Lecture 1:Information Processing and Signal Detection

Lecture 01 Part 03

Signal Detection Theory

Relevance

Fan, S., Ng, T.T., Herberg, J.S., Koenig, B.L. and Xin, S., 2012. Real or fake? Human judgments about photographs and computer-generated images of faces. In *SIGGRAPH Asia 2012 technical briefs* (pp. 1-4).

Williams, J.J., Tien, R.N., Inoue, Y. and Schwartz, A.B., 2016, August. Idle state classification using spiking activity and local field potentials in a brain computer interface. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1572-1575). IEEE.

Ponce, P., Molina, A. and Grammatikou, D., 2016. Design based on fuzzy signal detection theory for a semi-autonomous assisting robot in children autism therapy. *Computers in Human Behavior*, *55*, pp.28-42.

Valderrama, A.T., Paclik, P., Vansteensel, M.J., Aarnoutse, E.J. and Ramsey, N.F., 2012. Error probability of intracranial brain computer interfaces under non-task elicited brain states. *Clinical neurophysiology*, *123*(12), pp.2392-2401.

Learning Objectives

To extend the understanding of the key features of a psychometric function.

To develop an understanding of how individual differences can affect task performance.

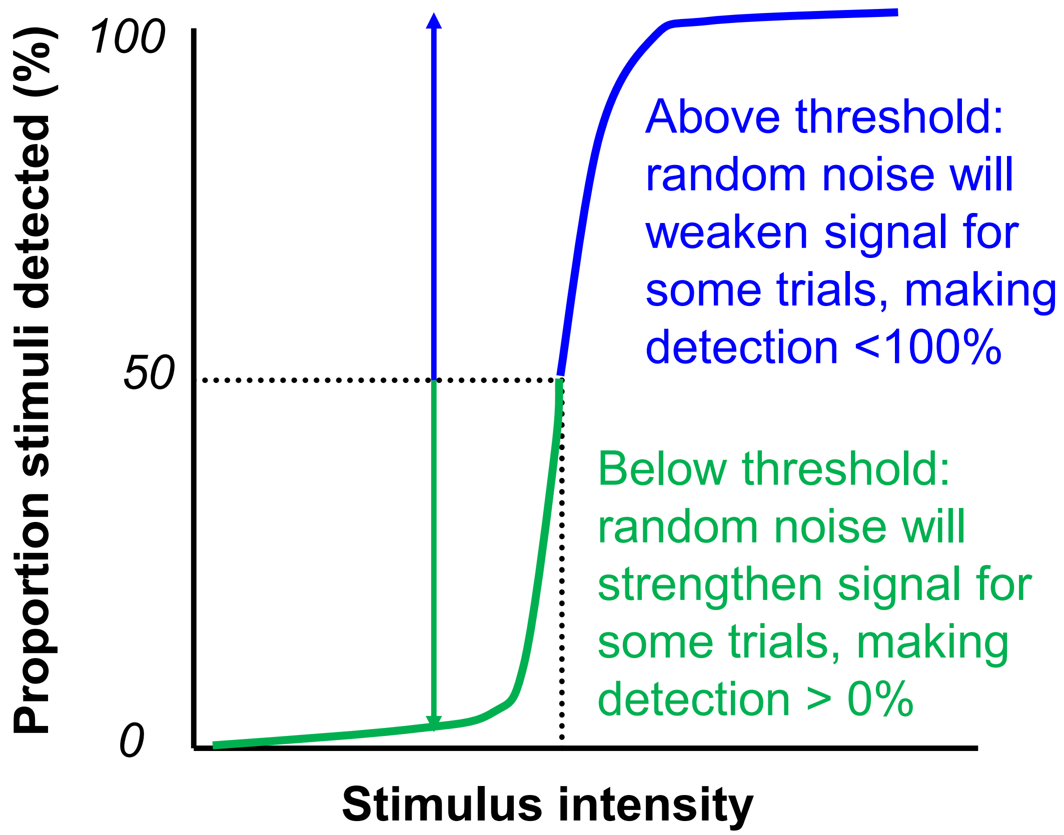
To develop an understanding of how signal detection theory can be used to separate the effects of sensitivity and criterion (which may not always be obvious from analysing the psychometric function alone).

To develop an appreciation of how psychometric functions and signal detection theory have been applied to interface design and evaluation.

Learning Outcomes

To be able to describe the main features of a psychometric function.
To be able to describe how individual differences can affect data interpretation.
To be able to describe signal detection theory and receiver operating characteristics and their role in evaluating individual differences.
To be able to provide examples of how psychometric functions, signal detection theory, and receiver operating characteristics have been applied to interface design and evaluation.

In the Last Lecture We Noted The Effect of Noise on Psychometric Functions



Threshold % is chosen by experimenter

The detection (or discrimination) of stimulus is always subject to noise. Noise can be due to:

Neural activity

Stimulus (physical)

Attention

On any trial, noise will randomly increase or decrease the perceived signal intensity.

Participant perceives *signal+ noise* (cannot tell the difference).

<u>This changes step function to a sigmoid (logistic) function which is what is obtained.</u>

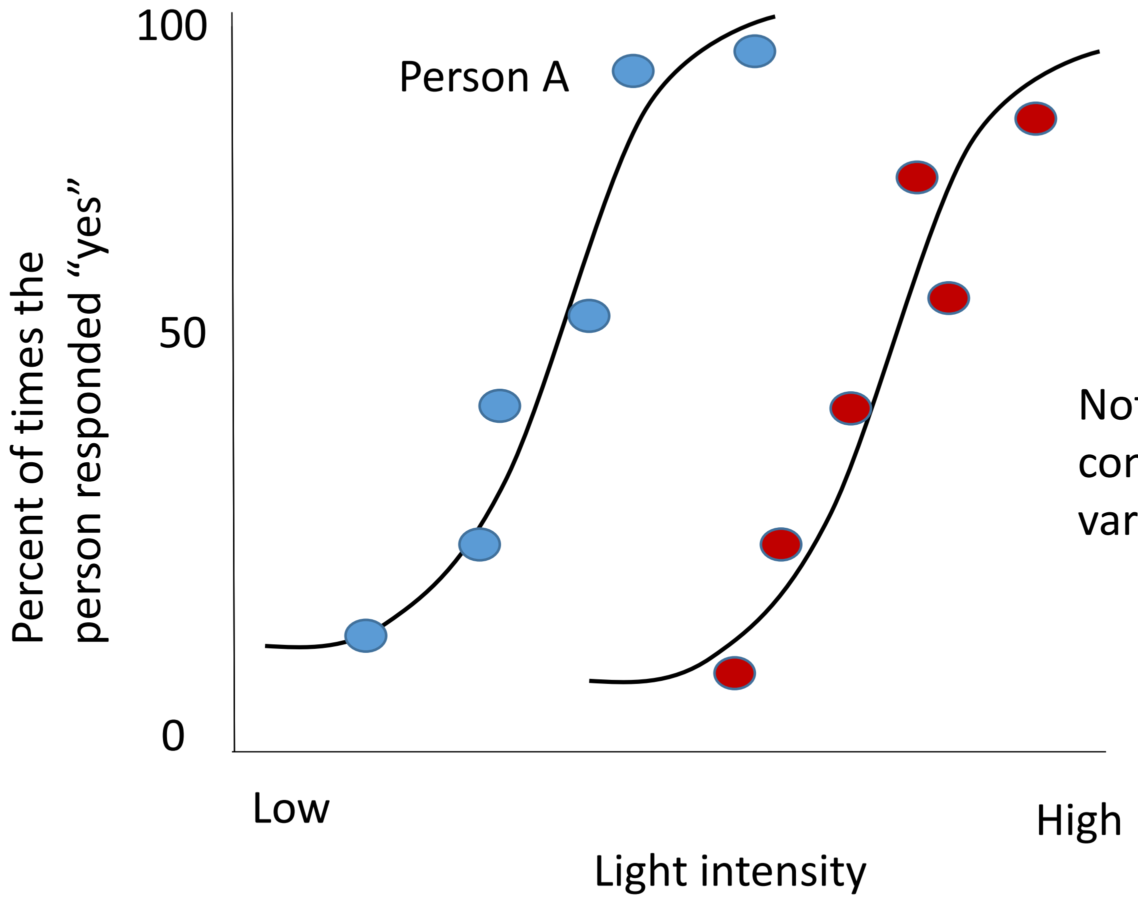
How Can Individual Differences Affect Psychometric Functions?

Example: Use the method of constant stimuli to obtain threshold:

Present lights of different intensities in random order and observer is requested to say "yes" if they see the light.

Person A: Wants to be sure does not miss a light, so will say "yes" even if there is a possibility that there **could** have been a light.

Person B: Responds more conservatively, wants to be totally sure there is a light if they say "yes", so will only respond if the light is **clearly visible**.



Note: we are considering a single variable case.

What do These Differences Mean?

Results for <u>method of constant stimuli</u> experiment:

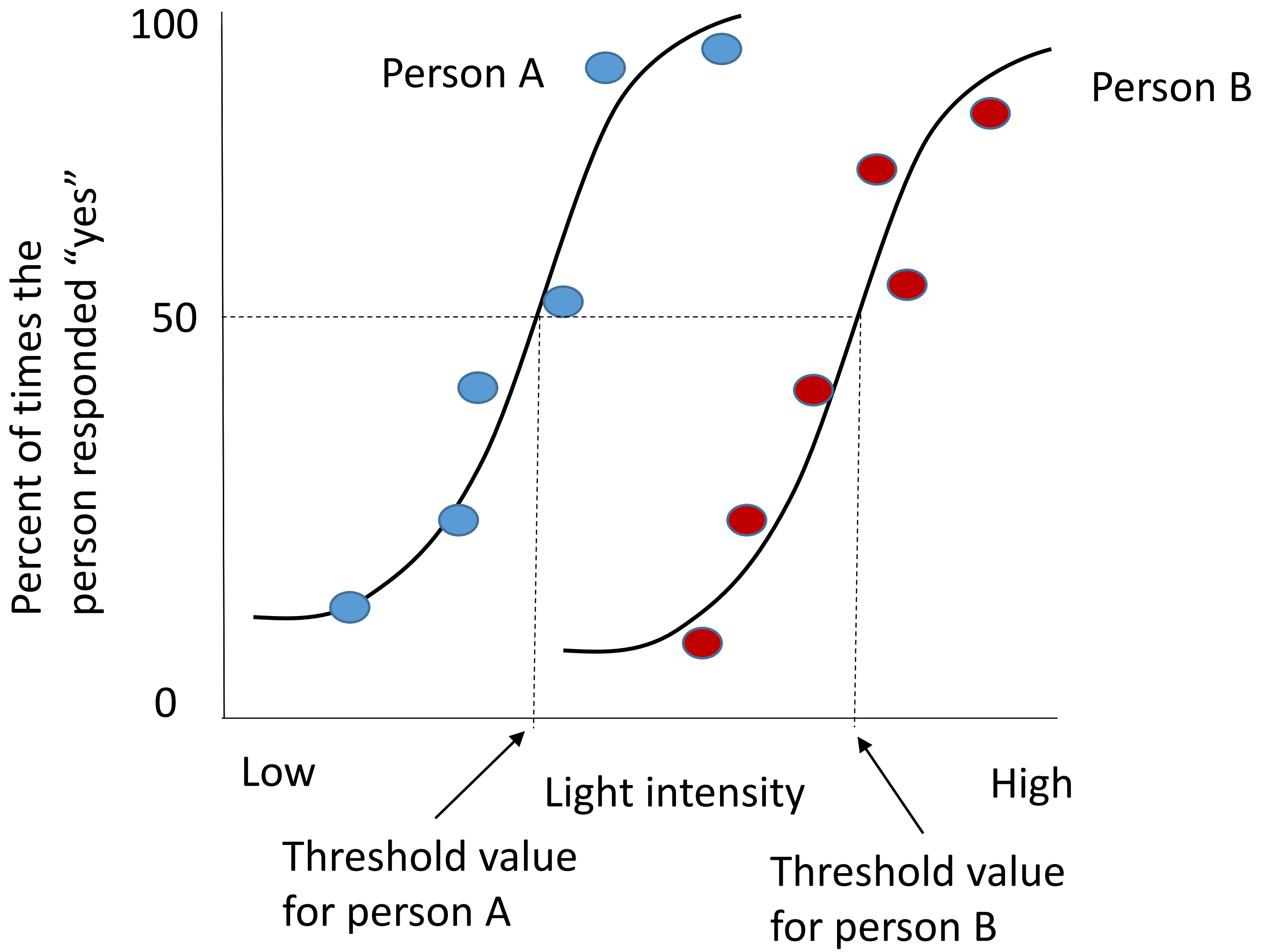
2 sigmoid fits to the data: one for each person.

