CS 598 ACK
Experimental HCI & Interactive Technologies

Defining the research (Part 1 of 2)
- *The research question:* a clear question that succinctly states the aim of the research;
- *Conditions:* the ideas of interest – these will be compared against each other;
- *The independent variable:* the set of conditions to be used in the experiment – there will always be more than one condition;
- *The population:* all the people who might use the idea; *the sample:* the set of people who will take part in the experiment;
- *Generalisability:* the extent to which experimental results can apply to situations not explicitly included in the experiment itself;
- *Experimental objects*: the way in which the ideas are presented to the participants – experimental objects embody the conditions so that they can be perceived;
- *Experimental stimulus*: the combination of an experimental object and a condition;
- *Experimental tasks*: what the participants will actually do with the experimental objects;
- *Experimental trial*: the combination of a condition, an experimental object, and a task.
The Research Question

EVEN THOUGH: Experiments or often (if not always) conducted as parts of larger, more substantial research projects, each particular experiment must begin with first defining a specific research question (or specific questions).

A useful heuristic for doing so is to ensure that the experiment is stated as having the goal of providing yes or no answers to research questions phrased (for the purpose of this course, in English) as well formed (grammatically correct) statements ending with “?”s.

E.g., Are our intended users more able, on average, in learning how to successfully connect a Bluetooth device to Bluetooth speaker while using interface A or while using interface B for the speaker’s design?
Examples of inappropriately phrased research questions are as follows:

- “To investigate the use of a visual mouse in a text reading task”; 
- “Asking people to draw graphs using a visual mouse and seeing if they like it”; 
- “Seeing if the visual mouse works.”

These could be better stated as follows:

- “Is reading a piece of text using a visual mouse more efficient than when using a physical mouse?” 
- “Do users prefer a visual mouse to a physical mouse when drawing graphs?” 
- “How accurate is the use of a visual mouse when performing fine-grained interaction tasks?”
Defining the research (Part 1 of 2)

The Research Question

“Ideally, of course, the experiment or evaluation should be conducted by someone (or a team) who has not been associated with developing the HCI idea, although this is seldom the case for academic research projects. Lieberman (2003) points out that it would be unthinkable for a new medical technique to be evaluated by the person who developed it.”

(where have we heard this before?!)
“Some researchers, especially those with a psychology background, prefer to express their experiment aims in terms of a null hypothesis statement that they will ultimately try to reject as being false (e.g., “There will be no efficiency difference when reading a piece of text between a visual mouse and a physical mouse”). Although this is a valid approach, I find that starting off with clear, focused research question is a better (and often less confusing) starting point.”
Defining the research (Part 1 of 2)

The Research Question

“There are some researchers, especially those with a psychology background, who prefer to express their experiment aims in terms of a null hypothesis statement that they will ultimately try to reject as being false (e.g., “There will be no efficiency difference when reading a piece of text between a visual mouse and a physical mouse”). Although this is a valid approach, I find that starting off with a clear, focused research question is a better (and often less confusing) starting point.”

Yet, in actuality both types of statements (of research questions) will be necessary, at least if we wish to employ inferential statistical methods to analyze data resulting from an experiment. – Why?
Defining the research (Part 1 of 2)

The Research Question

“Why?”

First, we will discuss this issue much more thoroughly when we get to the course material on statistical methods.

Yet, it is useful now to appreciate that in the vast majority of HCI experiments, no research questions, so as those we have discussed, can be answered with absolute certainty. Why?
The Research Question

“Why?”

First, we will discuss this issue much more thoroughly when we get to the course material on statistical methods.

Yet, it is useful now to appreciate that in the vast majority of HCI experiments, no research questions, so as those we have discussed, can be answered with absolute certainty. Why?

There will always be unavoidable risks in generalizing conclusions from samples (of subjects, of tasks, etc.) to populations, and there will always be some level of unpreventable noise or random error.
Conditions for Comparison

“The key to HCI experiments is the notion of comparison:

We compare the performance of one HCI idea against another. One or more alternative ideas need to be identified. Importantly, the alternatives must offer the same functionality as the idea you want to test; otherwise, a comparison is unfair (as in “comparing apples to oranges”).
Defining the research (Part 1 of 2)

Conditions for Comparison

“The key to HCI experiments is the notion of comparison:

We compare the performance of one HCI idea against another. One or more alternative ideas need to be identified. Importantly, the alternatives must offer the same functionality as the idea you want to test; otherwise, a comparison is unfair (as in “comparing apples to oranges”).

