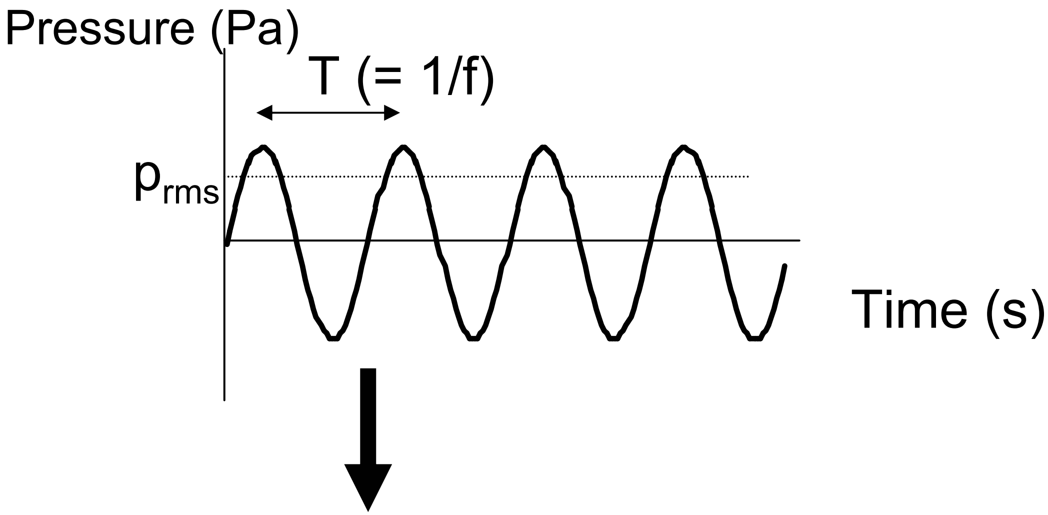
Frequency Selectivity and Masking

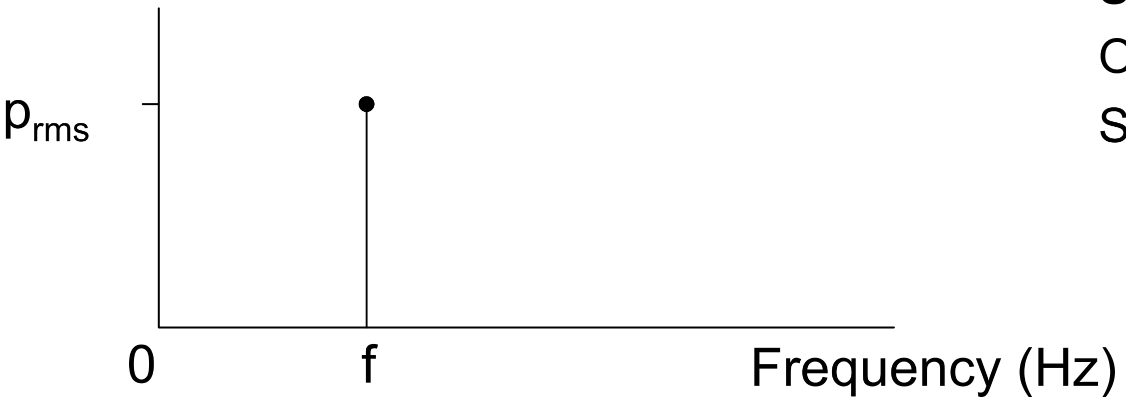
HST.723 Neural Coding and Perception of Sound

Measuring Sound

Time Domain



Frequency Domain



Sound Pressure Level

Measured in Pascals, relative to 2×10^{-5}Pa.

dB SPL = $20 \log10(p/p_0)$

0 dB SPL = 2×10^{-5}Pa

120 dB SPL = 20Pa

Spectrum level (dB/Hz)

Overall level =

Spectrum level + 10log(BW)

What is masking?