Person A responds with a “yes” for lower light intensities than person B.

Is person A’s visual system more sensitive than person B’s?

Sensitivity to light may be same for Person A and B, but Person A’s threshold is lower because they are more willing to respond by saying “yes”.

I.e., Each person has a different response criterion.



So How Can We Compare These Two Sets of Responses?

So how do we compare responses across participants (for the case when there are a few participants) when they may have different response criteria?

<u>Signal detection theory</u> is a framework that can be used to take differing response criteria into account.

**What we would want to know is if person A and person B are equally sensitive to the stimuli even though their response criteria are different.

Detecting stimuli in noise: Signal Detection Theory (SDT)

Signal Detection Theory is a framework that can be used to understand how stimuli are detected/discriminated.

SDT explains why shape of psychometric function varies with noise.

SDT explains how a subject's *criterion* (response bias) affects decisions and how to measure it.

SDT allows measurement of *sensitivity* (ability to make correct responses/decisions) regardless of criterion/bias.

Let’s Consider an Example: a Visual Task and Responses

We run an study to measure the responses of person A and person B on a task of light detection:

Use the method of constant stimuli: Different visual stimulus intensities are presented and a stimulus is presented on every trial.

It is a signal detection experiment: Use 1 low-intensity visual stimulus and present this on some of the trials and present no stimulus on the rest of the trials.

how SDT works

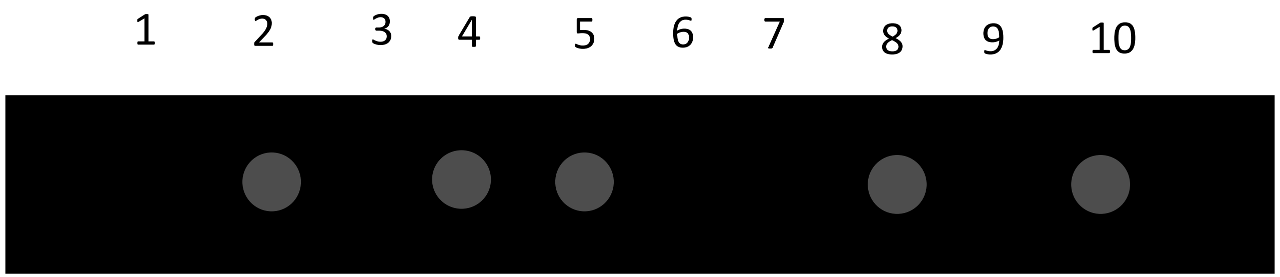
Let’s Consider an Example: a Visual Task and Responses

We run an study to the responses of person A and person B on a task of light detection:

Use the method of constant stimuli: Different stimulus intensities are presented and a stimulus is presented on every trial.

It is a signal detection experiment: Use 1 low-intensity stimulus and present this on some of the trials and present no stimulus on the rest of the trials.

Experimental Trials



Continues for say a total of 200 trials:

Stimulus is presented on 100 trials, and no stimulus is presented on 100 trials

Results-Response Comparisons

Person A's response:

When a signal was presented:

Said "yes" on 90 trials, when the signal was present. [HIT]

Said "no" on 10 trials, when the signal was present. [MISS]

When a signal was NOT presented:

Said "yes" on 40 trials, when the signal was not present. [FALSE ALARM]

Said "no" on 60 trials, when the signal was not present. [CORRECT REJECTION]

Person A is <u>more willing to say "yes" </u>(has a low criterion), and has a high hit rate of 90% and a 40% false alarm rate.

Results-Response Comparisons

Person B's response:

When a signal was presented:

Said "yes" on 60 trials, when the signal was present. [HIT]

Said "no" on 40 trials, when the signal was present. [MISS]

When a signal was NOT presented:

Said "yes" on 10 trials, when the signal was not present. [FALSE ALARM]

Said "no" on 90 trials, when the signal was not present. [CORRECT REJECTION]

Person B is <u>less willing to say "yes" </u>(has a higher criterion), and has a lower hit rate, 60% and lower false-alarm rate, 10%.

So How do we Compare Person A with Person B?

Person A and person B appear to be using different criteria for responding in the visual task, but are their underlying sensitivities to the visual signal in the task similar or not?

We can Use Signal Detection Theory (SDT)

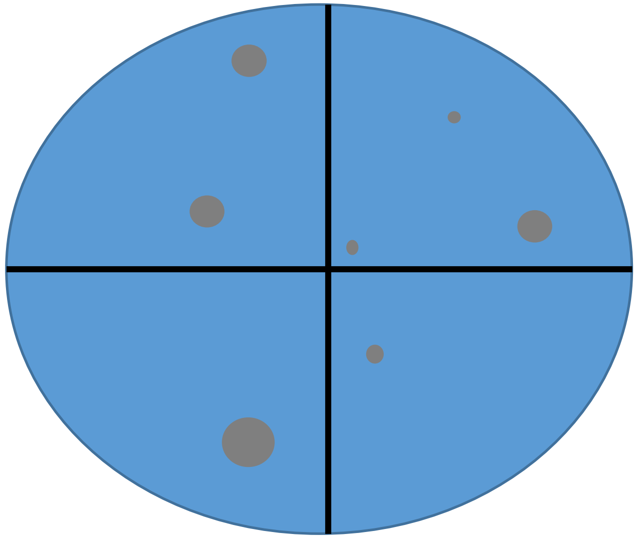
Origin of SDT was in radar operation ~ 70 years ago.

Task: To correctly identify aircraft.

But which of these shapes on the screen is an aircraft and which is not?

Requires correct identification.

Radar Screen



SDT: Analyse the Possible Decisions and Outcomes

For the visual study, <u>on each trial</u>:

The visual stimulus was either presented by the experimenter (yes) or not presented by the experimenter (no).

The participant made a <u>decision</u> as to whether the visual stimulus was present (yes) or not (no).

SDT: Analyse the Possible Decisions and Outcomes

What was the participant's decision?		Was the visual signal actually present on the trial: Yes	No
	Yes	HIT	FALSE ALARM
	No	MISS	CORRECT REJECTION

For the visual study, on each trial:

The visual stimulus was either presented by the experimenter (yes) or not presented by the experimenter (no).

The participant made a decision as to whether the visual stimulus was present (yes) or not (no).

SDT: Analyse the Possible Decisions and Outcomes

For the visual study, on each trial:

The visual stimulus was either presented by the experimenter (yes) or not presented by the experimenter (no).

The participant made a decision as to whether the visual stimulus was present (yes) or not (no).

Was the visual signal actually present on the trial

What was the participant's decision?		Yes	No
	Yes	HIT	FALSE ALARM
	No	MISS	CORRECT REJECTION

Hit: The participant correctly decides that a signal was present when it was actually present.
False Alarm: The participant decides that a signal was present when it was actually not present.

Miss: The participant decides that a signal was not present when it was actually present.
Correct Rejection: Participant correctly decides that a signal was not present when it was actually not present.

SDT: Effect of Criterion

As we know from the study results:

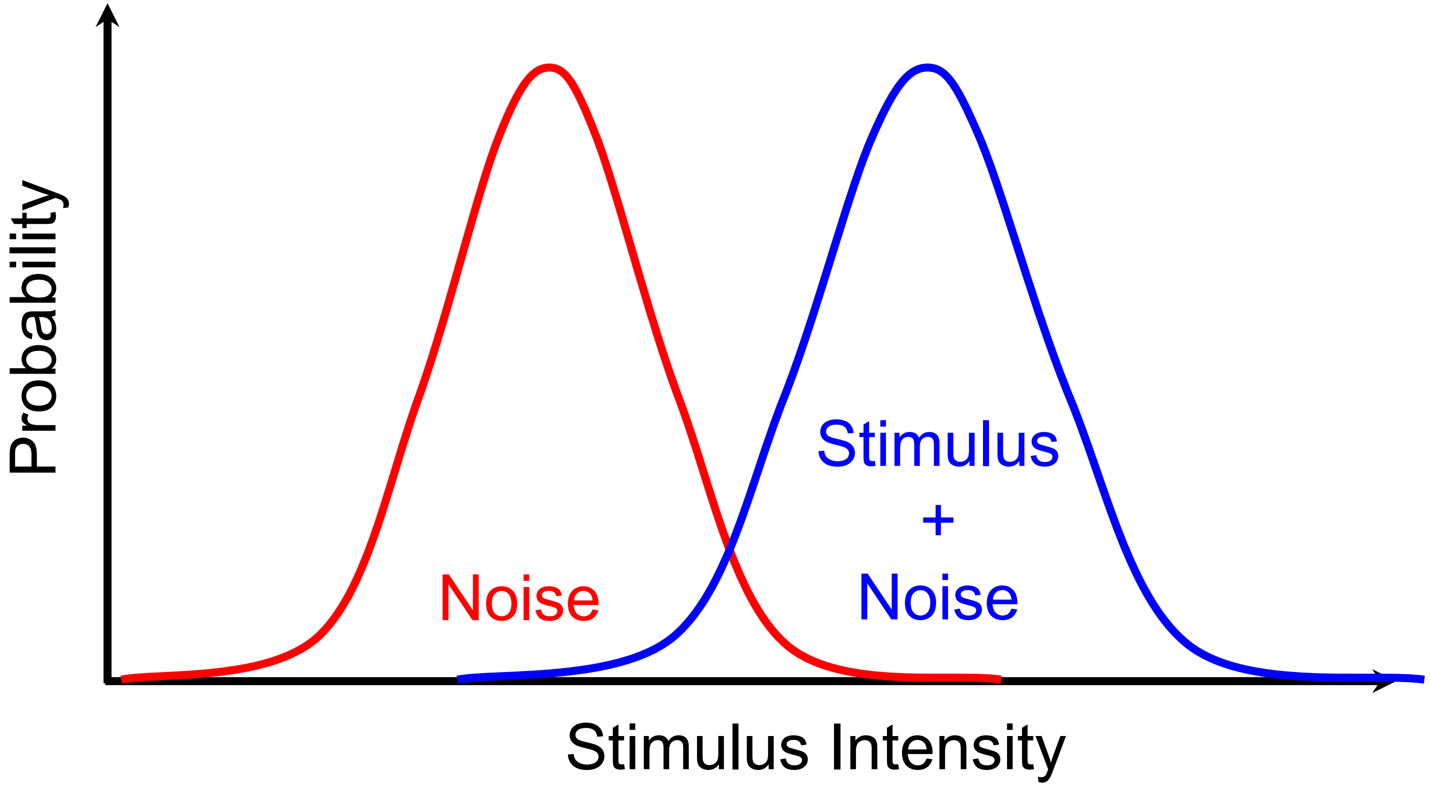
Person A is more willing to say “yes” (has a low criterion), and has a high hit rate of 90% and a 40% false alarm rate.
Person B is less willing to say “yes” (has a higher criterion), and has a lower hit rate, 60% and lower false-alarm rate, 10%.