YES, but let’s recall that we are involved in a venture that includes both discovery and invention. Increasingly, HCI researchers are involved with both inventing new system functionality as well as novel interfaces for existing system functionality. This can present challenges to designing experiments (examples?) yet defining the research question precisely is almost always the key to overcoming these challenges.
“Conditions for comparison” can be expressed in experimental design terms as choosing values of “independent variables,” as we have already touched on previously.

Independent variables must have at least 2 values or there wouldn’t be a research question being put to test.

These variables must be defined prior to experimentation.

After the data have been analyzed, the researcher may discover what appears to be an additional independent variable, at work during the experiment. Yet this information can only sometimes be validly reported among the “results.” Mainly, it should instead serve as fodder for additional experiments where such variables are in fact stated up front.
Defining the research (Part 1 of 2)

Conditions for Comparison

Getting a bit ahead with this, but always (try to) graph your data first!
Similar to other forms of design, experimental design is about making decisions. There are many choices to be made, and all alternatives should be considered. The concept of design space is useful here – a design space is an imaginary representation of the design options possible in a multidimensional space. Design is the process of analysing that multidimensional space to identify the best set of options, while taking into account constraints and the experimental aims.
However, it is important to note that no experiment can be perfect. The aim of the experiment is to investigate whether the values of the different conditions (the independent variable) have an effect on performance, satisfaction, enjoyment, usability, or anything that we can observe and record. The ideal experiment would ensure that only the independent variable is ever varied and that everything else in the experimental environment stays exactly the same. This is impossible – we cannot eliminate all extraneous factors that might affect the data collected.
Defining the research (Part 1 of 2)

Designing Experiments

“Experiments may be interrupted, software may crash, the sun might be shining on one day and not the next (affecting the perception of colors on a screen), participants might be in a bad mood (and not willing to give their best performance), and/or some might have useful background knowledge not possessed by others.

Extraneous factors such as these that might affect the reliability of the data are called confounding factors: some will be easy to identify (e.g., a computer that works slower than expected for some participants, the different content in Web page stimuli), and some will not (e.g., prior knowledge of each participant, whether a participant is tired or motivated that day).”
Designing Experiments

“Those aspects of the experimental environment that the experimenter can keep the same throughout the experiment are called experimental controls.

Thus, different factors of the experiment are (1) deliberately varied (the independent variable); (2) deliberately kept the same (the experimental controls or “control conditions”); or (3) accidentally varied (the confounding factors).”
HCl experiments are all assumed to have a target population in mind.

Experiments typically use a set of experimental participants (the sample), collect responses to the idea from these participants, and make generalizations from these responses (i.e., speculate about how the wider population would respond).

“At one extreme, a sample of just one person will produce results that cannot reasonably be generalized to the population; at the other extreme, a large sample that includes all members of the population is complete, and does not need to be generalized. The larger and the more diverse the sample of people used, the more likely it is that the generalization will be valid.”
Defining the research (Part 1 of 2)

Generalizability

The Graph shows the number of boys and girls playing sport in an English town in 2012.
Defining the research (Part 1 of 2)

Generalizability

“It is useful to consider other relevant generalizations, in particular, what the participants are asked to do (the experimental tasks) and the means by which the abstract idea is instantiated in a concrete form (the experimental objects).

If we ask our participants to perform only one task using the new HCI idea, then the results can only be representative of performance with that one task. And, as with the previous sample/population difference, it is typically impossible to ask the participants to perform every possible task (apart from when testing very limited systems or ideas).”
Defining the research (Part 1 of 2)

Generalizability

“Just as we have a population of people (from which we select a sample), we also therefore have a population of possible tasks (from which we select a sample) and a population of experimental objects (from which we also select a sample).

In the same way that we aggregate results over people (to generalize to the human population), we can also aggregate results over sample tasks and sample experimental objects to generalize the results to apply to more than one type of task and more than one type of experimental object.”

