system in place. This may range from a simulation of the system's interactive capabilities, without its underlying functionality (for example, the *Wizard of Oz* technique).
- However, some of these methods can also contribute to the earlier design stages, such as requirements capture, where observation and surveying users are important.



STYLES OF EVALUATION

- We will distinguish between two distinct evaluation styles:
 - Those performed under laboratory conditions.
 - Those conducted in the work environment or ‘in the field’.



STYLES OF EVALUATION: LABORATORY STUDIES

- In the first type of evaluation studies, users are taken out of their normal work environment to take part in controlled tests, often in a specialist usability laboratory.
- A well-equipped usability laboratory (sophisticated audio/visual recording and analysis facilities, two-way mirrors, instrumented computers, etc.), which cannot be replicated in the work environment. In addition, the participant operates in an interruption-free environment. However, the lack of context and the unnatural situation may mean that one accurately records a situation that never arises in the real world.



STYLES OF EVALUATION: LABORATORY STUDIES

- There are, however, some situations where laboratory observation is the only option, for example, if the system is to be located in a dangerous or remote location, such as a space station. Also some very constrained single-user tasks may be adequately performed in a laboratory.
- Finally, and perhaps most commonly, we may deliberately want to manipulate the context in order to uncover problems or observe less used procedures, or we may want to compare alternative designs within a controlled context.
- For these types of evaluation, laboratory studies are appropriate.

STYLES OF EVALUATION: FIELD STUDIES

- The second type of evaluation takes the designer or evaluator out into the user's work environment in order to observe the system in action.
- High levels of ambient noise, greater levels of movement and constant interruptions, such as phone calls, all make field observation difficult.
- However, the very 'open' nature of the situation means that you will observe interactions between systems and between individuals that would have been missed in a laboratory study.
- The context is retained and you are seeing the user in his 'natural environment'. In addition, some activities, such as those taking days or months, are impossible to study in the laboratory (though difficult even in the field).



STYLES OF EVALUATION: FIELD STUDIES

- On balance, field observation is to be preferred to laboratory studies as it allows us to study the interaction as it occurs in actual use. Even interruptions are important as these will expose behaviors such as saving and restoring state during a task.
- However, we should remember that even in field observations the participants are likely to be influenced by the presence of the analyst and/or recording equipment, so we always operate at a slight remove from the natural situation.



EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- One of the most powerful methods of evaluating a design or an aspect of a design is to use a controlled experiment. This provides empirical evidence to support a particular claim or hypothesis. It can be used to study a wide range of different issues at different levels of detail.
- Any experiment has the same basic form. The evaluator chooses a hypothesis to test, which can be determined by measuring some attribute of participant behavior.
- A number of experimental conditions are considered which differ only in the values of certain controlled variables. Any changes in the behavioral measures are attributed to the different conditions. Within this basic form there are a number of factors that are important to the overall reliability of the experiment, which must be considered carefully in experimental design. These include the participants chosen, the variables tested and manipulated, and the hypothesis tested.



EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- Participants:
 - The choice of participants is vital to the success of any experiment. In evaluation experiments, participants should be chosen to match the expected user population as closely as possible.
 - A second issue relating to the participant set is the sample size chosen. Often this is something that is determined by pragmatic considerations: the availability of participants is limited or resources are scarce.



EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- Variables.
 - Experiments manipulate and measure variables under controlled conditions, in order to test the hypothesis. There are two main types of variable: those that are 'manipulated' or changed (known as the independent variables) and those that are measured (the dependent variables).



EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- Hypothesis.
 - A hypothesis is a prediction of the outcome of an experiment. It is framed in terms of the independent and dependent variables, stating that a variation in the independent variable will cause a difference in the dependent variable.
 - The aim of the experiment is to show that this prediction is correct. This is done by disproving the null hypothesis, which states that there is no difference in the dependent variable between the levels of the independent variable.

EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- *Experimental design.*
 - In order to produce reliable and generalizable results, an experiment must be carefully designed. We have already looked at a number of the factors that the experimenter must consider in the design, namely the participants, the independent and dependent variables, and the hypothesis.
 - The first phase in experimental design then is to choose the hypothesis: to decide exactly what it is you are trying to demonstrate.
 - The next step is to decide on the *experimental method* that you will use. There are two main methods: *between-subjects* and *within-subjects*.

EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- In a between-subjects (or *randomized*) design, each participant is assigned to a different condition.
 - The advantage of a between-subjects design is that any learning effect resulting from the user performing in one condition and then the other is controlled: each user performs under only one condition.
 - The disadvantage is that a greater number of participants are required.
- The second experimental design is within-subjects (or *repeated measures*).
 - Here each user performs under each different condition.
 - Within-subjects is less costly than between-subjects, since fewer users are required

EMPIRICAL METHODS: EXPERIMENTAL EVALUATION

- *Statistical measures*
 - The first two rules of statistical analysis are to *look* at the data and to *save* the data. It is easy to carry out statistical tests blindly when a glance at a graph, histogram or table of results would be more instructive.
 - In particular, looking at the data can expose *outliers*, single data items that are very different from the rest. Outliers are often the result of a transcription error or a freak event not connected to the experiment.
 - Our choice of statistical analysis depends on the type of data and the questions we want to answer.

INTERFAZ HOMBRE-MAQUINA



EVALUATION TECHNIQUES: CHOOSING AN EVALUATION METHOD



INTRODUCTION

- A range of techniques is available for evaluating an interactive system at all stages in the design process. So how do we decide which methods are most appropriate for our needs? There are no hard and fast rules in this – each method has its particular strengths and weaknesses and each is useful if applied appropriately.

FACTORS DISTINGUISHING EVALUATION TECHNIQUES

- We can identify at least eight factors that distinguish different evaluation techniques and therefore help us to make an appropriate choice:
 - the stage in the cycle at which the evaluation is carried out
 - the style of evaluation: laboratory or field study
 - the level of subjectivity or objectivity of the technique
 - the type of measures provided: quantitative or qualitative
 - the information provided: low-level information or higher-level information
 - the immediacy of the response: now or later
 - the level of interference implied
 - the resources required.

A CLASSIFICATION OF EVALUATION TECHNIQUES

Classification of analytic evaluation techniques

	Cognitive walkthrough	Heuristic evaluation	Review based	Model based
Stage	Throughout	Throughout	Design	Design
Style	Laboratory	Laboratory	Laboratory	Laboratory
Objective?	No	No	As source	No
Measure	Qualitative	Qualitative	As source	Qualitative
Information	Low level	High level	As source	Low level
Immediacy	N/A	N/A	As source	N/A
Intrusive?	No	No	No	No
Time	Medium	Low	Low–medium	Medium
Equipment	Low	Low	Low	Low
Expertise	High	Medium	Low	High

A CLASSIFICATION OF EVALUATION TECHNIQUES

Classification of experimental and query evaluation techniques

	Experiment	Interviews	Questionnaire
Stage	Throughout	Throughout	Throughout
Style	Laboratory	Lab/field	Lab/field
Objective?	Yes	No	No
Measure	Quantitative	Qualitative/ quantitative	Qualitative/ quantitative
Information	Low/high level	High level	High level
Immediacy	Yes	No	No
Intrusive?	Yes	No	No
Time	High	Low	Low
Equipment	Medium	Low	Low
Expertise	Medium	Low	Low

EVALUATION TECHNIQUES

9.4 Choose an appropriate evaluation method for each of the following situations. In each case identify:

- (i) the participants
 - (ii) the technique used
 - (iii) representative tasks to be examined
 - (iv) measurements that would be appropriate
 - (v) an outline plan for carrying out the evaluation.
-
- (a) You are at an early stage in the design of a spreadsheet package and you wish to test what type of icons will be easiest to learn.
 - (b) You have a prototype for a theatre booking system to be used by potential theatre-goers to reduce queues at the box office.
 - (c) You have designed and implemented a new game system and want to evaluate it before release.
 - (d) You have developed a group decision support system for a solicitor's office.
 - (e) You have been asked to develop a system to store and manage student exam results and would like to test two different designs prior to implementation or prototyping.