Low criterion: More willing to say yes and have a high hit rate but also have a relatively high false alarm rate.

High criterion: More conservative, less willing to say yes unless very sure that a signal is present, so have a low false alarm rate, but will have more misses.

So which criterion is “best” (optimal)? That depends on the costs of making errors...
(and which errors are acceptable)

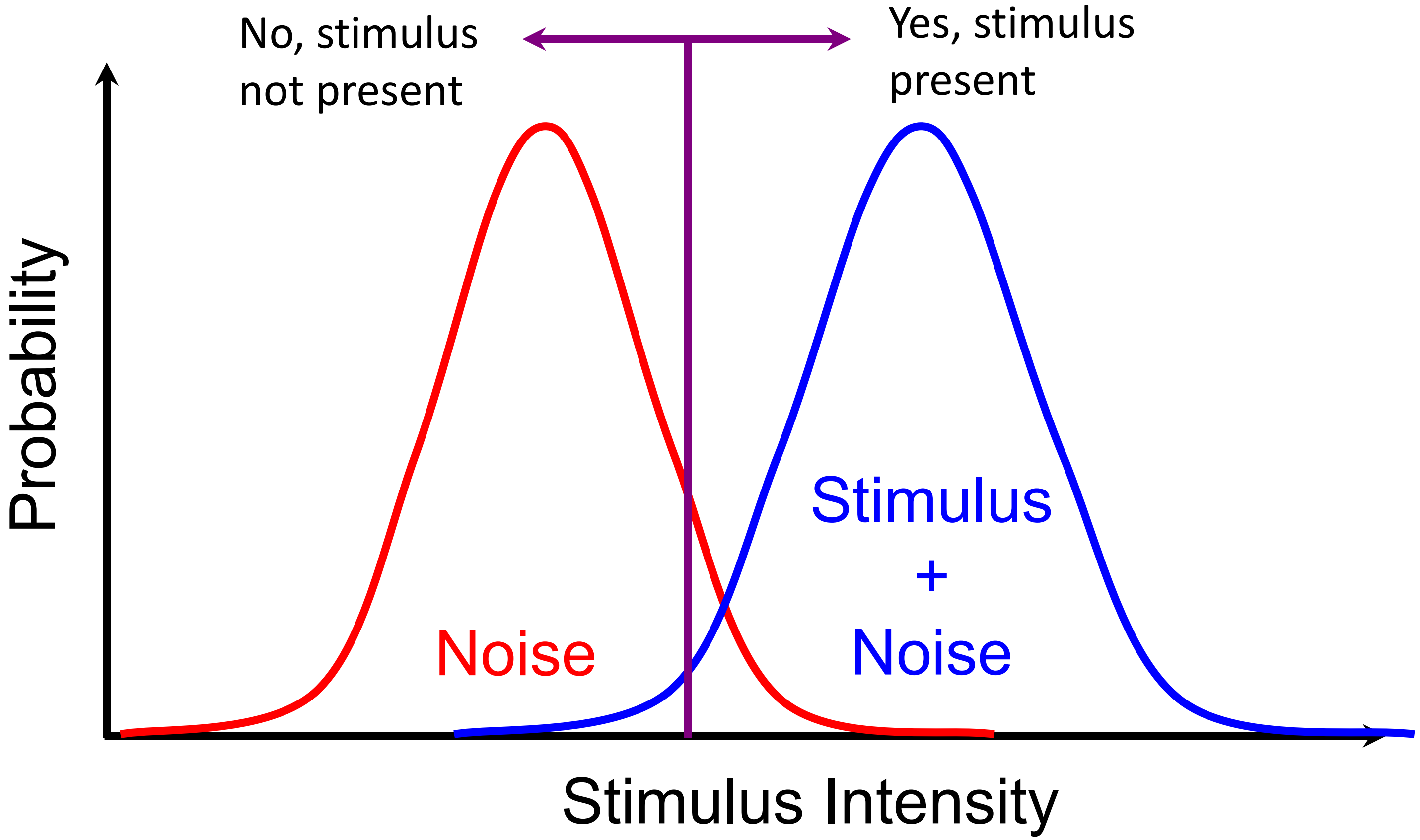
SDT: Probability distributions of "Noise" and "Stimulus + Noise"



