16.400 / 16.453J / 2.181J

Human Factors Engineering
Fall, 2002

Signal Detection Theory and Vigilance

Jim Kuchar

Two In-Class Experiments Today

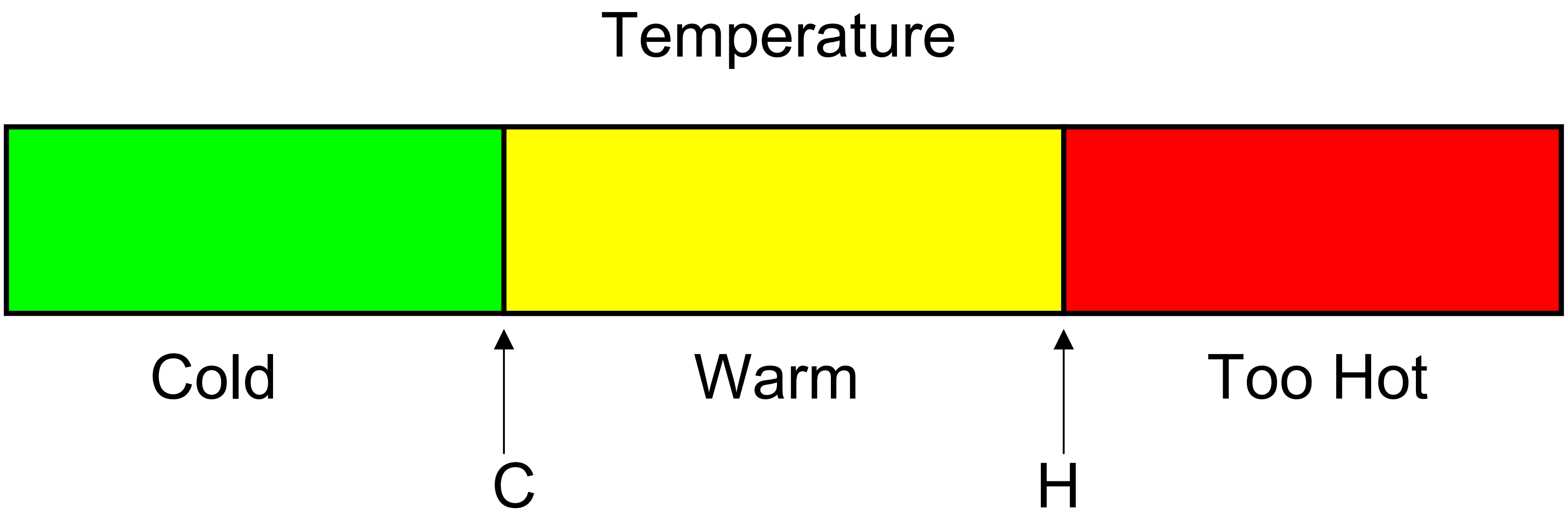
1. Hot / Cold decision-making (signal detection theory)

2. Vigilance

   Throughout today's lecture:

   Quietly keep a count of each time you see a sudden "x" or "+"

   Keep separate counts for "x" and "+"



Measurement contains noise (p=.5 C, p=.5 H)

Decide if temperature is truly too hot (**H**) or cold (**C**)

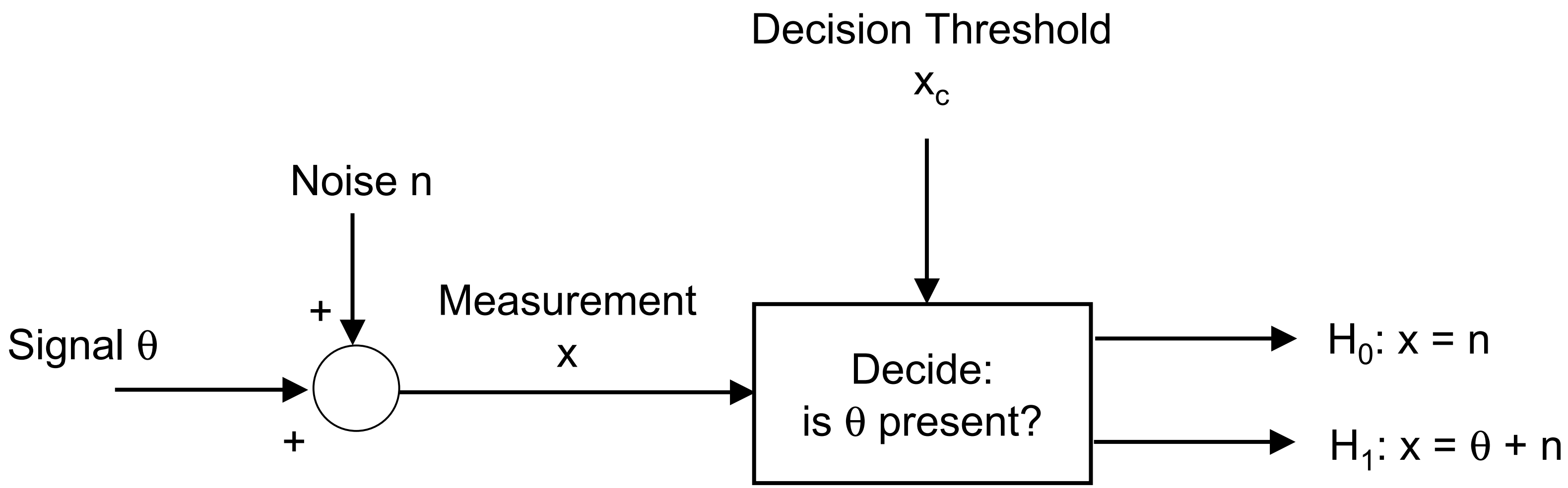
Lose $1 if you decide **H** but actually is **C**

Lose $2 if you decide **C** but actually is **H**

Signal Detection Theory (SDT)

- Formal model of information and decision-making criteria
- Allows for optimal decision-making under certain conditions
- Provides model of factors affecting human decision-making behavior
- History
  - Originally developed in late 1940s for radar detection problems
  - Extended to human perception and decision-making in 1960s
  - 1990s extended to complex human/automation alerting problems

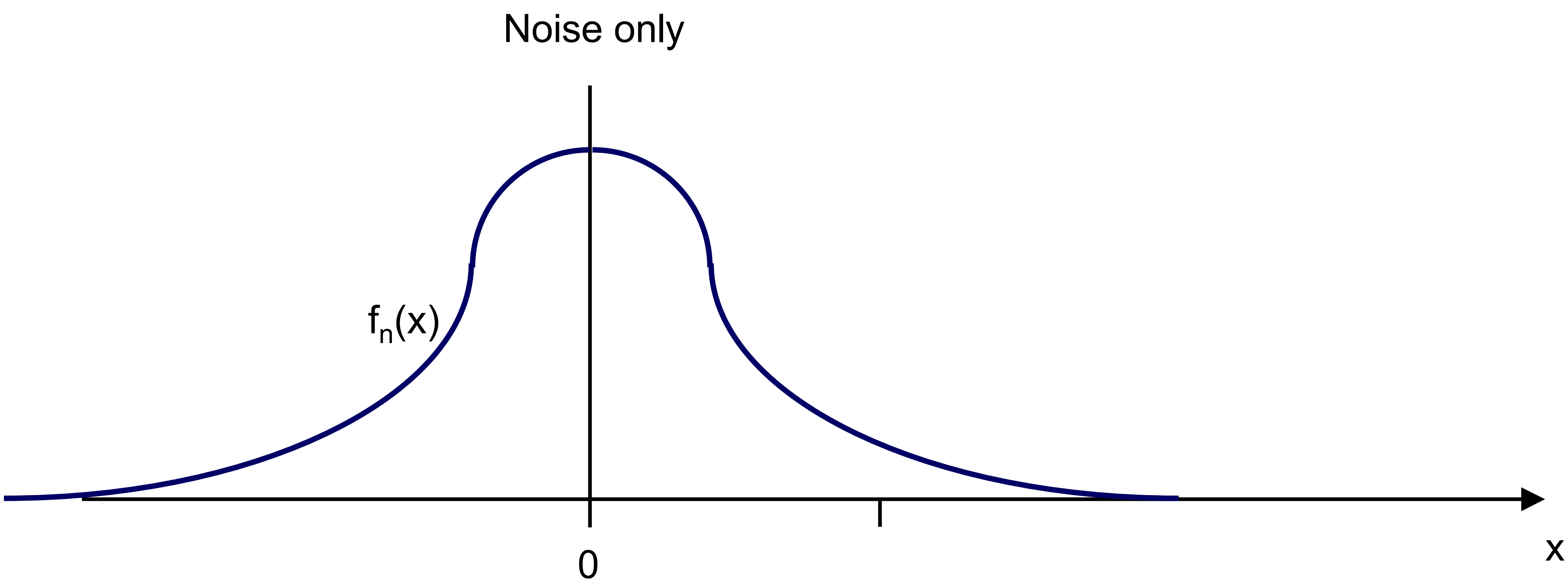
Information Flow



Basic SDT Problem

- Single known value of signal θ
- Random zero-mean normally-distributed additive noise n
- Single observable measurement x is taken
- Must decide either H_0 or H_1

Model of Signal and Noise

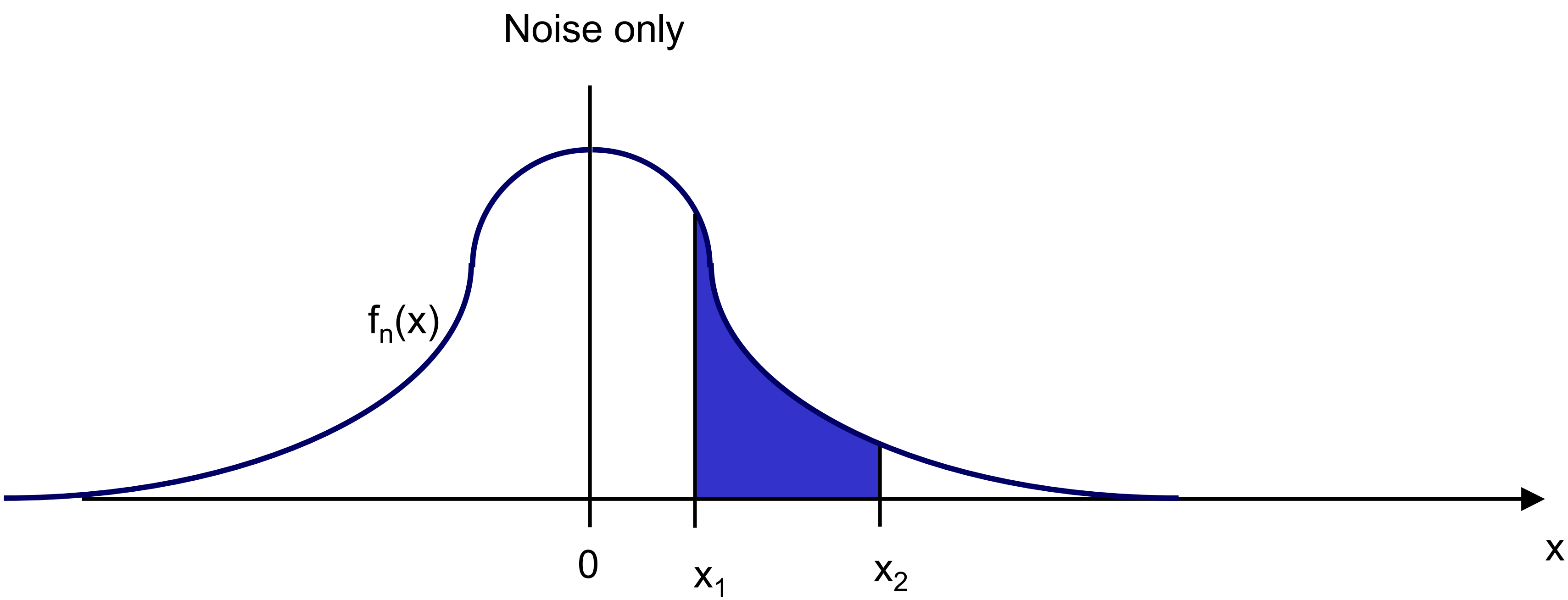


$f_n(x)$: probability density function of noise

Normally-distributed, zero mean (σ = standard deviation):

