1 Definition

Psychophysics comes from the words *psycho*, meaning some sort of mental process or subjective experience, and *physics*, the measurement of things. Together it means the science of making objective measurements of subjective experiences. Typical objective physical observations of anything would be weight, length, intensity. The subjective observations of these three would be heaviness, size, and brightness respectively. Because of the subjective nature of these observations, the measurements must be inferred by a personal response, rather than an actual physical measurement scale.

2 Methods

Methods are tools for measuring these subjective observations, intersecting with technology. By standardizing the methodology used to measure we can come up with an universal understanding of these complex notions. For example, in a virtual reality vs a real world, how can we show that the two are equivalent? When asked various questions such as "How big is the building?" and "How far away is an item to pick up?" we need ways to analyze the subjective answers to show the two worlds are or are not equivalent. Two psychophysical quantities are proposed.

2.1 Thresholds

*Thresholds* come in two kinds. The first is the *absolute threshold*, defined as the intensity at which a physical thing becomes detectable to an observer. A simple example for when determining the absolute threshold for light, a black slide with a dark square in the center would be shown, and the square would be brightened until detection. The absolute threshold is measured with the *Detection Task*.

The *Difference Threshold*, also known as the *Just Noticeable Distance* (JND), is the difference in intensities at which an observer can notice the differences. Back to the simple light example, two squares would be shown to the observer with the differences modified until the observer notes that the intensities are different. The difference threshold is measured with the *Discrimination Task*.

2.2 Scales

*Scales* are the application of a comparison between two different intensities and attempting to apply a scale to it. The question "Is light A twice as bright as light B" would be an example of a scale in use. This is also called the *Suprathreshold Appearance*. A question in class was raised to the point of: how can we actually and accurately say something is "twice" as much for subjective measurements? Two researchers in the 1800’s developed a series of methods to do just this.

3 Weber’s Law

Ernst Heinrich Weber, a researcher in the 1800’s, asked about a person’s perception of heaviness. Moreover, he asked if people are able to tell the difference between two weights. The task to determine this required a person to lift one weight (\(W_1\)), then lift another weight (\(W_2\)), and then report if the weights are different. This continued for various different weight distinctions. Weber then proposed that, if the weight difference was noticeable, then the just noticeable difference is the smallest difference between the two weights when the person could tell the difference. Further, Weber discovered that the JND was proportional to the weight observed. The larger the weight being compared, the larger the JND must be to be detected, and the smaller the weight, the smaller JND that can be detected. Another example would be that a very soft sound does
not need much more volume for the difference to be noticed, but a very loud initial sound must change volume by a proportionally larger amount for a difference to be heard. Both of these equations are written as follows:

\[ JND = |W_1 - W_2| \]
\[ JND \propto W \]

*Weber’s Law* is then stated as for the JND (\(\Delta I\)) is equivalent to some multiplier constant \(k\) times the intensity \(I\):

\[ \Delta I = kI \]

The question then is raised, is \(k\) universal over all things? Is it universal for all mediums but different per individual (A weight lifter feels weights differently than others)? Is it universal for all people but different for medium (Lights and weights have different \(k\))? The consensus is the latter; that \(k\) is different for different input mediums but the same across individuals.

4 Fechner’s Scale

Gustav Theodor Fechner, also a researcher in the 1800’s, developed a scale for measuring sensations. In order to measure these subjective aspects, we need two physical aspects: a zero and an incremental unit. For objective standards, temperature (in degrees Celsius) has a zero at the point which water freezes, and the incremental unit is the temperatures between zero and the temperature at water boiling, divided by 100, a degree.

For subjective scales, the zero is the detection threshold and the increment is the JND. However, since the JND is proportional to the weight, the increment size must change along the scale.

4.1 Method of Adjustment

Fechner also introduced the *Method of Adjustment* to measure the zero and incremental unit. First, a standard is selected. Then a value above the standard is presented side by side to the standard to a subject. The subject is then directed to adjust the second value until it seemingly is the same as the standard. This is called a *Descending Trial* because the second value is brought down. Then a value below the standard is presented and the subject is directed to adjust the value until it matches the standard. This is called an *Ascending Trial* because the second value is brought up. An example with color is shown below where the standard is an intensity of 128.
For all the trials the starting value and the stopping value is noted down. The ascending and descending trials are repeated in alternating order for multiple trials. An example of the data collected for the color value above could look like the following:

<table>
<thead>
<tr>
<th>Trial Number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>...</td>
</tr>
<tr>
<td>Start Intensity</td>
<td>108</td>
<td>135</td>
<td>102</td>
<td>140</td>
<td>...</td>
</tr>
<tr>
<td>End Intensity (A)</td>
<td>126</td>
<td>---</td>
<td>129</td>
<td>---</td>
<td>...</td>
</tr>
<tr>
<td>End Intensity (D)</td>
<td>---</td>
<td>125</td>
<td>---</td>
<td>130</td>
<td>...</td>
</tr>
</tbody>
</table>

The next step is to compute the mean of the ascending ($\mu_A$) and the descending ($\mu_D$) trials’ ending values. These two means should be the *same value*, otherwise there is some bias occurring between the two trials and further inspection is required. If the two means are the same, then compute the ground mean, or the mean of all ending values ($\mu$), also known as the *Point of Subjective Equality* (PSE). The PSE is the subjective equal value to the standard, meaning that users cannot tell the PSE and the standard values apart.

