

# 햅틱 시스템의 평가 방법

## Evaluating Methods for Haptic Systems

2014-02-12

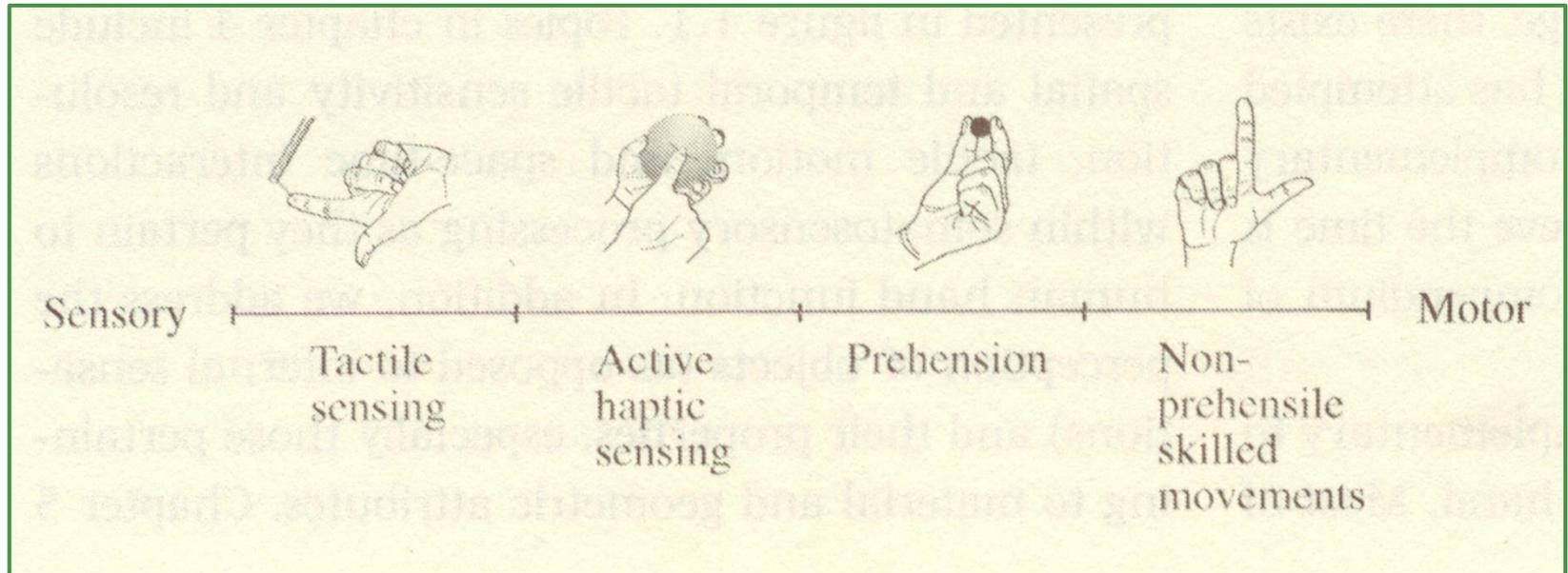
포항공과대학교 컴퓨터공학과 최승문

실제 튜토리얼에 사용될 슬라이드는 이 슬라이드에 바탕을 두고 있으나 상당한 변화가 있을 수 있습니다.

# INTRODUCTION

# Sensorimotor Continuum For Human Hand Function

---



[L. A Jones, and S. J Lederman, Human Hand Function, Oxford University Press, 2006]

# Motivations of This Talk

---

- ▶ Haptics is hard for many reasons.
- ▶ Haptic technology and science are tightly coupled.
- ▶ Knowledge on sensorimotor perception, cognition, and neurophysiology is mandatory for the appropriate designs of haptic devices and rendering algorithms.
- ▶ Engineers often have troubles in evaluating their haptic systems in proper ways.
- ▶ These slides are a VERY condensed version taken from the course materials of POSTECH EECE 617 Advanced Haptics ([http://hvr.postech.ac.kr/?page\\_id=773](http://hvr.postech.ac.kr/?page_id=773)).

# Haptic Sensations

---

- ▶ Tactile (Cutaneous) Sensation
  - ▶ Requires physical contacts between the skin and outer objects
  - ▶ Mediated by receptors under the skin
  - ▶ Primary sensations
    - ▶ Spatial Tactile Sensation: e.g., shape
    - ▶ Temporal Tactile Sensation: e.g., vibration
    - ▶ Temperature
    - ▶ Pain
- ▶ Kinesthetic (Proprioceptive) Sensation
  - ▶ No physical contact is required
  - ▶ Mediated by various reception systems inside the body
  - ▶ Primary sensations
    - ▶ Position, velocity, and acceleration of the body
    - ▶ Orientation of the body
- ▶ Haptic Information = Tactile + Kinesthetic Information

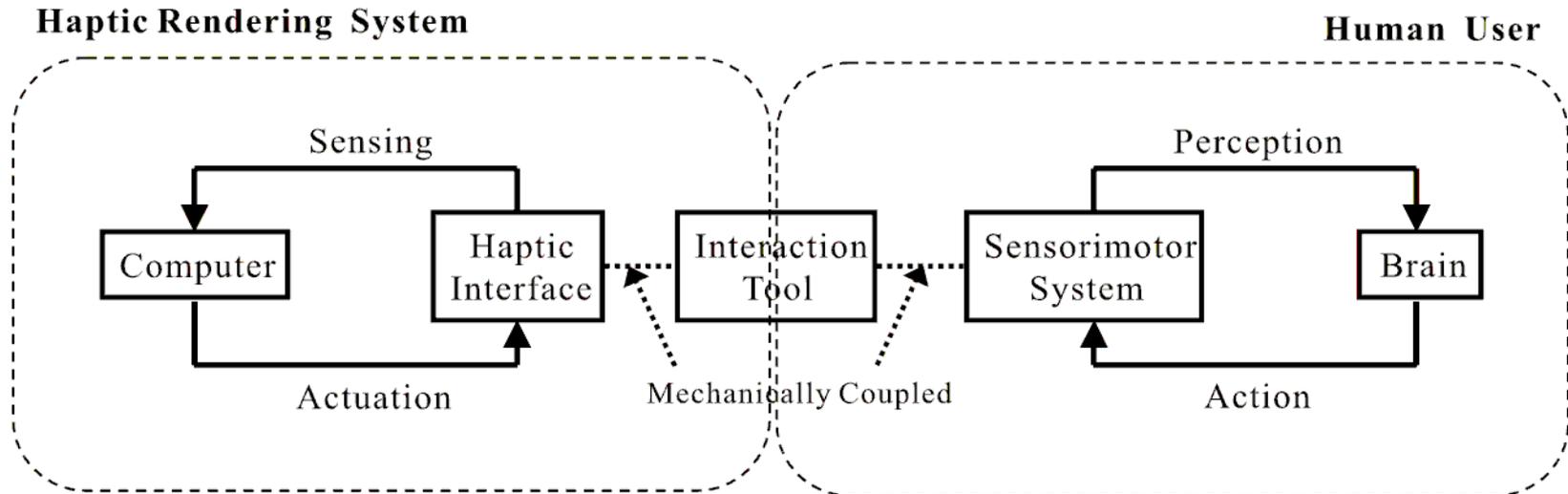
# Passive vs. Active Touch

---

- ▶ The haptic sensory channel is bidirectional.
  - ▶ Perception + Action
  - ▶ Good example: Somatosensory Cortex
- ▶ Passive Touch
  - ▶ Haptic perception occurs while motor commands to muscles are absent
- ▶ Active Touch
  - ▶ Haptic perception occurs while the user intentionally explores objects with body parts.
  - ▶ Often leads to better perceptual performance

# Haptic Interface

- ▶ Device that intervenes a user and a computer for haptic interaction
- ▶ In general, a haptic interface can sense the user's motion, create haptic stimuli, and transmit it to the user.
- ▶ Haptic interaction is symmetric and bidirectional.



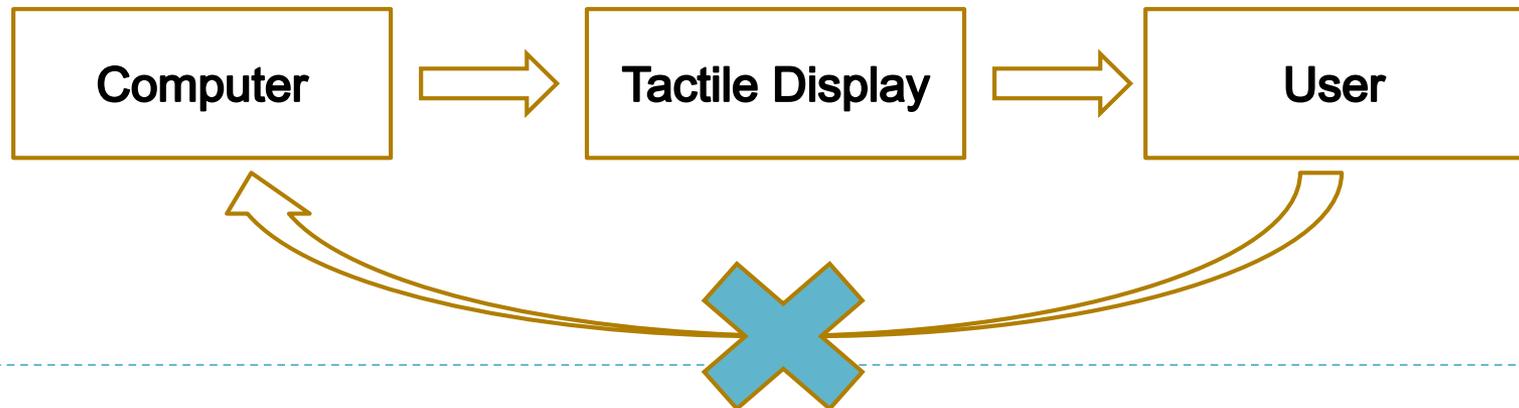
# Classification of Haptic Interfaces

---

- ▶ A variety of haptic interfaces have been developed for their own purposes, mostly due to the infeasibility of designing a “universal” haptic interface.
- ▶ Force-feedback (Kinesthetic) Interface
  - ▶ Can provide force to the user, inducing net movements of the user’s body
  - ▶ Example: A point-contact, 3 Degrees-of-Freedom (DoF), desktop force-feedback device
- ▶ Tactile Interface
  - ▶ Can provide tactile stimuli to the user, without net movements of the user’s body or negligible momentums
  - ▶ Example: Vibrotactile actuators in mobile devices.

# Why Tactile Rendering?

- ▶ Tactile rendering can deliver various haptic properties, such as
  - ▶ Pressure
  - ▶ Texture
  - ▶ Pulse
  - ▶ Vibration
  - ▶ Skin stretch
  - ▶ Temperature
- ▶ Most tactile interfaces are feed-forward, information-presenting displays.



# Tactile Displays

---

- ▶ To date, various kinds of tactile displays have been researched.
  - ▶ Depending on the purposes, tactile displays differ in skin site, contact area, types of tactile stimuli, actuation technology, and spatial and temporal resolutions.
  - ▶ The specifications of tactile displays must be based on relevant human perception.
  - ▶ The performance of tactile displays must be evaluated in terms of human perception and usability.
- ▶ Common Actuation Technology
  - ▶ DC motor (Vibration motor)
  - ▶ Voice-coil actuator (Linear resonant actuator)
  - ▶ Piezoelectric actuator
  - ▶ Electrocutaneous display
  - ▶ Tactile pin array
  - ▶ Temperature display (using Peltier elements)

# Research Steps for Tactile Rendering

Step	Task	Example
1	Determine a goal of your application	Create button click sensations for mobile devices
2	Design proximal stimuli that are supposed to be effective for the goal	Design torque and displacement profiles similar to real buttons
3	Find suitable actuation technology and implement your application using it	DC Motor, Vibration motor, LRA, Piezo, Electricity...
4	Repeat steps 2 and 3 until the best solution is found.	Motor? – Too big for mobile devices Vibration motor? – Too long lag LRA? – Fast response, but no displacement Piezo? – Fast response, displacement can be rendered depending on design, electronics is hard, weak to shock... Electricity? – Easy to make, but difficult to control loading conditions (sweat...)
5	Evaluate your system	Can user feel button click sensations? How faithful are they? Do they really help typing?

# Characteristics of Tactile Rendering

---

- ▶ Understanding human tactile perception plays a critical role for successful tactile rendering.
- ▶ Tactile stimuli that can be rendered are totally dependent upon actuation technology.
- ▶ Tactile stimuli that can be rendered are often very limited.
- ▶ Computational algorithms are usually simple (easy programming).
- ▶ Knowledge on signal I/O, signal analysis, and physical systems is mandatory.
- ▶ Real usage scenarios often come with multi-modality (vision and audio), cognitive workload (getting a phone call alarm while watching DMB), and noisy environment (in a bus or subway).
- ▶ In most cases, researchers need to conduct empirical studies for haptic perception and usability evaluations designed for their own tactile rendering applications.

# Relevant Topics for Evaluation

---

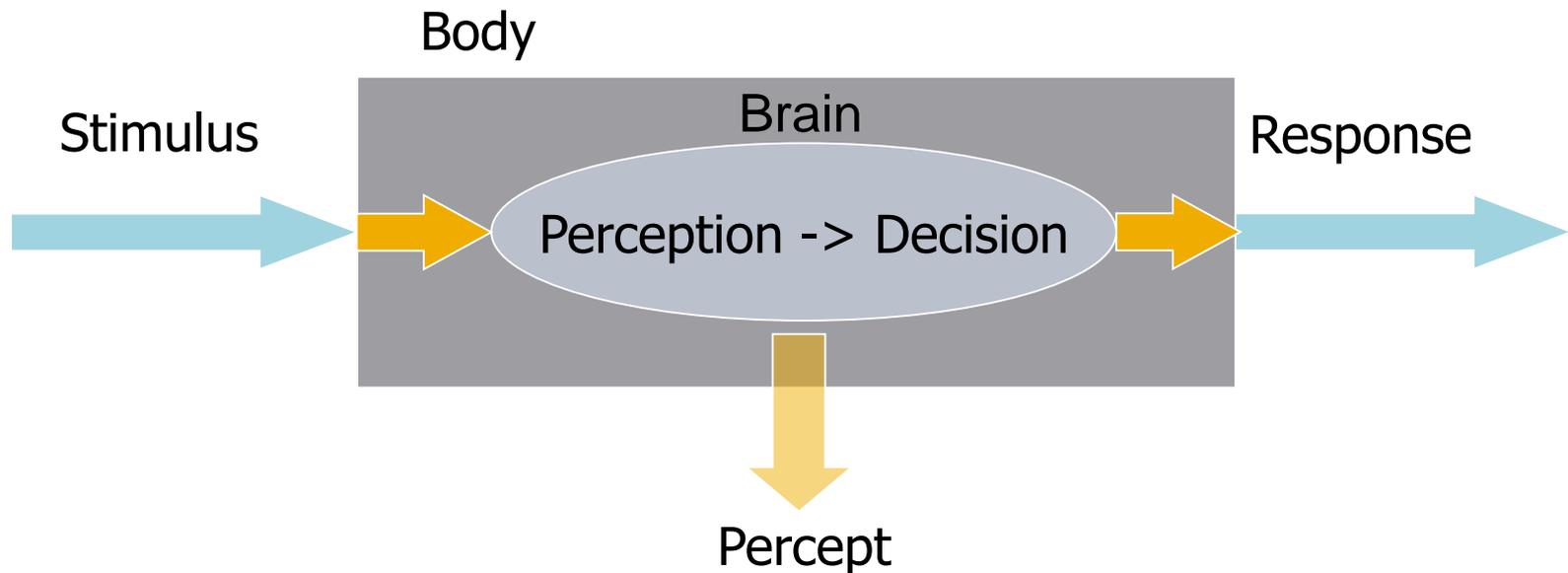
- ▶ Tactile Perception and Cognition
- ▶ Psychophysics
- ▶ Task Performance
- ▶ Cognitive Workload
- ▶ Subjective Evaluation
- ▶ Perceptual Space
- ▶ Affection

# PSYCHOPHYSICAL METHODS

# What is Psychophysics?

---

- ▶ Methodology for investigating relationships between stimuli in the physical domain and sensations in the psychological domain
- ▶ Central to experimental psychology



# Why Psychophysics for Haptics Research?

---

- ▶ Common Research Steps
  - ▶ Determine specifications for a haptics application based on the human capability of perception and action
  - ▶ Design and construct an appropriate haptic interface
  - ▶ Develop adequate software library including haptic rendering algorithms
  - ▶ Assess the developed system based on perception, task performance, subjective assessments, and usability
- ▶ Haptic perception is complex; receptors are distributed all over the body.
- ▶ Haptic perception is relatively less studied compared to visual and auditory perception.
- ▶ In many cases, we need to obtain perceptual data on our own.
- ▶ For tactile rendering, haptic perception is a key player.

# Topics to Cover for Psychophysics

---

- ▶ Basic Concepts and Laws
- ▶ Classical Psychophysical Methods
- ▶ Adaptive Methods
- ▶ Magnitude Estimation
- ▶ Decision Model for Psychophysics
- ▶ Information Theory
- ▶ Perceptual Space
- ▶ And many more...

