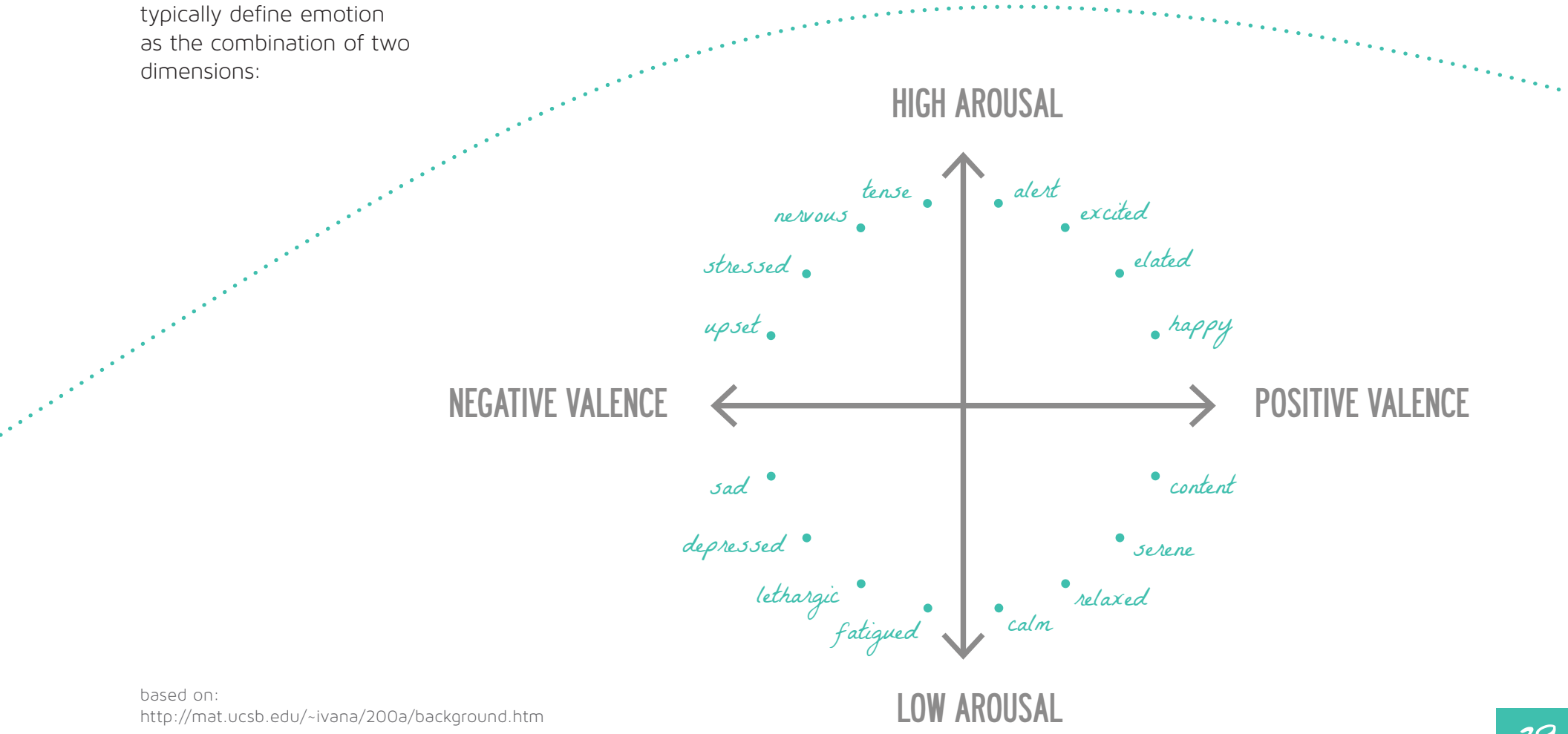
PHYSIOLOGICAL RESEARCH

In order to understand the current landscape of psychophysiological evaluation methods, we conducted a survey of academic literature. We explored several different experimental paradigms, including electromyography, electroencephalography, functional magnetic resonance imaging, galvanic skin response, computer vision, and electrocardiography. These methods were compared based on quality of results, feasibility of use for evaluating interfaces, and ease of experimental setup.

