SUBJECTIVE MEASURES (CONTINUED)
QUESTION DESIGN

Main challenge is response bias

A set of principles to avoid response bias
DESIGNING GOOD QUESTIONS

Principle

Avoid “leading” or “loaded” questions

Example

“Don’t you agree that social workers should earn more money than they currently earn?”

Yes, they should earn more

No, they should not earn more

Don’t know/no opinion
DESIGNING GOOD QUESTIONS

Principle

Avoid double negatives

Example

“Do you agree or disagree with the following statement?

Teachers should not be required to supervise their students during recess.”
DESIGNING GOOD QUESTIONS

Principle

Always aim at capturing firsthand experiences and beware of asking about information that is acquired only secondhand

Example

People are very good at describing criminal activity directed at them but terrible at describing how much crime happens in their neighborhood
DESIGNING GOOD QUESTIONS

Principle

Beware of asking hypothetical questions — people are not good at predicting what they will do as they have limited direct experience with future situations.
DESIGNING GOOD QUESTIONS

Principle

Beware of asking about causality — events mostly have more than one reason + people are not good at describing why they do the things they do

Example

“Were you limited in your daily activities because of your back problem?”

“What is the main reason why you did not vote?”

“Were you homeless because of high cost of housing?”
DESIGNING GOOD QUESTIONS

Principle

Beware of asking about solutions to complex problems — people in general do not have informed opinions about complex issues.
DESIGNING GOOD QUESTIONS

Principle

Avoid asking more than one question at a time — the answers to two questions can be dramatically different

Example

“Would you like to be rich and famous?”

“Are you physically able to do things like swim and run without difficulty?”
DESIGNING GOOD QUESTIONS

Principle

Avoid asking questions that impose unwarranted assumptions — double-barreled or one-and-a-half-barreled questions

Example

“Would you vote for or against a candidate who supports reducing federal spending on education and welfare?”

“With the economy the way it is, do you think investing in the stock market is a good idea?”
DESIGNING GOOD QUESTIONS

Principle

Beware of questions that include hidden contingencies — questions must apply to the majority of your sample

Example

To measure social activity:

“How often did you attend religious services or participate church-related activities during the past month?”
DESIGNING GOOD QUESTIONS

Principle

The words in questions should be chosen so that all respondents understand their meaning, and all respondents have the same sense of what the meaning is.
DESIGNING GOOD QUESTIONS

Principle

When words or terms that have meanings that are likely no to be shared, definitions should be provided to all respondents.

Example

In the past 12 months, how many times have you seen or talked with a medical doctor about your health?

If asked: “Include visits to psychiatrists, ophthalmologists, and any other professional with a medical degree.”
DESIGNING GOOD QUESTIONS

Principle

The time period referred to by a question should be unambiguous. Questions about feelings or behaviors must refer to a period of time.

Example

“Are you able to run half a mile without stopping?”

“How many drinks do you usually have on days when you drink any alcoholic beverages at all?”
DESIGNING GOOD QUESTIONS

Principle

If what is to be covered is too complex to be included in a single question, ask multiple questions.
DESIGNING GOOD QUESTIONS

Principle

Use multiple questions to measure the same thing
DESIGNING GOOD QUESTIONS

Principle

If definitions are provided, they should be given before the question itself is asked.

Example

“How many days in the past week have you done any exercise? When you consider exercise be sure to include walking, work around the house, or work on a job, if you think they constituted exercise.”

**Better:** “The next question is going to ask you about how often you’ve engaged in exercise. We want to you to include walking, anything you may do around the house, or work around the house, or work on a job, if you think they constituted exercise. Using this definition, in the last week, on how many days did you do any exercise?”
DESIGNING GOOD QUESTIONS

Principle

A question should end with the question itself. If there are response alternatives, they should constitute the final part of the question.

Example

“Would you say that you are very likely, fairly likely, or not likely to move out of this house in the next year?”

**Better:** “In the coming year, how likely are you to move to another home? Would you say very likely, fairly likely, or not very likely?”
DESIGNING GOOD QUESTIONS

Principle

A question should end with the question itself. If there are response alternatives, they should constitute the final part of the question.

Example

“Would you say that you are very likely, fairly likely, or not likely to move out of this house in the next year?”

Better: “Which one of these categories best describes how likely you think you are to move in the next year: very likely, fairly likely, or not likely?”
DESIGNING GOOD QUESTIONS

Principle

Clearly communicate to all respondents the kind of answer that constitutes an adequate answer to a question

Example

“When did you move to this community?”