$$f_n(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

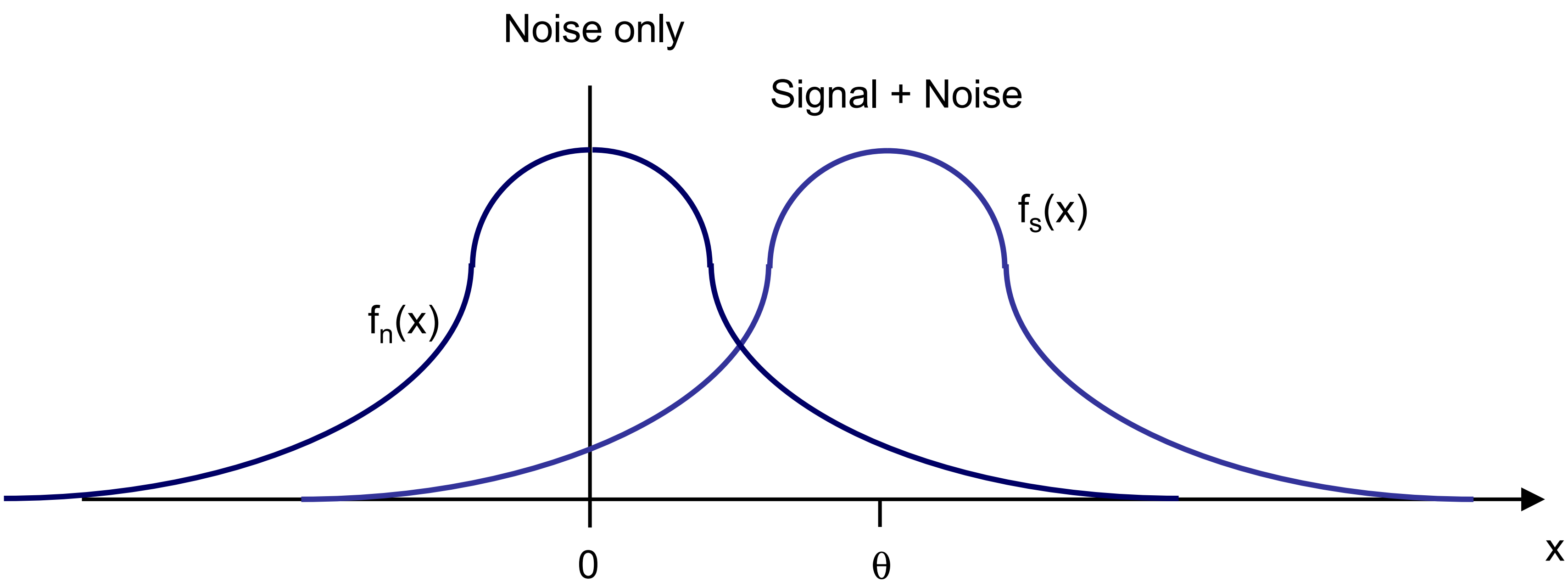
Model of Signal and Noise



probability that a single random noise value would be between x_1 and x_2:

$$P(x_1 \leq x < x_2) = \int_{x_1}^{x_2} f_n(x)\,dx$$

Model of Signal and Noise

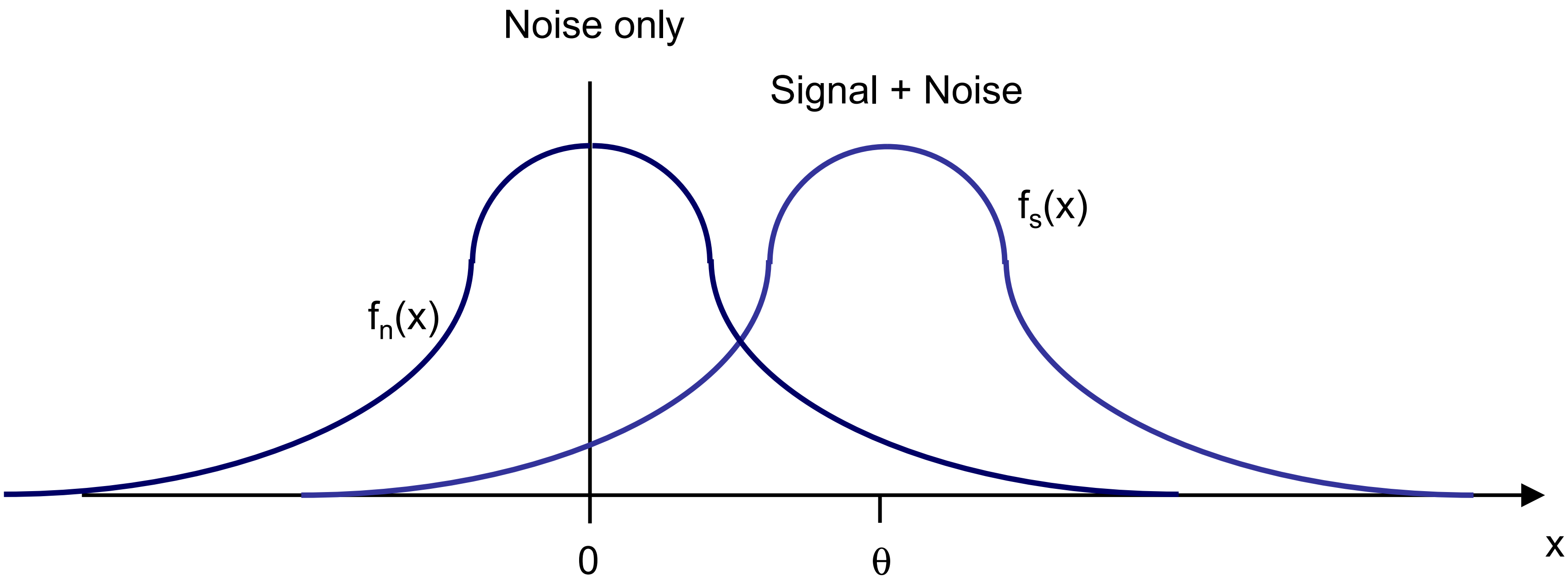


$f_s(x)$: probability density function of noise + signal

Normally-distributed, mean θ (σ = standard deviation):

$$f_s(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\theta)^2}{2\sigma^2}}$$

Sensitivity d’

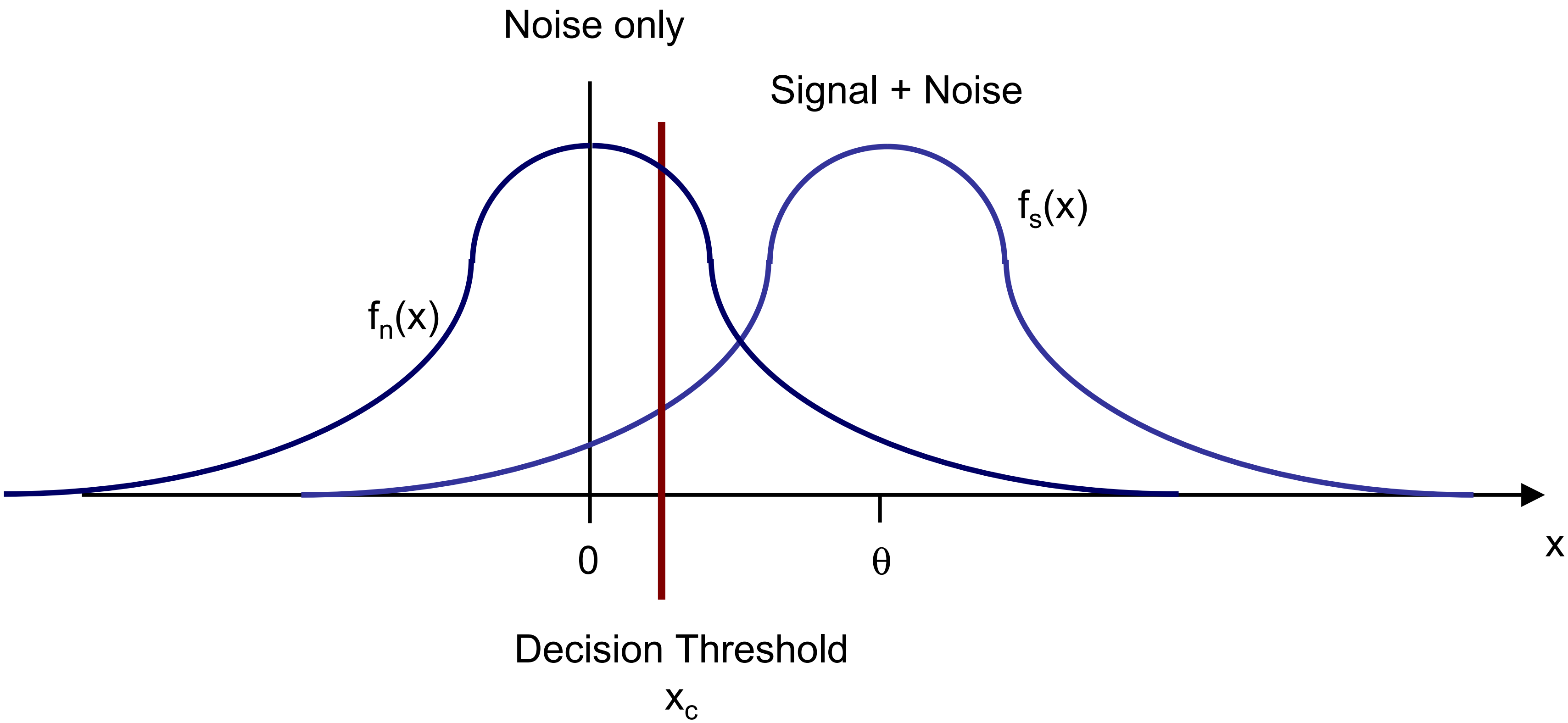


Sensitivity (signal-to-noise ratio):