# Threshold Level vs. Suprathreshold Level

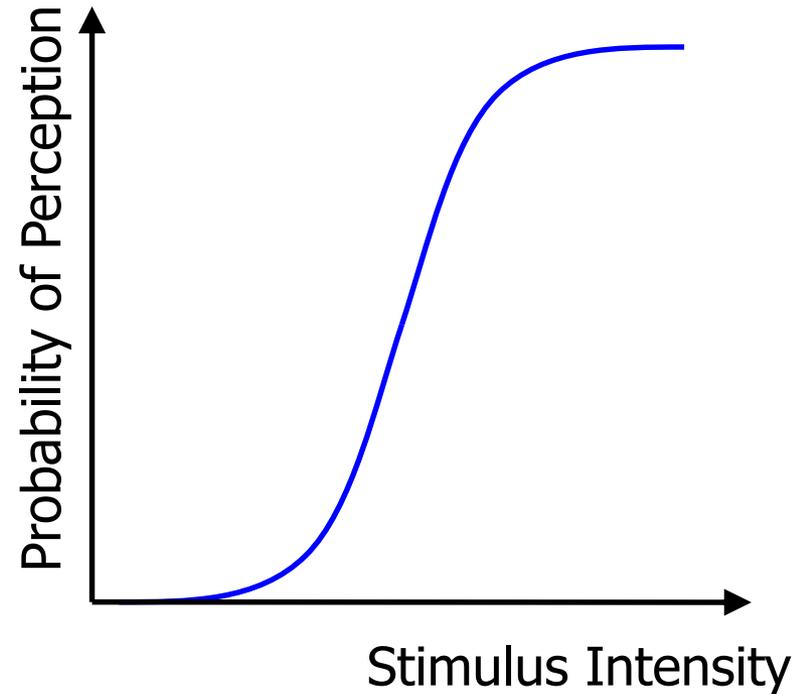
---

- ▶ Increasing the physical energy of stimulus increases the perceived magnitude of sensation.
- ▶ Perceptual Threshold
  - ▶ The smallest amount of stimulus energy required to produce a sensation
  - ▶ Also called Limen (originally German)
- ▶ Perceptual Magnitude
  - ▶ The magnitude of sensation we feel from a physical stimulus
  - ▶ Also called perceptual intensity, perceived magnitude, perceived intensity, sensation magnitude, sensation intensity ...
- ▶ Two Classes of Psychophysical Methods
  - ▶ Threshold Level: stimulus intensity  $\approx$  threshold
  - ▶ Suprathreshold Level: stimulus intensity  $\gg$  threshold

# Psychometric Function

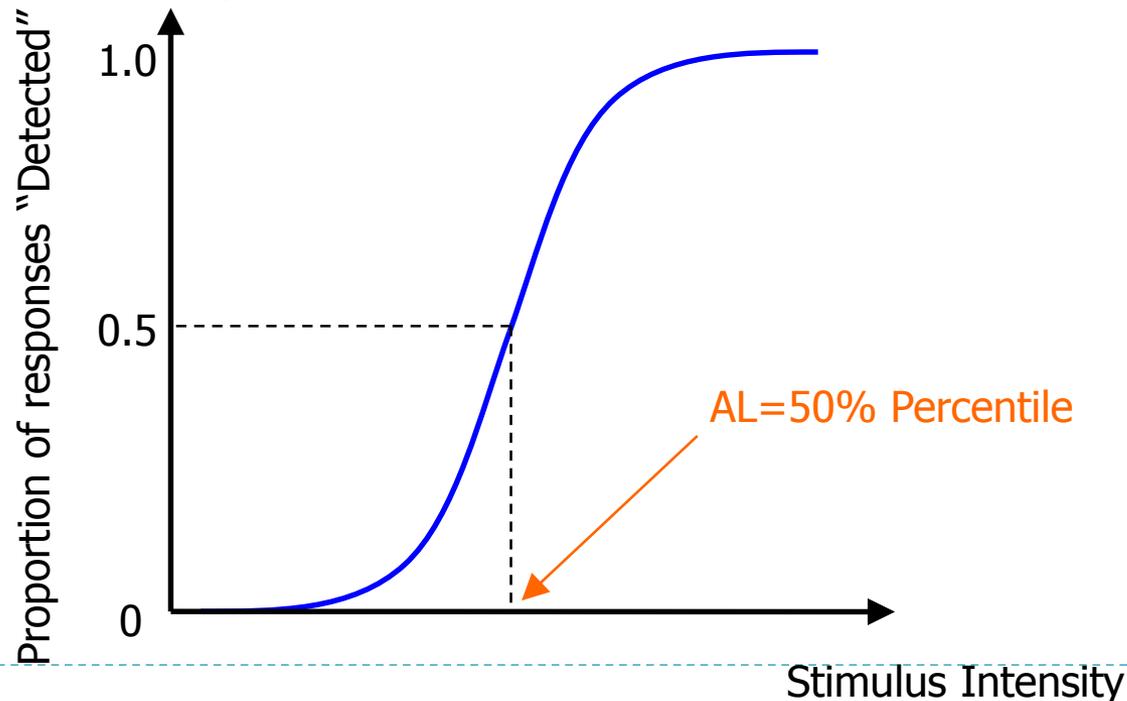
---

- ▶ A central concept of psychophysics for the threshold level
- ▶ A function from stimulus intensity to probability of perceiving the stimulus
- ▶ Explicitly models the sensory noise of perception process
- ▶ Usually a S-shaped ogive
- ▶ Modeled as cumulative normal or logistic distribution



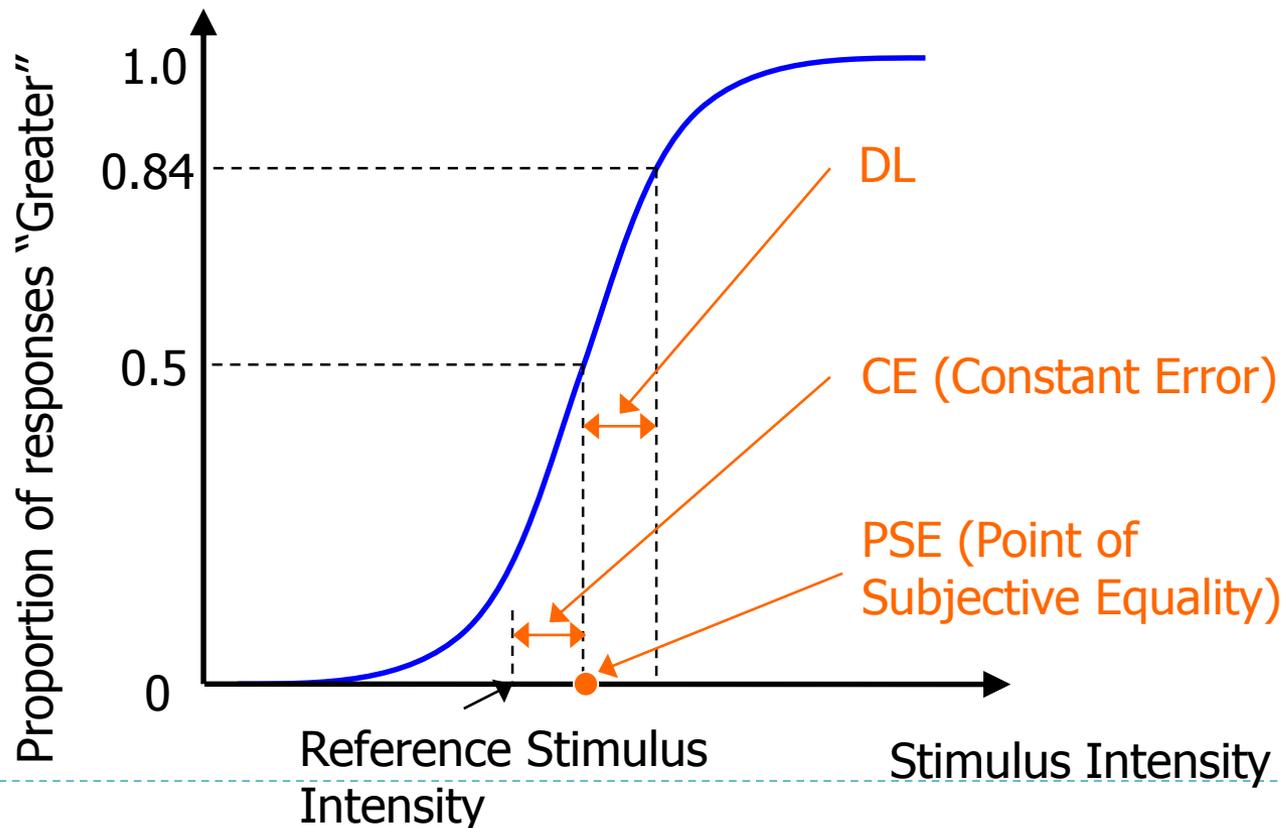
# Absolute Threshold

- ▶ The smallest amount of stimulus energy to produce a sensation that can be reliably detected.
- ▶ Also called detection threshold
- ▶ Often abbreviated as AL (absolute limen) or RL (Reiz Limen)
- ▶ 50% percentile is mostly widely used, but other higher percentiles are also frequently used.



# Difference Threshold

- ▶ The smallest amount of stimulus energy difference to produce a sensation that can be reliably discriminated
- ▶ Also called discrimination threshold
- ▶ Often abbreviated as DL (difference limen or differenz limen)



# Weber's Law

---

- ▶ Ernst H. Weber (1795 – 1878)
  - ▶ A founder of experimental psychology
- ▶ Weber's Law
  - ▶ Empirical law discovered in weight discrimination
  - ▶ Difference threshold is proportional to reference stimulus intensity

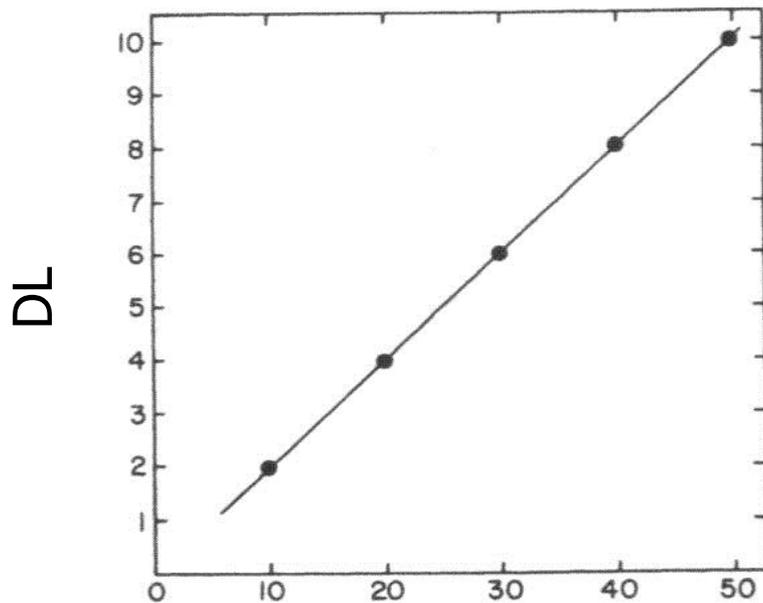
$$JND = wI_r \qquad w = \frac{JND}{I_r}$$

- ▶ JND: Just Noticeable Difference (= DL)
  - ▶ w: Weber Fraction
  - ▶  $I_r$ : Reference Stimulus Intensity
  - ▶ Mostly true, but exceptions also exist.
- 

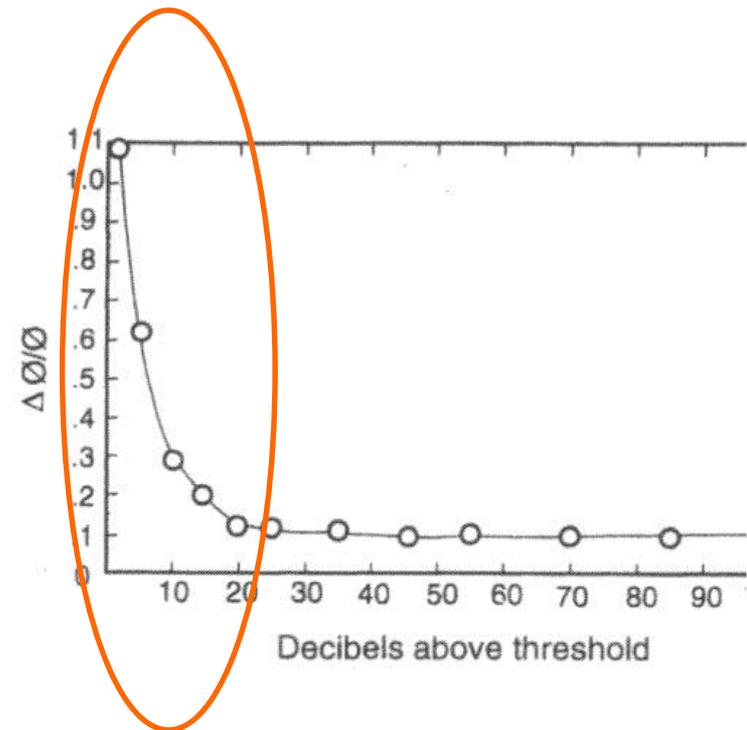


# Weber's Law – Graphical Illustration

---



Reference Stimulus Intensity



Common when reference stimulus intensities are close to AL

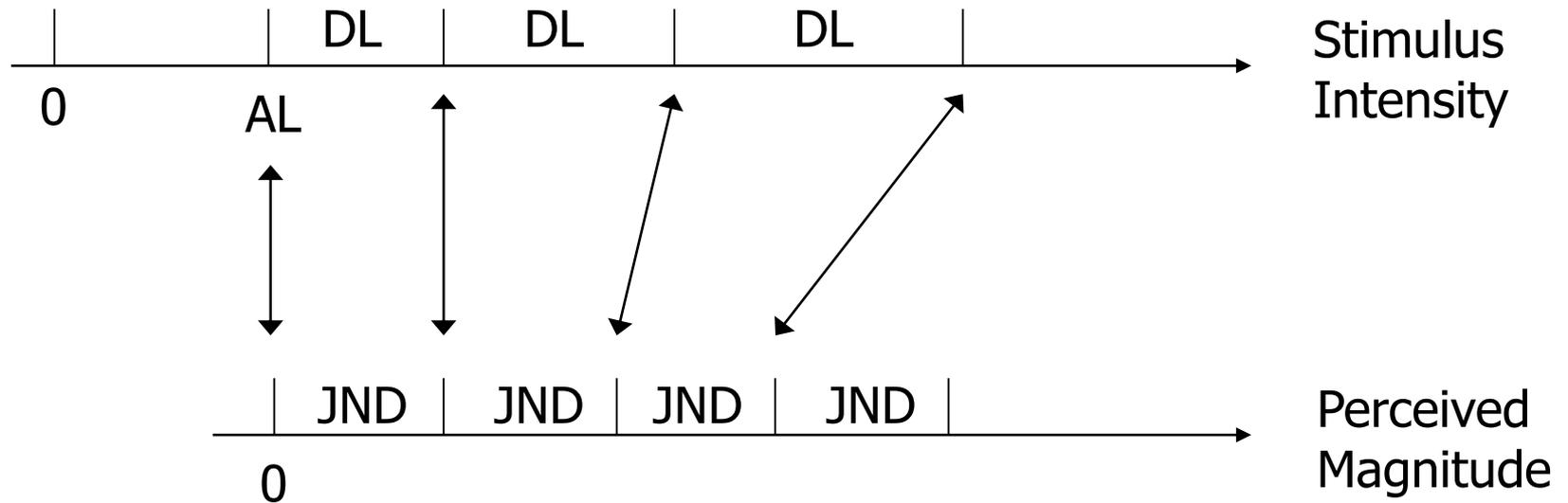
# Perceived Magnitude

---

- ▶ Psychophysical Magnitude Function
  - ▶ A mapping from stimulus physical intensity to perceived intensity
- ▶ Two empirical laws exist for the psychophysical magnitude function.
- ▶ Fechner's Law
  - ▶ Proposed by Fechner as a natural extension of Weber's law
  - ▶ Only applicable to very limited cases
  - ▶ Included for historical reasons
- ▶ Steven's Power Law
  - ▶ One of the best established empirical laws in psychology
  - ▶ Use this form to obtain a psychophysical magnitude function

# Fechner's Law – Idea

---



JND: Just Noticeable Difference (a Unit of Perceived Magnitude)

\* Note that DL and JND are used interchangeably in the most literature.

# Fechner's Law – Equation

---

- ▶ “Elements of Psychophysics,” 1860.
- ▶ P: Perceived magnitude  
I: Stimulus intensity  
I<sub>0</sub>: Absolute threshold
- ▶ From Weber's law,

$$dP = c \frac{dI}{I}$$

$$\int_{P_0}^P dP = c \int_{I_0}^I \frac{dI}{I}$$

$$P - P_0 = c(\log I - \log I_0)$$



$$P_0 = 0$$

Fechner's Law

$$P = c \log \left( \frac{I}{I_0} \right)$$

- ▶ Examples where Fechner's law holds well
  - ▶ Sound intensity: decibel
  - ▶ Pitch perception

# Steven's Power Law

---

- ▶ By Stanley S. Stevens (1906-1973)
- ▶ Can be viewed as a general form of Fechner's Law

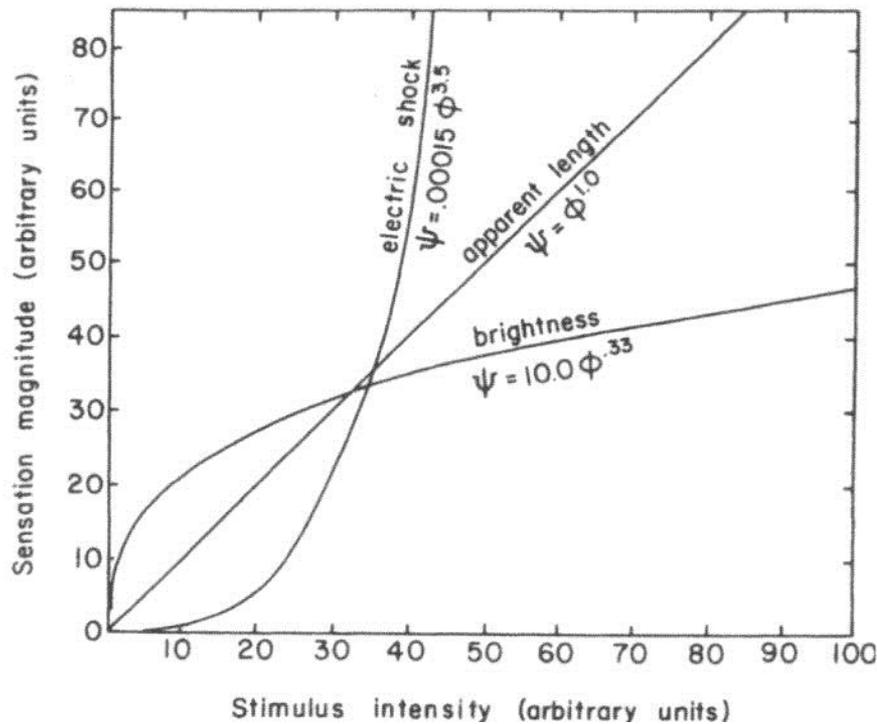
$$P = cI^n \quad \log P = n \log I + c'$$

$$\text{or } P = c(I - I_0)^n$$

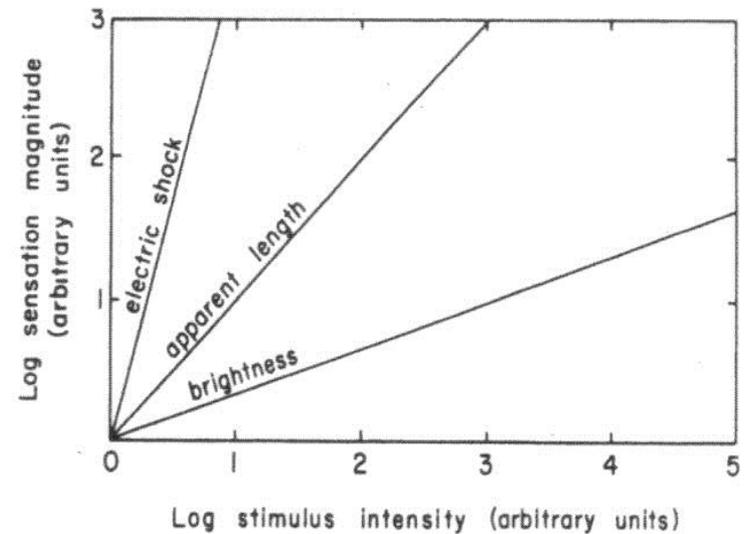
- ▶ The exponent  $n$  depends on stimulus conditions.
  - ▶ Brightness: 0.5 (Point source)
  - ▶ Vibration: 0.6 (250 Hz on finger)
  - ▶ Visual length: 1 (Projected line)
  - ▶ Electric shock: 3.5 (Current through fingers)
  - ▶ See [http://en.wikipedia.org/wiki/Stevens'\\_power\\_law](http://en.wikipedia.org/wiki/Stevens'_power_law) for a comprehensive list
- ▶ Use the Power law whenever you need to get a magnitude function