Possible answers:

When I was sixteen.
Right after I was married.
In 1953.

Better: “In what year did you move to this community?”
DESIGNING GOOD QUESTIONS

Principle

Specify the number of responses to be given to questions for which more than one answer is possible

Example

“What was it about the brand you bought that made you buy it rather than some other brand?”
DESIGNING GOOD QUESTIONS

Principle

Design survey instruments to make the tasks of reading questions, following instructions, and recording answers as easy as possible for interviewers and respondents.
DESIGNING GOOD QUESTIONS

Principle

Measurements will be better to the extent that people answering questions are oriented to the task in a consistent way — training respondents
CATEGORICAL QUESTIONS

Principles

Mutually exclusive response categories
Exhaustive list of categories
Don’t know/no answer category

Example

“What is your current age?”

10 or less
10 to 20
20 to 30
30 to 40
DESIGNING QUESTIONS

What is wrong with these questions?
How can we fix them?
EXERCISE 1

“How many cups of coffee or tea do you drink in a day?”

**Principle:** Ask for an answer in only one dimension.

**Fix:** Separate the question into two.

“How many cups of coffee do you drink during a typical day?”

“How many cups of tea do you drink during a typical day?”
EXERCISE 2

“What brand of computer do you own?”
[ ] IBM PC  [ ] Apple

**Principles:** Avoid hidden assumptions. Make sure to accommodate all possible answers.

**Fix 1:** Make each response a separate dichotomous item

“Do you own an IBM PC?”  [ ] Yes  [ ] No
“Do you own an Apple computer?”  [ ] Yes  [ ] No

**Fix 2:** Add necessary response categories and allow for multiple responses.

“What brand of computer do you own?” (Circle all that apply)
[ ] Do not own computer  [ ] IBM PC  [ ] Apple  [ ] Other
EXERCISE 3

“How have you had pain in the last week?”

- Never
- Seldom
- Often
- Very often

**Principle:** Make sure question and answer options match.

**Fix:** Reword either question or answer to match.

“How often have you had pain in the last week?”

- Never
- Seldom
- Often
- Very Often
EXERCISE 4

“Where did you grow up?”

- Country
- Farm
- City

**Principle:** Avoid questions having non-mutually exclusive answers.

**Fix:** Design the question with mutually exclusive options.

Where did you grow up?
- House in the country
- Farm in the country
- City
EXERCISE 5

“Are you against drug abuse?”
☐ Yes  ☐ No

Principle: Write questions that will produce variability in the responses.

Fix: Eliminate the question.
EXERCISE 6

“Which one of the following do you think increases a person’s chance of having a heart attack the most?”

- Smoking
- Being overweight
- Stress

**Principle:** Encourage the respondent to consider each possible response to avoid the uncertainty of whether a missing item may represent either an answer that does not apply or an overlooked item.

**Fix:** Which of the following increases the chance of having a heart attack?

Smoking:  
- Yes
- No
- Don’t know

Being overweight:  
- Yes
- No
- Don’t know

Stress:  
- Yes
- No
- Don’t know
EXERCISE 7

“Do you currently have a life insurance policy?”

☐ Yes    ☐ No — If no, skip the next question.

“How much is your annual life insurance premium?”

**Principle:** Avoid branching as much as possible to avoid confusing respondents.

**Fix:** If possible, write as one question.

“How much did you spend last year for life insurance?”
(Write 0 if none)
TYPES OF ANSWERS
OPEN QUESTIONS

Advantages

More free response — less bias due to assumptions of the investigator
Can lead to entirely new interpretations/unanticipated answers
Allows to study thought process of the respondent

Disadvantages

Often requires coding afterward (e.g., content analysis) — expensive
CLOSED QUESTIONS

Advantages

Easy and quick to fill in
Easy to code for quantitative analysis
More consistent across respondents
More precisely related to concept of interest

Disadvantages

Respondent might feel constrained or frustrated
Respondents with no answer tend to still give one
ORDERING OF QUESTIONS

From general to particular
From easy to difficult
From factual to abstract
Start with closed format
Start with questions relevant to the main subject
Do not start with demographic and personal questions
SENSITIVE QUESTIONS

Affects nonresponse

National Survey of Families (females)

- Total household income* 8.15%
- Lifetime male sexual partners† 3.05%
- Ever received public assistance* 2.22%
- Highest grade completed† 0.04%

Affects answers

10th graders’ smoking past month:

- 2 surveys at school = 29.8, 35.3%
- 1 survey in home = 23.4%

Marijuana past month:

- 2 surveys at school = 20.5%, 25%
- 1 survey in home = 12.5%

* ACASI = Audio Computer-Assisted Self Interview
† CAPI = Computer-Assisted Personal Interview
STANDARDIZING RESPONSES

Rating scales

Allows respondents to specify their level of agreement to a statement

Mostly used as 5-, 7-, and 9-point scales

“How well do you think Bucky Badger represents Wisconsin?”