$$\mathrm{d}' = \frac{\theta}{\sigma}$$

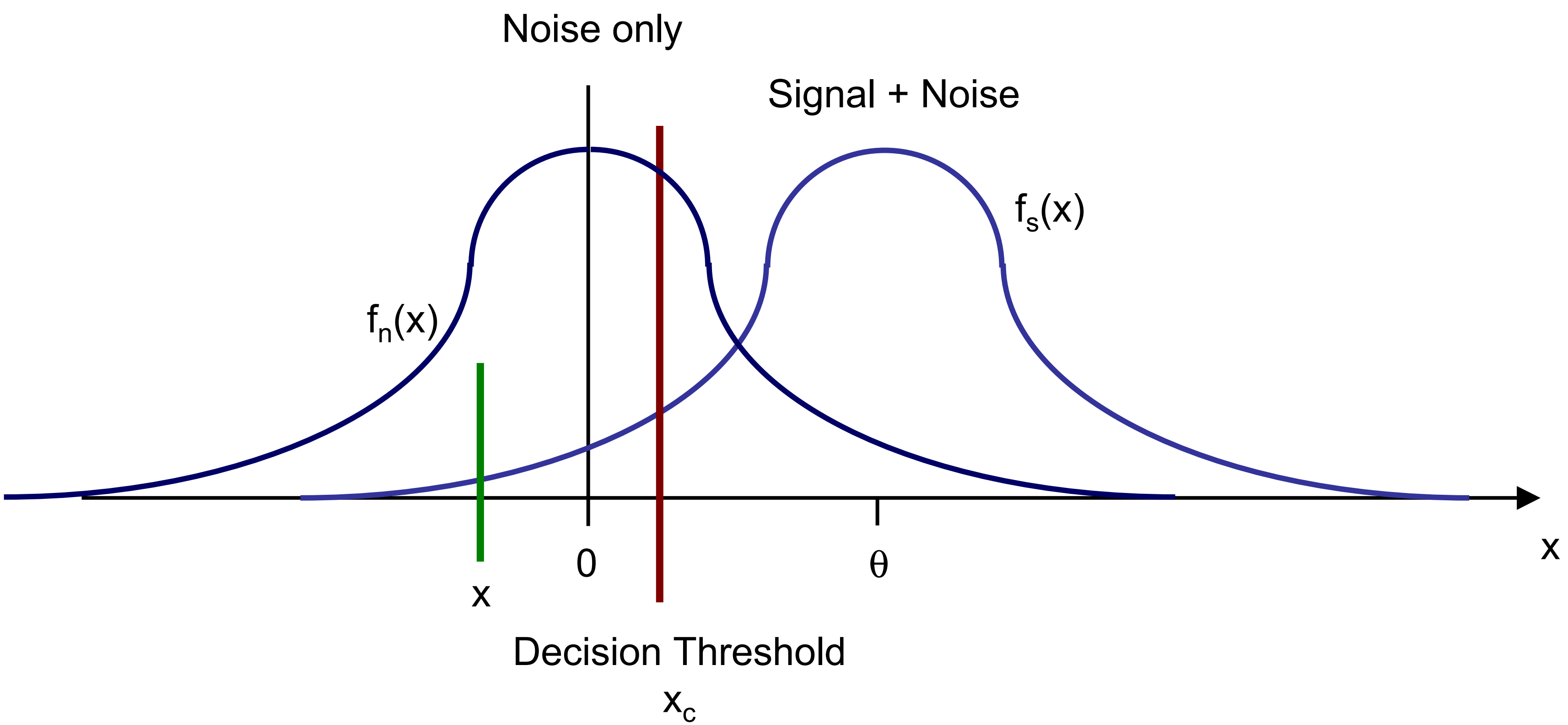
(separation between noise-only and signal+noise measured in standard deviations)
If $\sigma = 1$, then $d' = \theta$

Model of Thresholded Decision



Decision criterion:
If $x > x_c$, decide H_1 (signal + noise)
Otherwise, decide H_0 (noise only)

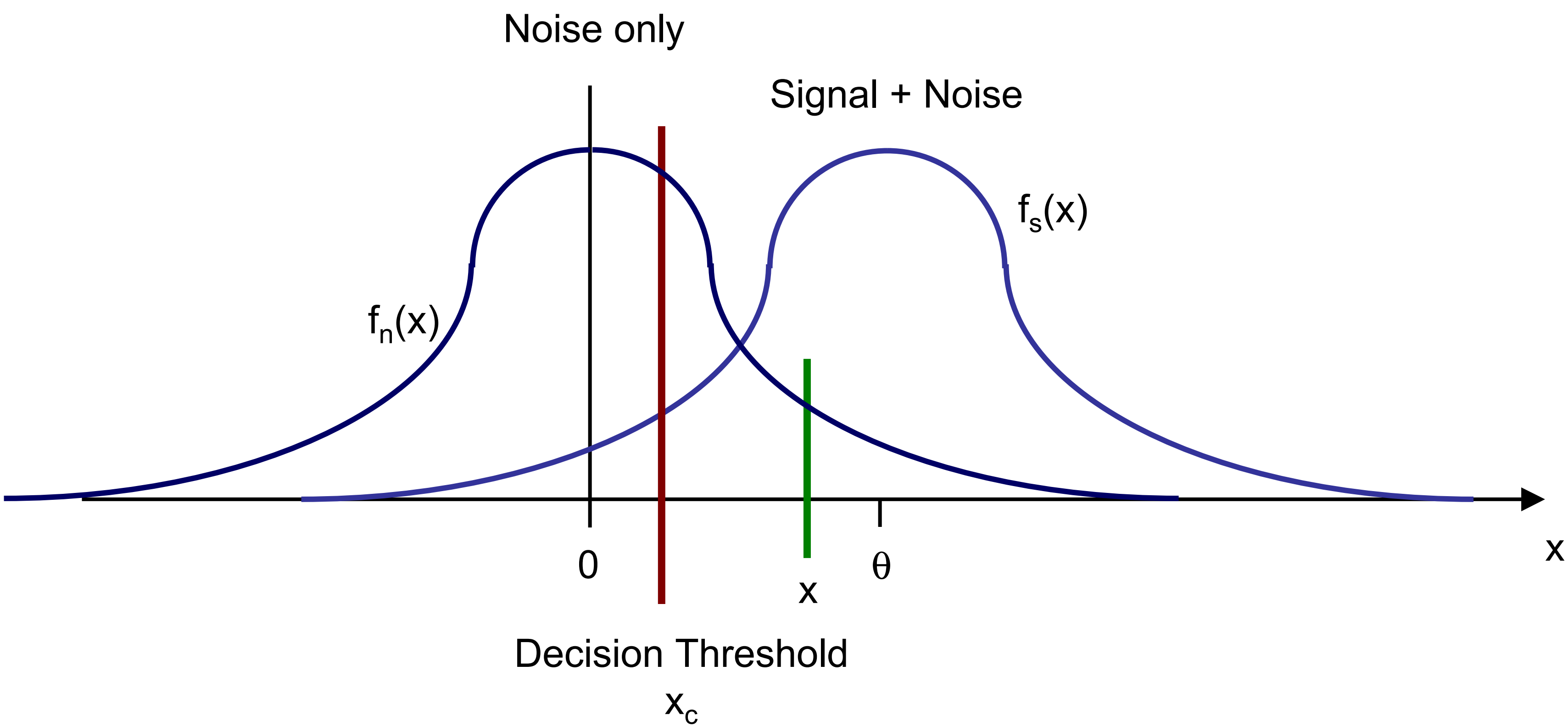
Model of Thresholded Decision



Decide Noise only

Decision criterion:
If $x > x_c$, decide H_1 (signal + noise)
Otherwise, decide H_0 (noise only)