# Steven's Power Law - Examples

Linear vs. Linear



Log vs. Log



# How to Measure Thresholds

---

## ▶ Classical Psychophysical Methods

	Accuracy	Efficiency
Method of constant stimuli	Best	Worst
Method of limits	Middle	Middle
Method of adjustment	Worst	Best

## ▶ Adaptive Methods

- ▶ PEST: Parameter Estimation by Sequential Testing
- ▶ Maximum-Likelihood Procedures
- ▶ Staircase Procedures

# Classical Psychophysical Methods

---

- ▶ Invented by Fechner
- ▶ For measuring perceptual thresholds

Method	Accuracy	Efficiency	Notes
Method of constant stimuli	Best	Worst	Only method that allows to obtain a whole psychometric function
Method of limits	Middle	Middle	Balanced Can be used for research purposes
Method of adjustment	Worst	Best	Good for pilot experiments or clinical trials Avoid using this for research purposes if possible

# Method of Constant Stimuli

---

## ▶ Stimuli

- ▶ A set of stimulus intensities (5-9) that are evenly spaced

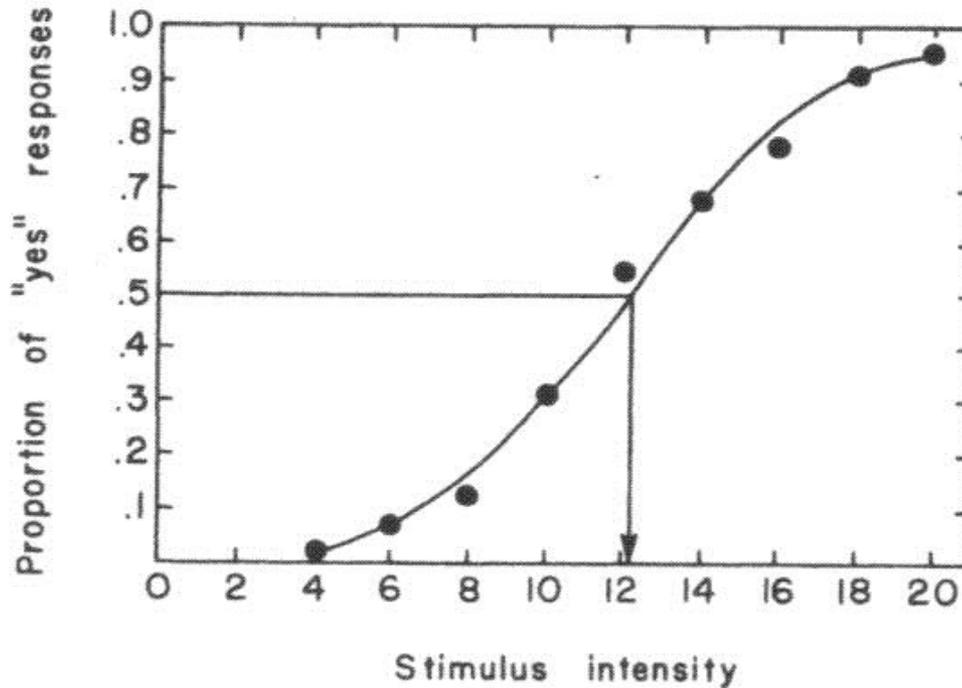
## ▶ Procedures

- ▶ A stimulus is randomly selected from the stimulus intensity set and presented to the participant
- ▶ The participant is asked to answer
  - ▶ Whether the stimulus was detected (for AL), or
  - ▶ Whether the test (or comparison) stimulus was greater than the reference stimulus (for DL)
- ▶ Each stimulus intensity should be repeated in a large number of trials (e.g., 100)
- ▶ The time necessary to finish an experiment also needs to be considered to determine the number of trials.

# Data Analysis

---

- ▶ Record the proportions of “yes” and plot them against stimulus intensity
- ▶ Fit a psychometric function to the recorded data



# Psychometric Function Fitting

---

## ▶ Widely-Used Probability Distributions

### ▶ The normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad F(x) = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{x-\mu}{\sqrt{2}\sigma} \right) \right), \text{ where } \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

### ▶ The logistic distribution

$$f(x) = \frac{e^{-\frac{x-\mu}{s}}}{s \left( 1 + e^{-\frac{x-\mu}{s}} \right)^2} \quad F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{s}}} = \frac{1}{2} \left( 1 + \tanh \left( \frac{x-\mu}{2s} \right) \right)$$

$$\text{mean} = \mu, \quad \text{s.d} = \frac{\pi}{\sqrt{3}} s$$

- ▶ Fit a function of your choice to measured data and determine function parameters (probit analysis for the normal distribution or regular regression analysis)

# Finding Thresholds from a Psychophysical Function

---

- ▶ Determine the “level” of correctness in perception
- ▶ Detection Experiment
  - ▶ 50% detectability (half right and half wrong):  $AL = \text{mean}$
  - ▶ 84% detectability:  $AL = \text{mean} + \text{standard deviation}$
  - ▶  $x\%$  detectability:  $AL = x\%$  percentile (stimulus intensity that results in the detection probability of  $x$ )
- ▶ Discrimination Experiment
  - ▶  $PSE = \text{mean}$
  - ▶ 84% discriminability:  $DL = \text{standard deviation}$
  - ▶  $x\%$  discriminability:  $DL = x\%$  percentile – mean (PSE)

# Remarks

---

## ▶ Pros and Cons

- ▶ The method of constant stimuli produces more reliable threshold estimates compared to the methods of limits or adjustment.
- ▶ It allows to measure a whole psychometric function.
- ▶ It, however, requires a more number of trials and takes longer to complete.

## ▶ Tips for stimulus design

- ▶ Use stimulus intensity levels symmetric to an expected threshold.
- ▶ Do not include many stimulus intensities that result in nearly zero or perfect perceptibility (only one in each side).
- ▶ Use the method of adjustment to find a rough threshold and use it to determine stimulus intensities.

# Methods of Limits: Procedures for Measuring Absolute Thresholds (1)

---

- ▶ The method of limits consists of the same number of ascending and descending series
- ▶ Ascending series
  - ▶ A series begins with a stimulus intensity well below the threshold.
  - ▶ A subject is asked to answer if s/he detected the stimulus.
  - ▶ If the answer is “no”, the stimulus intensity is increased by a small step.
  - ▶ Steps 2 and 3 are repeated until the subject changes his response from “no” to “yes.”
- ▶ Descending series
  - ▶ A series begins with a stimulus intensity well above the threshold.
  - ▶ A subject is asked to answer if s/he detected the stimulus.
  - ▶ If the answer is “yes”, the stimulus intensity is decreased by a small step.
  - ▶ Steps 2 and 3 are repeated until the subject changes his response from “yes” to “no.”

# Methods of Limits: Procedures for Measuring Absolute Thresholds (2)

---

## ▶ Cautions

- ▶ The order of series (ascending or descending) should be randomized.
- ▶ It is advised to randomize initial stimulus intensity (of course well above or below an expected threshold). Otherwise, participants tend to rely on memory.

## ▶ Data Analysis

- ▶ In each series, a threshold estimate is a mean between two stimulus intensities right before and right after a subject changes his answers from “no” to “yes” (ascending series) or from “yes” to “no” (descending series).
- ▶ The AL is the mean of the threshold estimates.

# AL Measurement – Example

for hearing by the Method of Limits\*

Stimulus intensity (dB)	A	D	A	D	A	D	A	D	A	D
10						Y				
9		Y				Y				Y
8		Y				Y				Y
7		Y		Y		Y				Y
6		Y		Y	Y	Y		Y		Y
5	Y	Y		Y	N	Y		Y		Y
4	N	Y	Y	N	N	N	Y	Y	Y	N
3	N	N	N		N		N	Y	N	
2	N		N		N		N	N	N	
1	N		N		N		N	N	N	
0	N		N				N		N	
-1	N		N				N			
-2	N						N			
-3	N						N			
-4	N									
-5	N									
-6	N									
-7	N									
-8	N									
-9	N									
-10	N									
Transition points =	4.5	3.5	3.5	4.5	5.5	4.5	4.5	2.5	3.5	4.5

\*Mean threshold value = 4.1

# Response Bias

---

- ▶ In a psychophysical experiment,
  - ▶ Stimulus → Perception → Decision → Response
- ▶ Response bias
  - ▶ The effect of decision, not perception, reflected in the responses of a subject
  - ▶ The tendency of a subject to favor one response over another determined by factors other than signal intensities
- ▶ Examples
  - ▶ Tendency to detect “non-present” signals in a detection experiment
  - ▶ Paying off a subject
- ▶ All classical psychophysical methods can suffer from the problem of response bias

# Constant Errors in the Method of Limits

---

## ▶ Errors of Habituation

- ▶ Refers to a tendency of a subject developing a habit of repeating the same response.
- ▶ The subject may continue to change responses a few trials after the threshold.
- ▶ Results in larger transition points in ascending series and smaller transition points in descending series.

## ▶ Errors of Expectation

- ▶ Refers to a tendency of a subject anticipating a premature arrival at the threshold.
- ▶ The subject may change responses a few trials before the threshold.
- ▶ Results in smaller transition points in ascending series and larger transition points in descending series.

# Measuring Difference Threshold – Procedures

---

- ▶ Ascending series
  - ▶ A series begins with a test stimulus the intensity of which is well below the reference stimulus intensity.
  - ▶ A subject is asked to answer if the test stimulus felt “less than”, “equal to”, or “greater than” the reference stimulus.
  - ▶ If the answer is “less than” or “equal to”, the stimulus intensity is increased by a small step.
  - ▶ Steps 2 and 3 are repeated until the subject changes his response from “equal to ” to “greater than.”
- ▶ Descending series
  - ▶ A series begins with a test stimulus the intensity of which is much above the reference stimulus intensity.
  - ▶ A subject is asked to answer if the test stimulus felt “less than”, “equal to”, or “greater than” the reference stimulus.
  - ▶ If the answer is “greater than ” or “equal to”, the stimulus intensity is decreased by a small step.
  - ▶ Steps 2 and 3 are repeated until the subject changes his response from “equal to” to “less than.”

# Measuring Difference Threshold – Data Analysis

---

- ▶ In each series, two points are recorded.
  - ▶ Upper limen ( $L_u$ ): A point where “greater than” responses changes to an “equal to” response.
  - ▶ Lower limen ( $L_l$ ): A point where “less than” responses changes to an “equal to” response.
- ▶ Then,

$$\text{PSE} = \frac{1}{2}(\bar{L}_u + \bar{L}_l), \text{ and}$$

$$\text{DL} = \frac{1}{2}(\bar{L}_u - \bar{L}_l),$$

where  $\bar{L}_u$  and  $\bar{L}_l$  are the means of upper and lower limens, respectively.

# DL Measurement – Example

for Hearing by the Method of Limits<sup>a</sup>

Stimulus intensity (dB)	A	D	A	D	A	D	A	D	A	D
24.5						G				
24.0		G				G		G		
23.5		G				G		G		G
23.0		G		G	G	G		G		G
22.5		G	G	G	E	G	G	G		G
22.0	G	E	E	G	E	G	E	G	G	E
21.5	E	E	E	E	E	G	E	E	E	E
21.0	E	E	E	E	E	E	E	E	E	E
20.5	E	E	E	E	E	E	E	E	E	E
20.0	E	E	E	E	E	E	E	E	E	E
19.5	E	E	E	E	E	E	E	E	E	E
19.0	E	E	E	E	E	E	E	E	E	E
18.5	E	L	E	E	E	E	E	E	E	E
18.0	E		E	L	E	L	E	L	E	E
17.5	L		L		E		L		L	L
17.0	L		L		L		L		L	
16.5	L		L		L		L		L	
16.0	L				L				L	
15.5	L				L				L	
Upper limen	21.75	22.25	22.25	21.75	22.75	21.25	22.25	21.75	21.75	22.25
Lower limen	17.75	18.75	17.75	18.25	17.25	18.25	17.75	18.25	17.75	17.75

<sup>a</sup>Interval of uncertainty =  $IU = \bar{L}_u - \bar{L}_l = 22.00 - 17.95 = 4.05$ . Difference limen =  $DL = \frac{1}{2} IU = \frac{1}{2} (4.05) = 2.025$ . Point of subjective equality =  $PSE = \frac{1}{2} (\bar{L}_u + \bar{L}_l) = \frac{1}{2} (22.00 + 17.95) = 19.97$ .

# Measuring Difference Threshold – Alternative

---

- ▶ There are both ascending and descending series.
- ▶ In each trial (either in ascending or descending series), the subject's task is to determine whether the test stimulus is “greater than” or “less than” the reference.
- ▶ The stopping rule is the same as that of AL measurement.
- ▶ Then,
  - ▶ PSE = the mean of all transition points
  - ▶ DL = the standard deviation of all transition points

# Summary

---

## ▶ Pros

- ▶ Balanced between accuracy and efficiency.
- ▶ Used frequently.
- ▶ Cannot estimate a psychometric function.

## ▶ Cons

- ▶ Two constant errors
- ▶ The subject is aware of the direction of stimulus intensity changes, which causes response bias.

## ▶ Notes

- ▶ Use the same number of ascending and descending series and randomize their presentation order.
- ▶ Also randomize initial stimulus intensity.

# Method of Adjustment – Procedures

---

- ▶ AL Measurement
  - ▶ In each trial, a subject is asked to adjust the stimulus intensity so that it is just barely detectable.
  - ▶ AL = the mean of the adjusted intensities.
- ▶ DL Measurement
  - ▶ In each trial, a subject is asked to adjust the test stimulus intensity so that it is perceptually identical to the reference stimulus.
  - ▶ PSE = the mean of the adjusted intensities.
  - ▶ DL = the standard deviation of the adjusted intensities.
- ▶ Pros and Cons
  - ▶ Most efficient
  - ▶ Generally considered too inaccurate for research purposes due to the too much control of a subject over stimulus intensities
  - ▶ Can be applied only when stimulus intensity can be adjusted continuously or in very small steps.

# Overview of Adaptive Methods

---

## ▶ Motivations

- ▶ The classical psychophysical methods tend to use test stimuli from which little information on thresholds can be obtained.
- ▶ This makes the classical methods inefficient in terms of experiment time.

## ▶ Adaptive Methods

- ▶ By “adaptive”, we mean that stimulus intensity is not predetermined, but calculated during an experimental run based on the previous history of stimulus/response pairs.
- ▶ Adaptive methods place stimulus intensities closely on a corresponding threshold level and thus require a significantly less number of trials than the classical methods.
- ▶ Adaptive methods are particularly useful for estimating a point in a psychometric function.

# Experiment Design for Adaptive Methods

---

- ▶ An adaptive method consists of the following four components.
  - ▶ Initial stimulus intensity
  - ▶ Step size
    - ▶ The difference of stimulus intensity between subsequent trials
  - ▶ Tracking algorithm
    - ▶ The process that guides the sequence of stimulus intensities in an experimental session
  - ▶ Stopping rule
    - ▶ A decision rule for ending the process

# Modern Adaptive Methods

---

- ▶ Categories
  - ▶ PEST: Parameter Estimation by Sequential Testing
  - ▶ Maximum-Likelihood Procedures
  - ▶ Staircase Procedures
- ▶ We will focus on the staircase procedures since they are much simpler than, yet their performances are comparable to the others.

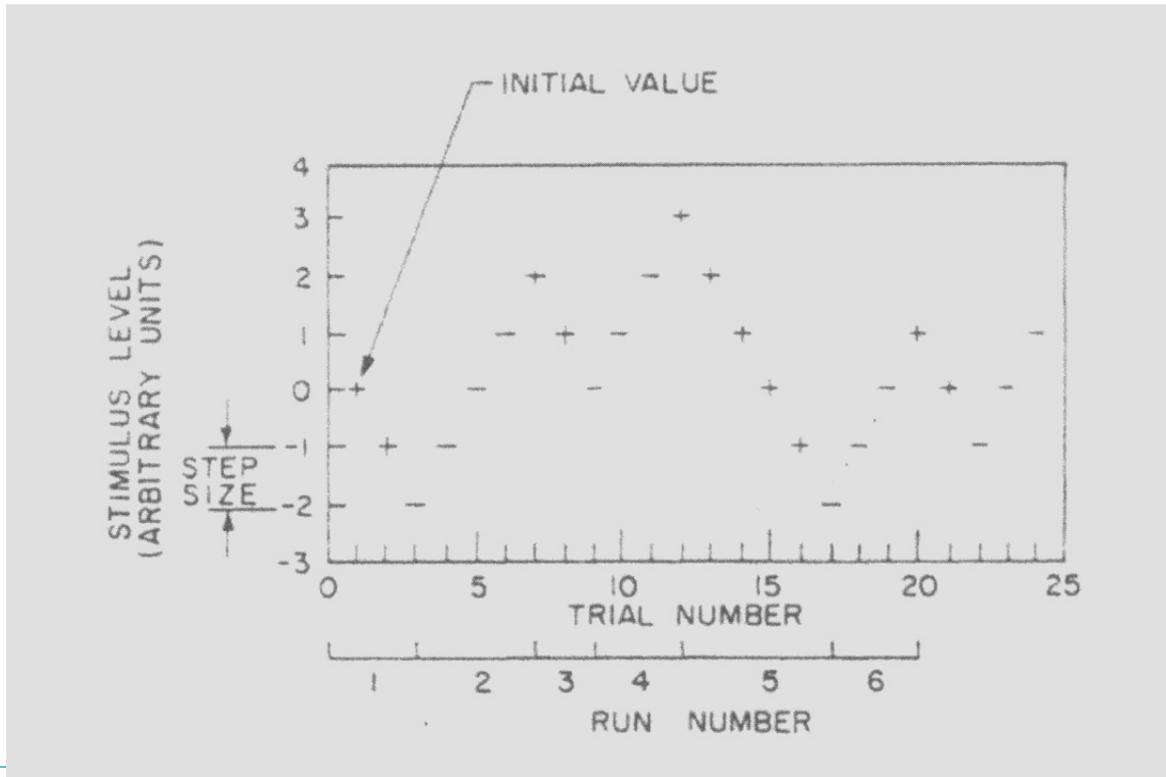
# Procedures for Staircase Methods

---

- ▶ An experimental run starts with a stimulus intensity well above (or sometimes below) a target performance level.
- ▶ Test stimulus intensity is increased after a negative response and decreased after a positive response by a step size.
- ▶ Rules for constituting the positive and negative responses are designed to meet the specified performance level of a threshold.
- ▶ Step sizes are relatively large in the beginning of a session for fast convergence, but are soon reduced to small values as the run proceeds.
- ▶ The session typically terminates after a specified number of response reversals.