Very poorly 1 2 3 4 5 Very well

Can also be used in forced choice tests

2- or 4-point scales

“Neither agree nor disagree” not provided
<table>
<thead>
<tr>
<th></th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Neither agree nor disagree</th>
<th>Agree</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel good about my work on the job.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>On the whole, I get along well with others at work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am proud of my ability to cope with difficulties at work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I feel uncomfortable at work, I know how to handle it.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can tell that other people at work are glad to have me there.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I know I'll be able to cope with work for as long as I want.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am proud of my relationship with my supervisor at work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am confident that I can handle my job without constant assistance.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel like I make a useful contribution at work.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can tell that my coworkers respect me.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**RATING SCALES**

“This part consists of a number of words that describe different characteristics of your partner. Read each item and mark the appropriate rating for it. Indicate to what extent you feel this way right now, that is, at the present moment.”

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>Unsociable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sociable</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kind</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Cruel</td>
</tr>
<tr>
<td>* Hard</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Soft</td>
</tr>
<tr>
<td>Successful</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Unsuccessful</td>
</tr>
<tr>
<td>* Foolish</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Wise</td>
</tr>
<tr>
<td>Strong</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>Weak</td>
</tr>
</tbody>
</table>
RATING SCALES

This scale consists of a number of words that describe different feelings and emotions. Read each item and then mark the appropriate answer in the space next to that word. Indicate to what extent you feel this way right now, that is, at the present moment. Use the following scale to record your answers:

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>very slightly or not at all</td>
<td>a little</td>
<td>moderately</td>
<td>quite a bit</td>
<td>extremely</td>
</tr>
</tbody>
</table>

interested  
distressed  
excited  
upset  
strong  
guilty  
scared  
hostile  
enthusiastic  
proud  
afraid  
irritable  
alert  
ashamed  
inspired  
nervous  
determined  
attentive  
jittery  
active
RATING SCALES

Please circle the picture below which best describes your relationship with your partner.
Below are two 2-dimensional grids representing a scale of how you perceived your partner’s attitude. Please mark where on the grid you think closest represents your partner’s attitude.

- **Low Pleasure** vs. **High Pleasure**
- **Low Arousal** vs. **High Arousal**
- **Stress** vs. **Excitement**
- **Hostile-Dominant** vs. **Friendly-Dominant**
- **Hostile** vs. **Friendly**
- **Hostile-Submissive** vs. **Friendly-Submissive**

RATING SCALES
ONLINE SURVEYS

Questions to ask yourself

Do all people in your population have access to the online survey?

Can all people in your population use an online survey with similar proficiency?

What are the risks if information is intercepted?

Are you providing an incentive?
COMMON ISSUES

Question complexity increases errors and nonresponse
Navigation complexity increases errors and nonresponse
Small field for open-ended question reduces words
No follow up for nonresponse increases nonresponse
Payment amount has little impact but equalizes groups
PHYSIOLOGICAL MEASURES
PHYSIOLOGICAL MEASURES

Eye tracking
Pupil dilation
Galvanic skin response (GSR)
Muscle movements — Galvanic Skin Response (GSR)
Muscle activity — Electromyogram (EMG)
Brain activity
  Electroencephalogram (EEG)
  Functional Magnetic Resonance Imaging (fMRI)
  Functional Near-infrared Imaging (fNIR)
EYE TRACKING

Head or desk mounted equipment tracks the position of the eye

Eye movement might indicate the amount of cognitive processing a display requires

Measurements:

**Fixations**: Eye maintains stable position. Number and duration indicate level of difficulty with display

**Saccades**: Rapid eye movement from one point of interest to another

**Scan paths**: Moving straight to a target with a short fixation at the target is optimal
Heat map showing male respondents’ view of Plan G incl. Ginger Draft Green.

Heat map showing male respondents’ view of Plan Z incl. Kirin Zero.

Heat map showing male respondents’ view of Plan G.

Heat map showing female respondents’ view of Plan G.