We can therefore speculate as to what results would have been obtained if a wider range of tasks or experimental objects had been used.”
Experimental Objects

“The HCI idea that is being tested, as well as the alternative conditions, need to be embodied in a perceivable experimental object: this is the way in which the idea is presented to the participants in a concrete manner, rather than as an abstract idea.

Experimental objects are the concrete means by which a participant can experience the HCI idea and its alternatives. If only one experimental object is used in the experiment, then the results of the experiment will only pertain to that particular experimental object.”

OFTEN, in my experience, achieving reliable experimental results or good HCI design requires a qualitative research phase (ethnography, etc.) to be done prior to experimentation to identify, select or design these objects.
Defining the research (Part 1 of 2)

Experimental Objects
Defining the research (Part 1 of 2)

Experimental Objects
Case Study: A major customer (left) of a major Point-of-Sale (POS) research and technology company (right) asked for our help in designing a next generation of POS terminals with a focus on database and interface design targeted specifically on the special needs and roles played by on-floor sales staff – and in particular, to put the customer “into a new pair of shoes today.”

Because we wanted to use experienced and knowledgeable employees in our user studies and experiments (to ensure immediately positive user acceptance and performance upon system implementation), we first conducted focus-group size interviews and discussions with them, and observed them in action.
Experimental Objects

A striking initial finding was that cases in which their inventory (the huge room of shoeboxes “behind the door” contained the exact shoe brought to them by the customer took up only a trivial fraction of their time.

Close to 90% of their time and energy went into efforts to still sell the customer a pair of shoes when the exact shoes initially requested were not in inventory, so workers needed to then find a “closest match” (or matches) to shoes they did have in immediate inventory.

Appreciating that the organization of our database and interface design would need to be compatible with their strategies for finding such near matches, we examined how these workers had learned to best organize their shoe inventory to help them readily find these close matches.
Defining the research (Part 1 of 2)

Experimental Objects

Think of the shoe inventory in terms of an N-dimensional space, with the (most obvious) dimensions of:

1. “Ladies shoes” versus “Mens shoes” – in their parlance.
2. Brand
3. Color
4. Price
5. Style
6. Size
7. Vegan or Not (just kidding, but learned that Zappos has this search option)
8. Anything else you can think of (??)

Questions: Given any 2 adjoining pair of shoe boxes chosen at random, on which of the above dimensions were they closest? Most distant?
Defining the research (Part 1 of 2)

Experimental Objects

Answer #1: Most distant: “Ladies” vs “Mens” – Separate locations in stores.

1. “Ladies shoes” versus “Mens shoes” – in their parlance.
2. Brand
3. Color
4. Price
5. Style
6. Size
7. Vegan or Not (just kidding, but learned that Zappos has this search option)
8. Anything else you can think of (??)

Questions: Given any 2 adjoining pair of shoe boxes chosen at random, on which of the above dimensions were they closest? Most distant?
Defining the research (Part 1 of 2)

Experimental Objects

Answer #Surprise! The “closest dimension” differed based on Answer #1

1. “Ladies shoes” versus “Mens shoes” – in their parlance.
2. Brand
3. Color
4. Price
5. Style
6. Size
7. Vegan or Not (just kidding, but learned that Zappos has this search option)
8. Anything else you can think of (??)

So, New Questions: What was the closest dimension for Mens shoes, and what was the closest dimension for Ladies shoes?
Defining the research (Part 1 of 2)

Experimental Objects

Answer for Mens shoes closest dimension: (5) Style

1. “Ladies shoes” versus “Mens shoes” – in their parlance.
2. Brand
3. Color
4. Price
5. Style
6. Size
7. Vegan or Not (just kidding, but learned that Zappos has this search option)
8. Anything else you can think of (??)
Defining the research (Part 1 of 2)

Experimental Objects

Answer for Women shoes closest dimension: (3) Color.

1. “Ladies shoes” versus “Mens shoes” – in their parlance.
2. Brand
3. Color
4. Price
5. Style
6. Size
7. Vegan or Not (just kidding, but learned that Zappos has this search option)
8. Anything else you can think of (??)