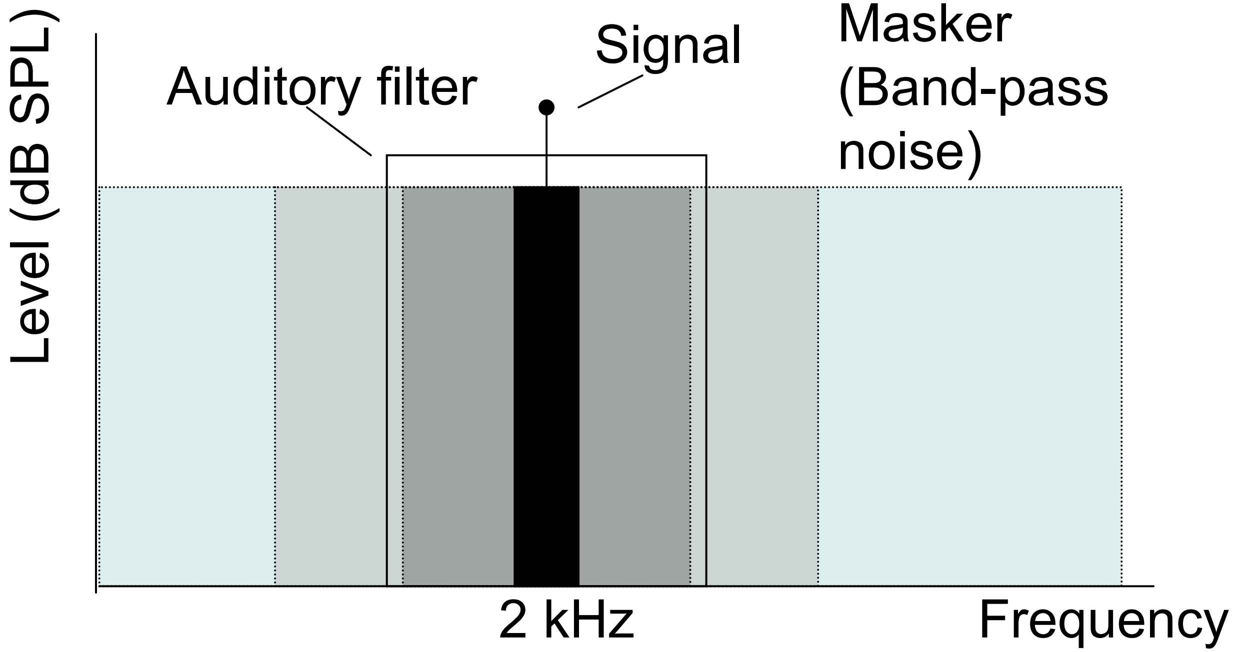
The process by which the threshold of audibility for one sound is raised by the presence of another (masking) sound.

(American Standards Association, 1960)

How can masking occur?

1) *Excitation:* Swamping of neural activity due to masker.
2) *Suppression:* Reduction of response to target due to presence of masker.

Critical bands in masking



Results suggest that we can “tune into” the region around 2 kHz, and that only masker energy around 2 kHz affects our ability to perceive the tone. The bandwidth of effective masking is the *critical band* (Fletcher, 1940).

Power spectrum model of masking

A signal is detected by an increase in power at the output of the auditory filter centered at the signal frequency:

$$P_s \quad K\int_0^{\infty=} W(f)\,N(f)\,\mathrm{d}f$$

where P_s is the power of the signal at threshold, W(f) is the filter shape, and N(f) is the masker’s power spectrum. K is the detector “efficiency”.

Assumptions:

- Filter is linear.
- Only one filter, centered at the signal frequency is used.
- Detection is based solely on overall power at filter output.

None of these assumptions are strictly true. However, they can often provide a reasonable first approximation.