# Simple Up-Down Method

- ▶ Increase stimulus intensity after one incorrect answer, and decrease stimulus intensity after one correct answer.
- ▶ Estimates the 50% percentile of a psychometric function.
- ▶ Estimate = mean of midpoints of 2nd (ascending) series



# Step Size

---

- ▶ In general, step size is gradually decreased during the course of an experiment.
- ▶ There are many rules for determining step sizes.
  - ▶  $c / n$  ( $c$ : constant and  $n$ : number of trials)
  - ▶ Divide the step size to a half after a fixed number of trials
  - ▶ Use a large step size in the beginning, and a small step size after a predetermined number of reversals

# Transformed Up-Down Method

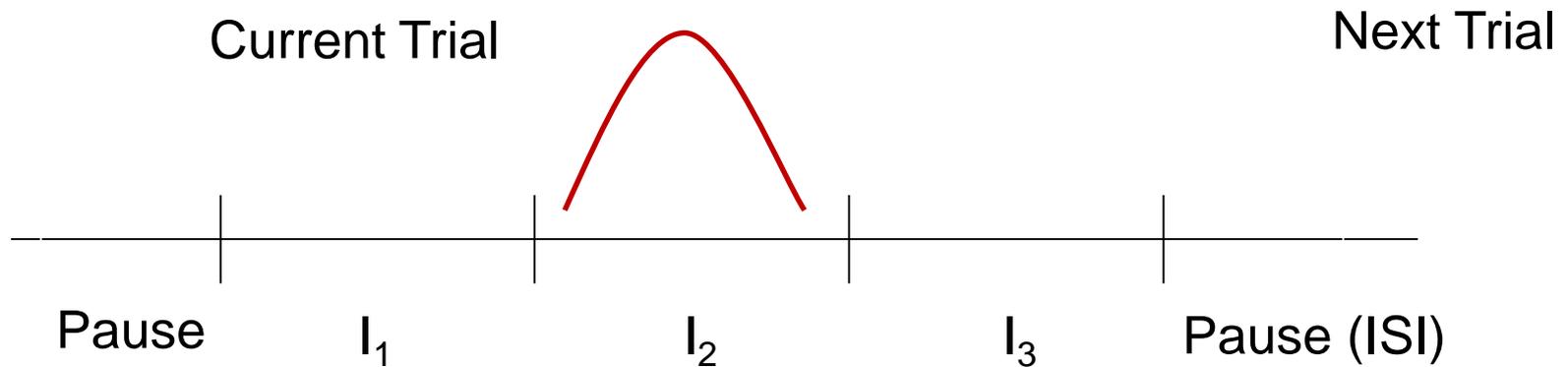
- ▶ Levitt (1971)
- ▶ Provides ways to measure, in a psychometric function, points other than a 50 % percentile.

	1-Up 1-Down	1-Up 2-Down	1-Up 3-Down	1-Up 4-Down
Up Group	–	+ – , –	++ – , + – , –	+++ – , ++ – , + – , –
Down Group	+	++	+++	++++
P(Down)	$P(X)$	$[P(X)]^2$	$[P(X)]^3$	$[P(X)]^4$
Probability Converged	0.5	0.707	0.794	0.841

# 3IFC 1-Up 3-Down Adaptive Staircase Procedure (1)

---

- ▶ 3 interval force-choice (IFC) or alternative force-choice (AFC)
  - ▶ A signal is presented in one randomly selected interval among three intervals.



- ▶ 2I vs. 3I
  - ▶ Two interval design is more common
  - ▶ Three interval design was shown more robust
- ▶ 1-Up 3-Down method
  - ▶ Estimates a 79 % percentile point.

# 3IFC 1-Up 3-Down Adaptive Staircase Procedure (2)

---

- ▶ Example rule for step size
  - ▶ 4 dB before first three reversals
  - ▶ 1 dB after that
- ▶ Stopping rule
  - ▶ An experimental run is terminated after 12 reversals occur at the 1 dB step size.
- ▶ A widely used procedure for threshold measurement due to its efficiency and robustness.
- ▶ Reference:
  - ▶ Brisben et. al. (1999)

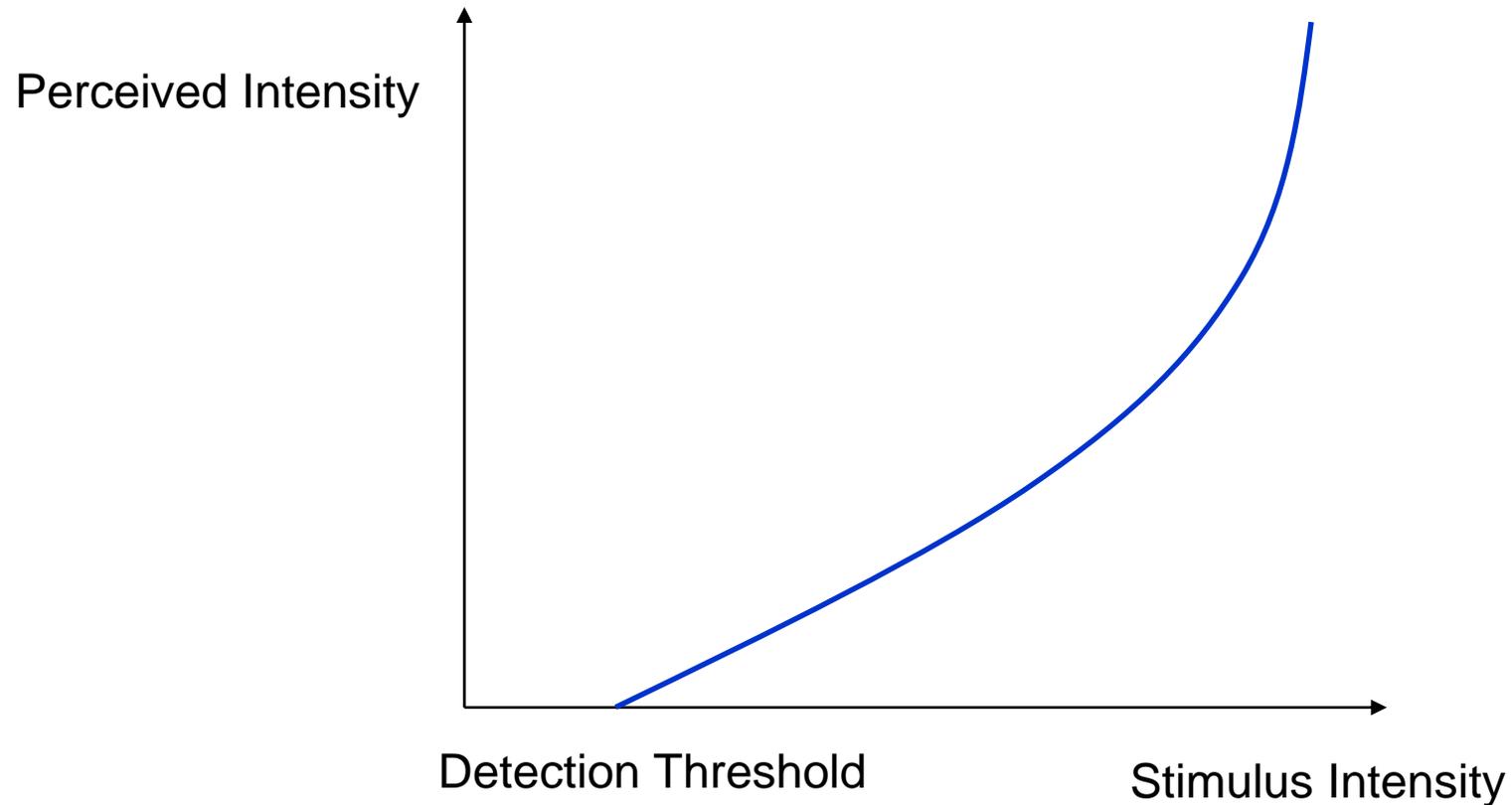
# Magnitude Estimation

---

- ▶ So far, we have focused on sensory perception in the threshold level.
  - ▶ The points where a percept can occur (i.e., detection and difference thresholds) received main attention.
  - ▶ The output of the perception process is essentially binary (i.e., can feel it/cannot feel it), although it was formulated in the probabilistic framework.
- ▶ We now move to how to explain the perception process in the supra-threshold level
  - ▶ The output of the perception process (perceived intensity) varies continuously with respect to the continuous change of the input (stimulus intensity).
  - ▶ A central concept is a psychophysical magnitude function.

# Hypothetical One-Dimensional Psychophysical Magnitude Function

---



The actual function shape follows Steven's power law.

# Measurement Scales

---

- ▶ Purpose

- ▶ To represent the different degrees of correspondence between the number system and the property of objects or events.

- ▶ Four Common Measurement Scales

- ▶ Nominal Scale: Identity
- ▶ Ordinal Scale: Order
- ▶ Interval Scale: Interval
- ▶ Ratio Scale: Origin
- ▶ Nominal Scale  $\subset$  Ordinal Scale  $\subset$  Interval Scale  $\subset$  Ratio Scale

# Nominal Scale and Ordinal Scale

---

## ▶ Nominal Scale

- ▶ Used to classify or identify objects.
- ▶ Example: Back numbers of soccer players
- ▶ Numbers in a nominal scale only label objects and do not have any quantitative information.
- ▶ Therefore, a nominal scale cannot provide a means of measurement.

## ▶ Ordinal Scale

- ▶ A set of measurements in which the amount of a property of objects or events can be ranked (greater-than or less-than relation)
- ▶ Example: Ranks in the heights of five soccer players.
- ▶ The numerical differences of the property amounts between ranks are not available.
- ▶ An ordinal scale simply provides an order of objects or events.

# Interval Scale

---

- ▶ Scale values represent differences or distances between the amounts of property.
- ▶ Example
  - ▶ Temperature scale
  - ▶ The difference between  $20^\circ$  and  $40^\circ$  is the same to that between  $60^\circ$  and  $80^\circ$ .
  - ▶ The difference between  $20^\circ$  and  $40^\circ$  is twice larger than that between  $60^\circ$  and  $70^\circ$ .
- ▶ The differences between numbers, as well as their ordinal relation, has meanings.

# Ratio Scale

---

- ▶ A ratio scale has the properties of order and distance, and also has a natural origin to represent zero amount of a property.
- ▶ Example
  - ▶ Temperature in Centigrade/Fahrenheit: Interval scale
  - ▶ Temperature in Kelvin: Ratio scale (0 K represent the absolute state in which no heat exists)
- ▶ A ratio scale is most useful and highly desirable measurement system.
- ▶ Therefore, many psychophysical scaling techniques were designed to construct ratio scales for sensory attributes.

# Psychophysical Scales

---

- ▶ There are mainly three kinds of methods for constructing and validating scales of sensory attributes.
- ▶ Discrimination Scales (Confusion Scales)
  - ▶ Can provide an interval scale
  - ▶ Based on indirect scaling procedures in which sensory magnitudes of stimuli are inferred from measures of stimulus discriminability.
- ▶ Partition Scales
  - ▶ Can provide an interval scale
  - ▶ Uses direct scaling procedures in which the participant must make direct judgments of psychological differences among stimuli.
- ▶ Ratio Scales
  - ▶ Can provide a ratio scale
  - ▶ Relies on the ability of the participant to make direct judgments of the ratio relationships between the magnitudes of sensations.

# Psychophysical Ratio Scaling

---

- ▶ Methods for constructing ratio scales of sensations have been extensively studied in the past 35 years.
- ▶ Ratio Scaling Methods
  - ▶ Ratio Production
  - ▶ Ratio Estimation
  - ▶ Magnitude Estimation
  - ▶ Magnitude Production
  - ▶ Absolute Magnitude Estimation
- ▶ Magnitude estimation is that the most widely used method for psychophysical ratio scaling

# Ratio Production and Estimation

---

- ▶ Ratio Production
  - ▶ The participant is asked to adjust a stimulus intensity so that the perceived intensity of the final stimulus is a fraction (e.g., half) of the perceived intensity of the standard stimulus.
- ▶ Ratio Estimation
  - ▶ The participant's task is to tell the estimated ratio between the perceived intensities of two stimuli.
- ▶ Ratio production and ratio estimation were actively used in the past.
- ▶ However, they are no longer used frequently, because they have been found to be particularly sensitive to stimulus context effects (contrast effect, sequential effect – assimilation bias, and stimulus range effect – stimulus-range equalizing bias ).

# Magnitude Estimation with Modulus

---

- ▶ The subject is presented with a standard stimulus, and instructed that the sensation it produces has a certain numerical value (modulus), e.g., 10.
- ▶ On subsequent trials, other stimuli are presented, and the subject assigns numbers to their sensations relative to the value of the modulus.
- ▶ The participant is instructed to make his judgment reflect how many times greater one sensation is than another (the ratio between the two sensations).

# Magnitude Estimation without Modulus

---

- ▶ No standard modulus is defined by the experimenter.
- ▶ A subject can define his own modulus in the first trial and use it as a modulus in the subsequent trials.
- ▶ Other procedures are the same to those of magnitude estimation with modulus.
- ▶ Magnitude estimation without modulus is used more frequently.
- ▶ In no modulus design, the data of different subjects are combined using the geometric mean.

$$\text{Geometric Mean} = \log^{-1} \left( \frac{\sum \log X}{N} \right)$$

# Magnitude Production

---

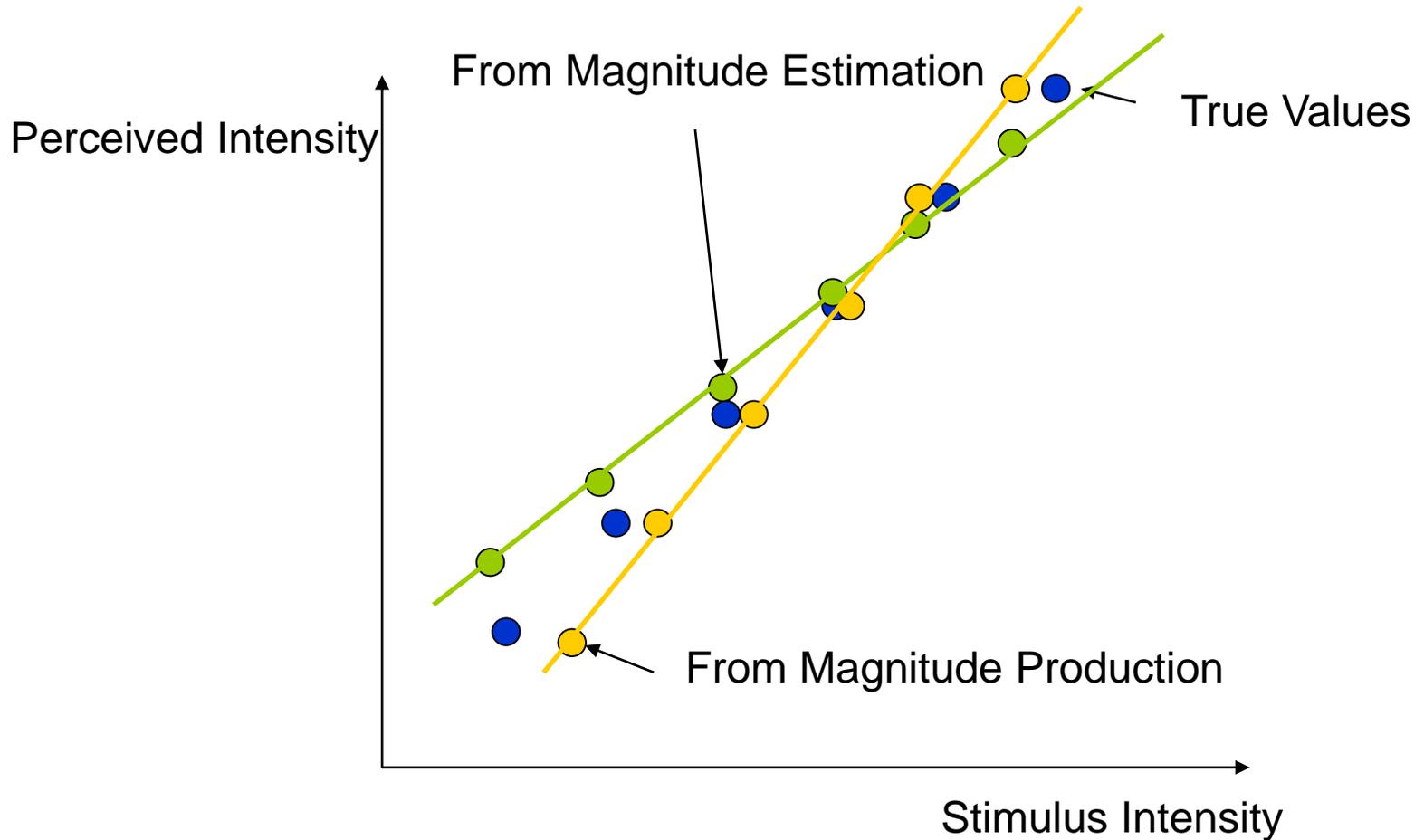
- ▶ An inverse of magnitude estimation.
- ▶ The experimenter tells the numerical value of some sensory magnitude to the subject, and the subject adjusts stimulus intensity to produce the number.
- ▶ Magnitude production can be used to confirm the soundness of magnitude estimation.
- ▶ Using magnitude estimation and production together can offset any systematic errors inherent in either method.

# Method of Numerical Magnitude Balance

---

- ▶ Regression bias
  - ▶ A regression of the subject's judgments toward the mean of his judgments.
  - ▶ The subject tends to be reluctant to make radical responses.
- ▶ In magnitude estimation, the subject answers perceived intensity higher than its true value for low stimulus intensities, and lower for high stimulus intensities. This decreases the slope of the psychophysical magnitude function.
- ▶ In magnitude production, the subjects adjust stimulus intensity higher than its true value in low perceived intensities, and lower for high perceived intensities. This increases the slope of the psychophysical magnitude function.
- ▶ Using the two methods together and taking the geometric means can be an effective way to cancel off the effects of regression bias.

# Example of Regression Bias



# Absolute Magnitude Estimation

---

- ▶ An absolute scale is a restricted case of ratio scales in which scale values cannot be transformed in any way.
- ▶ There has been ample evidence that we use an absolute scale for sensory magnitudes, which may have been fixed at an early age.
- ▶ If the relation between a standard stimulus and a modulus arbitrarily assigned by the experimenter is different from what a subject would assign in absolute scaling, the use of the standard stimulus may bias the resulting psychophysical scale.
- ▶ At present, the method of absolute magnitude estimation is recommended to prevent the potential biasing effects of a standard stimulus.

# Subject Instructions in Absolute Magnitude Estimation

---

- ▶ In this experiment, we would like to find out how intense various stimuli appear to you. For this purpose, I am going to present a series of stimuli to you once at a time. Your task will be to assign a number to every stimulus in such a way that your impression of how large the number is matches your impression of how intense the stimulus is. We all have impressions of how large various numbers are, and impression of how intense various stimuli are. I would like you to assign a number to each stimulus so that your impression of the size of the number matches your impression of the intensity of the stimulus.
- ▶ You may use any positive number that appears appropriate to you – whole numbers, decimals, or fractions. Do not worry about running out of numbers – there will always be a smaller number than the smallest you use and a larger one than the largest you use. Do not worry about numbers you assigned to preceding stimuli. Do you have any questions?