When trying to quantify emotion, the main challenge is measuring and comparing emotional states. Researchers typically define emotion as the combination of two dimensions:

- **valence**, whether the emotion is positive or negative
- **arousal**, how strong the emotion is



based on:
http://mat.ucsb.edu/~ivana/200a/background.htm

FACIAL EMG

Description

Facial electromyography (EMG) measures the electrical activity of small micro-expressions made by facial muscles in response to an emotion.

Facial EMG records two main valences, which are commonly associated with positive and negative responses. However, looking more closely at what these actually indicate, it might be more accurate to refer to the positive response as the release response and the negative response as a tension response. A small amount of tension shows that people are engaged with the product, as opposed to being relaxed and bored. Examining when the release and tension responses occur can give a good indication of how someone is feeling, and probing further with qualitative questions can help give the complete picture.

Sample Protocol

Facial EMG data records both arousal and valence of emotion.

Have the participant sit in front of the interface and connect electrodes to the corrugator cupercilii and zygomaticus major muscles. Begin data collection, but wait a few minutes to give the participant the task in order to establish a baseline. Once the baseline calibration is complete, give the participant a task to complete using the interface. The researcher should keep an eye on the EMG readings as they occur and use any release or tension spikes to probe further into what the participant is thinking and feeling by asking questions about what has just happened. If the total experiment time will exceed 20 minutes, it is a good idea to take baseline readings every several minutes, to correct for any baseline drift.

Analysis of Feasibility

Facial EMG is easy to set up, unobtrusive, and able to be used in real time.

This method has several advantages, including being fast and easy to set up. Additionally, the experimental setup is less obtrusive than other methods, making it easier to collect data on realistic usage of an interface. The electrodes used in the experiment are relatively inexpensive, between 30–40 cents each. The main initial cost would be the EMG amplifier. Biopac, one of the more well-known companies for biological signal equipment, produces an EMG amplifier that currently sells for $595.

This method provides enough information about a participant's feelings in order to probe further and possibly predict preferences. However, it can be hard to compare this type of data across subjects; it is easier to identify a change within a subject.

Another significant advantage of this method is that it is possible to view the raw results in real time and use them to guide probes into the participant's emotions.

EEG

Description

Electroencephalography (EEG) records the electrical activity of the scalp in order to examine brain activity.

Looking at Event Related Potentials (ERP) can help identify brain responses that result from thought or perception. These ERPs are EEG data which are linked to specific events, such as the participant clicking a button or navigating to a certain section.

Sample Protocol

EEG data can reveal cognitive activity at a certain point in time.

Set up the EEG cap on the participant, using a blunt needle to insert the electrogel into the electrodes. Take care that channels F3 and F4, which are the emotional control area, and Fz and Pz, which are the decision-making areas, have strong connections. Begin data collection and establish a baseline reading. Assign the participant a task using the interface, with event codes sent to the data collection based on where the participant is within the interface. After the conclusion of the experiment, the data can be filtered and ERPs can be extracted from the relevant channels in order to analyze the participant's cognitive response at different points during the task.

Analysis of Feasibility

EEG takes some time to set up and feels a bit unnatural for the participant, but it can be used to look at many different factors in thought or perception.

One drawback to this method is that it can be quite time-consuming to filter and analyze the data and it is not typically done in real time, ruling out the possibility of qualitative probes based on the EEG results.

Additionally, a complete system can be rather expensive, costing thousands of dollars and varying quite a bit depending on the quality of the company and the number of electrodes available. The setup and materials required for each individual participant is in the range of a few dollars. There are professors at Carnegie Mellon with these systems already in their labs and it may be possible to collaborate and receive use of these systems.

However, an advantage of this method, especially if recording with a full 64 electrode EEG cap, is that it may be possible to go back and examine other aspects of the brain response by analyzing different electrodes, which can be useful if the focus of the study changes. EEG data can be quite complex, which can allow for more nuanced results but can also make it harder to extract the relevant information.

FMRI

Description

Functional Magnetic Resonance Imaging (fMRI) looks at changes in blood flow in order to measure neural activity.

fMRI has been used in research to identify people's preferences of soda brands and sports cars. It has been found that increased blood flow can be correlated to desirability of a product. It is less common in the field of HCI, possibly because of experimental limitations or accessibility of other techniques.