Measuring frequency selectivity

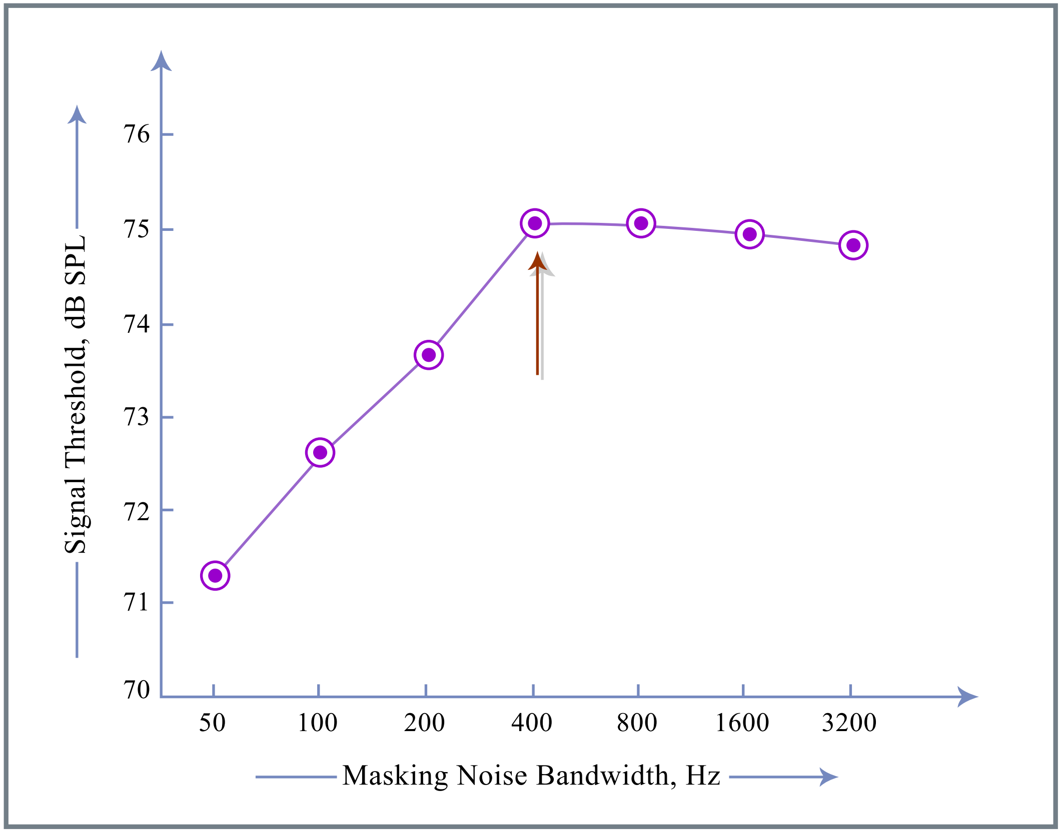


Figure by OCW. After Moore et al., 1993.

- The critical band is the point at which thresholds no longer increase.
- Conceptually very powerful, but not much use in providing an accurate estimate of filter bandwidth.
- Also, unable to discern filter *shape* from results.

Psychophysical tuning curves (PTCs)

Fixed signal; masker level adjusted to just mask signal.

Figure removed due to
copyright reasons.

From Moore (1997)

Advantages:
- Concept v. similar to neural tuning curves, allowing direct comparisons.

Potential problems:
- “Off-frequency listening”
- Detection of beats if using a sinusoidal masker

Notched-noise method

- Has similar advantages to Fletcher's band-widening method, but also enables a more accurate estimate of filter bandwidth and shape.

Figures removed due to copyright reasons.

From Moore (1986)

Does the notched-noise method reflect cochlear tuning?

- Filters are nonlinear
  - Change with level
  - Suppression effects may broaden apparent tuning
- Nevertheless, when these factors are accounted for, frequency selectivity does seem to match physiological measures of cochlear tuning quite well.

Figure removed due to copyright reasons.

See also Shera *et al.* (2002; PNAS) for human comparisons.

Frequency selectivity as a function of center frequency

Figure removed due to copyright reasons.

- *Absolute* bandwidths increase with increasing CF (important for harmonic resolvability)

- *Relative* bandwidths decrease or stay roughly constant.

Masking patterns and Excitation patterns

Figures removed due to copyright considerations. Please see:
Egan, J. P., and H. W. Hake. “On the masking pattern of a simple auditory stimulus.” *J Acoust Soc Am* 22 (1950): 622-630.

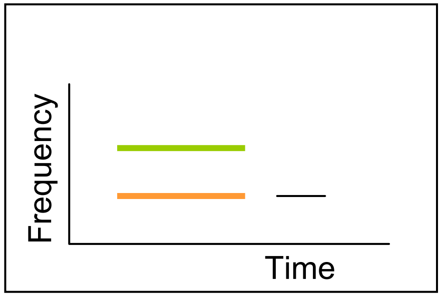
- Given auditory filter shapes, it is possible to derive masking patterns for any arbitrary stimulus.
- Under the power spectrum model assumptions, a masking pattern is equivalent to an *excitation pattern* – the internal representation of a sound’s spectrum.

But does masking pattern = excitation pattern?

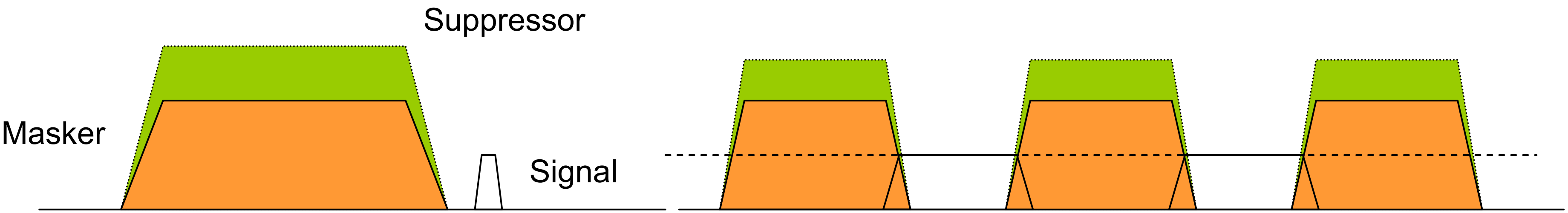
Suppression in hearing

Houtgast pioneered the search for evidence of “lateral suppression” in psychoacoustic tasks.

