Fundamentals of Psychophysics

John Greenwood
Department of Experimental Psychology

NEUR3045

Contact: john.greenwood@ucl.ac.uk
Visual neuroscience

• How do we see the world?
• The brain is a complex system with many different levels
• So we need approaches with different scales of analysis
  • Physiology, e.g. recording from individual neurons
  • Neuroimaging, e.g. fMRI scans to see the brain regions activated
  • Psychophysics: the system as a whole, e.g. reading, seeing colours
Today

- An overview of psychophysics and its methodological approaches

**Stimulus**
- Luminance patch
- Gabor etc.

**Task**
- detection
  - yes/no
  - forced choice
- discrimination
  - yes/no
  - forced choice

**Method**
- Limits
- Adjustment
- Constant stimuli
- SDT

**Outcome**
- performance
- appearance
  - Reaction times
  - Percent correct
  - Thresholds
  - d' and c
- Matching
- Scaling
Psychophysics

• Originates from Fechner (1860)

• Investigates the relationship between physical stimuli and psychological quantities (‘psyche’)
  • We can’t measure the mind directly, so we measure behaviour
  • Requires linking hypotheses between subjective and objective phenomena
  • Requires precise control over physical stimuli and testing procedures
  • If you know the properties of a stimulus, and how a person responds to that stimulus, you can infer the underlying perceptual operations of the brain
From behaviour to function

• How can we infer neural processes from behaviour?
The stimulus

- **Luminance patch**
  For brightness or contrast perception

- **Oriented Gabor**
  For orientation perception/spatial vision

- **Moving Gabor**
  For motion perception

- **Stereoscopic stimuli**
  For depth perception

- **Letters**
  To study reading and/or acuity

- **Faces**
  For face vs. object recognition
From behaviour to function

- How can we infer neural processes from behaviour?

**Stimulus**
- Luminance patch
- Gabor etc.

**Task**
- detection
- discrimination

**Method**

**Outcome**
- performance
- appearance
- Reaction times
- Percent correct
- Thresholds
Thresholds

• A major concern of psychophysics: thresholds
  • The lowest stimulus quantity that can be reliably seen
  • e.g. for size, brightness/luminance, motion, etc.

• Thresholds measure sensitivity, which can be related to the tuning of a neural detector (our linking hypothesis)

• Two types of thresholds:
  • Absolute / detection thresholds
  • Difference / discrimination thresholds
Detection thresholds

- Definition: The minimum intensity at which a stimulus is “just detectable”

- e.g. brightness
  - The lowest brightness value you can see

- e.g. motion
  - The slowest speed that you can see
Discrimination thresholds

• The smallest *difference* in intensity that is just detectable
• Requires comparison between two or more stimuli, or between one stimulus quantity and a standard/reference

  e.g. brightness
  The lowest difference in brightness that you can see

  e.g. orientation
  The smallest orientation offset from vertical that you can see
From behaviour to function

- How can we infer neural processes from behaviour?

Stimulus:
- Luminance patch
- Gabor etc.

Method:

Task:
- Detection
- Discrimination

Outcome:
- Performance
- Appearance
- Reaction times
- Percent correct
- Thresholds
Task

• Let’s select detection thresholds for a luminance patch
  • Why do this? e.g. Hecht, Haig & Chase (1937)
  • Measured detection thresholds after different durations in the dark
• How do you ask the question?
  • Yes/no methods
  • e.g for detection: “Can you see it?”
    Yes / no
Method / sampling procedure

• On the range of intensity values, where do you select the ones to show the observer?

• Method of Limits
  • Intensity gradually increased/decreased until response changes
Method of limits

• Errors of habituation
  • Giving the same response continually and don’t change

• Errors of anticipation
  • Know the threshold is coming and change response too soon
Method of limits

• Minimise these errors by approaching in both directions ($\infty$ to 0, and 0 to $\infty$)

• Threshold is then the average of these measurements
  • performance measure: average setting $=$ brightness threshold
Method of adjustment

• As with the method of limits, but the observer adjusts the stimulus levels themselves until their report changes from visible to invisible or vice versa
Issues with adjustment/limits

• Advantage of method:
  • Rapid estimation of threshold

• Disadvantage:
  • Errors of habituation and anticipation
  • Although these errors can be partly overcome with different directions of measurement, there is an alternative
Method of Constant Stimuli

- Intensities presented in a random order, with repeats
  - Removes issues of anticipation/habituation
Method of Constant Stimuli

• But where is the threshold?
  • Performance varies from 0 to 100%
• Can fit a ‘psychometric curve’
  • Cumulative form of a Gaussian function

![Graph showing the relationship between brightness and percentage "yes" responses.](image)
Method of Constant Stimuli

• At what point do we call the threshold?
  • It should be above 0% (never seen) and below 100% (always seen)
  • The midpoint (50%) represents the ‘tipping point’ between predominantly “yes” and predominantly “no” (and the point with the fastest rate of change)
Issues with MCS

• Advantages:
  • Avoids issues of habituation/anticipation

• Disadvantages:
  • Slower estimation of threshold
  • Need to test a predetermined range of intensity values
How can we infer neural processes from behaviour?