# Detection Theory

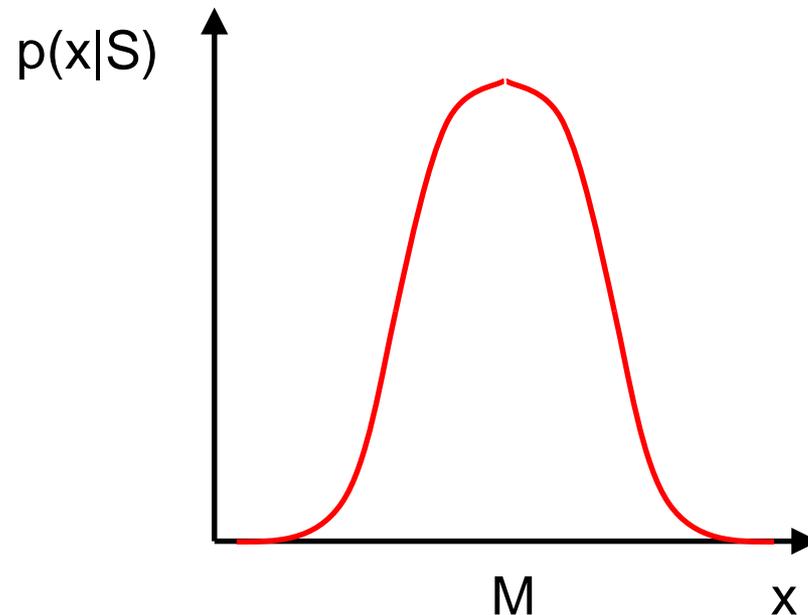
---

- ▶ A more modern approach
- ▶ Based on the signal detection theory in communication
- ▶ Provides means to separate the decision process (e.g., response bias) from perception process
  
- ▶ Models the human as a noisy communication channel
- ▶ Uses the metrics of sensitivity index ( $d'$ ) and response bias ( $c$ )

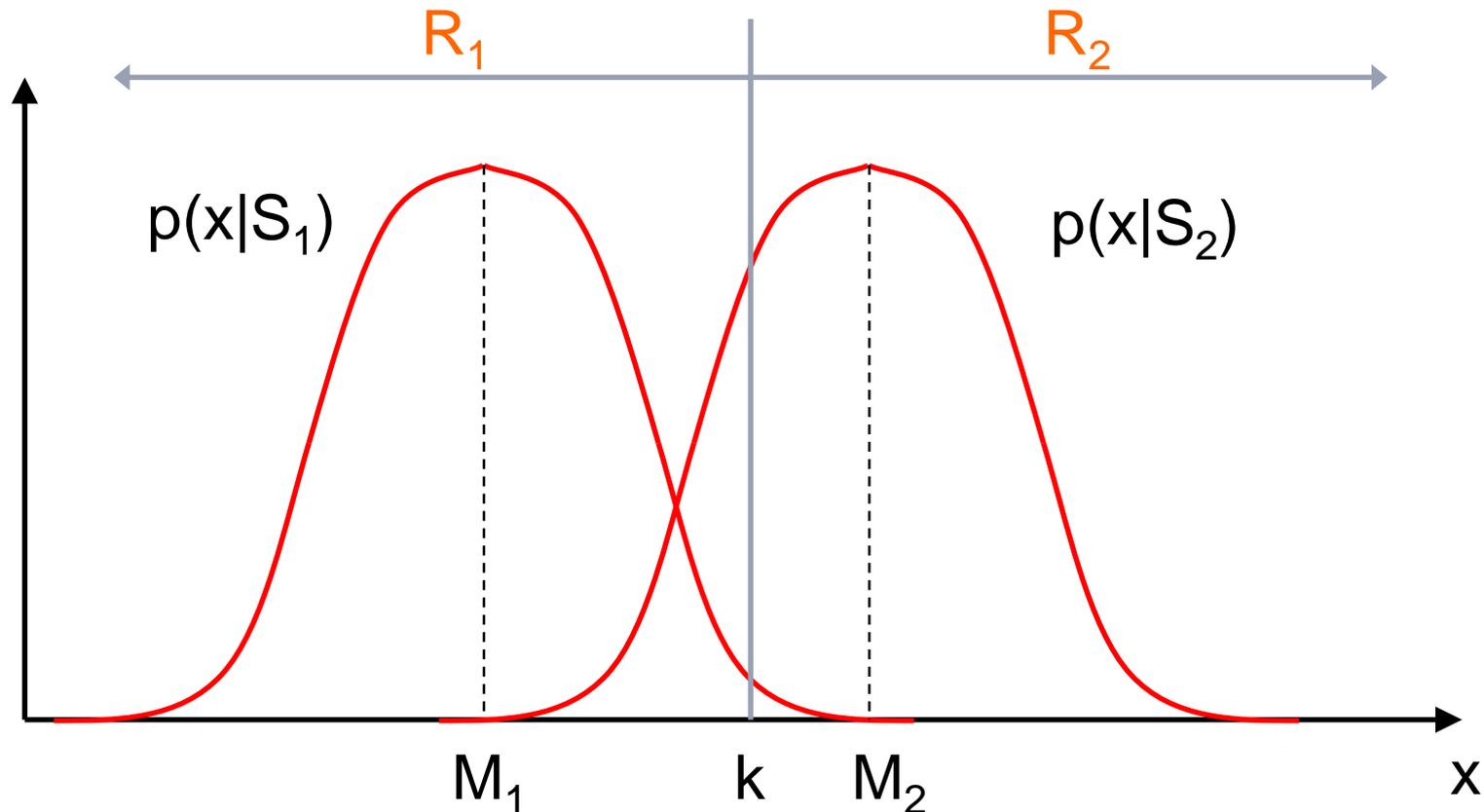
# Decision Model for One Stimulus

---

- ▶ **S**: Stimulus
- ▶ **X**: Random variable for perceived magnitude
- ▶  $p(x|S)$ : Conditional probability density function for perceived magnitude of  $x$  to occur as a response of stimulus **S**
- ▶ **M**: Mean of  $p(x|S)$  (Perceptual) Decision Space Model



# Decision Model for Two Stimuli



$x < k \rightarrow$  Subject response =  $R_1$

$x \geq k \rightarrow$  Subject response =  $R_2$

# Sensitivity Index and Response Bias

---

## ▶ Sensitivity Index $d'$

$$d' = \frac{M_2 - M_1}{\sigma}$$

- ▶ A measure of discriminability
- ▶ Normalized distance between the two means
- ▶ Note that  $d'$  reflects pure perception performance, without the effect of the decision process represented by  $k$ .

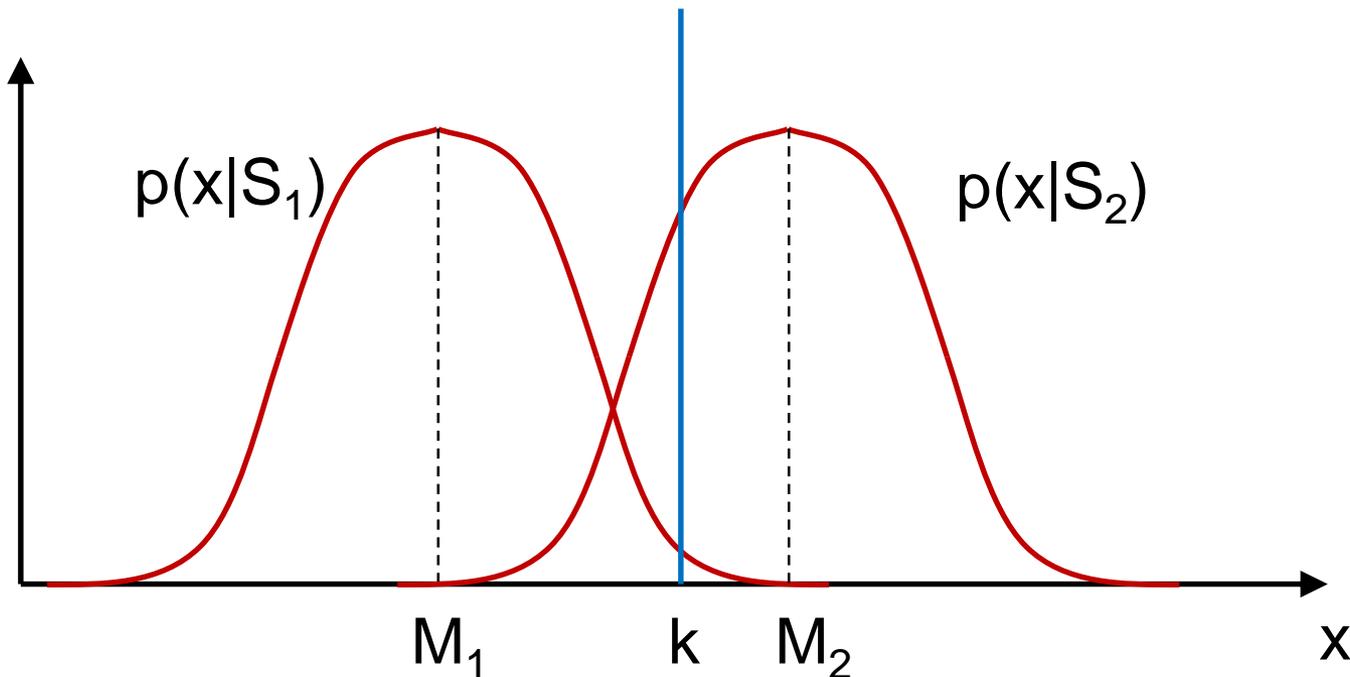
## ▶ Response Bias $c$

$$c = \frac{1}{\sigma} \left( k - \frac{M_1 + M_2}{2} \right)$$

- ▶ A representation of the decision rule
- ▶ Normalized distance between  $k$  and the average of the two means
- ▶ Response bias  $c$  can change independently from  $d'$ .

# Definitions of Four Probabilities

- ▶ Consider that  $S_1$  = signal absent and  $S_2$  = signal present.
- ▶ Hit Rate:  $P(R_2|S_2)$                       Miss Rate:  $P(R_1|S_2)$
- ▶ False Alarm Rate:  $P(R_2|S_1)$         Correct Rejection Rate:  $P(R_1|S_1)$



# Procedures for 1I 2AFC Experiment

---

- ▶ One interval, two alternatives, forced-choice
- ▶ The same framework for both detection and discrimination
- ▶ Prepare two stimuli: S1 and S2
  - ▶ S1 =“quieter tone”, S2 =“louder tone”
  - ▶ S1 =“softer spring”, S2 =“harder spring”
  - ▶ S1 =“noise”, S2 =“signal embedded in noise”
- ▶ On each trial,  $S_i$  is presented with a probability of  $P(S_i)$ , where  $P(S1) + P(S2)=1$ .
- ▶ The subject is asked to feel the signal and to answer which signal s/he felt, i.e., only two answers (R1 and R2) are allowed.
- ▶ Trial-by-trial correct answer feedback is optional.

# Data Analysis: Confusion Matrix

---

- ▶ Compile the data of each subject in a 2x2 matrix.

	$R_1$	$R_2$
$S_1$	$n_{11}$	$n_{12}$
$S_2$	$n_{21}$	$n_{22}$

- ▶ Hit Rate:  $H$

$$H = \frac{n_{22}}{n_{21} + n_{22}}$$

- False Alarm Rate:  $F$

$$F = \frac{n_{12}}{n_{11} + n_{12}}$$

- ▶ From  $H$  and  $F$ , we can estimate  $d'$  and  $c$ .

# Examples of Confusion Matrix

---

- ▶ For each confusion matrix, think over their underlying probability distributions.

	$R_1$	$R_2$
$S_1$	48	2
$S_2$	1	49

	$R_1$	$R_2$
$S_1$	5	45
$S_2$	1	49

	$R_1$	$R_2$
$S_1$	2	48
$S_2$	49	1

# Data Analysis: Computing $d'$ and $c$ (1)

---

- ▶ Compute  $H$  and  $F$  from a confusion matrix.
- ▶ Calculate  $z(H)$  and  $z(F)$ .
- ▶ Then,

$$d' = z(H) - z(F) \qquad c = -\frac{z(H) + z(F)}{2}$$

- ▶ Proof:

$H = P(X \geq k | S_2) = P(X < 2M_2 - k | S_2)$  by symmetry.

Use a transform of  $Z = \frac{X - M_2}{\sigma}$ . Then,

$$H = P\left(Z < \frac{M_2 - k}{\sigma} \mid S_2\right).$$

## Data Analysis: Computing $d'$ and $c$ (2)

---

Then,  $z(H) = \frac{M_2 - k}{\sigma}$  and similarly,  $z(F) = \frac{M_1 - k}{\sigma}$ .

$M_1 = \sigma z(F) + k$  and  $M_2 = \sigma z(H) + k$ .

Therefore,

$$d' = \frac{M_2 - M_1}{\sigma} = z(H) - z(F)$$

$$c = \frac{1}{\sigma} \left( k - \frac{M_1 + M_2}{2} \right) = -\frac{z(H) + z(F)}{2}.$$

- ▶ It follows that  $d'$  and  $c$  can be estimated from experimentally obtained hit and false alarm rates.

# Response Bias Revisited

---

- ▶ Response bias ( $c$  thus  $k$ ) can be systematically controlled by changing some factors, such as:
  - ▶ Probability of presenting S1 (or S2)
    - ▶ Increasing  $P(S2)$  in a detection experiment increases the subject's expectation of signal appearance, thus decreasing  $k$ .
  - ▶ Payoff for hits and misses
    - ▶ Rewarding hits and penalizing misses encourage the subject to answer "yes", thereby decreasing  $k$ .
- ▶ This problem was a primary motivation for the birth of the decision model.

# More Topics on Decision Theory

---

- ▶ ROC (Receiver Operating Characteristics)
- ▶ Relation between  $d'$  and a perceptual threshold
- ▶ Decision theory for multiple alternatives (stimuli)

# Information Theory

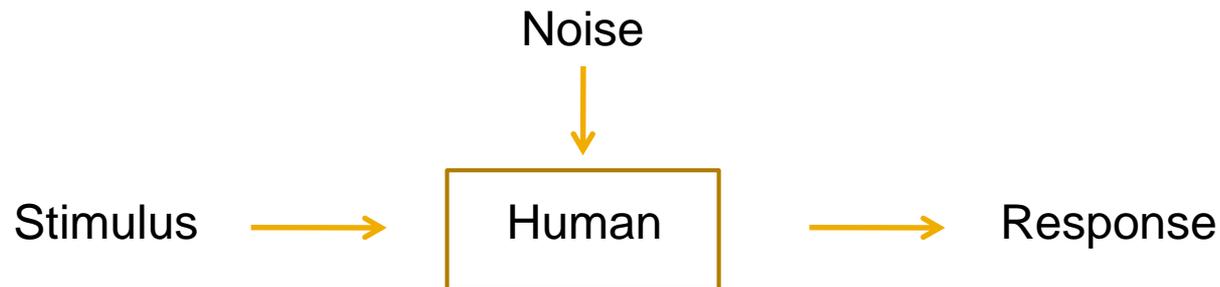
---

- ▶ Proposed by Claude E. Shannon (the father of information theory; 1916-2001)
- ▶ Can we count the amount of information?
- ▶ Foundations for all modern communication theories
- ▶ What is communication?
  - ▶ Phones
  - ▶ Internet
  - ▶ Speech
  - ▶ Perception
  - ▶ Any form of information flow can be regarded as communication

# Information Theory Applied to Perception

---

- ▶ Motivations
  - ▶ Peripheral to central limitations
  - ▶ Miller's "magic number  $7 \pm 2$ "
  - ▶ Humans as noisy communication channels
- ▶ Advantages
  - ▶ Allows to quantify our perceptual capacity
  - ▶ Context-free measure of information flow





# Definition of Uncertainty

---

- ▶ Uncertainty of outcomes with uniform probability distribution

$$U = \log_2 k$$

- ▶ k: # of alternatives

- ▶ Given a random variable X, the uncertainty of a given outcome  $X_i$  is

$$U_i = \log_2 \frac{1}{P(X_i)} = -\log_2 P(X_i)$$

- ▶ The average uncertainty of a random variable X is

$$U_i = \sum P(X_i)U_i = -\sum P(X_i) \log_2 P(X_i) = -E[\log_2 P(X)]$$

- ▶ Unit: bits

# Absolute Identification (AI) Experiment

---

- ▶ One-interval experiment
- ▶ Stimuli:  $S_i$ ,  $i \in [1, k]$  ( $k > 2$ )
- ▶ Responses:  $R_j$ ,  $j \in [1, k]$
- ▶ One-to-one mapping ( $S_i \Leftrightarrow R_i$ )
- ▶ On each trial, one of the stimuli  $S_i$  is presented with an a priori probability of  $P(S_i)$
- ▶ Subject makes a response with  $R_j$
- ▶ Trial-by-trial correct-answer feedback is optional

# S-R Confusion Matrix (e.g., $k=5$ )

	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	
$S_1$	14	3	2	0	1	20
$S_2$	0	13	2	3	1	19
$S_3$	4	3	11	1	0	19
$S_4$	2	0	2	15	1	20
$S_5$	5	3	2	0	12	22
	25	22	19	19	15	100

*# of times  $S_2$  was Presented.*

*# of times the joint event ( $S_3, R_4$ ) occurred.*

*# of times  $R_5$  was called.*

© Hong Z. Tan & Zygmunt Pizlo

# IS and IR

---

- ▶ IS (Information in Stimulus)

- ▶ IS is the average uncertainty in stimulus

$$IS = -\sum_{i=1}^k P(S_i) \log_2 P(S_i)$$

- ▶ If all stimuli are equally likely, then

$$IS = \log_2 k$$

- ▶ IR (Information in Response)

- ▶ IR is the average uncertainty in response

$$IR = -\sum_{j=1}^k P(R_j) \log_2 P(R_j)$$

# IT (Information Transfer)

---

- ▶ Also called “mutual information”
- ▶ IT = reduction in uncertainty
- ▶ For a particular ( $S_i, R_j$ ) pair:
  - ▶ U( $S_i$ ) before:  $-\log_2 P(S_i)$ 
    - ▶ Assuming that  $P(S_i)$  is constant throughout the exp.
  - ▶ U( $S_i$ ) after:  $-\log_2 P(S_i | R_j)$

$$IT(S_i, R_j) = -\log_2 P(S_i) - [-\log_2 P(S_i | R_j)]$$

$$IT(S_i, R_j) = \log_2 \frac{P(S_i | R_j)}{P(S_i)}$$

- ▶ Average IT =  $\sum \sum P(S_i, R_j) IT(S_i, R_j)$
- ▶ IT: the degree of correlation between S's and R's

# Randomization with Replacement

---

- ▶ Imagine that you have  $k$  containers for the  $k$  stimulus alternatives
- ▶ The  $i$ -th container has a fixed number of copies ( $n_i$ , proportional to  $P(S_i)$ ) of the  $i$ -th stimulus
- ▶ On each trial, one of the  $\sum n_i$  ( $i=1, \dots, k$ ) stimuli is selected to be presented to the subject
- ▶ That stimulus is immediately replaced in its corresponding container
- ▶ Then, the a priori probability for  $S_i$  ( $i=1, \dots, k$ ) remains the same for all trials
- ▶ The stimulus uncertainty remains the same on all trials

$$IS = -\sum_{i=1}^k P(S_i) \log_2 P(S_i)$$

# Randomization without Replacement

---

- ▶ Imagine that you have  $k$  containers for the  $k$  stimulus alternatives
- ▶ The  $i$ -th container has a fixed number of copies ( $n_i$ , proportional to  $P(S_i)$ ) of the  $i$ -th stimulus
- ▶ On each trial, one of the  $\sum n_i$  ( $i=1, \dots, k$ ) stimuli was selected to be presented to the subject
- ▶ That stimulus is NOT replaced in its corresponding container
- ▶ Then, the a priori probability for  $S_i$  may change from trial to trial
- ▶ The stimulus uncertainty  $IS$  may change from trial to trial
- ▶ On the last trial, the subject knows exactly what stimulus to expect (whichever stimulus is the last one left in a container)

# More on Randomization

---

- ▶ We prefer the method of “randomization with replacement” because
  - ▶ It ensures constant IS for each trial
  - ▶ It makes data analysis easier
- ▶ With the method of “randomization with replacement,” equal a priori probability no longer guarantees equal number of occurrences for all stimulus alternatives.
- ▶ Note that frequency of occurrence  $\neq$  probability
  
- ▶ The advantage of “randomization without replacement” is that the experimenter controls the exact number of times each stimulus alternatives is presented.