Forward masking



Pulsation threshold



Using these techniques it is possible to show “two-tone suppression”. This is not possible with simultaneous masking, as the suppressor suppresses both the masker and the signal, giving zero net effect.

Example of suppression data

Figure removed due to copyright considerations. Please see: Shannon, R. V. "Two-tone unmasking and suppression in a forward masking situation." *J Acoust Soc Am* 59 (1976): 1460-1470.

Effects of changing the suppressor frequency. Masker and probe are always at 1 kHz. (Shannon, 1976)

Masking Patterns vs. Excitation Patterns

According to the power spectrum model, masking patterns and excitation patterns are essentially the same thing. But this is not true if masking is in part due to suppression.

Figure removed due to copyright considerations. Please see: Oxenham, A. J., and C. J. Plack. “Suppression and the upward spread of masking.” *J Acoust Soc Am* 104 (1998): 3500-3510.

101 Uses for Excitation Patterns

- Loudness: Transformed area under the excitation pattern
  - Suggested by Fletcher, formalized by Zwicker, refined by Moore.

- Timbre: Centroid, or center-of-gravity of an excitation pattern

- Pitch: Positions of peaks within the excitation pattern or amplitudes

- Masking: Predicting the masking effectiveness of an arbitrary stimulus.
  - Used (with modifications) in audio coding, e.g., MP3.

Limitations of excitation pattern model

- Nonlinearities, such as suppression and distortion products, are not accounted for.

Can *over*estimate masking:

- Ignores temporal information (envelope or fine structure)
  - Beats
  - Effects of masker modulation
  - Detection of tones in roving-level narrowband noise

Can *under*estimate masking:

- Stimulus uncertainty (e.g., Neff and Green, 1987) can produce large amounts of “informational” masking without any energy around the signal frequency.

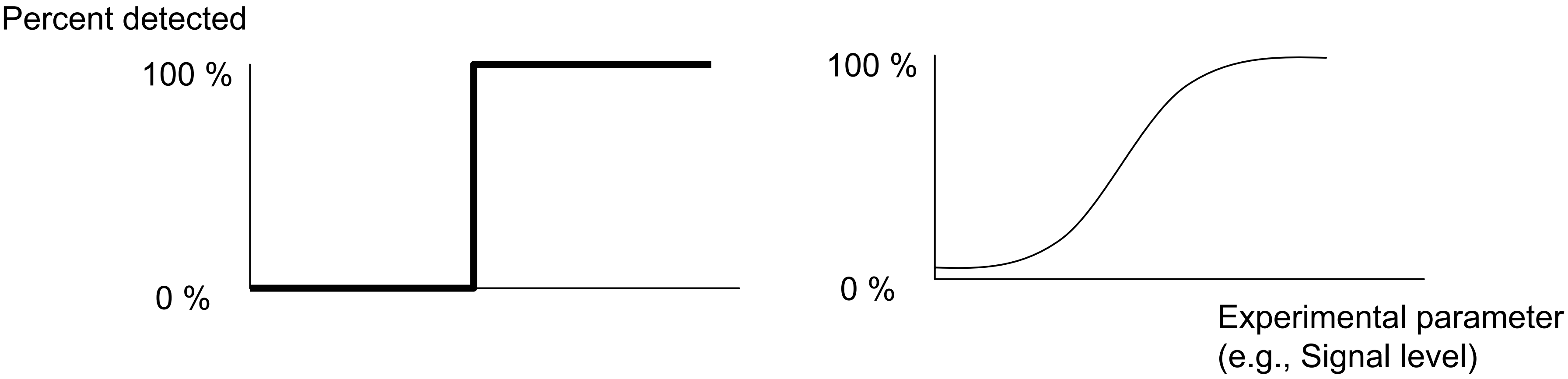
What is a Threshold?

A Brief Introduction to Signal Detection Theory

Historically two types of threshold:

- Absolute threshold: Minimum audible signal
- Differential threshold: Minimum perceptible change, aka difference limen (DL) or just noticeable difference (jnd).