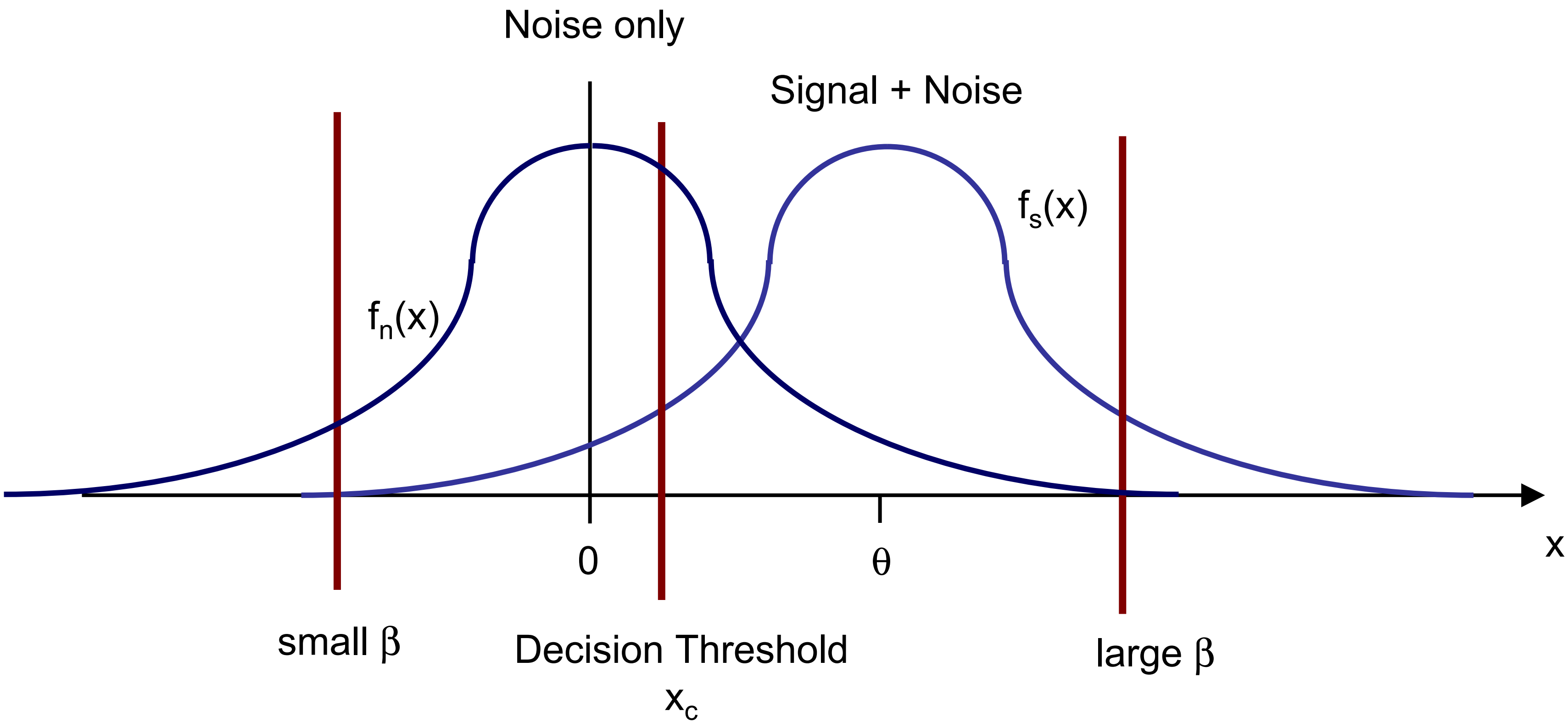
Model of Thresholded Decision



Decide Signal + Noise

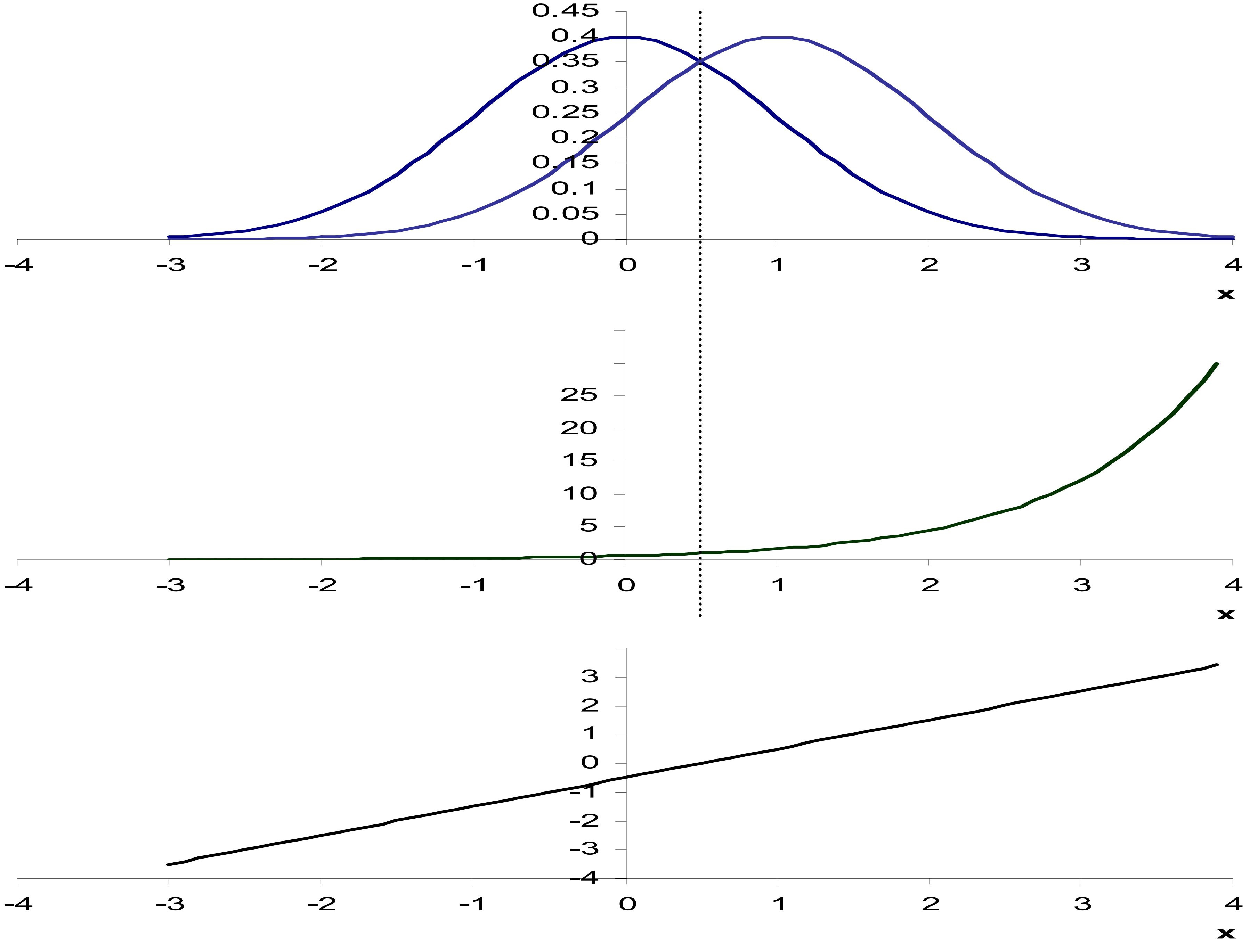
Decision criterion:
If $x > x_c$, decide H_1 (signal + noise)
Otherwise, decide H_0 (noise only)

Decision Bias: β



$$\beta = \frac{f_s(x_c)}{f_n(x_c)}$$

relative likelihood of observing the threshold value x_c if the signal is present to observing x_c if the signal is not present

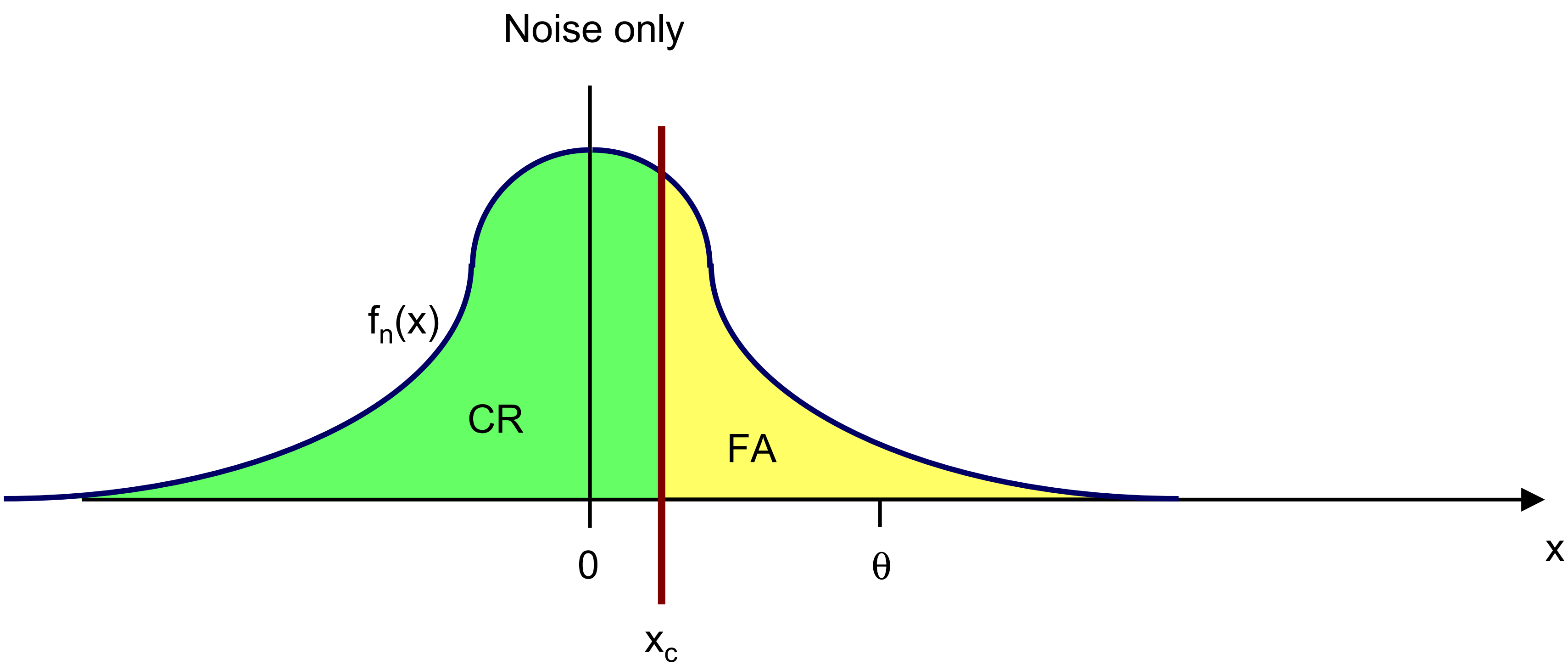


0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-4
-3
-2
-1
0
1
2
3
4
x
25
20
15
10
5
0
-4
-3
-2
-1
0
1
2
3
4
x
3
2
1
0
-1
-2
-3
-4
-4
-3
-2
-1
0
1
2
3
4
x

Decision Outcomes

True State		Decision: Noise only (H_0)	Decision: Signal + Noise (H_1)
True State	Noise only (H_0)	Correct Rejection (CR)	False Alarm (FA) Type I Error False Positive
True State	Signal + Noise (H_1)	Missed Detection (MD) Type II Error False Negative	Correct Detection (CD)

Decision Outcomes



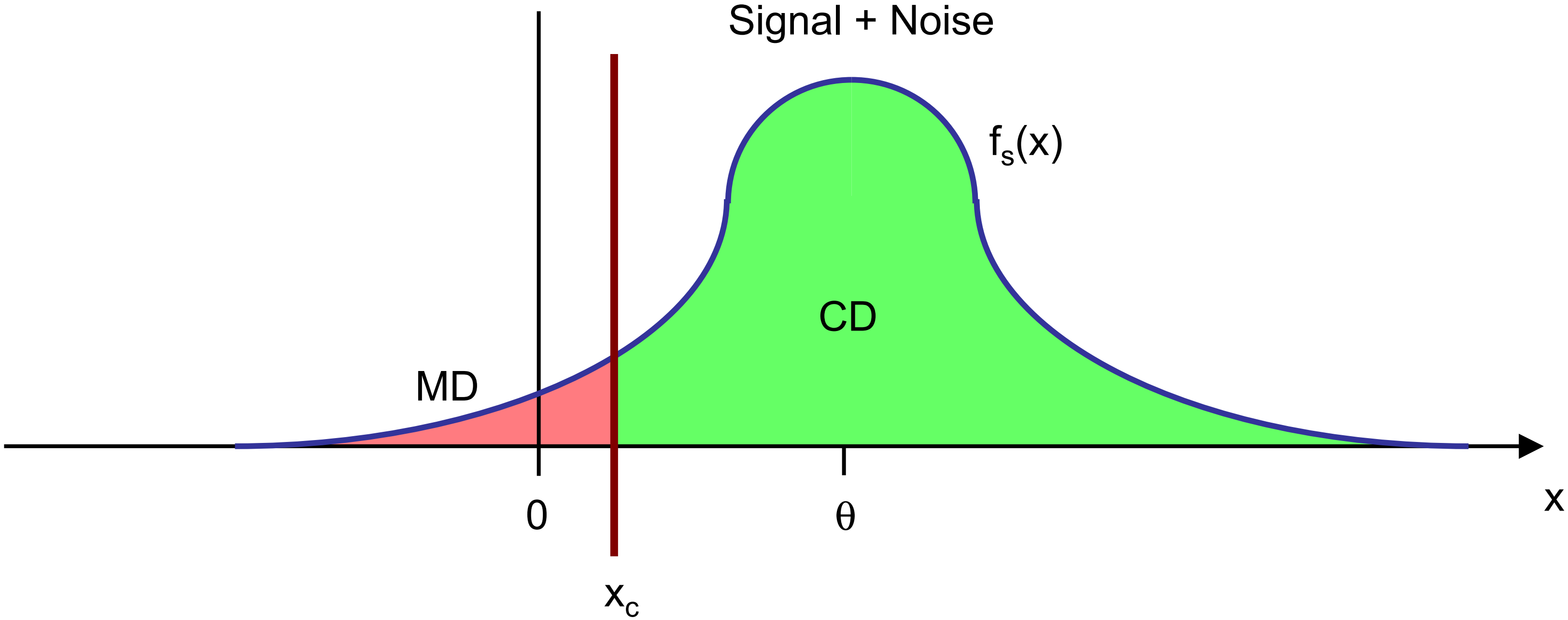
If truly noise only,

$$P(FA \mid N) = \int_{x_c}^{\infty} f_n(x)\, dx$$

$$P(CR \mid N) = \int_{-\infty}^{x_c} f_n(x)\, dx$$

(want x_c as large as possible)