# Example S-R Confusion Matrix for Randomization with Replacement

---

	$R_1$	$R_2$	$R_3$	$R_4$	$R_5$	
$S_1$	14	3	2	0	1	20
$S_2$	0	13	2	3	1	19
$S_3$	4	3	11	1	0	19
$S_4$	2	0	2	15	1	20
$S_5$	5	3	2	0	12	22
	25	22	19	19	15	100

# Estimation of IT – ITest

---

- ▶ Average information transfer:

$$IT = \sum_{j=1}^k \sum_{i=1}^k P(S_i, R_j) \log_2 \frac{P(S_i | R_j)}{P(S_i)}$$

- ▶ Its maximum-likelihood estimate:

$$IT_{est} = \sum_{j=1}^k \sum_{i=1}^k \left( \frac{n_{ij}}{n} \right) \log_2 \left( \frac{n_{ij} \cdot n}{n_i \cdot n_j} \right)$$

**where**

$$\begin{aligned} n_{ij} &= \sum_{j=1}^k n_{ij} & n_j &= \sum_{i=1}^k n_{ij} \\ n &= \sum_{j=1}^k \sum_{i=1}^k n_{ij} = \sum_{i=1}^k n_i = \sum_{j=1}^k n_j \end{aligned}$$

- ▶ Interpretation of  $2^{IT}$  or  $2^{IT_{est}}$  (compare with  $k=2^U$ )

# Percent-correct scores and ITest

$$IT_{est} = \sum_{j=1}^k \sum_{i=1}^k \left(\frac{n_{ij}}{n}\right) \log_2 \left(\frac{n_{ij} \cdot n}{n_i \cdot n_j}\right)$$

(A)

25	25
25	25

50%  
0 bits

(B)

25	25	25	25
25	25	25	25
25	25	25	25
25	25	25	25

25%  
0 bits

(C)

25	0	0	0
0	25	0	0
0	0	25	0
0	0	0	25

100%  
2 bits

(D)

0	0	0	25
0	0	25	0
0	25	0	0
25	0	0	0

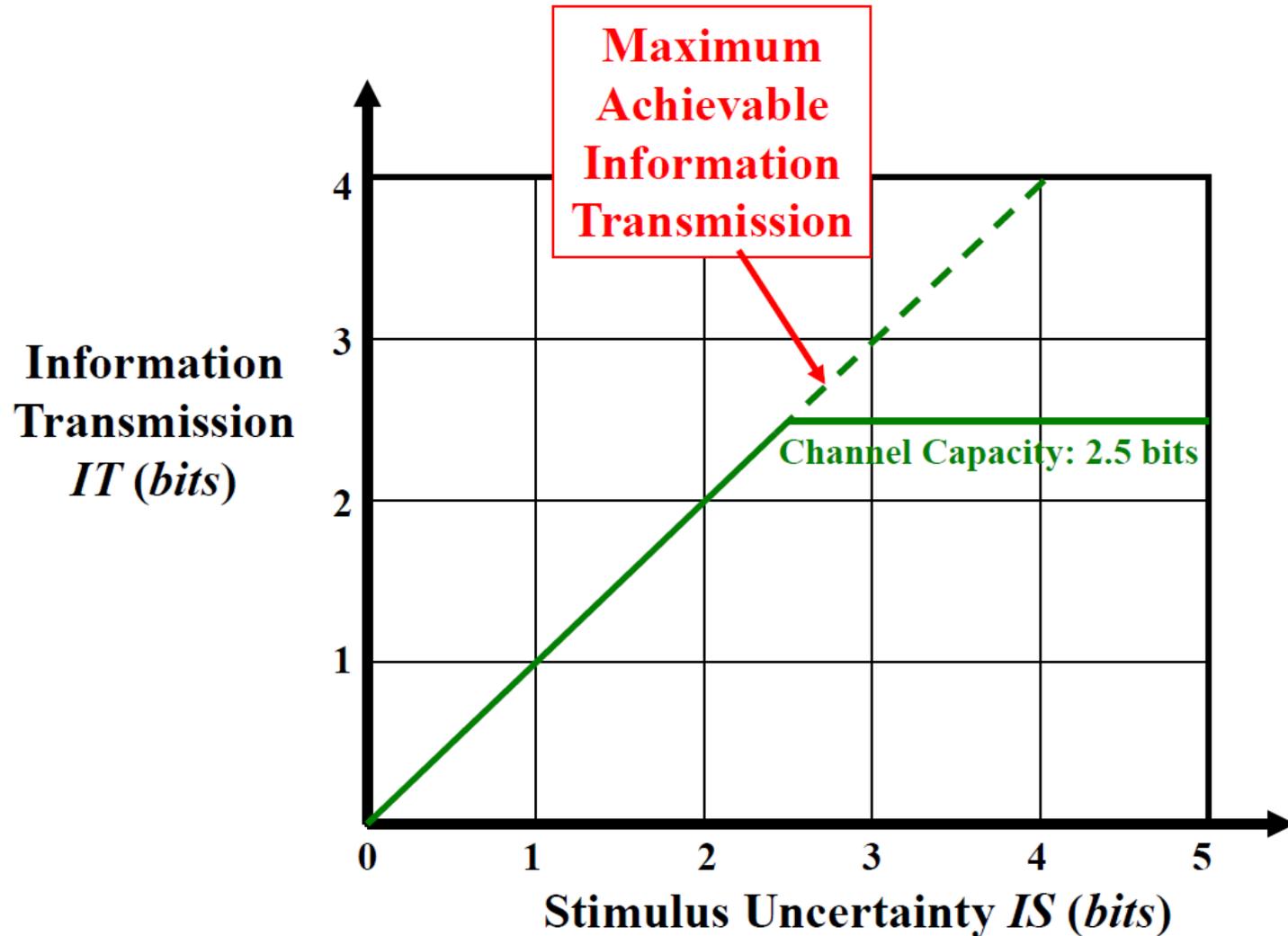
0%  
2 bits

# Maximum Information Transmission

---

- ▶ Mathematically,  $IT \leq IS$ .
- ▶ Intuitively, if the input and output are perfectly correlated, then  $IT = IS (= IR)$ .
- ▶ Assume that there exists a maximum information transmission
  - ▶ For small values of  $IS$ ,  $IT = IS$ .
  - ▶ As  $IS$  increases,  $IT = \text{constant}$  regardless of the value of  $IS$ .
- ▶ This maximum  $IT$  is accepted as the channel capacity.

# Channel Capacity – Example



# The Magic Number $7 \pm 2$

---

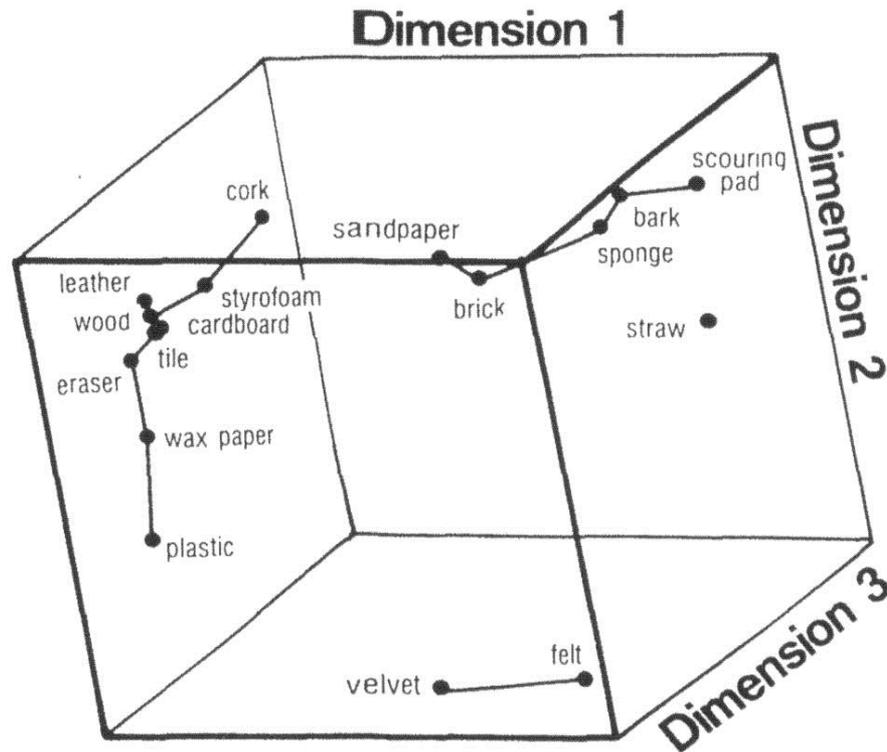
- ▶ The “magic number” is derived from an IT range of 2.3 –3.2 bits
- ▶ The “magic number” summarizes the typical channel capacity for uni-dimensional stimuli
- ▶ Uni-dimensional stimuli
  - ▶ Only one physical variables (target) is manipulated to form the stimulus set
  - ▶ Other physical variables (background) are either held constant or randomized
- ▶ How “Magic” is the Magic Number?
  - ▶ The “Magic Number” does NOT apply to
    - ▶ Absolute pitch (over-learnt stimuli)
    - ▶ Human face recognition (Multi-dimensional stimuli)

# Perceptual Space

---

- ▶ An abstract, mathematical space that visualizes the structure of perception
- ▶ Actively used in perception and HCI research
- ▶ Procedures of perceptual space estimation
  - ▶ Obtain dissimilarity data between stimuli
  - ▶ Apply Multi-Dimensional Scaling (MDS) to the data
  - ▶ Interpret the results of MDS based on the experimenter's intuition and additional analysis

# Perceptual Space – Haptic Texture



Perceptual Dimensions: Soft-Hard, Smooth-Rough, Slippery-Sticky

# Dissimilarity Measurement

---

- ▶ Prepare a large number of stimuli that vary in the physical parameters of your interest
- ▶ One of two common procedures for dissimilarity measurement.
- ▶ Direct estimation of dissimilarity
  - ▶ Participants provide direct estimations of pairwise dissimilarity between two stimuli
  - ▶ Provides interval-scale data
  - ▶ Useful for a relatively small number of stimuli
- ▶ Indirect estimation of dissimilarity from similarity
  - ▶ Participants sort similar stimuli into groups (clusters)
  - ▶ Similarity scores are computed and then transformed to dissimilarity scores
  - ▶ Provides data between ordinal and interval scales
  - ▶ Useful for a large number of stimuli

# Multi-Dimensional Scaling (MDS)

---

- ▶ A statistical method that estimates the dimension and positions of a set of points from the pairwise distances between the points.
- ▶ Direct problem
  - ▶ Given  $X = \{x_i\}$ , find  $d(x_i, x_j)$  where  $d()$  is a distance metric.
- ▶ Inverse problem
  - ▶ Given  $\{d(x_i, x_j)\}$ , find  $\{x_i\}$ .
- ▶ MDS solves the inverse problem iteratively
  - ▶ Increase the dimension of  $x_i$
  - ▶ Estimate the positions of  $\{x_i\}$
  - ▶ Evaluate the goodness of fit.
  - ▶ If the goodness of fit is not good enough, repeat from the first step.

# Classification of MDS

---

- ▶ By the type of dissimilarity data
  - ▶ Metric: Interval and ratio scale data
  - ▶ Non-metric: Nominal and ordinal scale data
- ▶ By the number of similarity matrices and the nature of MDS model
  - ▶ Classical MDS: One similarity matrix and unweighted model
  - ▶ Replicated MDS: Several similarity matrices and unweighted model
  - ▶ Weighted MDS (= INDSCAL, Individual Difference Scaling)  
Several similarity matrices and weighted model

# Classical Metric MDS

---

- ▶ **Input**
  - ▶ Input matrix is square and symmetric
  - ▶ Input data is metric
- ▶ **Model**
  - ▶ Transformation function  $f$  is assumed to be linear
- ▶ **Software tool**
  - ▶ Use CMDSCALE function in MATLAB
- ▶ **Advantage**
  - ▶ Relatively precise solution and little computation time
- ▶ **Limitations**
  - ▶ Only one symmetric matrix is allowed
  - ▶ Input data type restriction

# Classical Non-Metric MDS

---

- ▶ Input
  - ▶ Input matrix is square and symmetric
  - ▶ Input data is non-metric
- ▶ Model
  - ▶ Transformation function  $f$  is assumed to be monotonic
- ▶ Software tool
  - ▶ Use MDSCALE in MATLAB
  
- ▶ Advantage
  - ▶ High applicability
- ▶ Limitations
  - ▶ Only one symmetric matrix is allowed
  - ▶ A risk of suboptimal solution exists.

# INDSCAL MDS

---

- ▶ Input
  - ▶ Input matrix is square and symmetric
  - ▶ input data is metric
- ▶ Model
  - ▶ Transformation function  $f$  is assumed to be multiple linear functions for each subject
- ▶ Usage
  - ▶ Use ALSCAL MDS package in SPSS
- ▶ Advantage
  - ▶ The model accounts for systematic differences between subjects
  - ▶ Solution is unique, unrotatable -> often directly interpretable
- ▶ Limitations
  - ▶ Symmetry of the input matrices
  - ▶ Metric data

# Determining the Number of Perceptual Dimensions

---

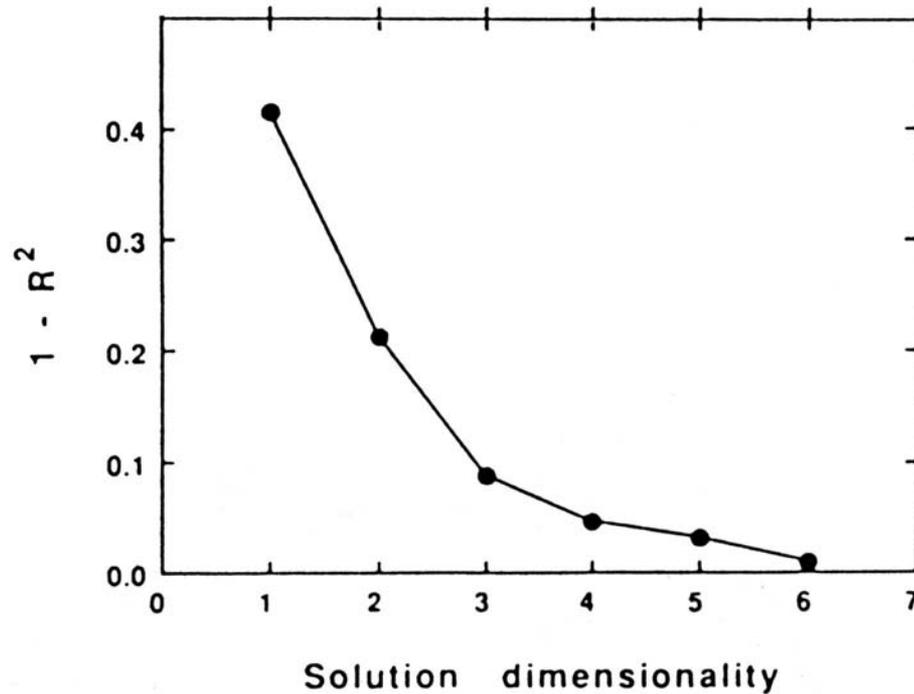
- ▶ Step 1: Perform MDS with # of dimensions = 1
- ▶ Step 2: Check out the goodness of fit using S-stress, stress, and  $1-R^2$
- ▶ Step 3: If the fitting is good, stop here. If no, go back to step 1 with # of dimensions increased by 1.

- ▶ Example of stress: Kruskal's stress
  - ▶  $d_{ij}$ : true distances between samples
  - ▶  $\delta_{ij}$ : estimated distances by MDS
  - ▶  $n$ : number of samples

$$S = \left[ \frac{\sum_{i=1}^n \sum_{j=1}^n (\delta_{ij} - d_{ij})^2}{\sum_{i=1}^n \sum_{j=1}^n d_{ij}^2} \right]^{\frac{1}{2}}$$

# Example of Goodness of Fitting

---



- ▶  $1 - R^2$  plot from Hollins et al.
- ▶ What value can use for the best number of dimensions?