Sample Protocol

Changes in blood flow recorded by fMRI can be mapped to certain emotions and cognitive activities.

Since the participant must remain motionless during an fMRI scan, the interaction possibilities are likely limited to pushing a button or simply passively viewing an interface. This makes it harder to study a person actually completing a task using the interface. A possible protocol could involve showing a subject two interfaces and having him push a button to indicate his preferred interface, while examining the decision-making areas of the brain.

Analysis of Feasibility

fMRI systems are expensive and require professional expertise, but can reveal changes in neural activity.

Purchasing an fMRI system can cost $1–3 million. It may be possible to find an institution that rents out time on a machine, but this could still cost thousands of dollars per hour. The machine will require a trained technician to operate. Additionally, it is probably necessary to hire a radiologist to interpret the results, which adds to the cost. It will also be difficult and expensive to acquire subjects for this procedure.

It is not clear that the added expense and difficulty of fMRI will result in more useful results. In fact, it significantly limits the experimental protocol and is the most obtrusive of the methods presented here.

GALVANIC SKIN RESPONSE

Description

Galvanic Skin Response (GSR) measures the electrical activity of the skin in order to identify an increased response in the sympathetic nervous system.

GSR measures the electrical activity of the skin in order to identify an increased response in the sympathetic nervous system. It gives general levels of arousal or tension but must be used in conjunction with qualitative methods in order to determine which of these emotions the user is experiencing at a particular point in time.

Sample Protocol

GSR measures changes in emotion strength, but qualitative probes are necessary to determine valence.

Similar to the facial EMG protocol, the GSR sensor would be hooked up and attached to the participant. Data collection would begin with a baseline reading to establish normal skin conductance levels. The participant would be asked to interact with the interface, probably through a specified task. The researcher would monitor the data output and watch for any deviation from the baseline. The researcher could then probe further in real time to understand the emotions underlying this change in electrodermal activity.

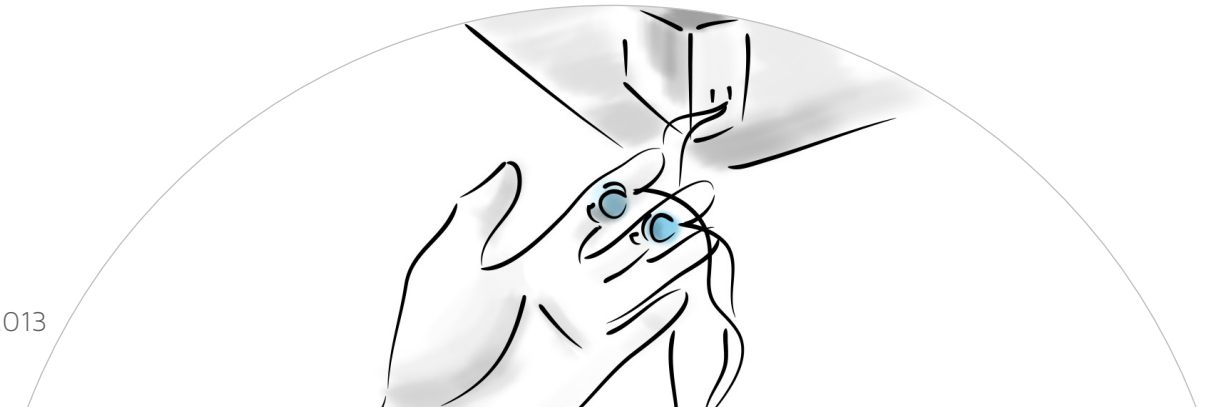
Analysis of Feasibility

GSR systems are inexpensive, easy to set up and can be analyzed in real time. However, the data interpretation is very dependent on qualitative probes.

A GSR sensor can cost around $300, but could be reused for all participants. An advantage of this method is that the sensor is generally unobtrusive. However, it is worth noting that data from this method is somewhat limited and requires more probing than other methods to find out a participant's actual emotions.

GSR has been used previously in HCI research and has particular application in evaluating games, due to the fact that it is unobtrusive and easy to ignore while using an interface.