- **Stimulus**
  - Luminance patch
  - Gabor etc.

- **Task**
  - detection
    - yes/no
    - e.g. “can you see it?”
  - discrimination
    - yes/no
    - e.g. “are they different?”

- **Method**
  - Limits
  - Adjustment
  - Constant stimuli

- **Outcome**
  - performance
  - appearance
  - Reaction times
  - Percent correct
  - Thresholds
Threshold vs. criterion

- Yes/no procedures confound the threshold with the observer’s subjective criterion
- Consider the effect of increased sensitivity vs. decreased sensitivity
Threshold vs. criterion

- Now consider the effect of criterion differences
  - Someone eager to indicate “yes” (a liberal criterion) vs. someone reluctant to do so (conservative criterion)
  - Impossible to distinguish from changes in sensitivity
Forced-choice measures

- Yes/no measures rely on a subjective criterion
- Forced-choice measures can minimise this influence
  - Force the observer to choose between 2 or more responses on each trial
  - Compare these judgements against an objective standard
  - e.g. two-alternative forced choice (2AFC)

Was the patch to the left or right?

Was the motion to the left or right?
Forced-choice MCS

- What happens with the psychometric function?
  - e.g. the detection task for luminance/brightness
  - With a 2AFC design the guess rate is 50%
  - The midpoint (threshold) is now taken as 75% correct
Forced-choice issues

• Advantages
  • Avoids issues of subjective criterion
  • Can use to test perception in animals / pre-verbal children

• Disadvantages
  • Not always possible to create ‘objective’ scoring

from Carandini & Churchland (2013)
From behaviour to function

Stimulus
- Luminance patch
- Gabor etc.

Method
- Limits
- Adjustment
- Constant stimuli
- SDT

Task
- detection
- discrimination

Outcome
- performance
- appearance

SDT

- yes/no
- forced choice
- yes/no
- forced choice

Thresholds
- Reaction times
- Percent correct
- d’ and c
- Percent correct
- d’ and c
- Thresholds
- Reaction times
Signal Detection Theory

• Derives from radar operators during World War II
  • Radar antenna direction given by line
  • Dots trailing this visible only briefly and could arise from objects in environment, weather patterns, noise, or enemy aircraft
  • Upon seeing a dot: should you raise the alarm or not?
**Signal Detection Theory**

**Consequences:**

- **Hit:** Enemy are engaged and turned away
- **Miss:** Enemy attack their target unscathed
- **False alarm:** Aircraft take off for nothing, fuel wasted, pilots fatigued
- **Correct rejection:** Crew able to rest and fuel is not wasted

<table>
<thead>
<tr>
<th>Signal: Is it actually an enemy plane?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision: Is there an enemy plane?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>Hit</td>
<td>False Alarm</td>
</tr>
<tr>
<td>No</td>
<td>Miss</td>
<td>Correct Reject</td>
</tr>
</tbody>
</table>
SDT for brightness

- Formalised for psychophysics by Green & Swets (1966)
- Easy to transpose this situation into a yes/no decision task, e.g. with our luminance patch
- Here we need two types of trials: signal present or absent
  - Decisions in each case: yes/no for each type of trial

<table>
<thead>
<tr>
<th>Signal: Is there a luminance patch?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Hit</td>
<td>False Alarm</td>
</tr>
<tr>
<td>No</td>
<td>Miss</td>
<td>Correct Reject</td>
</tr>
</tbody>
</table>
SDT and X-ray diagnosis

Radiologists examine chest X-rays and asked “is a tumour present or absent?” (Kundel & Nodine, 1975)

- What limits performance and how can we characterise this?
Noise

• Uncertainty on these tasks arises from two types of noise

  • **External noise:** e.g. imaging errors, variation in lung tissue
  
  • **Internal noise:** radiologist uses some neural response to detect a tumour - these responses are variable
Internal distributions

- Compare internal response probability of occurrence curves for noise alone vs. signal+noise trials
- Discriminability of the two possibilities set by separation/breadth of curves
- But decision also requires that we set a criterion value
Distributions to responses

• Signal present trials:
  • Response above the criterion = hit
  • Response below the criterion = miss
Distributions to responses