Kin-Mugi
Gaze concentration in a straight line in the middle; brand logo and message “Rich Malt” viewed.

Clear Asahi
Gaze concentration from mid-to-upper; indicates viewing of the brand logo and beer bubbles motif in the center, which includes a brand message.

Mugi-to-Hop
Gaze concentration at center; brand logo and the entirety of the message text seen.
Exceeding for the most sensitive skin.

Add the chemicals and moisture to your baby's sensitive skin. The chlorine-free materials and absorbent polymers are non-toxic and non-irritating. Clinically tested and pediatrician recommended for babies with allergies and sensitive skin.

If you are not satisfied with the baby leakage protection, you will get your money back. Read more about our leak-free guarantee at www.baby.com

BEFORE

AFTER
PUPIL DILATION

Pupils dilate in response to:

- Extreme emotional situations (fear, pain, contact with nerves)
- Loads on working memory, increased attention, sensory discriminations, cognitive load
- Intensifies perceptions of emotional states (sadness)
- Mirror responses to other people’s pupil sizes
- Might indicate empathy
Mean Pupillary Dilation

Group 1

p<0.01*

0

1.0

0.5

0.0

Mean Pupillary Dilation

Group 1

Group 2

Group 3

Pupillary dilation in response to eye-shaped patterns

(Coss, 1972)
Emotional response linked to physical changes
These may help determine a user’s reaction to an interface

Measurements:

- Heart activity, including blood pressure, volume and pulse
- Activity of sweat glands: Galvanic Skin Response (GSR)
- Electrical activity in muscle: electromyogram (EMG)
- Electrical activity in brain: electroencephalogram (EEG)

Difficult to interpret some physiological responses
Galvanic skin response (GSR)

Also called Electodermal response (EDR)

A.k.a. “sweat”

Sensitive to affective stimulation and attention

May measure long-term recall and effectiveness

Mandryk, et al. 2006
Climbing up / down the stairs

Medium pace walking

Running

Bending down

Loosing orientation (oscillatory movement)

Hitting an obstacle (wall, door handle, etc.)

Falling sideways (to the left - armband side)

Falling forward

Falling backward

Falling sideways (to the right)

Longitudinal acceleration (average)
Transverse acceleration (average)
GSR (average)

~ 30 seconds

A study of falls in the elderly
Jodi Forlizzi, Scott Hudson, Francine Gemperle, Bilge Mutlu

Bodymedia Armband Data Analysis
The electroencephalograph (EEG)

Alpha activity

Degree of activation

Associated with *inactivity, resting, & sleeping*

Hemispheric lateralization

Distinguishes between activity in the:

- Left hemisphere of the brain — processes visual stimuli
- Right hemisphere of the brain — processes verbal stimuli
Electromyography (EMG)

Measures electrical activity in the muscles

Facial expressions
Enabling Always-Available Input With Muscle-Computer Interfaces

T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, James A. Landay

1Microsoft Research
2University of Washington
3University of Toronto
4Microsoft Corporation

ACM UIST 2009
Functional Magnetic Resonance Imaging (fMRI)

- Measures in changes in blood flow related to neural activity
- Blood-oxygen-level dependence (BOLD)
- Measures activation in all regions with high spatial resolution
- Particularly useful in emotion research
Mutual and averted gaze stimuli & response in the brain (Pelphrey et al., 2004)
Functional near-infrared imaging (fNIR)

- Measures oxygen levels
- Indicates cognitive activity

Figure 4.

Example trials: 1st and 2nd row contains successful J and C construction respectively; the 3rd and 4th rows contain unsuccessful J and C construction respectively.

Future Work

The results presented here represent an initial study of the efficacy of letter construction using a fNIRS-based BMI. We expect a complete study that tests all letter construction efficacy and accuracy to be successful.

Several enhancements were noted during data gathering that will inform more sophisticated inference mechanisms for the interface. Currently, we are working with a more general interface that removes a priori biases in the drawing start state.

Finally, combining the continuous control shown here with discrete choice control will allow for a more robust drawing experience. Using additional sensors and mental tasks to represent discrete choices addresses the additional representational complexity. Target tasks include color palette selection and brush selection as well as other functionality common in conventional drawing interfaces.

References


PHYSIOLOGICAL MEASURES

Advantages

More direct measures of emotion, cognitive activity, etc.

Enables new research

Disadvantages

Expensive, needs mastery, technicians

High signal-to-noise ratio

Best to be used along with other measures