Although GSR is similar in feasibility to facial EMG, the depth of the data provided is not as great.



COMPUTER VISION

Description

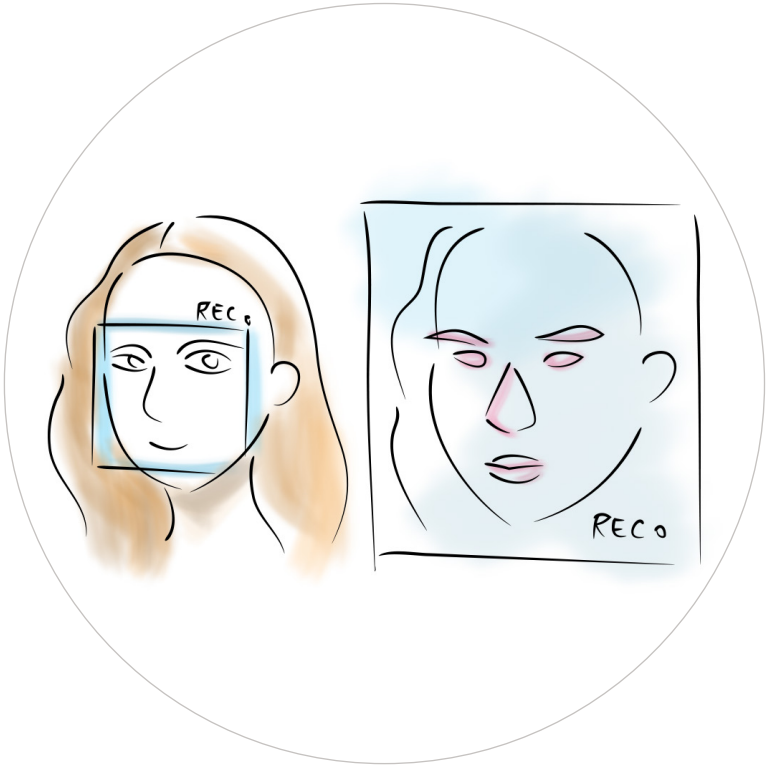
Computer vision uses a webcam and face reading software to determine emotional responses based on facial expressions.

There are several companies making software for emotion recognition using computer vision. One possible company is Affectiva, which is well regarded within the scientific community. It is an MIT Media Lab spinoff and is backed by the National Science Foundation. Other producers of computer vision software for these purposes include RealEyes and Noldus.

Sample Protocol

These systems usually classify the user's emotion into predetermined categories (happy, sad, surprised, frustrated, etc.) and also record valence and strength of emotion.

Calibrate the system and adjust the lighting as necessary. Ask the participant to complete a task using the interface and record video along with the data collection. The exact form of data collection and analysis will depend on the specific software package used.



Analysis of feasibility

Computer vision systems can be expensive, but have the advantage of being able to conduct experiments remotely.

One significant advantage of this method is that it is possible to complete studies from a distance and to test people outside a lab setting simply using a video feed from their webcams. It should be noted that lighting can influence the results and therefore a lab setting will provide the most reliable results, but nonetheless it is possible to complete these studies in other locations. The Noldus system is currently priced at $7590.

METHOD COMPARISON

	Facial EMG	EEG	fMRI	GSR	Computer Vision
affordability	4	2	1	5	3
ease of setup	4	2	1	4	5
quality of data	3	4	5	1	2
total	**11**	**8**	**7**	**10**	**10**

Rated on a scale of 1–5, with 5 being the most desirable.

These five methods were compared and evaluated based on three main factors:

- Affordability, which ranged from millions of dollars for fMRI to several hundred dollars for the more affordable methods like GSR and facial EMG.
- Difficulty of setting up each method, which included participant preparation and equipment calibration.
- Quality of data, which accounts for how much neural activity is recorded by a particular method and whether additional probes are necessary.

This evaluation showed three methods to have the biggest potential: facial EMG, GSR, and computer vision. Although any of these three methods would be appropriate for our purposes, **our recommendation is facial EMG**, as it strikes a good balance of being relatively affordable, easy to setup, and provides information about both dimensions of emotional state.