A question is raised in class asked the difference between the PSE and the standard. The difference is that the standard is an arbitrary value you assign for the test, and the PSE is a calculated value from the test subject results. The JND is found by the following formula, where $\sigma$ is the standard deviation of the end intensities. Also, the UL, LL, and IU are also as follows:

\[
JND = 0.67\sigma
\]

\[
UL = PSE + JND, \quad LL = PSE - JND
\]

\[
IU = UL - LL
\]

The *Upper Threshold* is also known as UL or the Upper Limit, the *Lower Threshold* is also known as LL or the Lower Limit, and the *Interval of Uncertainty* is also known as IU, and also could be represented as $2 \times JND$.

### 4.2 Additional Illustrative Example

Suppose you are a researcher whom wished to study and create a scale for electric guitar sounds. To find the Just Noticeable Difference you run ten trials in an ascending and descending pattern around a sample of a guitar playing a single, constant note of Middle C at a volume level of 100. You start the volume levels slightly above and below a volume of 100, and ask the participant to ascend or descend to the target volume (100). The hypothetical data you obtain can be seen in the following table:

<table>
<thead>
<tr>
<th>Trial</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
<td>A</td>
<td>D</td>
</tr>
<tr>
<td>Start Intensity</td>
<td>87</td>
<td>110</td>
<td>87</td>
<td>114</td>
<td>86</td>
<td>108</td>
<td>85</td>
<td>111</td>
<td>93</td>
<td>115</td>
</tr>
<tr>
<td>End Intensity (A)</td>
<td>97</td>
<td>---</td>
<td>100</td>
<td>---</td>
<td>98</td>
<td>---</td>
<td>102</td>
<td>---</td>
<td>101</td>
<td>---</td>
</tr>
<tr>
<td>End Intensity (D)</td>
<td>---</td>
<td>98</td>
<td>---</td>
<td>99</td>
<td>---</td>
<td>103</td>
<td>---</td>
<td>98</td>
<td>---</td>
<td>98</td>
</tr>
</tbody>
</table>

We can then calculate the statistics of the data collected:

\[
\mu_A = \frac{97+100+98+102+101}{10} = 99.602
\]

\[
\mu_D = \frac{98+99+103+98+98}{10} = 99.153
\]

\[
PSE = \mu = \frac{97+100+98+102+101}{10} = 99.378
\]

\[
\sigma = 1.874, \quad JND = 0.67 \times \sigma = 0.67 \times 1.874 = 1.255
\]

\[
UL = PSE + JND = 99.378 + 1.255 = 100.633
\]

\[
LL = PSE - JND = 99.378 - 1.255 = 98.122
\]

\[
IU = UL - LL = 100.633 - 98.122 = 2.511
\]

This process for developing a standard for subjective measurements is vital to the study of color science. It is used in research regarding film, paint, photos, graphics, and so on.
5 Signal Detection Theory

With signal detection, the best case scenario in theory is a simple "on" or "off" situation. However, in practice, the definition isn’t quite as clear-cut as to when a signal is on versus off.

![Signal Detection Theory](image)

Visual example of the theory vs reality in signal detection.

The role playing exercise done in class was to imagine you are a radar control operator and you must determine if a "blip" on the radar is actually an enemy aircraft or background noise. The goal is to make a detection: are the skies clear or not? Noise can come from many factors, from the reading equipment to the user’s eyes. An example of the kinds of sensations you could register is shown in the chart below.

![Signals and Noise](image)

At low intensities, the probability that it is simply noise is higher. At high intensities, the probability that the sensation is an actual signal is higher. The problem lies in the middle range: where do you make the decision to label something as simply noise to one side of the threshold, and the other is an actual signal, despite there being a non-zero chance that you are incorrectly labeling it?
The stimulus response matrix for visualizing the classifications of each correct and incorrect category.

The dream would be to maximize the number of \textit{Hits} and \textit{True Rejects}, however the most realistic solution needs some sort of compromise, and it all depends on the situation. The decision criteria can change with the job definition, such as being told that \textit{False Alarms} would be heavily punished, but \textit{Misses} would not.

A black decision line is shown. The red banded area is the \textit{Miss} category, while the blue banded area is the \textit{False Alarm} category. The area under the blue curve that is not banded is the \textit{True Reject}, and the area under the red curve that is not banded is the \textit{Hit} category.

Picking an arbitrary decision line, everything to the right of the black line in the image is reported as an object seen, and everything to the left is reported as noise. We can then develop two measures based on this decision line:

\[
d' = \frac{\mu_{SN} - \mu_N}{\sigma_N}
\]

\[
\beta = \frac{Y_s}{Y_n}
\]

The first measure is the \textit{True Sensitivity}. The larger the difference between the two means of the noise and signal, the more accurate or sensitive the responses will be. However, a larger standard deviation, $\sigma_N$, means the opposite and the responses will be less accurate. The second measure is the \textit{Response Bias}, utilizing $Y_s$ and $Y_n$ signified as green stars in the image. The response bias is how likely this particular observer’s decision line is to privilege hits over misses.

This examination of the signal detection led to a new task.
6 Two Alternative Force Choice Task

The *Two Alternative Force Choice Task* (2AFC Task), takes a slightly different approach. Instead of "did you see something", the question is to pick the closest of two options.

Which one is closest to the given color on the top?

A question was raised in class: why is this method better? In a dial ascending/descending task there is a response bias, and so the 2AFC task eliminates this bias. Another question is asked: What if you can’t tell the difference? That is the goal: to keep offering two pairs until the interval of uncertainty is found.