Decision Outcomes



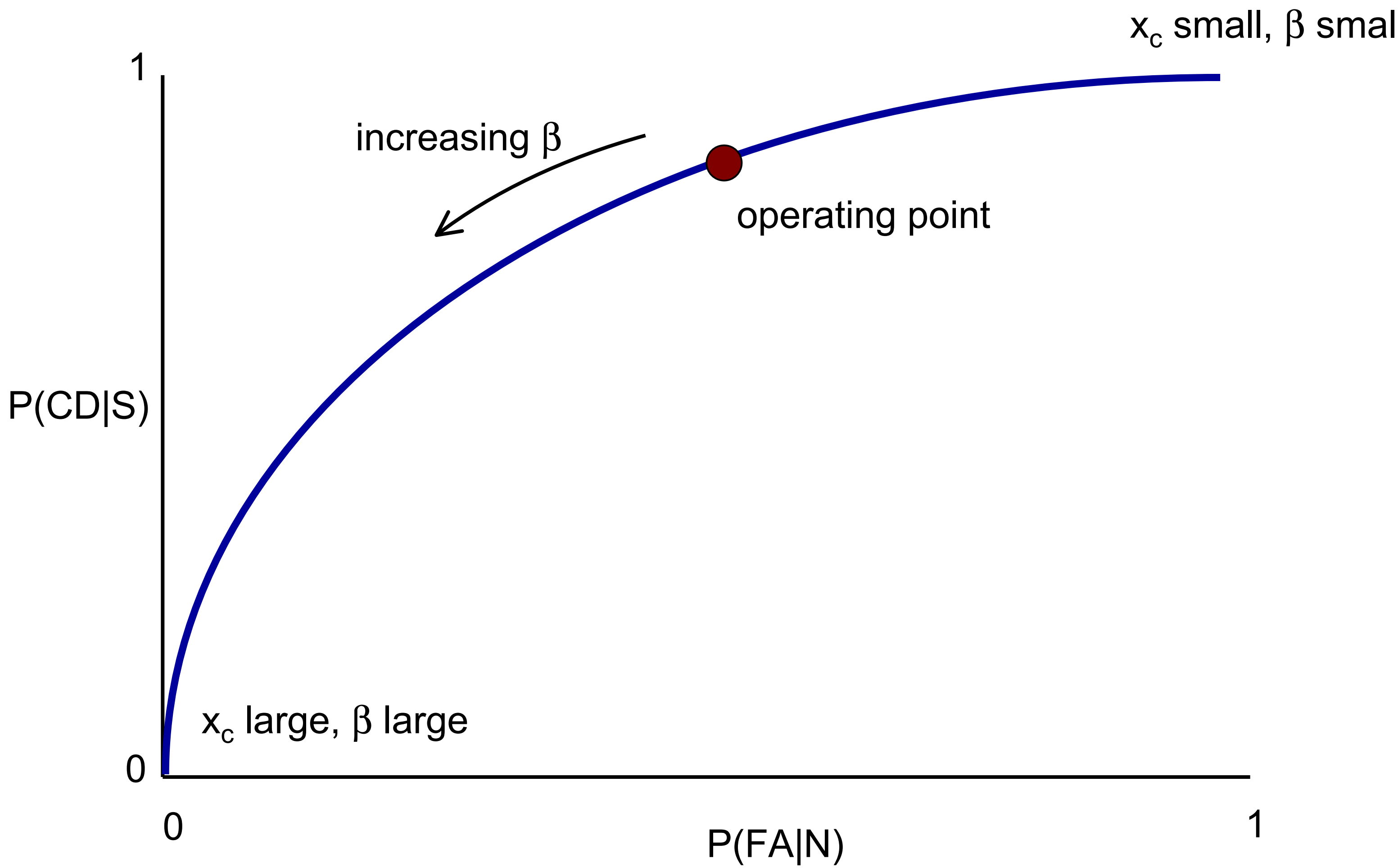
If signal is truly present,

$$P(CD \mid S) = \int_{x_c}^{\infty} f_s(x)\,dx$$

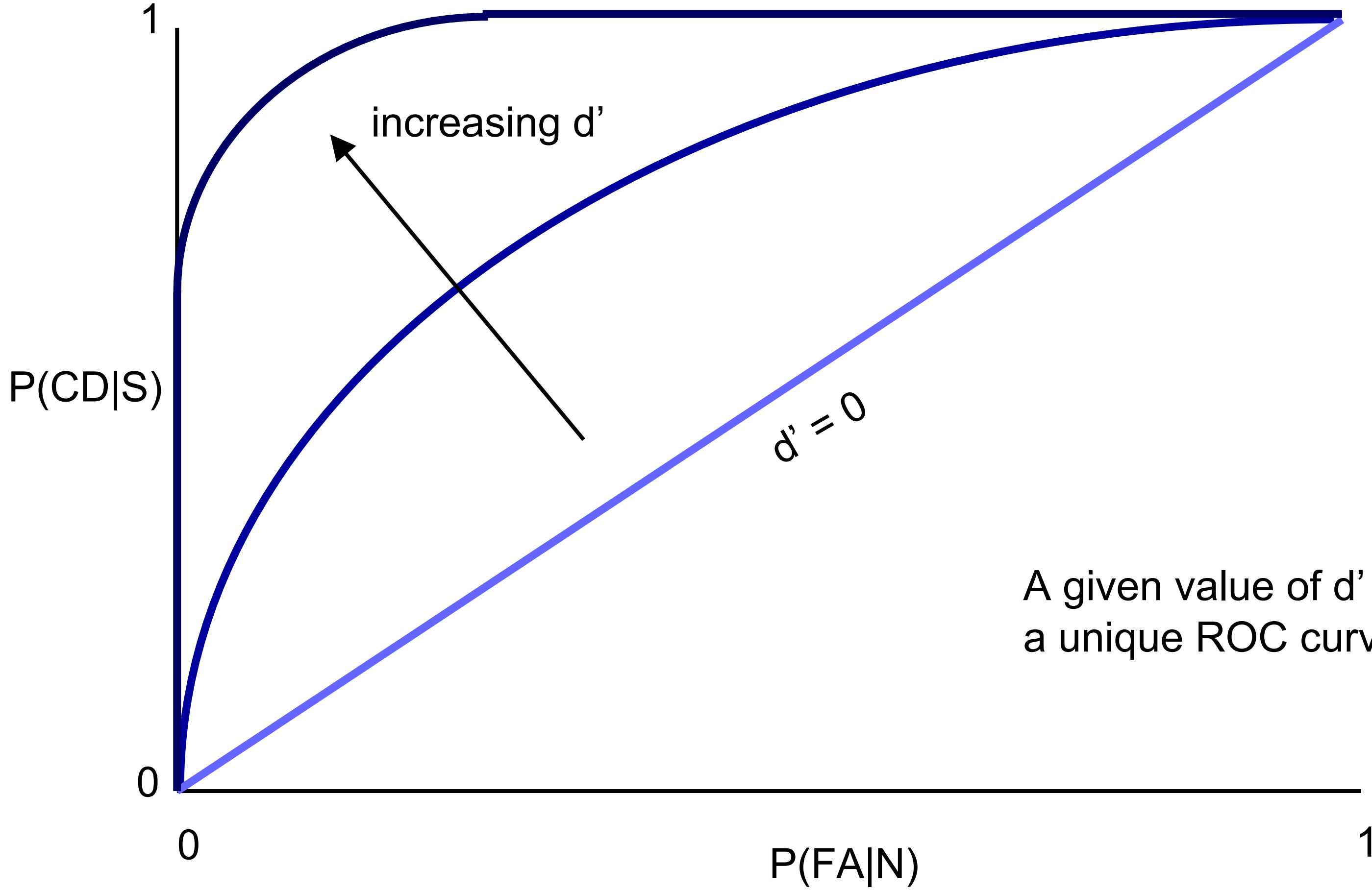
$$P(MD \mid S) = \int_{-\infty}^{x_c} f_s(x)\,dx$$

(want x_c as small as possible)

Receiver Operating Characteristic (ROC)



Receiver Operating Characteristic (ROC)



A given value of d' traces out a unique ROC curve as β varies

Optimal Threshold Placement

- Requirements:
  - Probability density functions for noise and signal+noise
  - Probability of signal being present, P(S)
  - Payoff matrix: costs & values associated with each of the 4 outcomes V(CR), V(CD), V(FA), V(MD)

Probabilities of Outcomes

Our earlier definitions of P(FA|N), P(MD|S), P(CR|N), P(CD|S) are **conditional probabilities**

That is, P(FA|N) is the probability of a false alarm occurring **if we are given the fact that no signal is present**

P(FA|N) is NOT the unconditional probability (or expected frequency) of a false alarm occurring when we don't know whether the signal is present

Example: say β is very small, so P(FA|N) $\approx$ 1 but the signal is almost always present. Then we would expect very few false alarms to actually occur even though P(FA|N) = 1.