• Signal absent trials:
  • Response below the criterion = correct rejection
  • Response above the criterion = false alarm
Measuring sensitivity

• Sensitivity is characterised by $d'$ (d prime)

$$d' = \frac{\mu_{S+N} - \mu_N}{\sigma}$$
Calculating d'

- Sensitivity is characterised by d' (d prime)

\[ d' = \frac{\mu_{S+N} - \mu_N}{\sigma} \]

- \( d' = z(\text{Hit}) - z(\text{FA}) \)
**d' examples**

**e.g. early stage tumour**

<table>
<thead>
<tr>
<th>Decision: Is there a tumour?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.84</td>
<td>0.50</td>
</tr>
<tr>
<td>No</td>
<td>0.16</td>
<td>0.50</td>
</tr>
</tbody>
</table>

**Signal:** Is there a tumour?

- Early stage tumour: \( d' = z(0.84) - z(0.5) = 1 \)
- Late stage tumour: \( d' = z(0.98) - z(0.33) = 2.5 \)

**e.g. late stage tumour**

<table>
<thead>
<tr>
<th>Decision: Is there a tumour?</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>0.98</td>
<td>0.33</td>
</tr>
<tr>
<td>No</td>
<td>0.02</td>
<td>0.77</td>
</tr>
</tbody>
</table>

**Signal:** Is there a tumour?
Criterion effects

- The criterion can also alter performance drastically
  - e.g. Radiologists may weigh errors differently - one considers missed diagnoses fatal, another minimises unnecessary procedures
  - Note there is no point that completely removes false alarms without missing many ‘signal present’ trials

\[ d' = 1.0 \]

- Low
  - Hits = 98%
  - False Alarms = 84%

- Med.
  - Hits = 84%
  - False Alarms = 50%

- High
  - Hits = 50%
  - False Alarms = 16%
Measuring the criterion

• Is there a way to characterise this criterion?

\[ c = \frac{-(z(\text{Hit}) + z(\text{FA}))}{2} \]

• Negative means many ‘yes’ responses; positive means ‘no’
Criterion examples

\[ c = \frac{-(z(\text{Hit}) + z(\text{FA}))}{2} \]

Low

\[ d' = 1.0 \]

Hits = 98%
False Alarms = 84%

\[ c = -1.5 \]

Med.

\[ d' = 1.0 \]

Hits = 84%
False Alarms = 50%

\[ c = -0.5 \]

High

\[ d' = 1.0 \]

Hits = 50%
False Alarms = 16%

\[ c = 0.5 \]
SDT summary

• We can characterise performance using two values
  • d’ - sensitivity
  • c - criterion
• Previously we sought to avoid the subjective criterion through the use of forced choice procedures
• SDT allows us to measure it
  • Through the separation of ‘signal present’ and ‘signal absent’ trials
From behaviour to function

- Two more approaches to measure stimulus appearance

**Stimulus**
- Luminance patch
- Gabor etc.

**Method**
- Limits
- Adjustment
- Constant stimuli
- SDT

**Task**
- Detection
- Discrimination

**Outcome**
- Performance
- Appearance

**Limits**
- Reaction times
- Percent correct
- Thresholds
- d’ and c

**Scaling**
- Matching
Appearance: matching

• A simple way to measure the perceived equivalence of two stimuli: ask observers to match their appearance

  • e.g. with two patches of colour: match the appearance of a narrowband yellow reference with a test patch made via superimposition of red & green lights

  • Allows the measurement of metamers - stimuli that are physically dissimilar but perceptually identical
Appearance: scaling

• A method to measure the perceived difference between two stimuli
  • Formalised by Stanley Smith Stevens
  • Show a reference stimulus and assign a value (e.g. “10”)
  • Then show a test stimulus at a different intensity
  • Observer must then assign a number to the test that is proportional to the reference
    • e.g. if the test appears twice as bright it is reported as 20
  • And so on…
Psychophysics provides tools to investigate the relationship between physical stimuli and psychological quantities.

**Stimulus**
- Luminance patch
- Gabor etc.

**Task**
- detection
- discrimination
- yes/no
- forced choice

**Method**
- Limits
- Adjustment
- Constant stimuli
- SDT

**Outcome**
- performance
- appearance
- Reaction times
- Percent correct
- Thresholds
- d’ and c
- Matching
- Scaling
Summary

• With linking hypotheses we then make inferences about the mechanisms underlying our visual perception

• We can divide psychophysical methods into four broad components: stimulus, method, task, and outcome

• In future lectures you’ll go into more detail about specific visual dimensions

• Reading for this lecture: Chapter 1 of Sensation & Perception by either Goldstein or Wolfe et al (or both)