The Psychometric Function

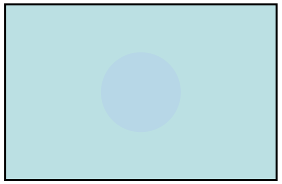


a) Original concept of “threshold”

b) Experimental data

One-interval, two-alternative paradigms (yes-no)

Question: Is there a signal present?
- Tone in a noise
- Aircraft on a radar screen
- Tumor on an x-ray image



Who is more sensitive?

	Observer 1			**Observer 2**		
	"Yes"	"No"	**P(C)**	"Yes"	"No"	**P(C)**
Signal present	90%	10%	**90%**	65%	35%	**65%**
Signal absent	90%	10%	**10%**	15%	85%	**85%**
Total percent correct			**50% (chance)**			**75%**

Hits and False Alarms

a) Signal detection

	"Yes"	"No"
Signal	**Hits (H)**	Misses
No signal	**False alarms (F)**	Correct rejections

b) General formulation

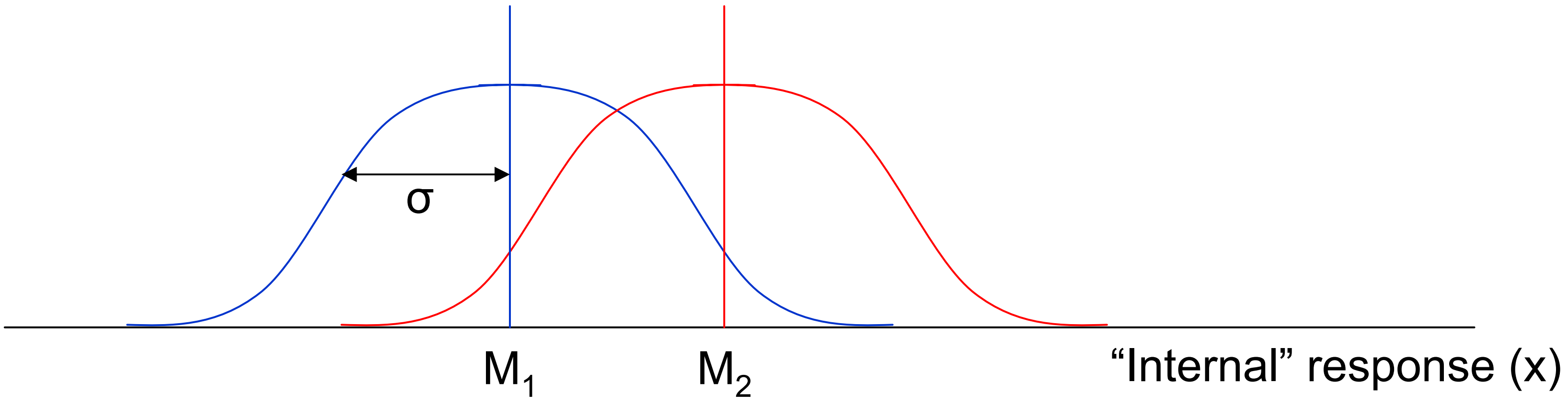
	R2	R1
S2	**R2|S2**	R1|S2
S1	**R2|S1**	R1|S1

A good measure of sensitivity must take into account both hits and false alarms:

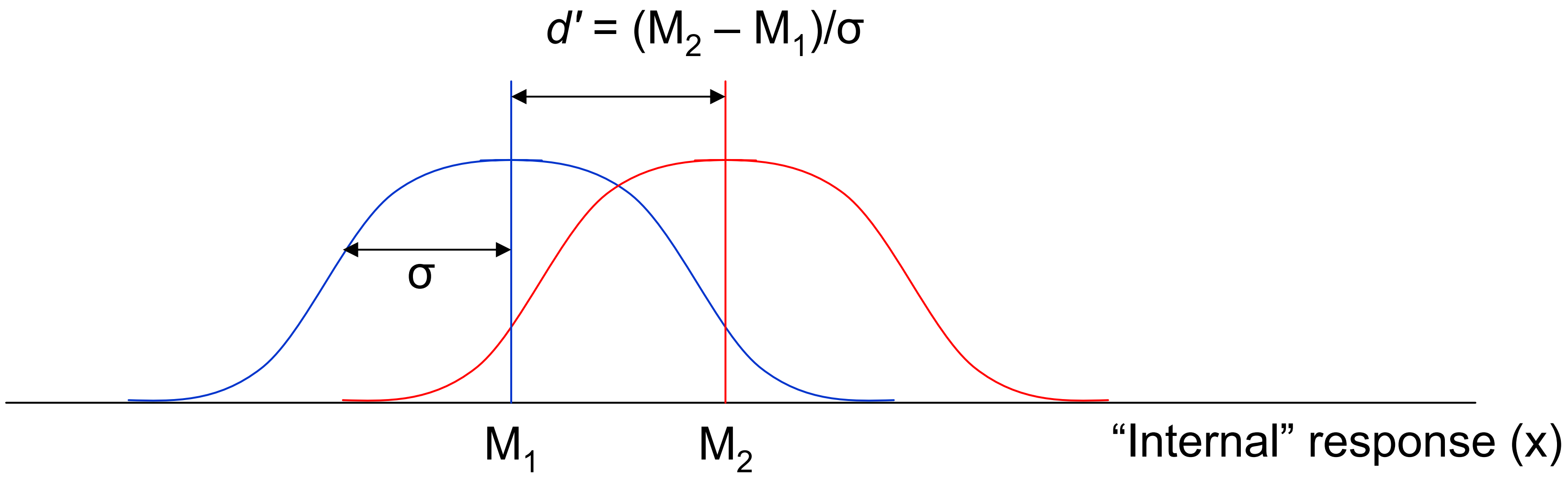
$$\mathit{Sensitivity} = v[u(H) - u(F)]$$

Signal Detection Theory

- The “internal response”, x, to a stimulus can be represented as a *random variable* [often assumed to have Gaussian (*normal*) distribution].
- So, two identical stimuli will *not* necessarily result in identical percepts.
- Detecting a signal (or discriminating between two stimuli) relies on deciding whether the percept arose from the distribution with mean M1 or the distribution with mean M2 (both have unit variance ($\sigma^2 = 1$).



Signal Detection Theory II



- The perceptual distance between M1 and M2, in units of standard deviations (σ), is called *d′*, pronounced “d-prime”.

$$d' = z(H) – z(F),$$

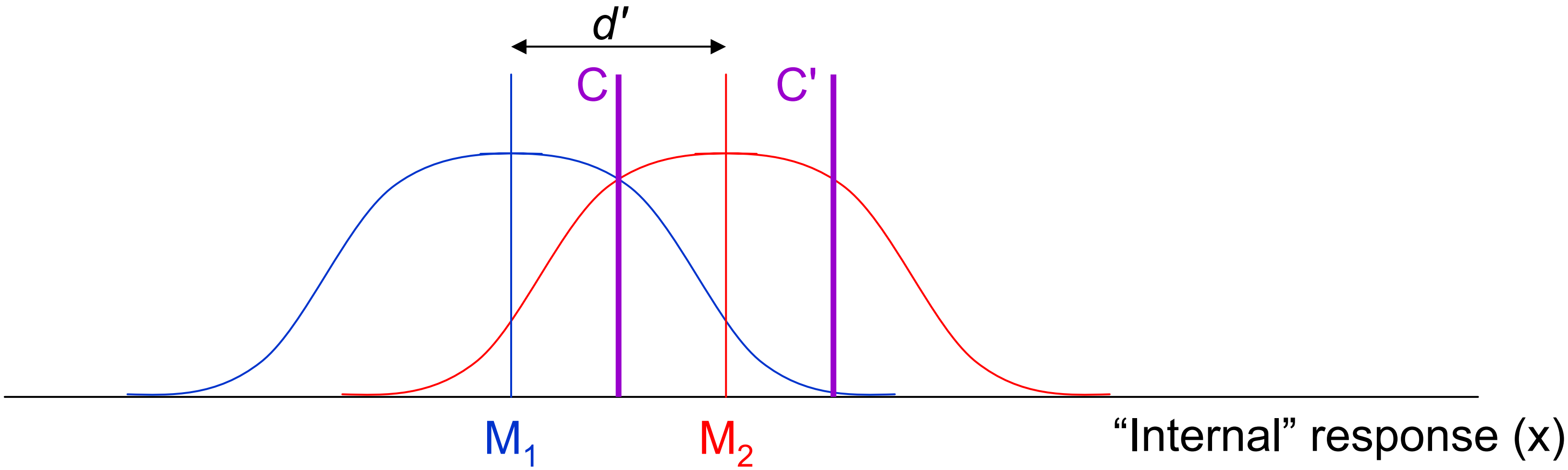
where z is the inverse of the normal distribution function.

- This implies there is no “threshold”.