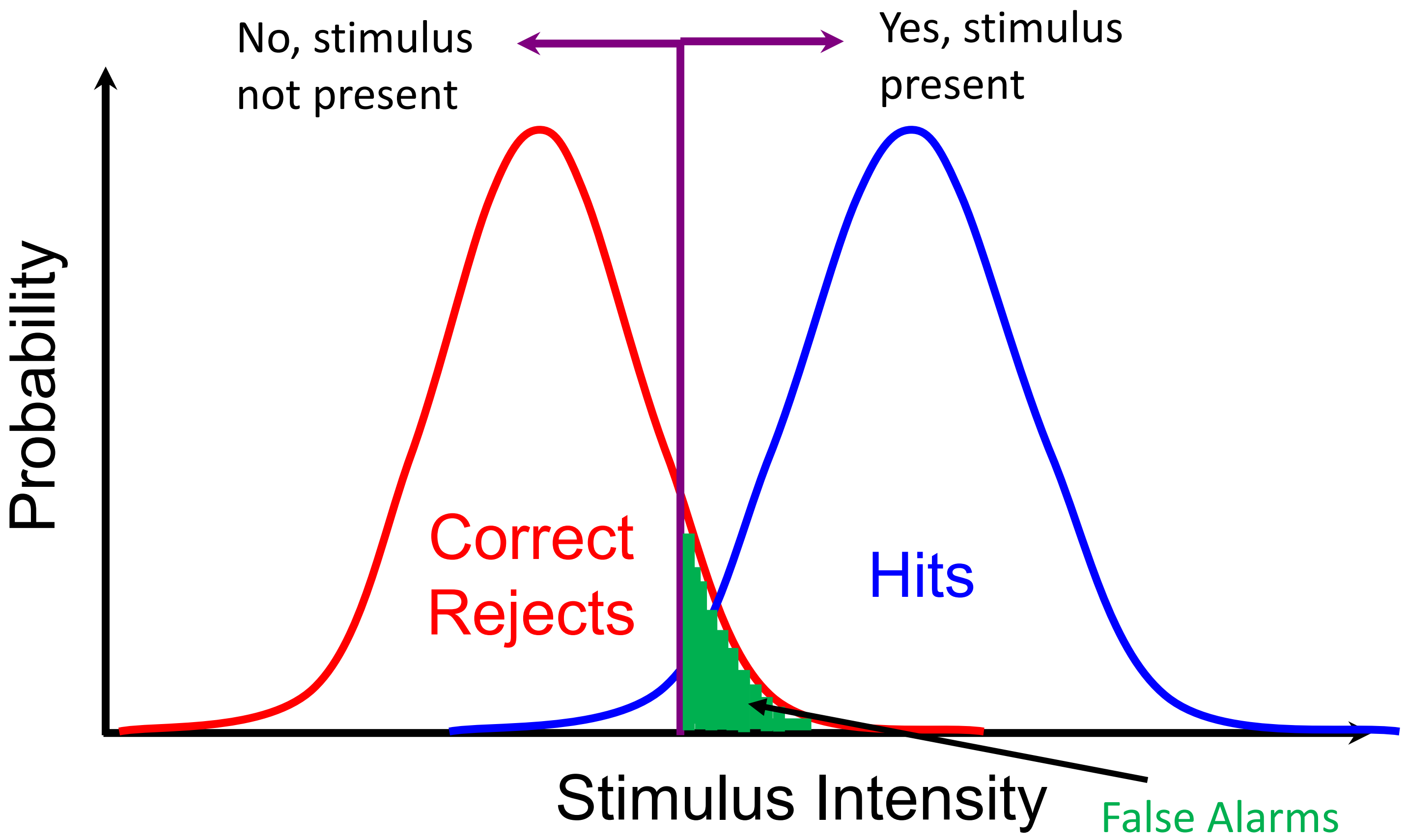
Noise: N

Stimulus + Noise: SN

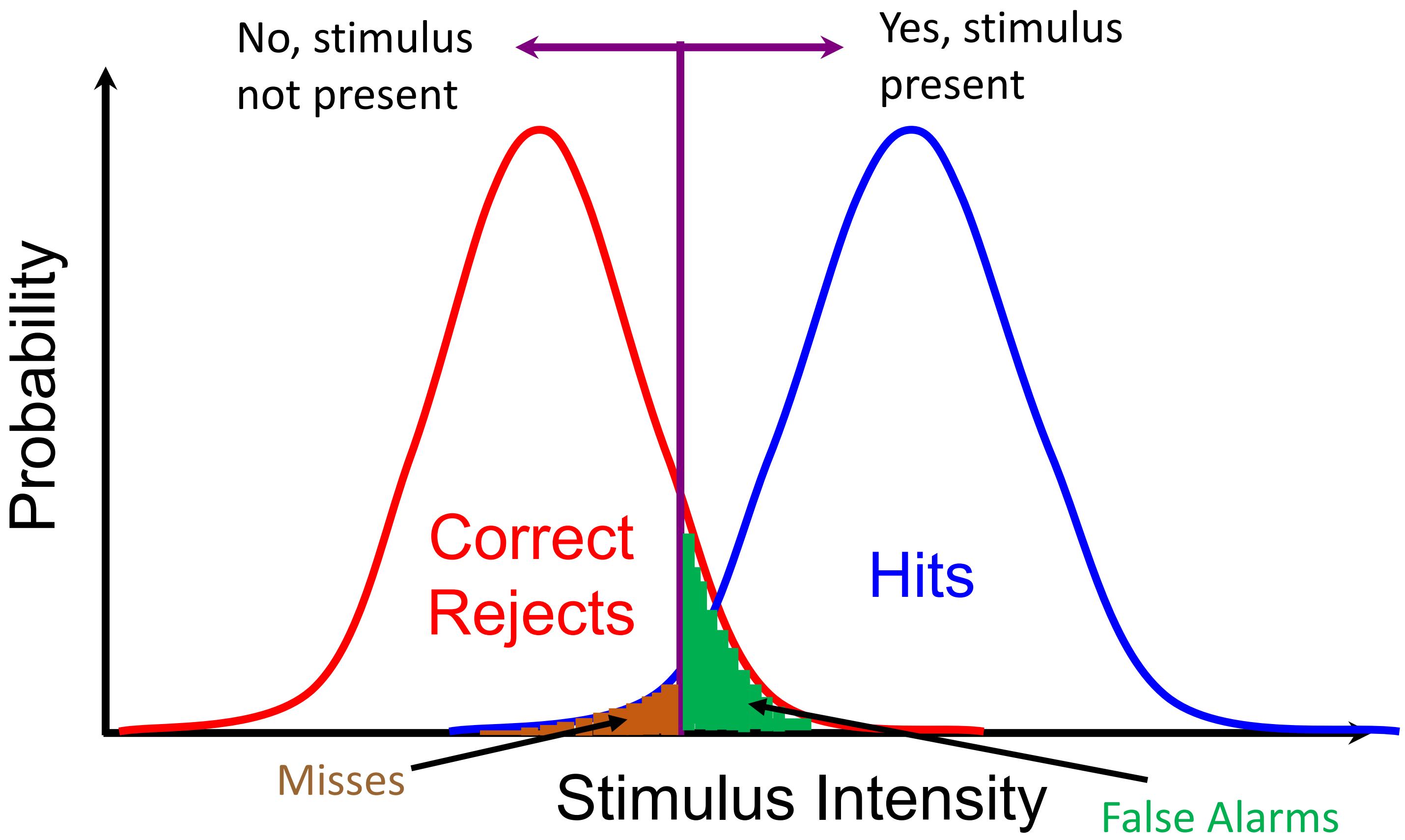
SDT: Effect of Criterion



SDT: Effect of Criterion



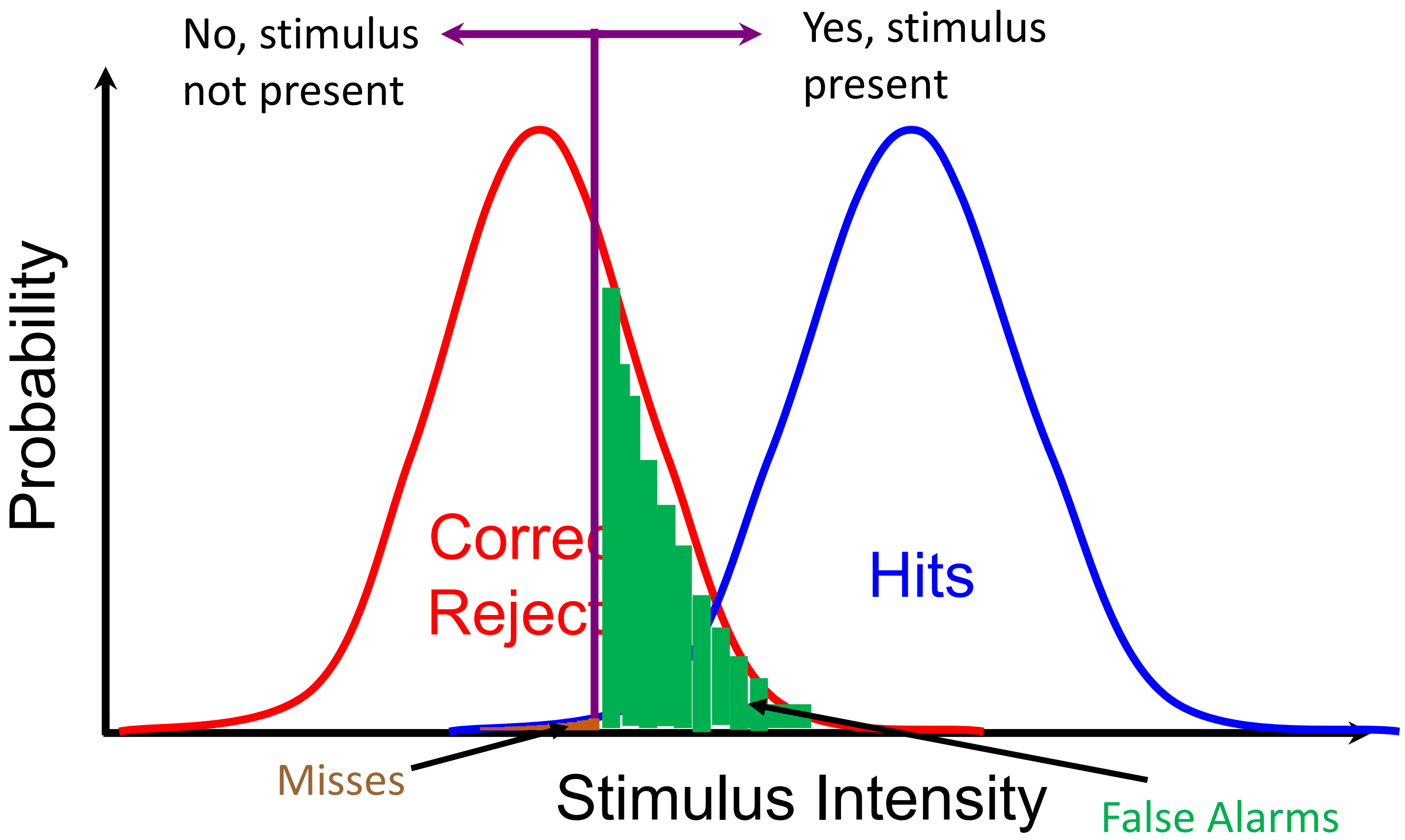
SDT: Effect of Criterion



Noise:

Stimulus + Noise:

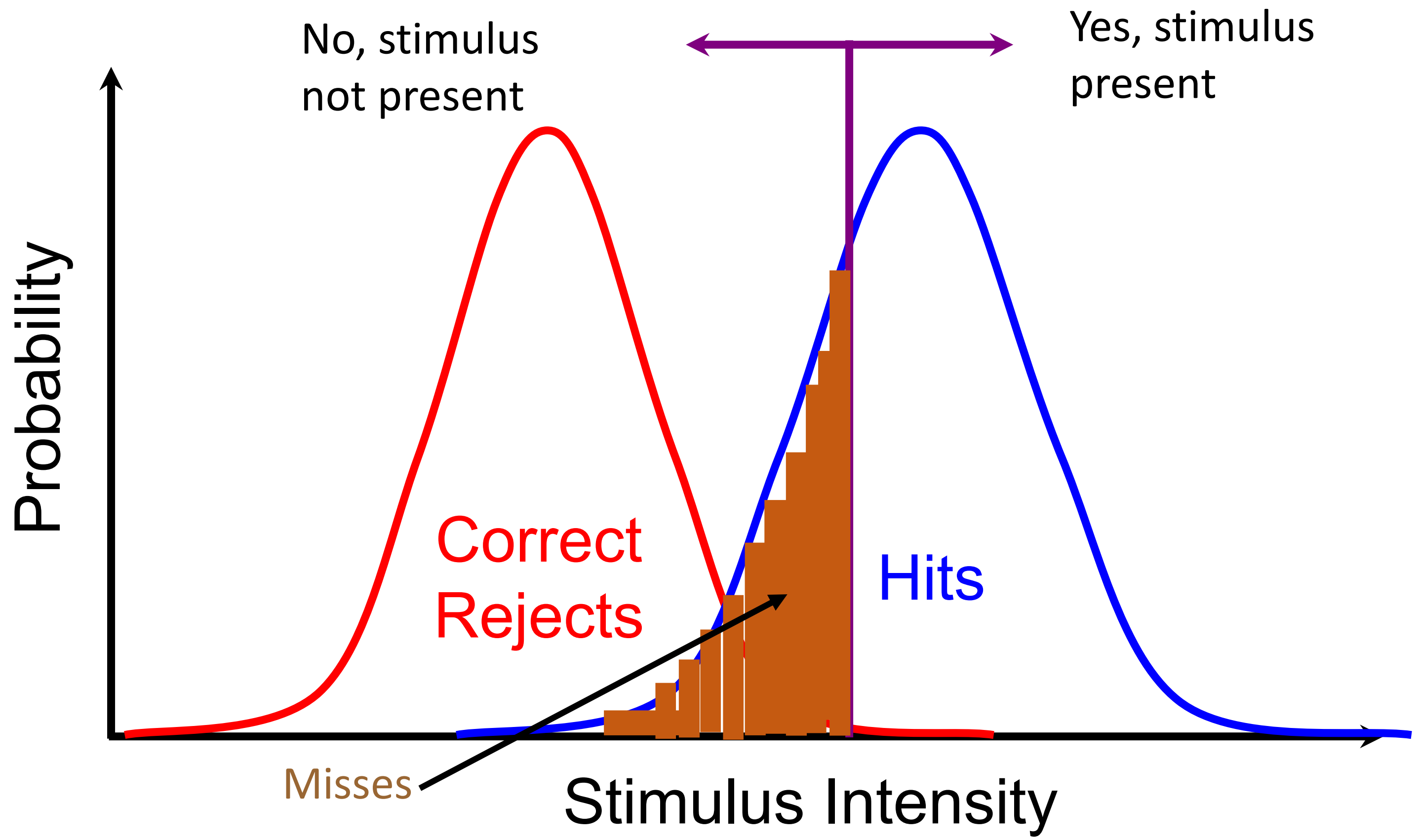
SDT: Effect of Criterion



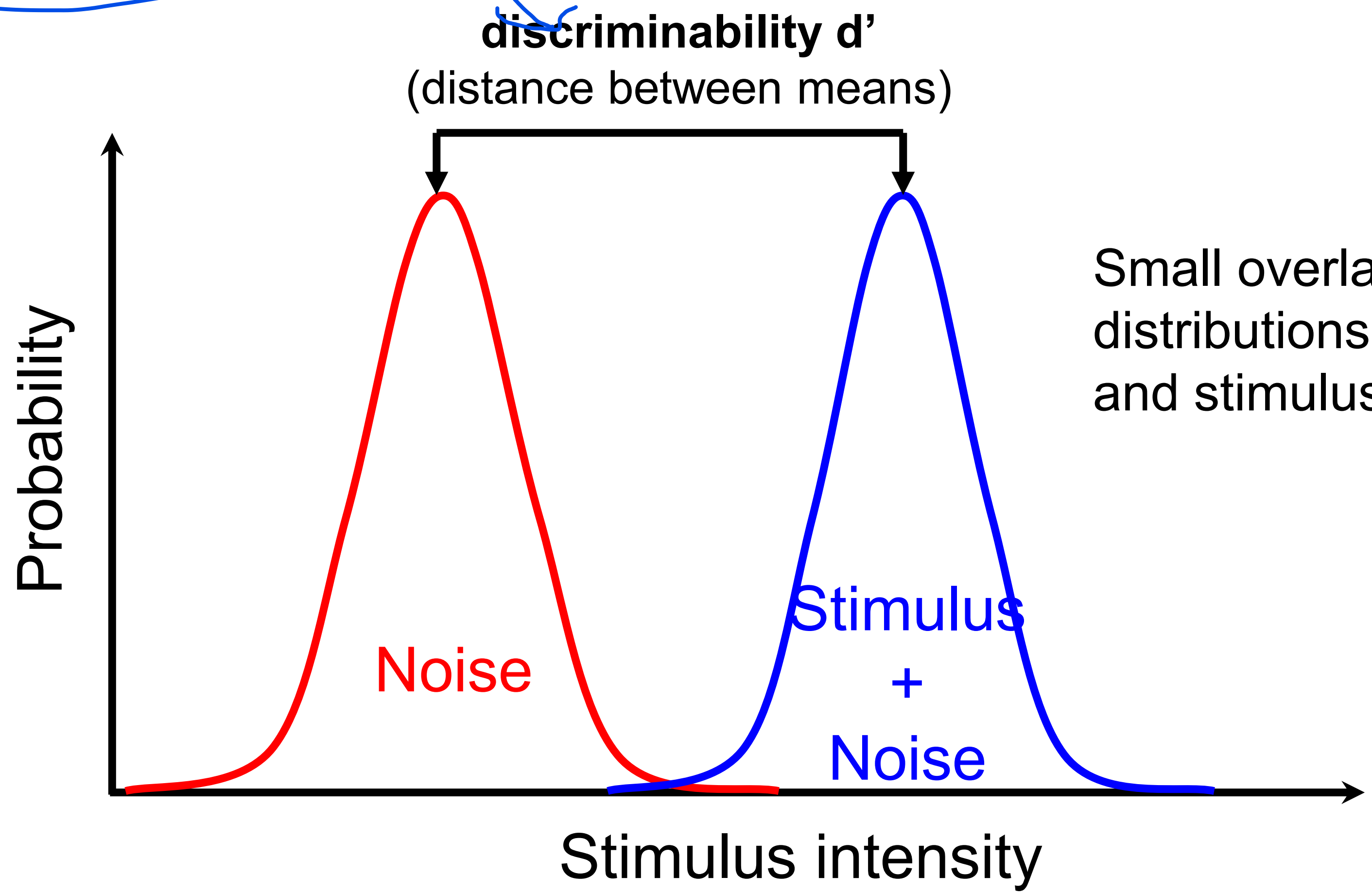
Noise:

Stimulus + Noise:

SDT: Effect of Criterion

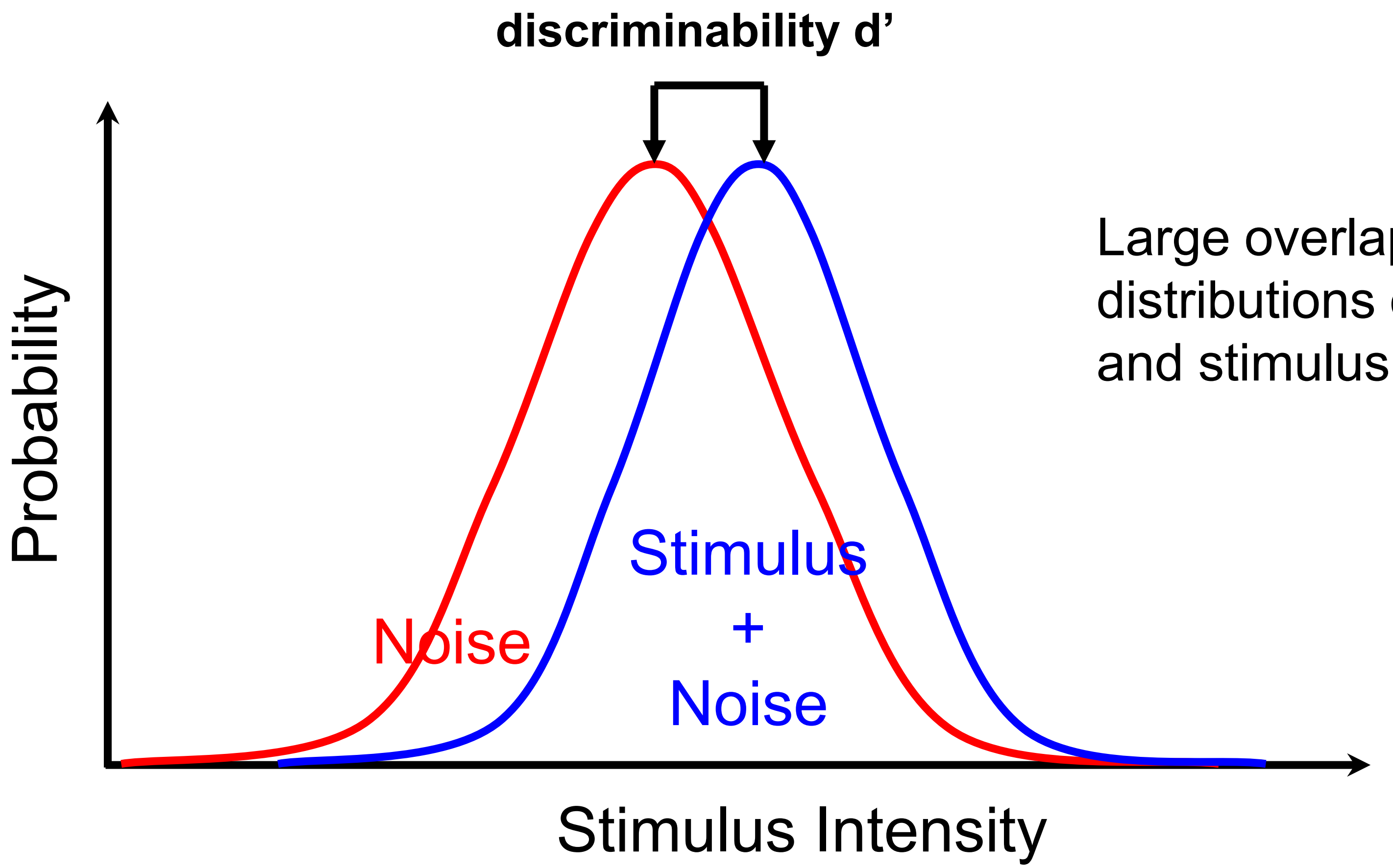


Low noise: High discriminability & sensitivity (few misses & false alarms)



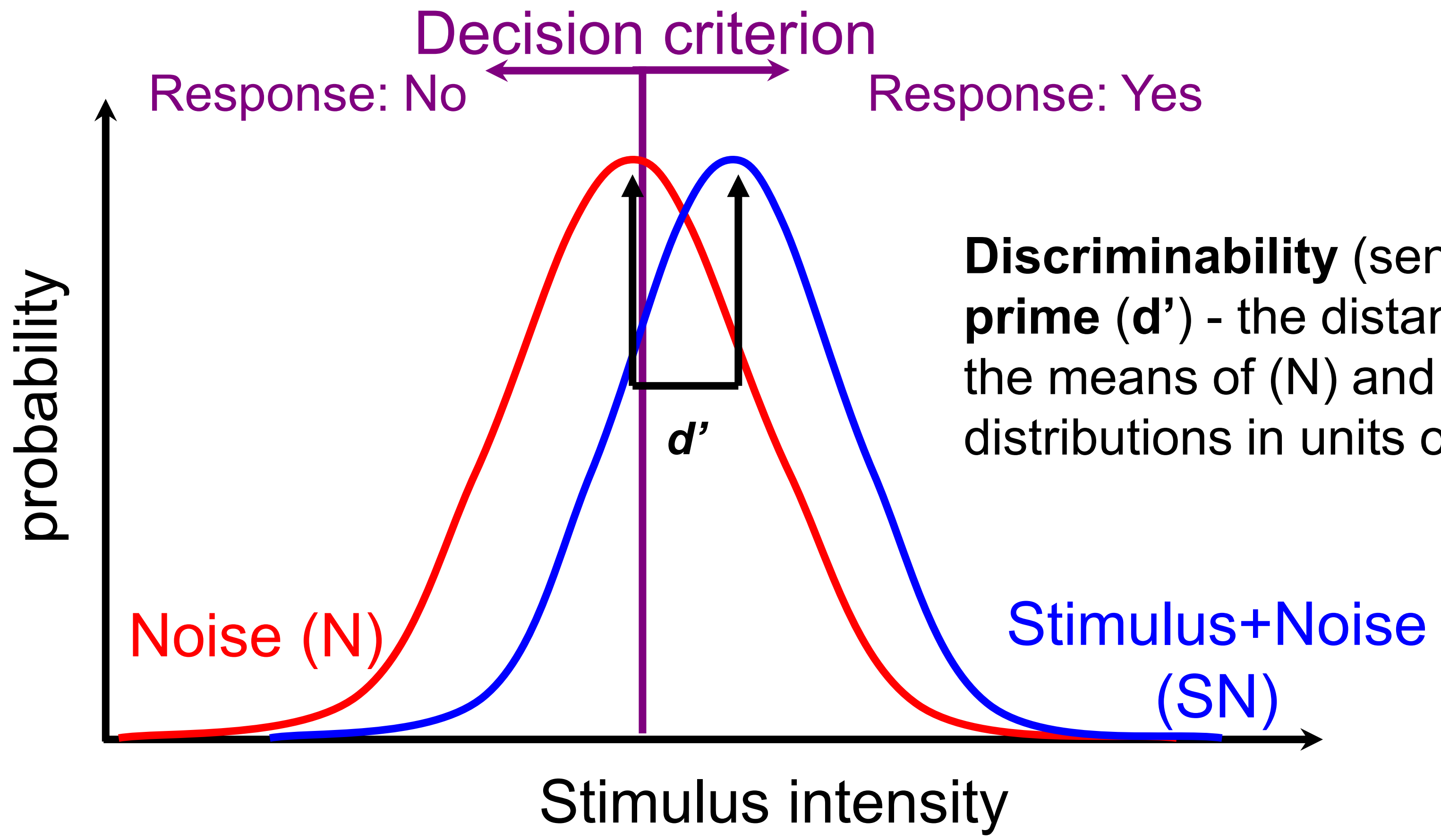
Small overlap between distributions of noise and stimulus + noise

High noise: Low discriminability & sensitivity (many misses & false alarms)