# Case Study 1 – Direct Dissimilarity Estimation

---

## ▶ Reference

- ▶ I. Hwang and S. Choi, “Perceptual Space and Adjective Rating of Sinusoidal Vibration Perceived via Mobile Device,” In *Proceedings of the Haptics Symposium (HS)*, pp. 1-8, 2010

## ▶ Theme

- ▶ Perceptual space of vibration transmitted to the hand holding a mobile device

## ▶ Approach

- ▶ Obtain a perceptual space by direct dissimilarity rating followed by MDS
- ▶ Find a set of adjectives that can explain the percept distribution and map them to the perceptual space

# Case Study 1: Perceptual Space – Methods

## ▶ Stimulus Design

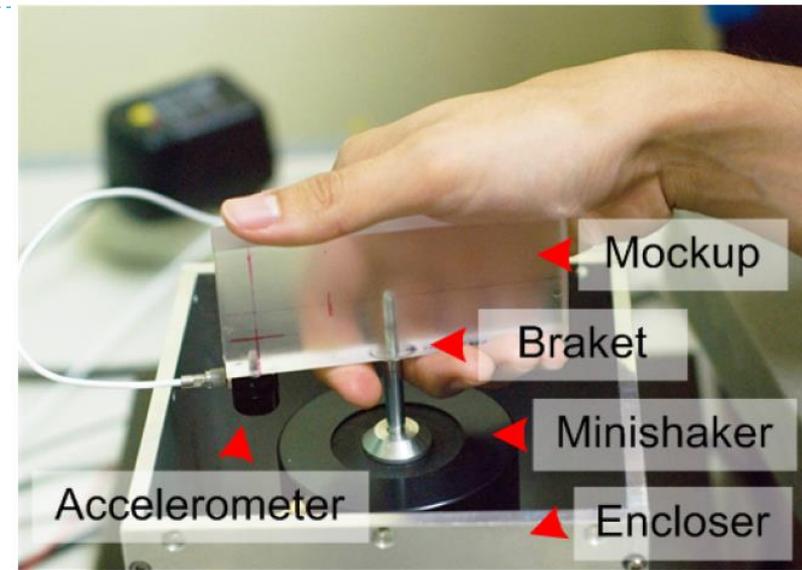
- ▶ Frequencies: 40, 80, 100, 120, 150, 200, and 250 Hz
- ▶ Amplitudes: 30 and 40 dB SL
- ▶ Four repetitions with 105 pairs
- ▶ Signal: 1s vibration – 1s rest – 1s vibration

## ▶ Participants

- ▶ Ten males (6 – right handed)
- ▶ Average 23 years old (20 – 28)

## ▶ Measurement

- ▶ Absolute magnitude estimation
- ▶ 0 ~ 100 ratings on each vibration stimulus (0: identical, 100: Completely Different)



<Experiment Setup>

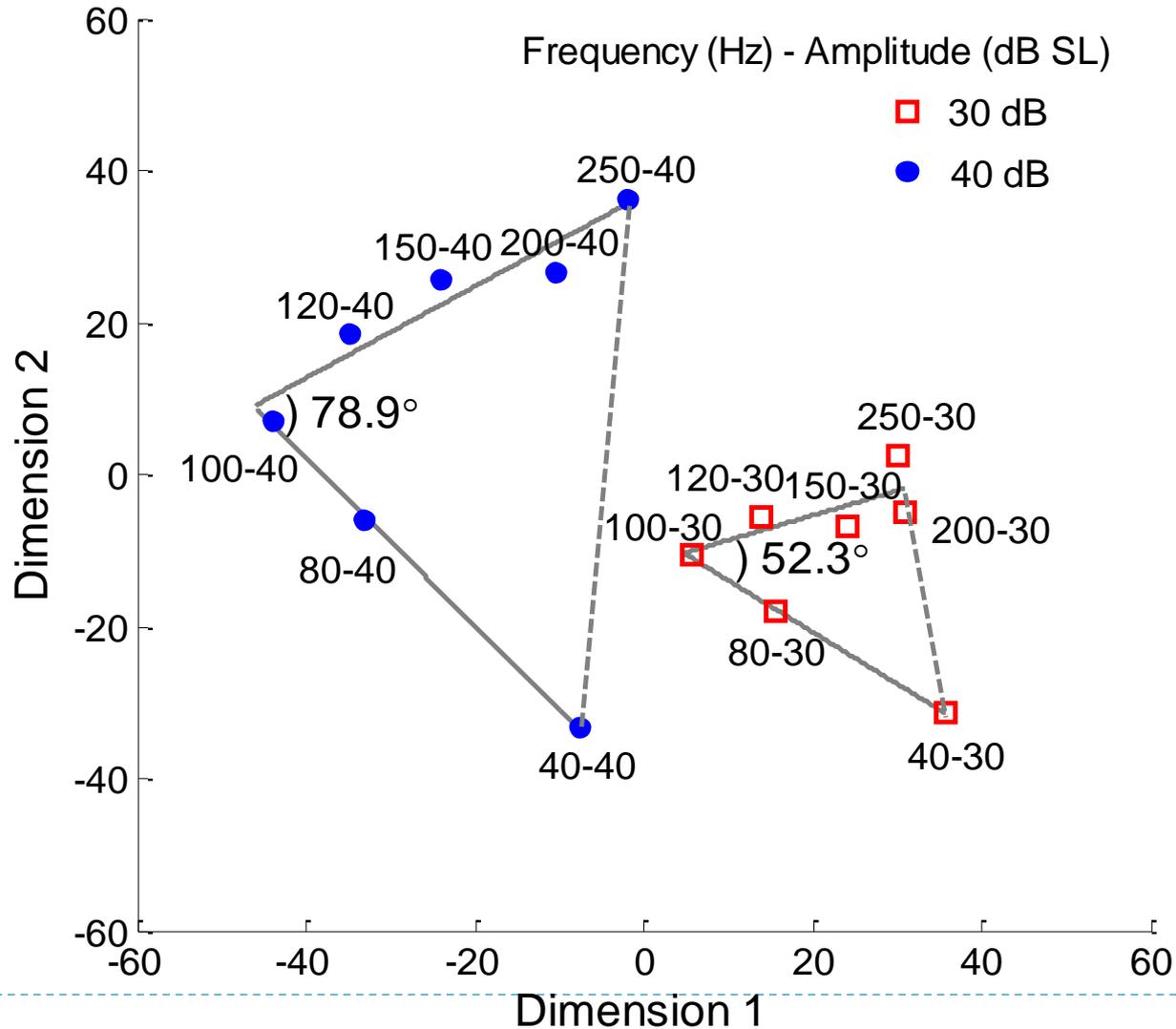
# Case Study 1: Perceptual Space – Dissimilarity Matrix

- ▶ Dissimilarity Matrix
  - ▶ One square, symmetric matrix

Hz-dB SL	40-40	80-30	80-40	100-30	100-40	120-30	120-40	150-30	150-40	200-30	200-40	250-30	250-40
40-30	46.3	31.0	69.4	38.9	87.3	33.1	84.1	34.7	77.0	29.0	68.7	31.0	77.4
40-40		32.8	39.0	35.2	61.2	41.0	63.4	44.6	57.0	43.4	56.0	42.5	59.2
80-30			55.0	21.4	63.3	16.6	60.0	18.3	57.6	20.1	47.9	27.6	54.4
80-40				46.7	17.4	46.4	22.8	56.2	37.7	61.8	39.4	58.9	50.7
100-30					57.0	14.7	49.9	21.8	49.0	29.1	47.5	27.9	47.8
100-40						53.6	23.0	71.2	26.1	70.0	41.2	73.8	46.0
120-30							52.3	14.8	53.6	17.9	45.9	20.8	50.9
120-40								64.8	22.7	68.5	30.1	66.3	34.9
150-30									56.9	18.0	51.4	16.1	54.0
150-40										62.7	20.3	58.9	30.5
200-30											51.7	13.1	51.8
200-40												53.3	21.7
250-30													47.2

- ▶ Goodness-of-fit was measured by Kruskal's stress
- ▶ Number of perceptual dimensions = 2

# Case Study 1: Perceptual Space – Result



# Case Study 1: Perceptual Space – Discussion

---

- ▶ Qualitative sensations of vibration [Tan, 1996]
  - ▶ Slow kinesthetic motion: 1 – 3 Hz
  - ▶ Rough motion or fluttering: 10 – 70 Hz
  - ▶ Smooth vibration: 100 – 300 Hz
- ▶ Difference in the perception channel
  - ▶ Most sensations are achieved from the PC channel when frequency  $> 100$  Hz.
  - ▶ When frequency  $< 100$  Hz, perception through the NP1 channel affects on the feeling of vibration.
  - ▶ At higher amplitudes, the NP1 channel is activated more with increased effects on the sensations.

# Case Study 1: Adjective Rating – Adjective Collection

## ▶ Stimulus Design

- ▶ Frequencies: 40, 80, 150, and 250 Hz
- ▶ Amplitudes: perceived magnitude level 11, approximately corresponds to 40 dB SL
- ▶ Duration: 1 s

## ▶ Participants

- ▶ Nine males, native Koreans

## ▶ Methods

- ▶ Adjectives were found on web pages and Korean dictionary
- ▶ Multiple-choice questionnaire and free answers

Pair No.	Adjective 1	Adjective 2
1	slow	fast
2	sparse	dense
3	blunt	sharp
4	bumpy	smooth
5	hard	soft
6	jagged	aligned
7	thick	thin
8	vague	distinct
9	heavy	light
10	deep	shallow
11	dark	bright
12	gentle	brisk
13	dull	clear

<Gathered Adjectives>

# Case Study 1: Adjective Rating – Methods

## ▶ Stimulus Design

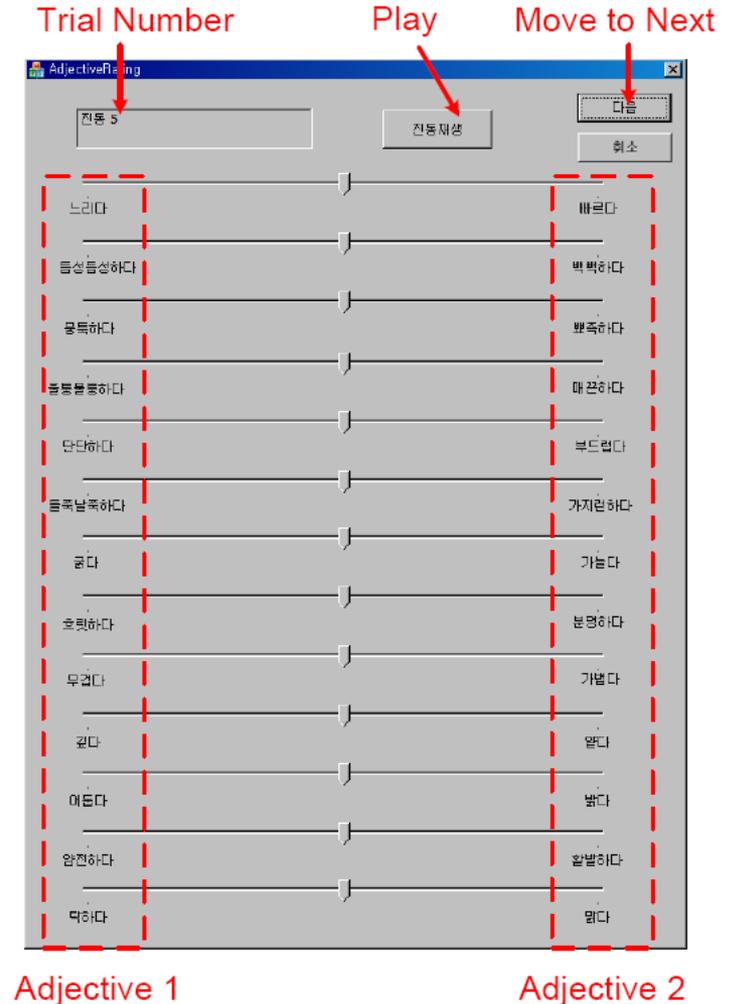
- ▶ Frequencies: 40, 80, 100, 120, 150, 200, and 250 Hz
- ▶ Amplitudes: perceived magnitude level 11, approximately corresponds to 40 dB SL
- ▶ Duration: 1 s

## ▶ Participants

- ▶ Seven males and three females, native Koreans

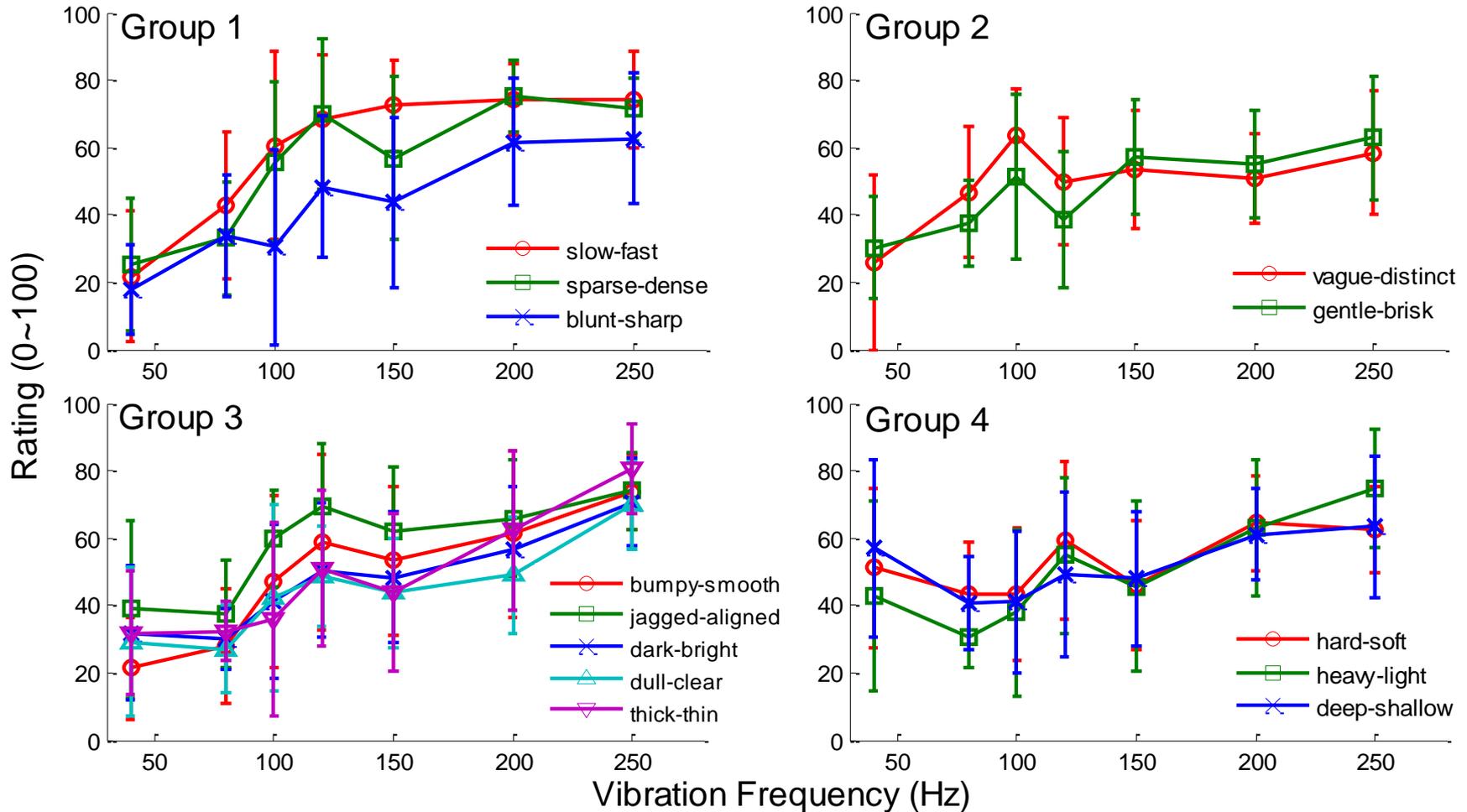
## ▶ Measurement

- ▶ Ratings on each adjective for each vibration stimulus



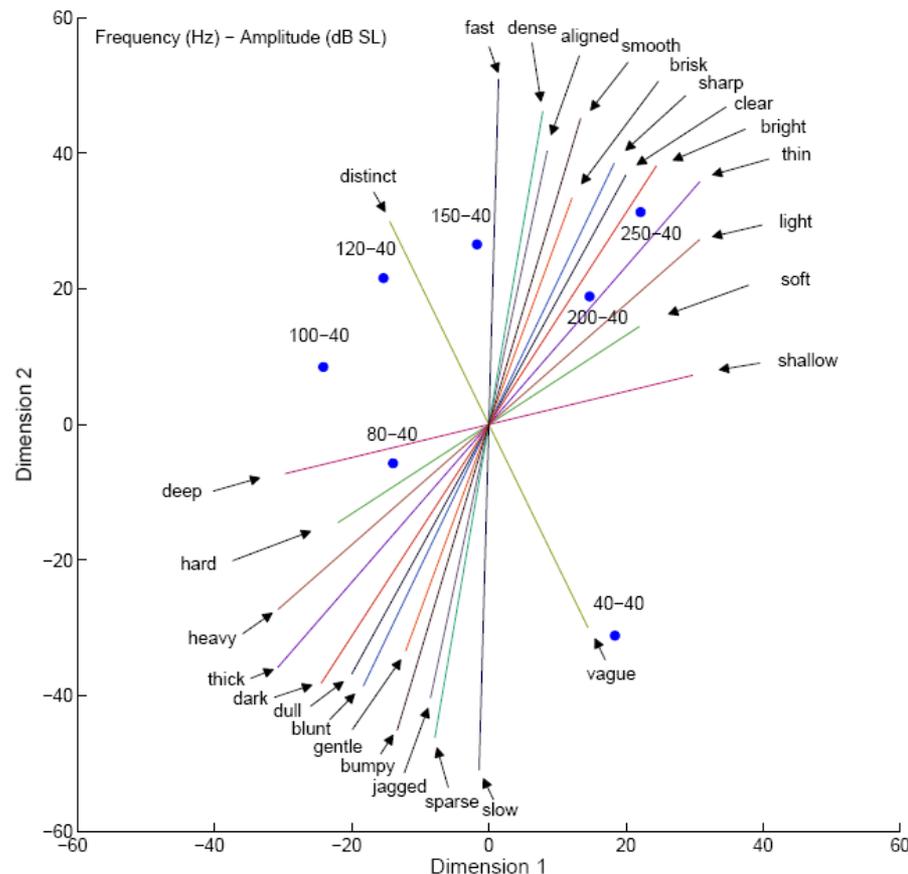
<GUI Used in Experiment>

# Case Study 1: Adjective Rating – Raw Data



# Case Study 1: Adjective Rating – Mapped into Perceptual Space

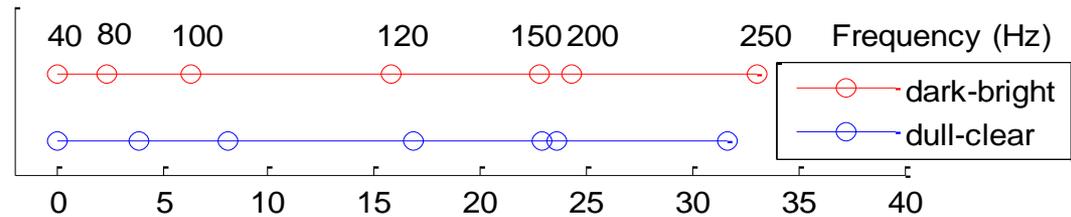
- ▶ The dissimilarity matrix for 40 dB SL amplitude was reprocessed by and, and the adjective pairs were projected into the perceptual space using multiple linear regression



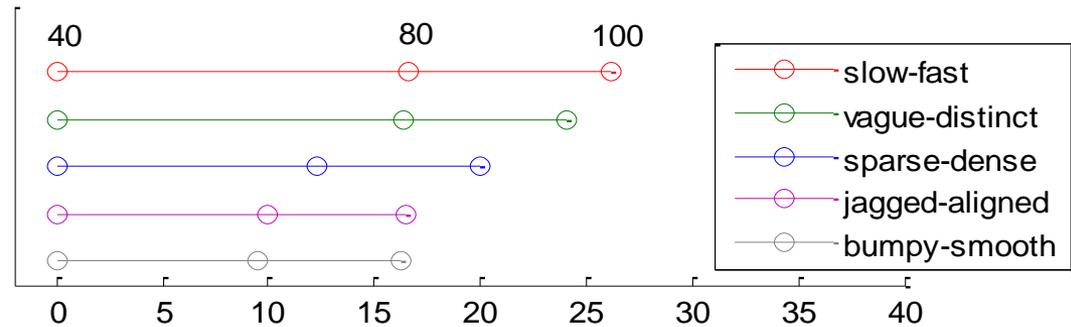
# Case Study 1: Adjective Rating – Good Adjectives

- ▶ Monotonic variations in a certain frequency band.