Unconditional Probabilities of Outcomes

$$P(FA) = P(FA \mid N)\, P(N) = P(FA \mid N)\, (1-P(S))$$

$$P(CR) = P(CR \mid N)\, P(N) = P(CR \mid N)\, (1-P(S))$$

$$P(MD) = P(MD \mid S)\, P(S)$$

$$P(CD) = P(CD \mid S)\, P(S)$$

Expected Value of Decision-Making

J = (Probability of outcome) x (Value of outcome)
summed over all outcomes

J = P(FA | N) (1-P(S)) V(FA) + P(CR | N) (1-P(S)) V(CR) +
P(MD | S) P(S) V(MD) + P(CD | S) P(S) V(CD)

Maximize value / minimize cost:

$$\frac{\partial J}{\partial x_c} = \frac{\partial P(FA|N)}{\partial x_c}(1-P(S))\,V(FA) + \frac{\partial P(CR|N)}{\partial x_c}(1-P(S))\,V(CR) + \frac{\partial P(MD|S)}{\partial x_c}P(S)\,V(MD) + \frac{\partial P(CD|S)}{\partial x_c}P(S)\,V(CD) = 0$$

Optimal Threshold

$$\frac{\partial P(FA \mid N)}{\partial x_c} = \frac{\partial}{\partial x_c} \int_{x_c}^{\infty} f_n(x)\, dx = -f_n(x_c)$$

$$\frac{\partial P(CR \mid N)}{\partial x_c} = \frac{\partial}{\partial x_c} \int_{-\infty}^{x_c} f_n(x)\, dx = f_n(x_c)$$

$$\frac{\partial P(MD \mid S)}{\partial x_c} = \frac{\partial}{\partial x_c} \int_{-\infty}^{x_c} f_s(x)\, dx = f_s(x_c)$$

$$\frac{\partial P(CD \mid S)}{\partial x_c} = \frac{\partial}{\partial x_c} \int_{x_c}^{\infty} f_s(x)\, dx = -f_s(x_c)$$

Optimal Threshold

$$\frac{\partial J}{\partial x_c} = -f_n(x_c)\,(1 - P(S))\,V(FA) + f_n(x_c)\,(1 - P(S))\,V(CR) +$$

$$f_s(x_c)\,P(S)\,V(MD) + -f_s(x_c)\,P(S)\,V(CD) = 0$$

$$\frac{f_s(x_c)}{f_n(x_c)} = \frac{1 - P(S)}{P(S)}\,\frac{V(FA) - V(CR)}{V(MD) - V(CD)}$$

β*

$$\beta^* = \left[\frac{1 - P(S)}{P(S)}\right]\left[\frac{V(FA) - V(CR)}{V(MD) - V(CD)}\right]$$

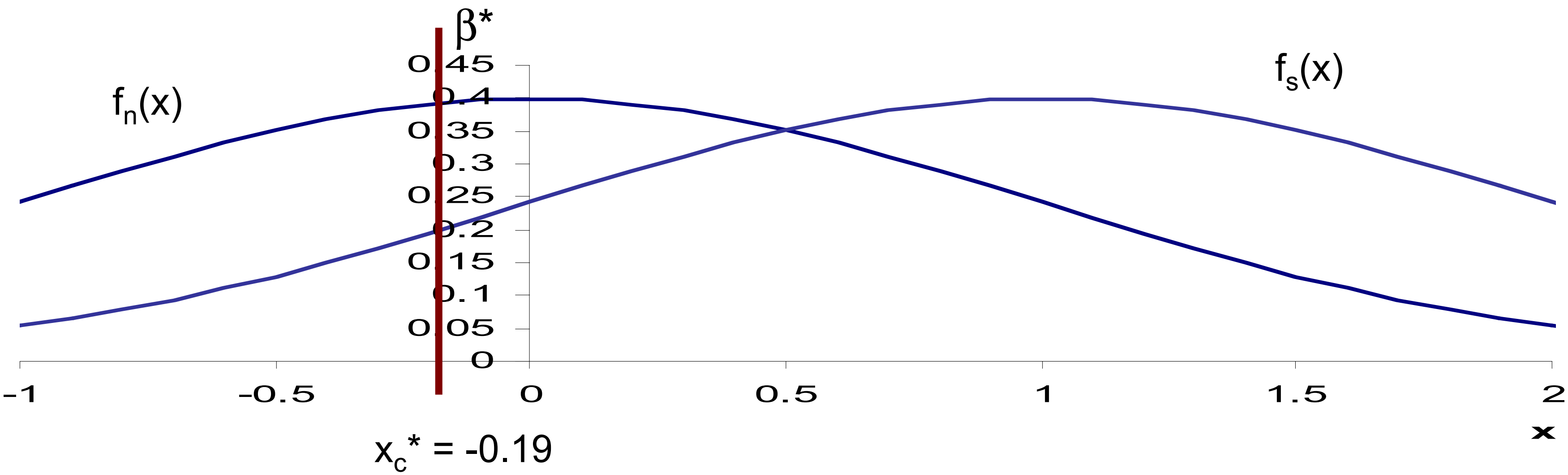
Optimal Threshold

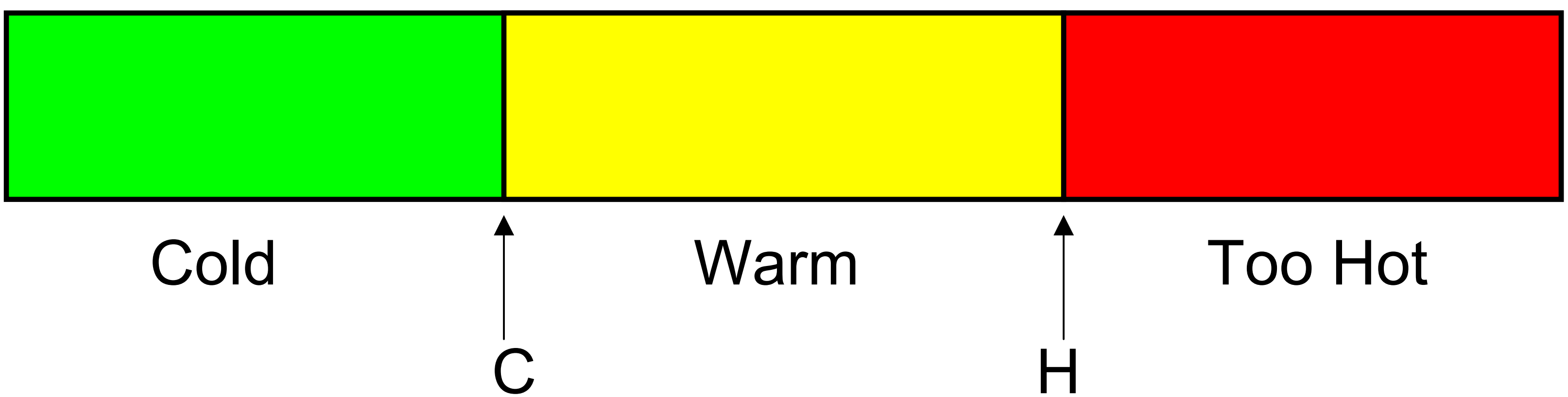
$$\beta^* = \left[\frac{1-P(S)}{P(S)}\right]\left[\frac{V(FA)-V(CR)}{V(MD)-V(CD)}\right]$$