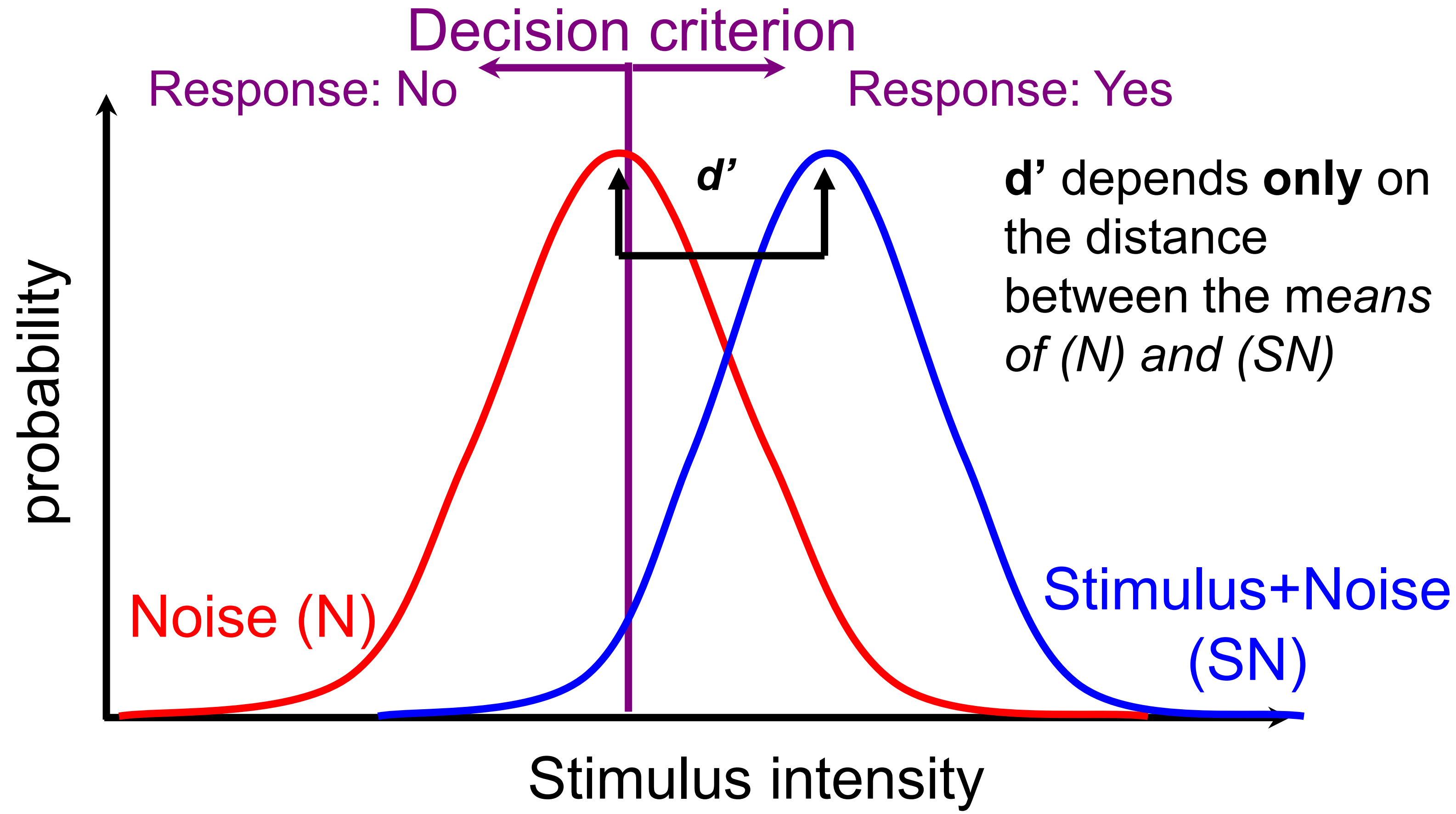
Large overlap between distributions of noise and stimulus + noise

SDT & Psychophysics



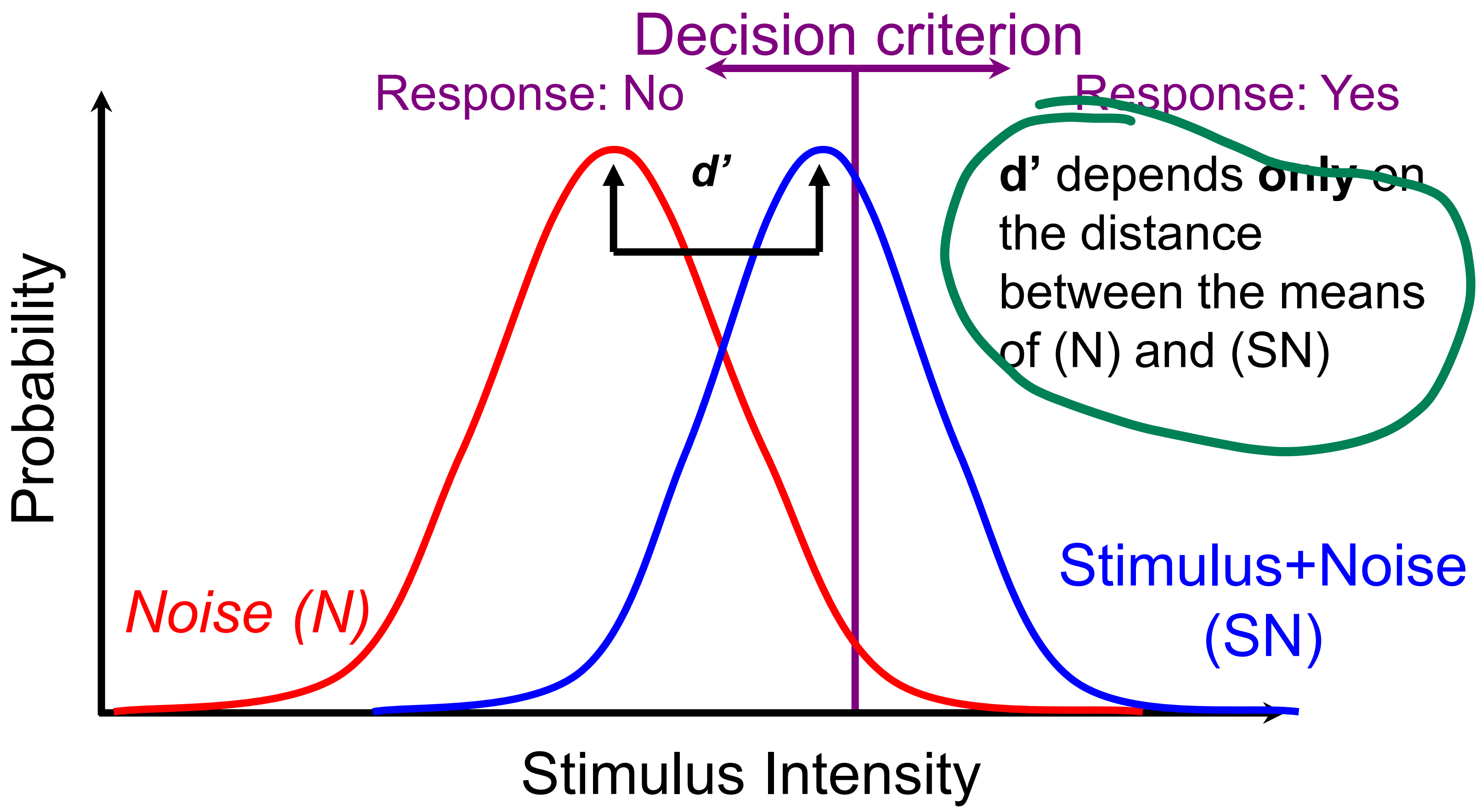
Discriminability (sensitivity): **d-prime** (**d'**) - the distance between the means of (N) and (SN) distributions in units of S.D.

Discriminability (d') is independent of criterion



d' depends **only** on the distance between the *means of (N) and (SN)*

Discriminability (d’) is independent of criterion



Estimation of d’

d’ is the difference between the *means* of the noise (N) and stimulus+noise (SN) distributions, in units of *standard deviations* of the noise (N) distribution:

$$\boldsymbol{d' = [m_{SN} - m_N] / s_N}$$

But these distributions are not usually known!

d’ is more easily computed from the *hit rate* (proportion of stimuli reported when present, [yes|SN]) and the *false alarm rate* (proportion of stimuli reported when not present, [yes|N]):

Convert hit & false alarm rates (which are probabilities) to *z* scores from tables of *z* distribution:

Hit rate = P(yes|SN) => *z*(yes|SN)

False alarm rate = P(yes|N) => *z*(yes|N)

***d’* = *z*(yes|SN) - *z*(yes|N)**

Decision criterion must be fixed!

Interpreting d’

Low *d’* means stimulus (signal) + noise (SN) distribution is highly overlapping with noise (N) distribution

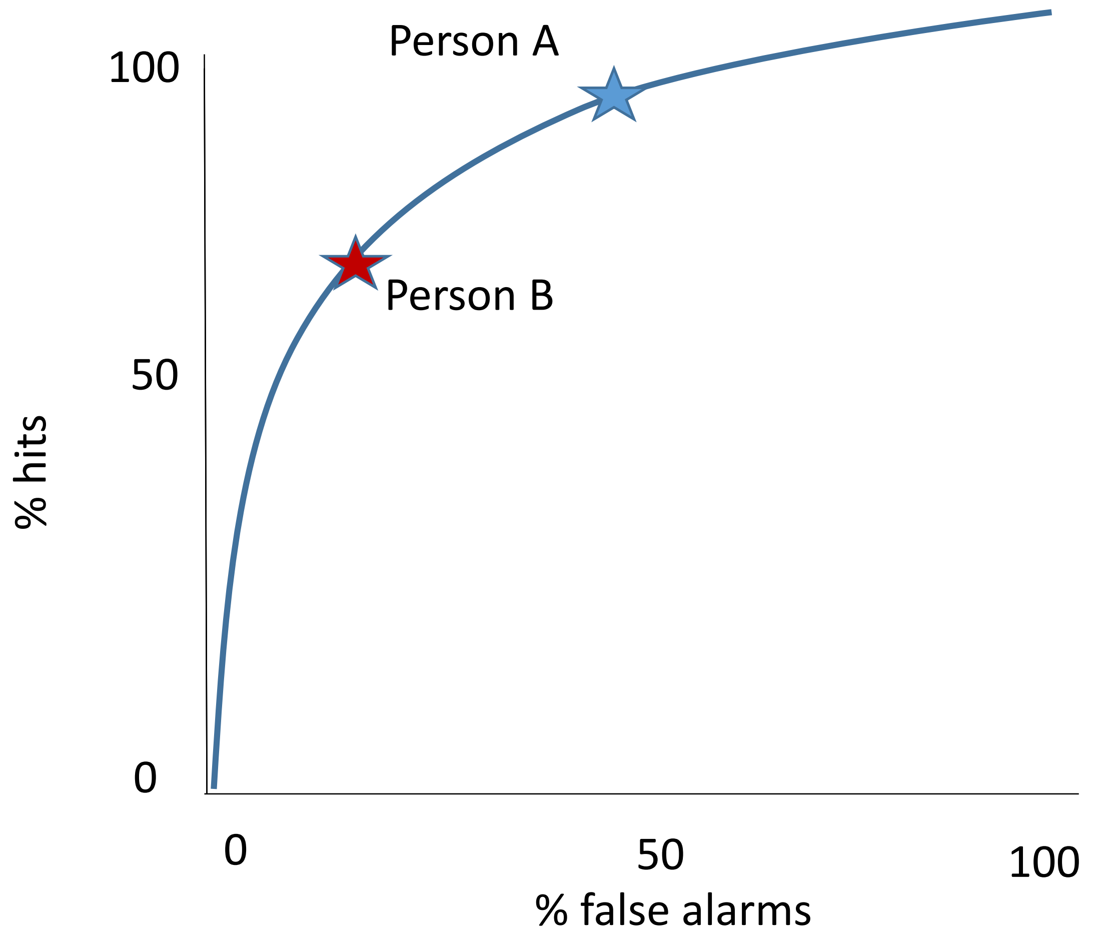
d’ = 0: chance level performance (N and SN overlap)