Sensitivity and Bias

- The optimal rule is to set a criterion ‘C’:

Percept	*Response*
$x < C$	“No signal” (R1)
$x \geq C$	“Signal” (R2)



Where the criterion is set depends on:

- *a priori* probabilities of presentation
- Motivation and instructions (reward vs. punishment)

A change in the criterion (C) does *not* mean a change in sensitivity (*d'*).

Receiver Operating Characteristic (ROC)

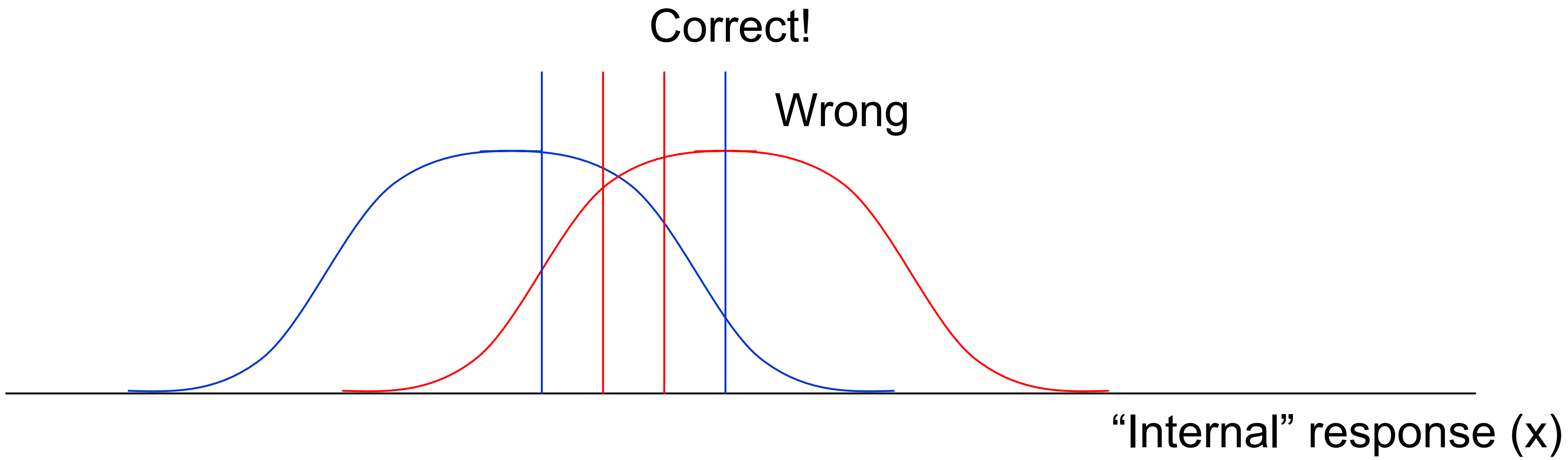
Plotting various combinations of Hit rates and False Alarm rates for a given sensitivity results in a Receiver Operating Characteristic (**ROC**).

Figure removed due to
copyright reasons.

m-interval, m-alternative forced-choice experiments

One way to reduce the effects of bias is to present both types of stimulus on each trial. Most popular is the 2-interval, 2-alternative forced-choice procedure (2AFC).

Each trial consists of either {S2 S1} or {S1 S2}, with an *a priori* probability of 0.5 for each. Subjects respond ‘1’ or ‘2’ after each trial, depending on which interval contained S2.



Note: No criterion is required – just a comparison of the two intervals.

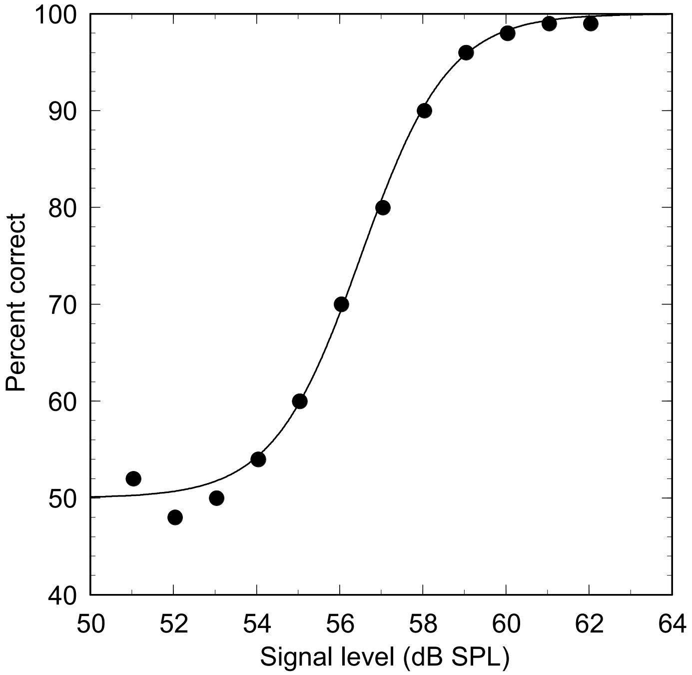
Forced-choice paradigms

Empirical results have generally shown only small biases in such experiments, meaning responses are generally symmetric. In this case, d’ can be directly calculated simply from percent correct.