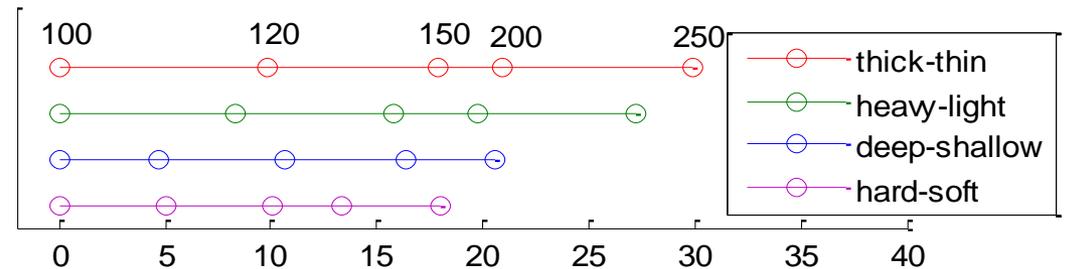
40 – 250 Hz



40 – 100 Hz



100 – 250 Hz



# Case Study 2: Cluster Sorting

---

## ▶ Reference

- ▶ Pasquero, J.; Luk, J. and Little, S. & Maclean, K., “Perceptual Analysis of Haptic Icons: An Investigation into the Validity of Cluster Sorted MDS,” In *Proceedings of the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, pp. 437-444, 2006.

## ▶ Theme

- ▶ Perceptual space of vibration generated by a lateral skin-stretch tactile pin array

## ▶ Approach

- ▶ Obtain a perceptual space by cluster sorting followed by MDS
- ▶ A nice set of discussions for the validity analysis of cluster sorting method

# Case Study 2: Cluster-Sorting Method

---

- ▶ Pair-wise Comparison Method
  - ▶ Long experiment time
  - ▶ 10 stimuli -> 45 comparisons, 20 stimuli -> 190 comparisons, ...
  - ▶ Lack of consistency
- ▶ Cluster-Sorting Method
  - ▶ To shorten the experiment time while preserves the characteristics of perceived similarity/dissimilarity
  - ▶ Lawrence M. Ward developed a cluster-sorting method (1977)
  - ▶ MacLean et al. applied similar techniques to haptic icon design (many papers)

# Case Study 2: UI for Cluster-Sorting Method

Haptic Icon MDS test - testing No Name

The interface displays five clusters of icons, each with a label and a list of numbers:

- Jerky**: 23
- fast**: 10, 22
- Slow**: 9, 28 (highlighted)
- Bad**: 13, 2
- Good**: 19
- Smooth**: 36

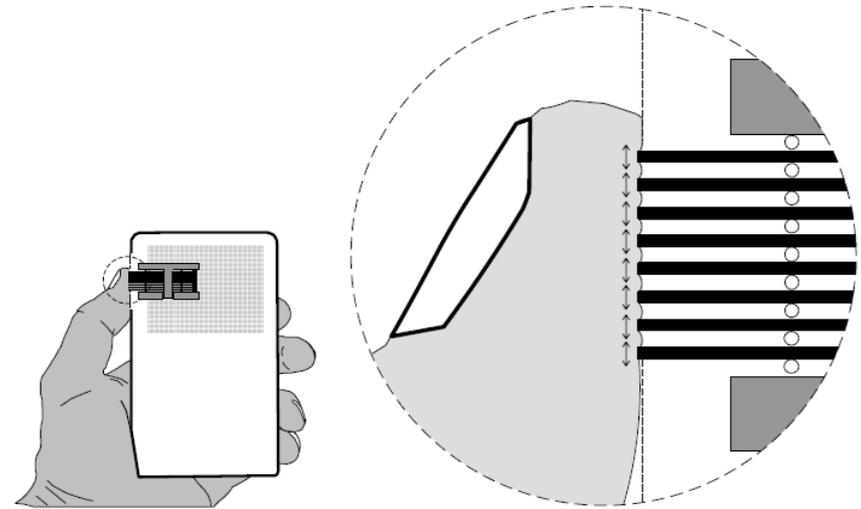
At the bottom, the 'Select Number of Boxes' control includes a minus sign (-) and a plus sign (+) button. To its right is a 3x10 grid of numbered boxes:

		33	8	1	30		31	12	11		14
24	20		32	21	29	15		17		34	5
3	6	18	26		7	35	27		4	16	25

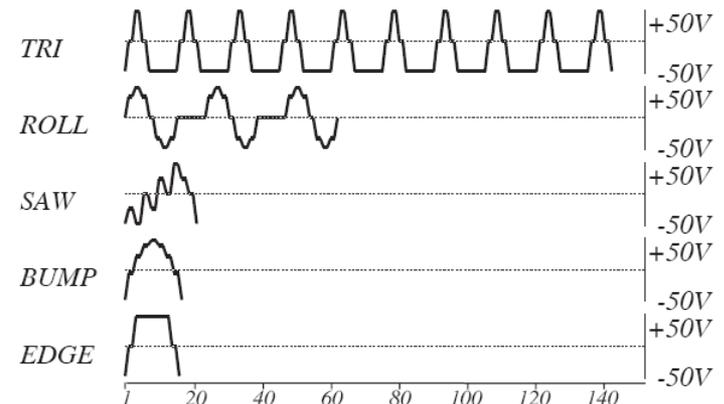
End This Sort. Continue to Next Step

# Case Study 2: Methods (1)

- ▶ Apparatus
  - ▶ Pin array for thumb
  - ▶ “Skin-stretch” type
  
- ▶ Stimulus Design
  - ▶ Waveform: TRI, ROLL, SAW, BUMP, and EDGE
  - ▶ Amplitude: full scale and half scale
  - ▶ Duration: differs for each stimulus
  
- ▶ Participants
  - ▶ Ten participants (seven males)



waveform (Fig. 3)	wavelength (in samples)	duration (ms) at speed		
		$\frac{1}{10}$	$\frac{1}{15}$	$\frac{1}{20}$
TRI	143	480	720	960
ROLL	62	221	331	442
SAW	20	86	130	173
BUMP	16	74	110	147
EDGE	16	74	110	147



# Case Study 2: Methods (2)

---

## ▶ Clusters

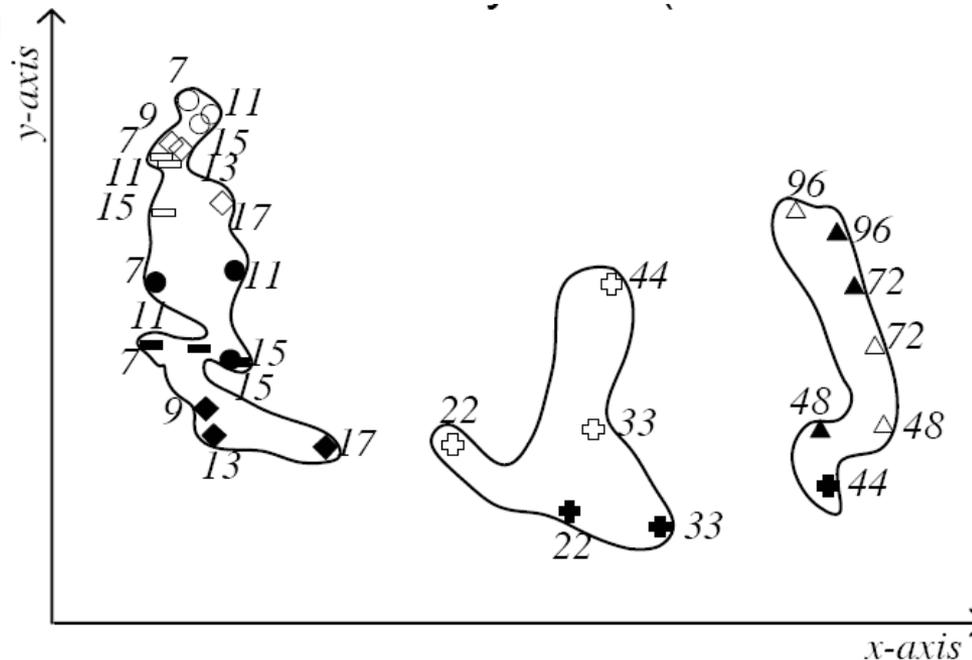
- ▶ Five clusters with 3, 6, 9, 12, and 15
- ▶ User select the first cluster size
- ▶ Each sorting box in clusters was assigned a name by four categories (duration, magnitude, multiplicity, and description)

## ▶ Similarity Computation

- ▶ If a pair of stimuli is sorted in a cluster with size  $n$ , this pair receives a similarity score  $n$ .
- ▶ The user-selected cluster score is adjusted to the closet number in  $\{3, 6, 9, 12, 15\}$
- ▶ The similarity score of this pair is accumulated over all the clusters.
- ▶ Then the score is linearly scaled to obtain a dissimilarity score between 0 (always grouped together) and 1000 (never grouped together)

# Case Study 2: Perceptual Space

- ▶ Dissimilarity matrix was analyzed by using 2D MDS analysis tool SPSS™ 13.0
- ▶ 3-dimensional MDS plot does not reveal any extra structural information



△ TRI    ⊕ ROLL    □ EDGE    ◇ SAW    ○ BUMP  
□ half amplitude    ■ full amplitude    number: duration (x10ms)

<Ordinal Untied MDS Plot from the Cluster-Sorting Experiment>

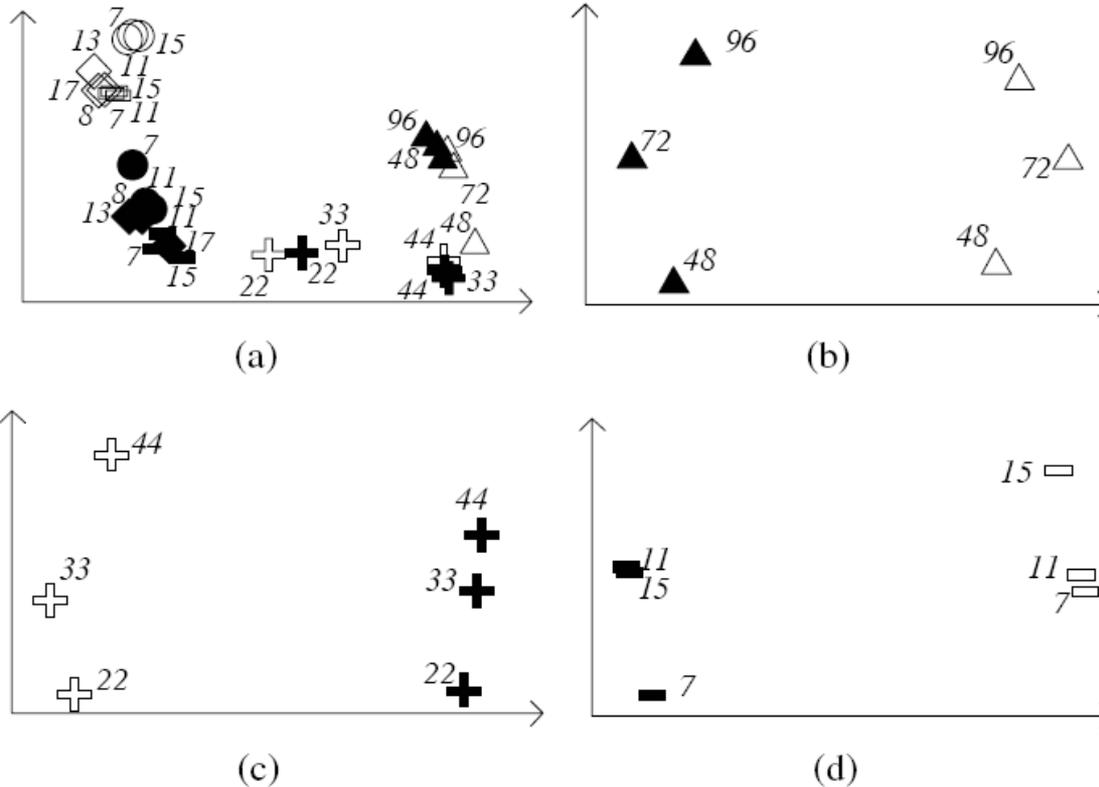
# Case Study 2: Discussion (1)

---

- ▶ Are the data obtained by cluster sorting metric or non-metric?
  - ▶ Non-metric MDS showed better results.
  - ▶ There were other evidence where metric MDS showed good results for cluster sorting.
  - ▶ Data from cluster sorting is clearly ordinal, but you need to careful to interpret them as interval data.
- ▶ The dissimilarity matrix obtained by cluster sorting is restricted.
  - ▶ In direct estimation of dissimilarity, each element in a dissimilarity matrix can be any number and independent from other elements.
  - ▶ In cluster sorting, however,
    - ▶ Each element obtained by cluster sorting can be dependent on other elements.
    - ▶ The resolution of dissimilarity scores is finite.
    - ▶ Therefore, the “power” of MDS is less than that of direct dissimilarity estimation, making it difficult to reveal “fine” differences.

# Case Study 2: Discussion (2)

- You often need additional MDS analysis using sub-dissimilarity matrices to see less dominant stimulus distinctions.



△ TRI    ⊕ ROLL    □ EDGE    ◇ SAW    ○ BUMP

□ half amplitude    ■ full amplitude    number: duration (x10ms)

# TASK PERFORMANCE

# Task Performance

---

- ▶ More application oriented
- ▶ Given a task, its performance is evaluated using several metrics
  - ▶ Task success rate
    - ▶ Accuracy
    - ▶ Error rate
    - ▶ Time on task
  - ▶ Task completion time
  - ▶ Efficiency
  - ▶ Learnability
- ▶ Widely used
- ▶ The values of performance metrics depend on the task context significantly.

# Issues in Designing Tactile Signals

---

## ▶ Discrimination

- ▶ Can we feel the differences between haptic signals?
- ▶ Pairwise discrimination and MDS are frequently used for discriminability evaluation.

## ▶ Identification

- ▶ Can we tell the identity, i.e., the name of a haptic signal?
- ▶ We can assess the identifiability of haptic signals, either with or without a learning phase.
  - ▶ Without a learning phase: Signal-identity table is given to participants. For example, signals can be presented graphically with their IDs. Do not enforce the participants to use their memory.
  - ▶ With a learning phase: In this case, the performance is affected by both signal identifiability and learnability of associated meanings.

# Learning Strategy (1)

---

- ▶ What is an effective way for users to learn stimulus-meaning associations?
- ▶ Designing an effective learning strategy is also important for abstract message delivery.
- ▶ Most research used neutral learning procedures to see the learnability between stimuli and meanings, not affected by the learning procedures.
- ▶ More sophisticated learning procedures, e.g., those using reinforcement learning (include rewards that are expected to improve desired behaviors) or multimodal signals (reminding the stimuli-meaning associations) can accelerate the learning, and can be very important in actual deployments.

# Learning Strategy (2)

---

- ▶ Evaluation of Learnability
  - ▶ Limit the participant's effort spent for training (e.g., time or number of repetitions). Then measure the accuracy of learning after the training.
  - ▶ Repeat training until the performance of stimuli-meaning association reaches a predefined level. Then, measure the effort (time or the number of repetitions) required to reach the goal performance.
  - ▶ Design of learning procedures is better to comply with the context of applications.

# Retention (Recall) Test

---

- ▶ It is also important how well users can retain the learned memory of stimuli-meaning associations.
- ▶ For this, we often need to conduct retention tests.
- ▶ Two standard retention tests
  - ▶ Immediate retention tests
    - ▶ Test right after learning is completed.
    - ▶ Other aids helping learning must be removed in retention tests. In this context, tests during or after learning to evaluate learning performance and immediate retention tests may need to be separate.
  - ▶ Delayed retention tests
    - ▶ Tests some time later after learning completion.
    - ▶ Depending on applications, the performance of delayed retention tests should be regarded as the ultimate performance.
    - ▶ An appropriate time gap between learning and delayed retention tests is contingent upon the context of applications.

# Field Test

---

- ▶ When we are satisfied with the performance of tactile display system, the last and most difficult step is to deploy the system to a larger population and observe the user responses for a long time.
- ▶ This step determines the true viability of your systems.
- ▶ However, field tests are difficult, in terms of money, time, and researchers' efforts, especially for university researchers, and are mostly neglected.
- ▶ Still very needed step for industrial adoption.

# Considerations for Participants

---

- ▶ In designing experiments for tactile display systems, we need to be very careful of selection of participants.
- ▶ Experiments mostly depending on haptic perception
  - ▶ A basic assumption is that our perceptual ability is more or less the same, except for a few factors such as age.
  - ▶ In limited cases, gender affects the perceptual data (e.g., the rate of growth in sensation magnitude of vibrotactile stimulus)
- ▶ Experiments including motor and cognitive factors
  - ▶ Related participant ability are individual-specific, thus participants need to be carefully controlled.
  - ▶ Factors to consider
    - ▶ Age, gender, sensory impairments, education (general and some specific skills), culture (e.g., applicable only in Korea?), experience (to similar tasks and experiments), and so on.
    - ▶ Every other participant factor that may affect the performance
  - ▶ Cases studies with a limited number of participants who are rare and difficult to recruit are often considered as valuable as well (e.g, with visually impaired participants).

# SUBJECTIVE EVALUATION

# Subjective Evaluation

---

- ▶ Asks the participant to rate a certain metric in numbers within a range
- ▶ Examples
  - ▶ Fun, Easiness to use, Difficulty to learn ...
- ▶ Mainly used as an accompanying evaluation in addition to quantitative assessments
  
- ▶ Mostly uses Likert scales

# ADVANCED TOPICS

# Mental Workload

---

- ▶ NASA TLX (Task Load Index)
  - ▶ A multidimensional scale consisting of 6 dimensions
  - ▶ Mental demand (MD): Related to perceptual activity such as thinking, deciding, calculating, remembering, looking, and searching
  - ▶ Physical demand (PD): Related to physical activity such as pushing, pulling, turning, controlling, and activating
  - ▶ Temporal demand (TD): Related to time pressure
  - ▶ Performance (OP; Own Performance): Related to personal goal accomplishment
  - ▶ Effort (EF): Related to energy expenditure in accomplishing the level of performance
  - ▶ Frustration (FR): Related to the feelings of insecurity, discouragement, irritation, stress, and annoyance

# Affection

---

- ▶ Cognitive Valence Theory
  - ▶ Theoretical framework that describes and explains the process of intimacy exchange within a dyad relationship
  - ▶ Uses subjective or biometric measurements

# REFERENCES

# References

---

- ▶ G. A. Gescheider, "Psychophysics: The Fundamentals," 3rd ed., Lawrence Erlbaum Associates, 1997.
  - ▶ Much more detailed topics can be found about psychophysics.
- ▶ T. D. Wickens, "Elementary Signal Detection Theory," Oxford University Press, 2002
  - ▶ Signal detection theory
- ▶ POSTECH EECE617 Introduction to Haptics Course notes, [http://hvr.postech.ac.kr/?page\\_id=773](http://hvr.postech.ac.kr/?page_id=773)