High *d’* means SN and N distributions are far apart

d’ = 1: moderate performance

d’ = 4.65: “optimal” (corresponds to hit rate=0.99, false alarm rate=0.01)

Receiver Operating Characteristics (ROC)

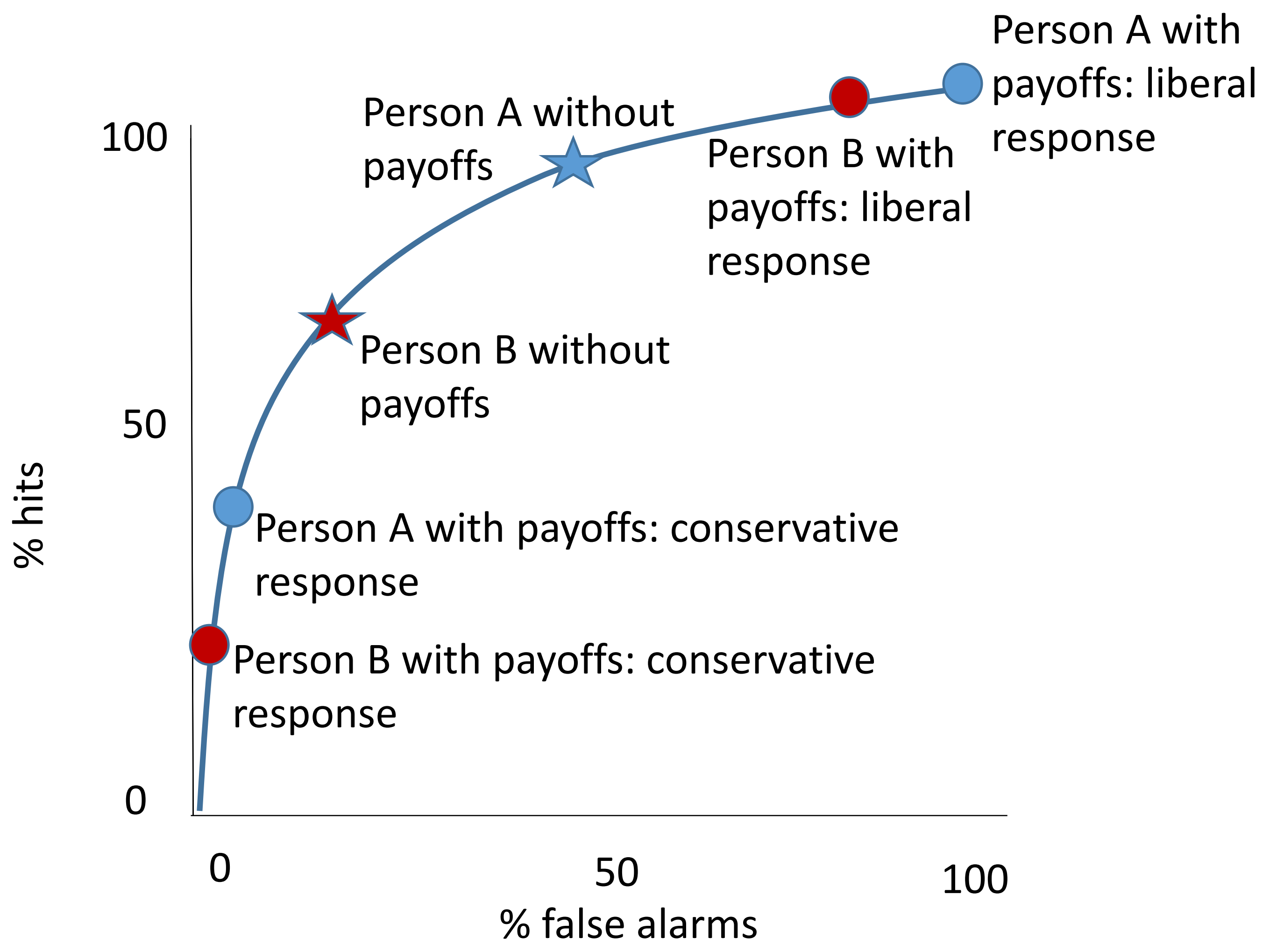


Person A has a low criterion, and has a high hit rate of 90% and a 40% false alarm rate. Person B has a higher criterion, and may have a lower hit rate e.g., 60% and lower false-alarm rate (e.g., 10%).

A receiver operating characteristic curve (ROC).

Person A and Person B's data points lie on the same curve, therefore they have the same sensitivity to the stimulus.

Receiver Operating Characteristics (ROC) and Effect of Payoffs



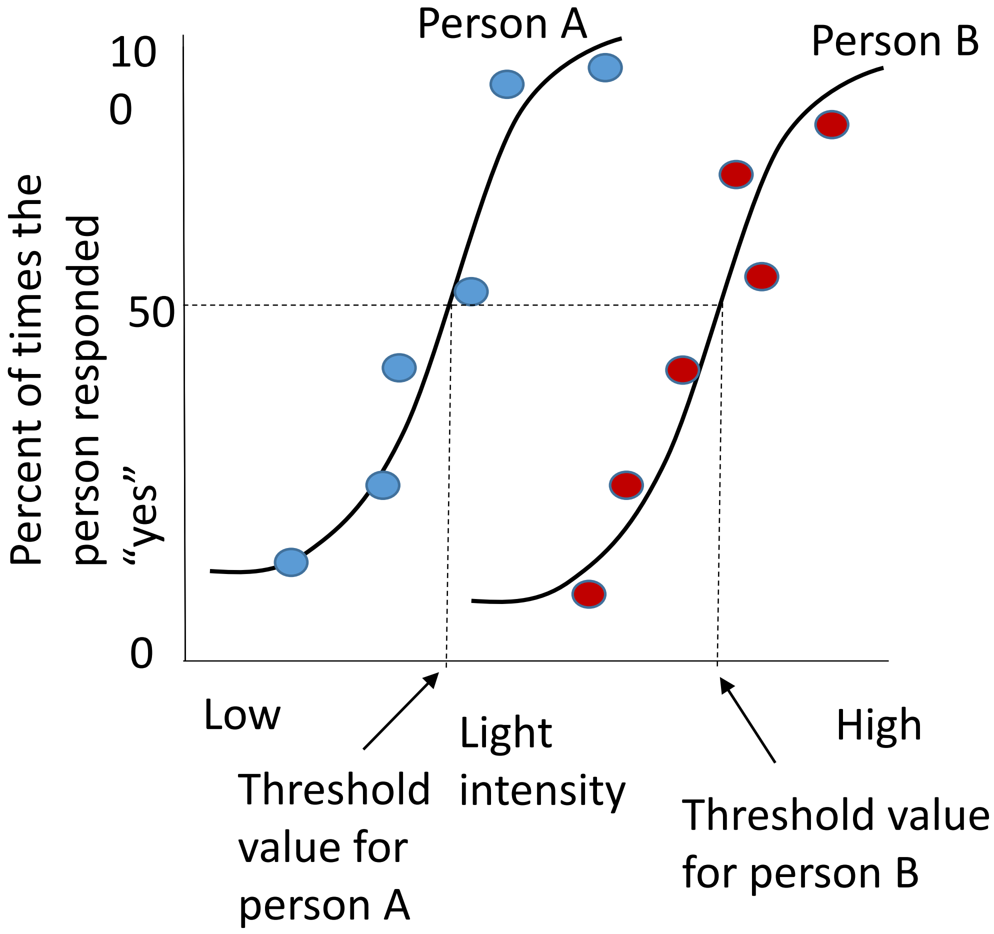
A receiver operating characteristic curve (ROC).

How can we cause Person A and B to change their percentages of hits and false alarm rates without changing characteristics of the stimuli?

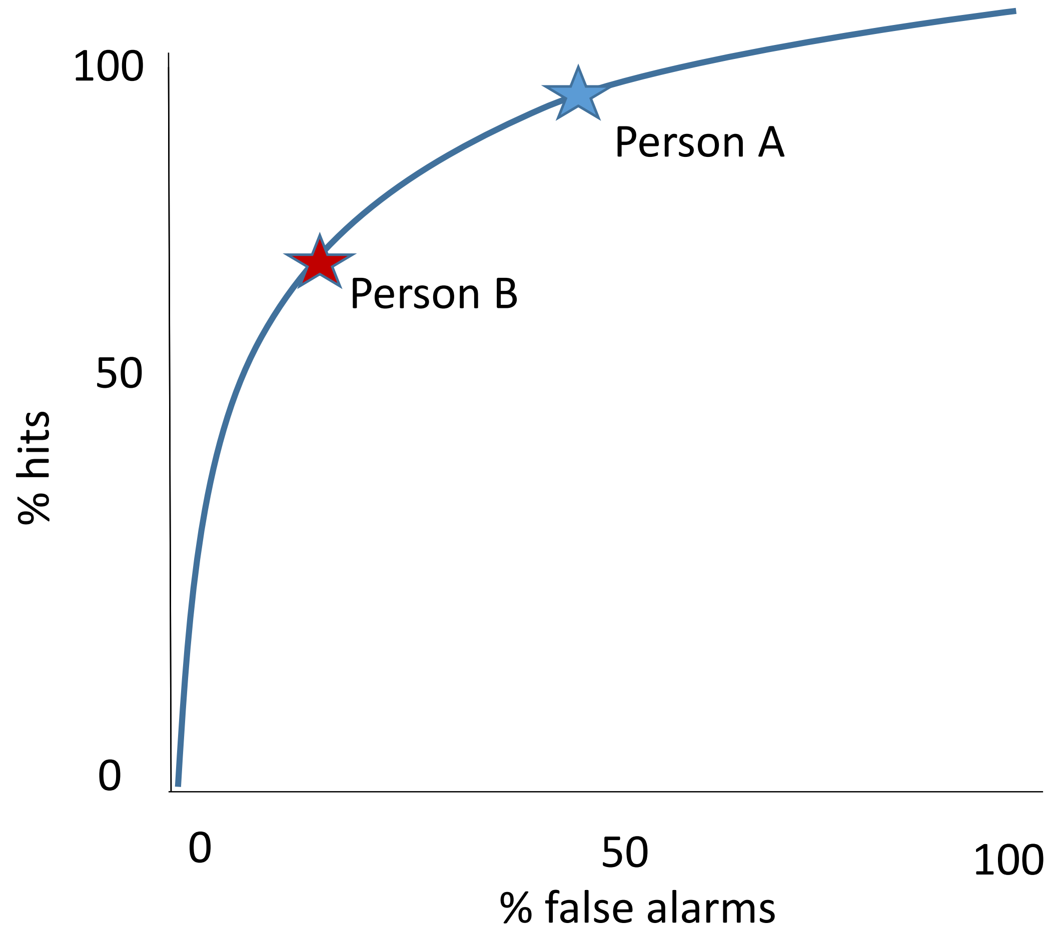
We can alter each person’s motivation by means of payoffs. E.g., can introduce a reward for making correct responses.