or, in terms of the threshold setting x_c

$$x_c^* = \frac{\ln \beta^*}{\theta} + \frac{\theta}{2}$$

when $\sigma = 1$

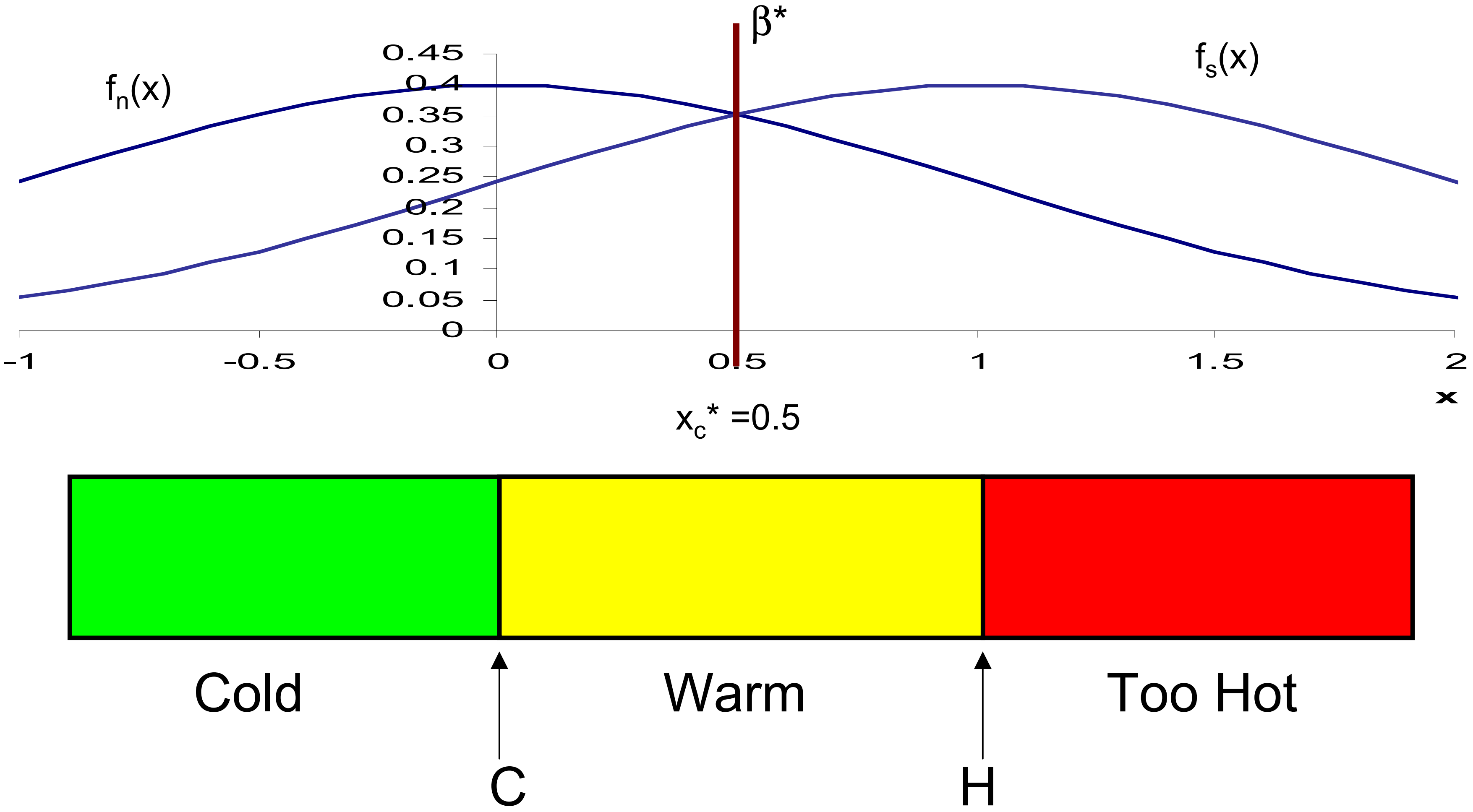




$P(S) = 0.5, \quad 1\text{-}P(S) = 0.5$

$V(FA) = 1, \; V(MD) = 2, \; V(CD) = V(CR) = 0$

$\beta^* = 0.5$, Expected payoff = $4

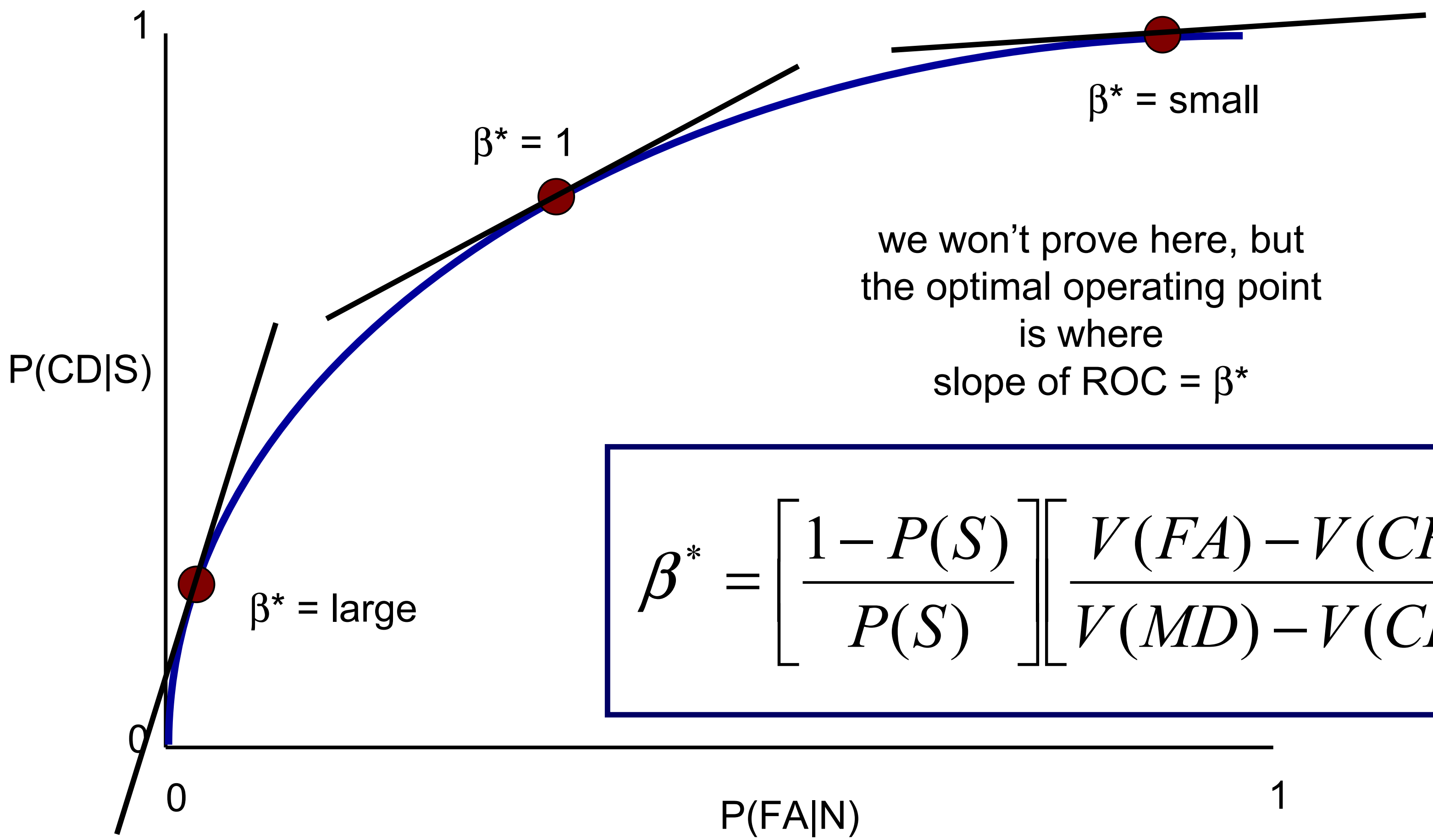


P(S) = 0.5, 1-P(S) = 0.5

IF V(FA) = 1, V(MD) = 1, V(CD) = V(CR) = 0

$\beta^* = 1.0$

Receiver Operating Characteristic (ROC)

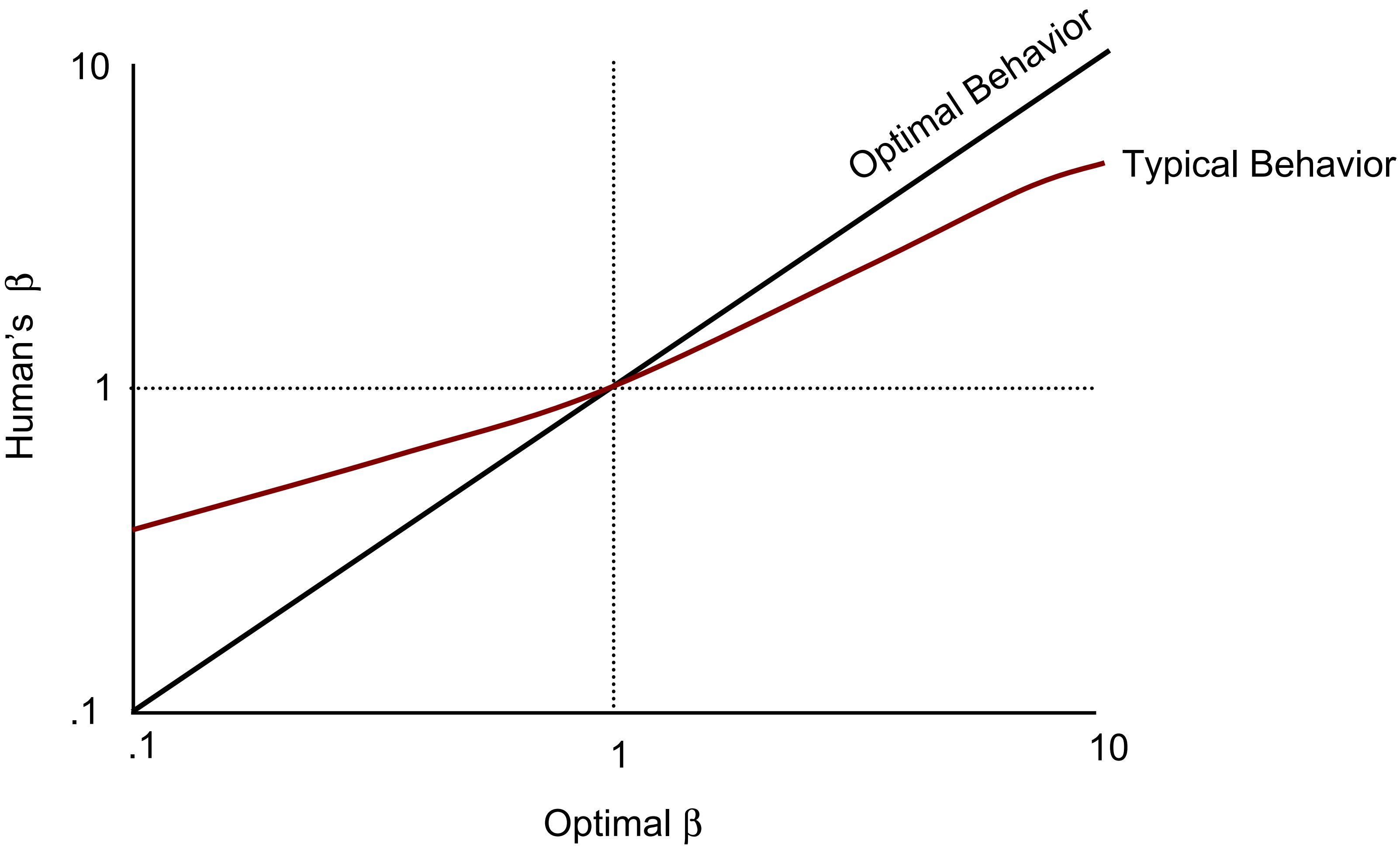


we won’t prove here, but
the optimal operating point
is where
slope of ROC = β*

$$\beta^* = \left[\frac{1 - P(S)}{P(S)}\right]\left[\frac{V(FA) - V(CR)}{V(MD) - V(CD)}\right]$$

Human Performance From SDT Point of View

Humans tend to set β closer to 1 than is optimal: “sluggish beta”

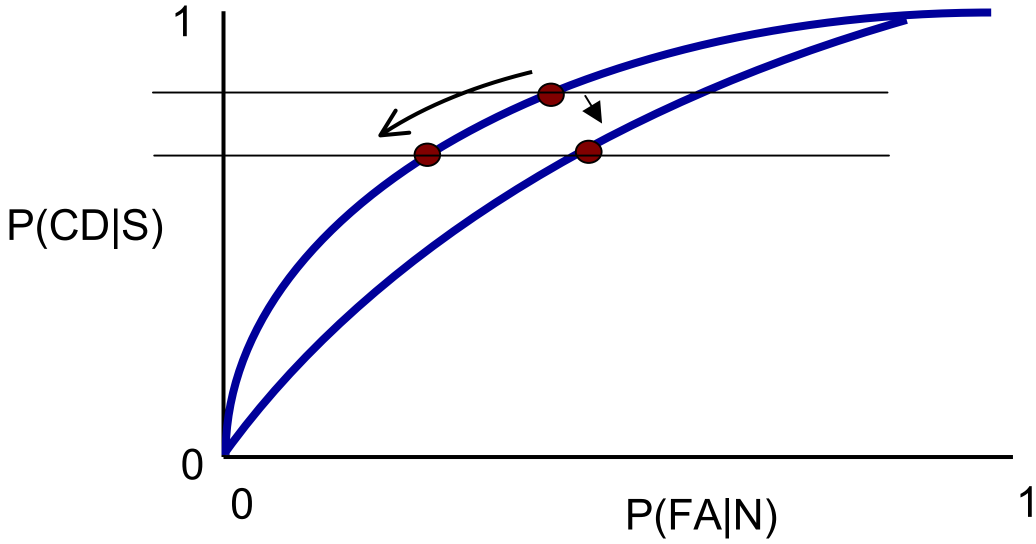


Application of SDT: Vigilance

Human observer must detect periodic (but random) rare event

Vigilance decreases over time: “vigilance decrement”

- Sensitivity decrement (d’ decreases)
- Beta increment (CD rate reduces)



Sensitivity Decrement

In a sustained attention task,

- Fatigue increases
- Subject looks away more often, misses signals
- Mental workload increases, which may result in reduced sensitivity

SDT Explanation of β Increment in Vigilance

$$\beta^* = \left[\frac{1-P(S)}{P(S)}\right]\left[\frac{V(FA)-V(CR)}{V(MD)-V(CD)}\right]$$

Human misses a signal

Estimate of P(S) decreases

Results in increase in β*

Human less likely to have CD or FA
more likely to miss another signal, leads to vicious circle

“expectancy theory”

Mitigating Losses in Vigilance

- Sensitivity loss
  - Refresh with examples of the target
    - reduces memory load remembering what the target is
  - Increase target salience
    - blinking, audio cues, etc.
    - may distract or otherwise interfere with other activities
    - automation may not be entirely accurate
  - Reduce event rate or allow subjects to vary event rate
    - e.g., assembly line
  - Training
    - automaticity reduces workload (but don’t overdo it!)