A forced-choice paradigm does not rule out bias effects. Theoretically, it is preferable to record hits and false alarms. However, in practice most investigators only report percent correct.

Threshold estimation I

- *Method of constant stimuli.* Sometimes, it is necessary to find out how sensitivity (*d'*) changes as a function of a stimulus parameter (e.g., signal level). In this case, a psychometric function can be generated by repeated measurements at a number of fixed values.

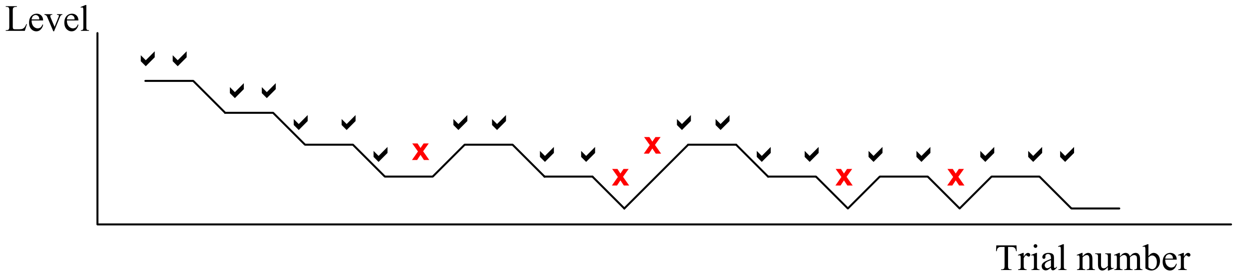


Threshold estimation II

- Adaptive procedures. **x-down y-up** adaptive procedures converge on a fixed level of performance. This allows more flexible and rapid measurement of performance.
- 2-down 1-up procedure:

X : ↑=

✓ ✓ : ↓=



Tracks the 70.7% correct point on the psychometric function [$p(\checkmark\ \checkmark) = p(\checkmark)^2 = 0.5$].

Further Reading in Signal Detection Theory

Wickens, T.D. (2001) “Elementary Signal Detection Theory,” (Oxford University Press).

Macmillan and Creelman (1991) “Detection Theory: A User’s Guide.” Cambridge Univ. Press. (Out of print)

Green and Swets (1966) “Signal Detection Theory and Psychophysics,” (Reprinted 1974 and 1989).

References

- Egan, J. P., and Hake, H. W. (**1950**). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. **22**, 622-630.
- Evans, E. F. (**2001**). "Latest comparisons between physiological and behavioural frequency selectivity," in Physiological and Psychophysical Bases of Auditory Function, edited by J. Breebaart, A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs and R. Schoonhoven (Shaker, Maastricht).
- Evans, E. F., Pratt, S. R., and Cooper, N. P. (**1989**). "Correspondence between behavioural and physiological frequency selectivity in the guinea pig," Brit. J. Audiol. **23**, 151-152.
- Glasberg, B. R., and Moore, B. C. J. (**1990**). "Derivation of auditory filter shapes from notched-noise data," Hear. Res. **47**, 103-138.
- Houtgast, T. (**1972**). "Psychophysical evidence for lateral inhibition in hearing," J. Acoust. Soc. Am. **51**, 1885-1894.
- Moore, B. C. J. (**1986**). Frequency Selectivity in Hearing (Academic, London).
- Neff, D. L., and Green, D. M. (**1987**). "Masking produced by spectral uncertainty with multi-component maskers," Perception and Psychophysics **41**, 409-415.
- Oxenham, A. J., and Plack, C. J. (**1998**). "Suppression and the upward spread of masking," J. Acoust. Soc. Am. **104**, 3500-3510.
- Shannon, R. V. (**1976**). "Two-tone unmasking and suppression in a forward masking situation," J. Acoust. Soc. Am. **59**, 1460-1470.
- Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proc. Natl. Acad. Sci. USA **99**, 3318-3323.