Person A and Person B’s data points lie on the same curve, therefore they have the same sensitivity to the stimulus.

Method of constant stimuli suggested different sensitivities

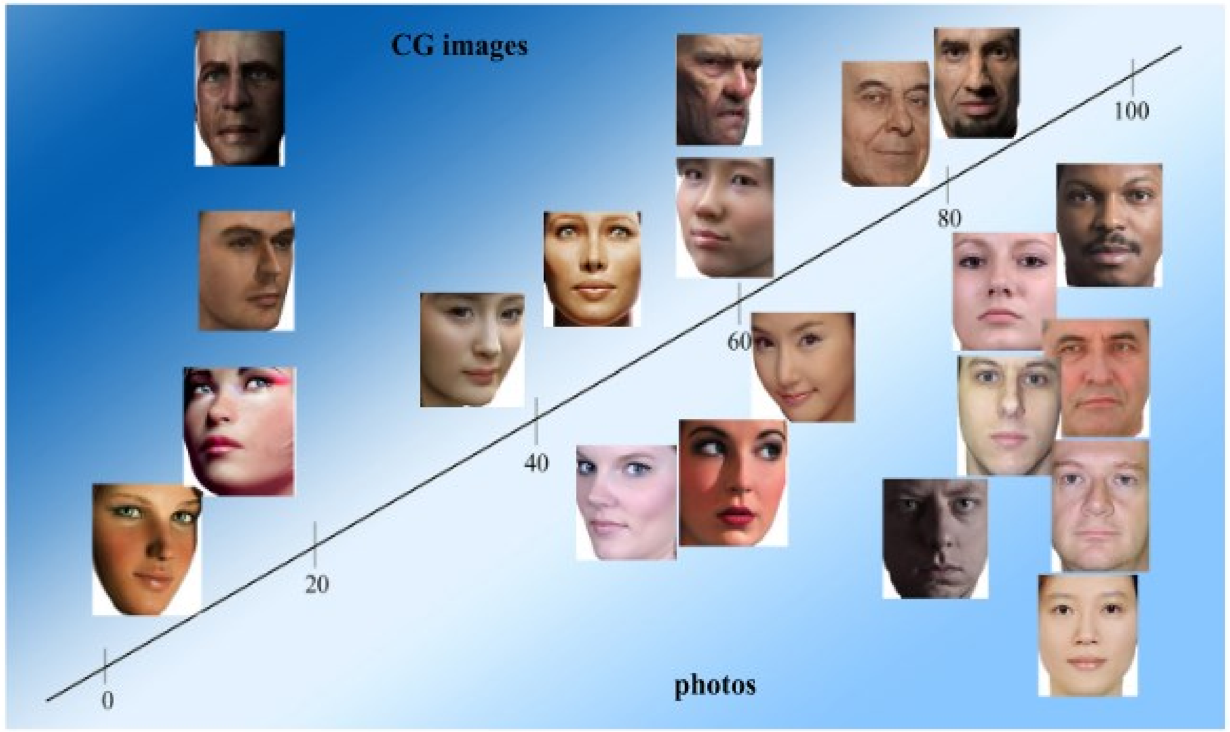


Signal detection theory shows that sensitivities are the same (data points lie on the same curve), but person A has a lower criterion for saying "yes" compared to person B.



An Example Application (see slide 3 for more examples)

Fan, S., Ng, T.T., Herberg, J.S., Koenig, B.L. and Xin, S., 2012. Real or fake? Human judgments about photographs and computer-generated images of faces. In *SIGGRAPH Asia 2012 technical briefs* (pp. 1-4).



Previously know: Determining how real an image appears – its visual realism, is important in the field of computer graphics.

This study investigated:
1) Intrinsic image properties: reflectance and shading.
2) Individual differences among participants in their image-processing expertise.

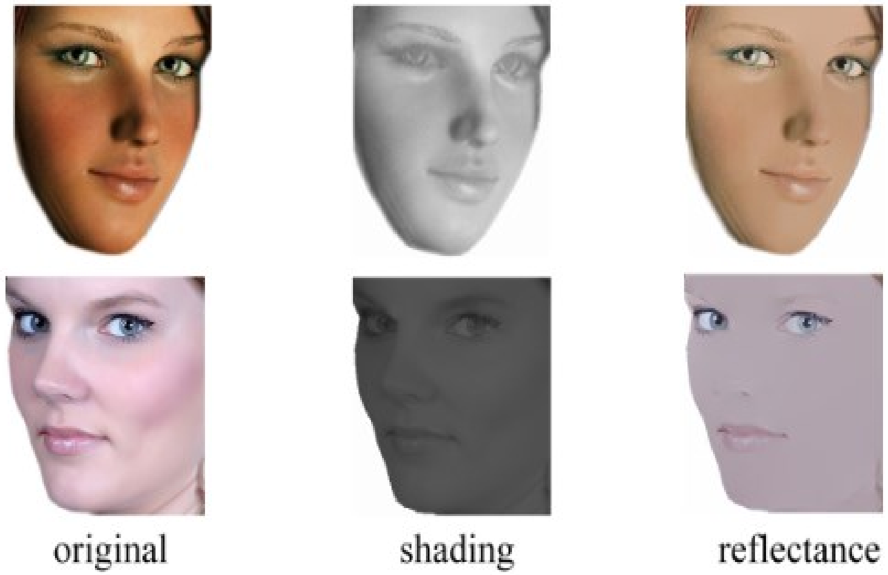


Figure 3: *Two examples of intrinsic image decomposition, the above is a CG image, below is the photo in the pair.*

Psychophysics Experiment:

261 laypersons and 122 image-processing experts viewed photos and computer generated (CG) images of faces and judged whether it was a photo or a CG image.

Stimuli were 500 pairs of images of faces. Each pair contained one CG image and one photograph.
The stimuli were i) unaltered colour images, ii) images modified to show only a grayscale version, and iii) images modified to show only a reflectance version.

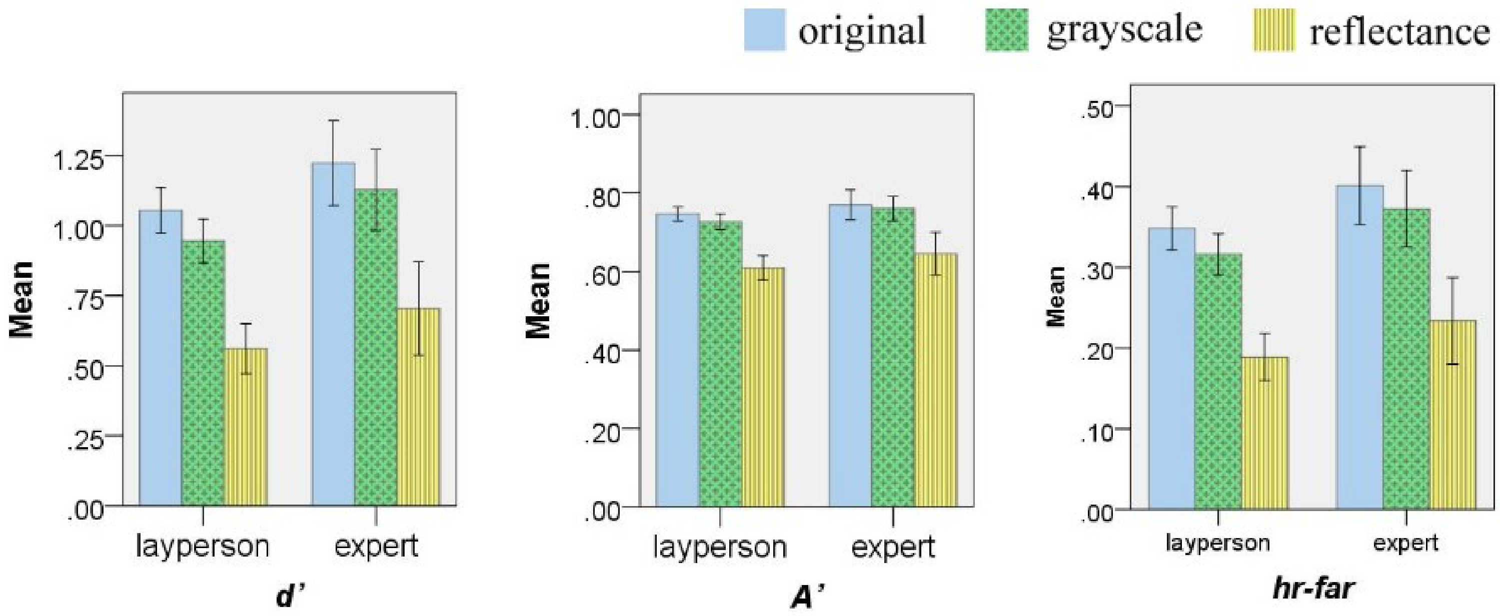


Figure 4: *SDT values of three image types.*

Analysis:

Photos were classed as the signal and CG as the noise.

d' measure obtained.

For both participant categories both sensitivity and accuracy were highest on original colour images, moderate for grayscale images, and lowest on reflectance images.

Questions to Consider (5 mins)

Q1: How could a payoff scheme have been introduced into the Fan et al (2012) study?

Q2: How could the payoff have affected the results?

(We can discuss this in the Question and Answer Sessions)

Overall Summary

Data can be plotted as psychometric functions.

Individual differences in performance can influence the psychometric function.

To separate out effects of criterion from sensitivity signal detection theory can be used.

ROC can be used to further explore differences in sensitivity and criterion

Psychometric functions, signal detection theory, and receiver operating characteristics have been applied to interface design and evaluation.

Resources

Essential:

Sensation and Perception 8th edition (book) page 3 to 20 and page 401-406

Fan, S., Ng, T.T., Herberg, J.S., Koenig, B.L. and Xin, S., 2012. Real or fake? Human judgments about photographs and computer-generated images of faces. In *SIGGRAPH Asia 2012 technical briefs* (pp. 1-4).

Williams, J.J., Tien, R.N., Inoue, Y. and Schwartz, A.B., 2016, August. Idle state classification using spiking activity and local field potentials in a brain computer interface. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 1572-1575). IEEE.

Ponce, P., Molina, A. and Grammatikou, D., 2016. Design based on fuzzy signal detection theory for a semi-autonomous assisting robot in children autism therapy. *Computers in Human Behavior*, *55*, pp.28-42.

Supplementary:

Valderrama, A.T., Paclik, P., Vansteensel, M.J., Aarnoutse, E.J. and Ramsey, N.F., 2012. Error probability of intracranial brain computer interfaces under non-task elicited brain states. *Clinical neurophysiology*, *123*(12), pp.2392-2401.