Mitigating Losses in Vigilance

- Response bias shift
  - Consistent instructions
  - Feedback on success / failure at the task
  - Introduce false signals
  - Allow use of confidence level in decision

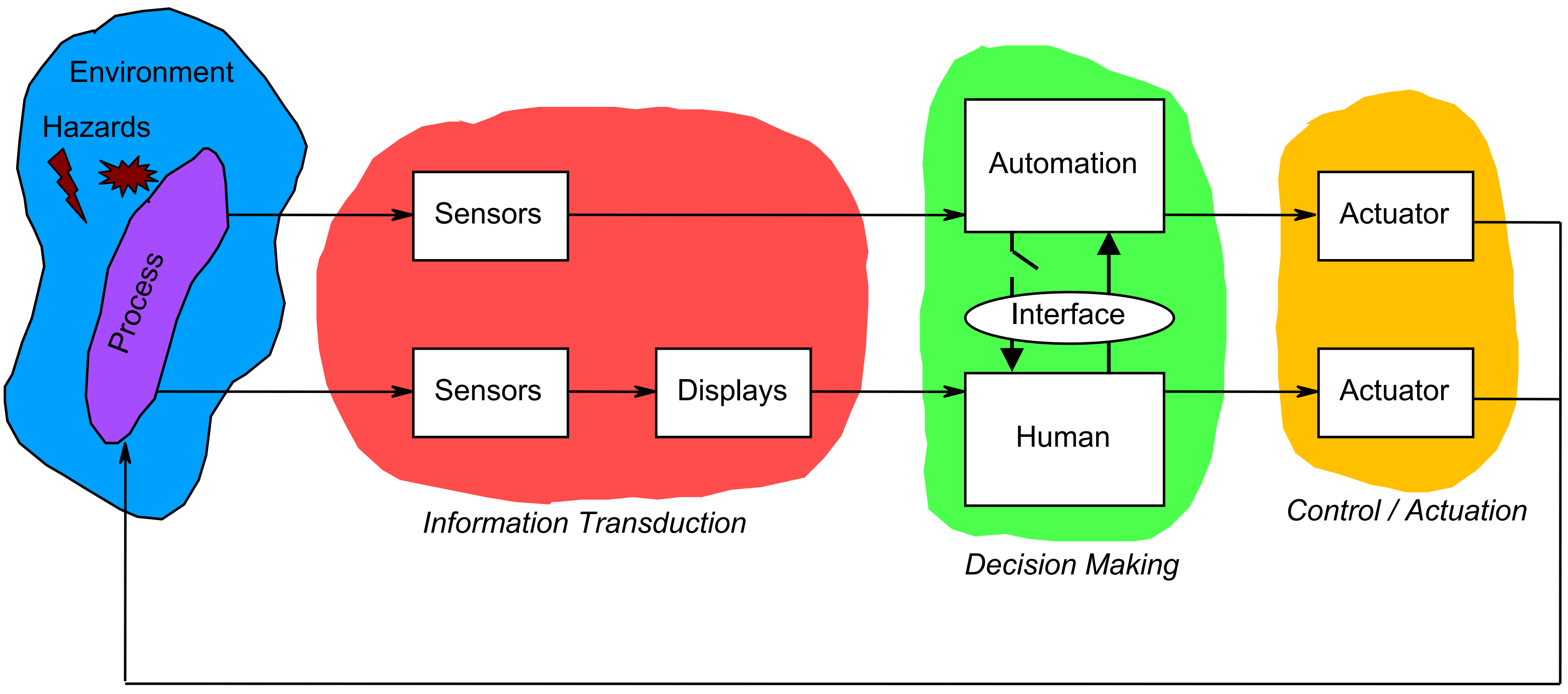
16.400 / 16.453J / 2.181J

Human Factors Engineering
Fall, 2002

Alerting and Warning Systems

Jim Kuchar

Alerting System Block Diagram



Why Have an Alerting System?

- Human can’t observe a threat
  - seeing mountains through clouds
- Human isn’t observing a threat
  - physically not attending to information (looking down instead of out the window)
  - limited / no spare attention (sensory or mental load)
- Human may not understand level of risk
  - deliberate low-level flight
- Human may be taking the wrong action
  - turning left instead of right

Alerting System Functions

- Detect event or condition that warrants awareness or action
- Attract attention of the human
- Convey the nature of the situation
- Convey the urgency of the situation
- Provide information to help in correcting the situation
- Have a means for inhibiting or acknowledging alerts

Example: Fire Alarm

- Detect event or condition that warrants awareness or action
  - threshold level of smoke particulate concentration
- Attract attention of the human
  - audible bell + flashing strobe lights
- Convey the nature of the situation
  - implicit / encoded in the alarm bell + lights
- Convey the urgency of the situation
  - implicit
- Provide information to help in correcting the situation
  - implicit
- Have a means for inhibiting or acknowledging alerts
  - only for authorized personnel (fire fighters)

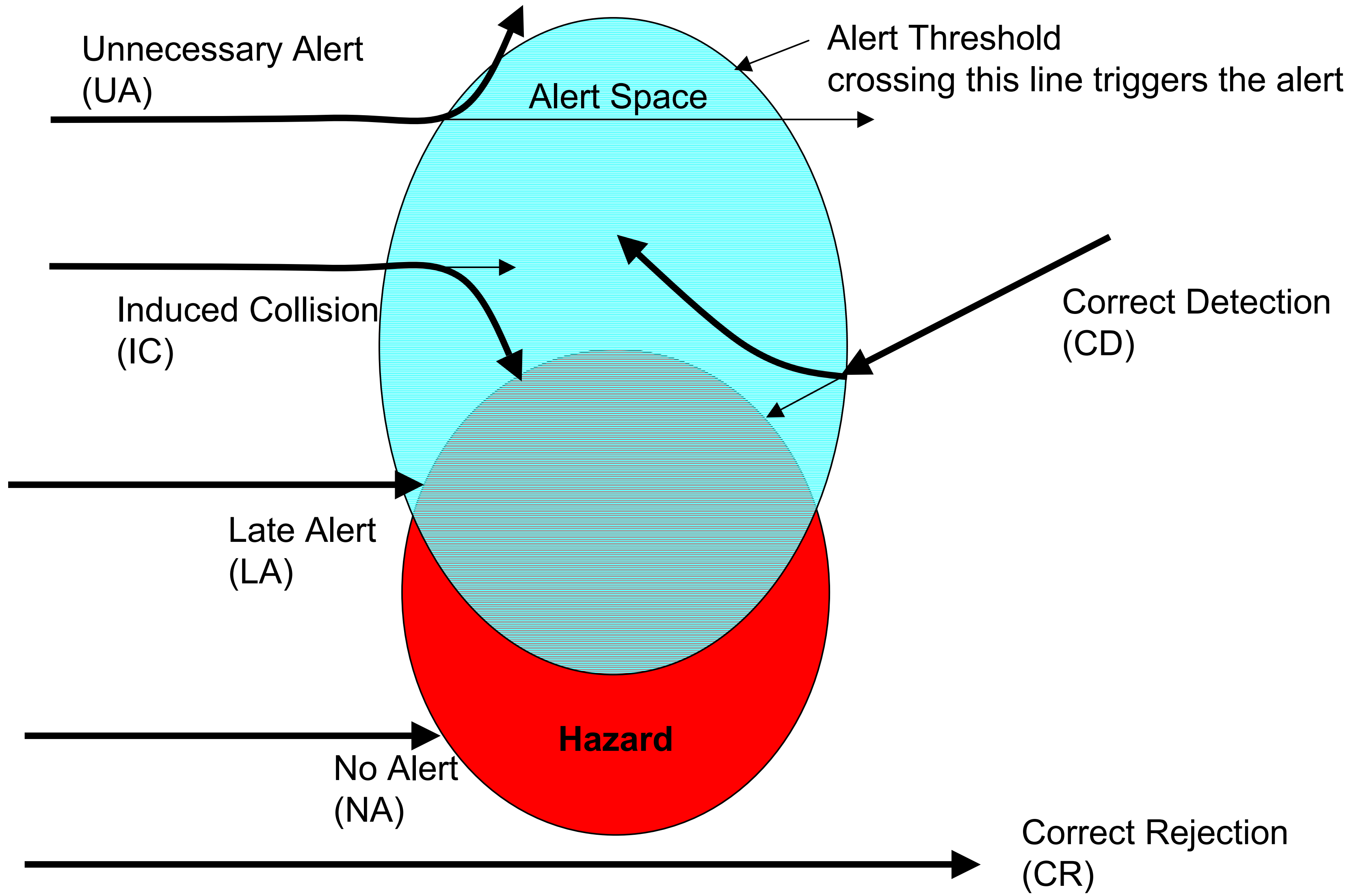
Example: Traffic Collision Alerting System

- Detect event or condition that warrants awareness or action
  - projected collision with another aircraft within 20 seconds
- Attract attention of the human
  - “Traffic Traffic”, “Climb! Climb!”
  - Visual traffic display with colored icons for other aircraft
- Convey the nature of the situation
  - implicit / encoded in the display
- Convey the urgency of the situation
  - “Traffic”, yellow colors = caution, action may be needed
  - “Climb!”, red colors = act immediately
- Provide information to help in correcting the situation
  - pitch guidance given to pilot on display
- Have a means for inhibiting or acknowledging alerts
  - pilot can inhibit the system or pull the circuit breaker

Engineering Design Issues

- Tension between who knows best: human or automation?
  - is the human already aware of the threat?
  - does the human need assistance in resolving the threat?
  - what can't the automation observe that the human can?
  - varying levels of personal risk acceptance or judgment of risk
- How to most efficiently transmit information?
  - Time-critical
  - Safety / life-critical
  - High stress, high workload
  - False alarms have large negative effect

6 Alerting Outcomes



Attracting Attention

- Modes
  - Visual
  - Auditory
  - Tactile
  - Olfactory

- To be more effective at attracting attention
  - Increase signal level
  - Use multiple modalities

- BUT: Tradeoff between being really good at attracting attention vs. being too distracting
  - False alarms - disturb other tasks, reduce confidence
  - Multiple / concurrent alarms - information overload

Levels of Urgency

- Supplemental
  - system status
- Advisory
  - require crew awareness, may require action
- Caution
  - require immediate awareness and prompt action. If uncorrected, a warning will occur
- Warning
  - require immediate corrective action, but secondary to maintaining flight
- Time-Critical Warning
  - require unconditionally immediate corrective action
  - insufficient time for use of other information than that needed to successfully correct the problem

Typical Threat Categorization

- Advisory
  - Brake overheat
  - Cabin call
  - Electrical bus off in galley
- Caution
  - Low oil pressure
  - Landing gear disagree
  - Open door
- Warning
  - Gear not down at low altitude
  - Collision avoidance
  - Ground proximity
  - Overspeed

Conveying Levels of Urgency

- Integrated
  - Attention-getting signal itself conveys the type and urgency of the threat (e.g., fire alarm)
  - Good for rapid processing
  - Limit on number of signals human can remember (7 +/- 2)

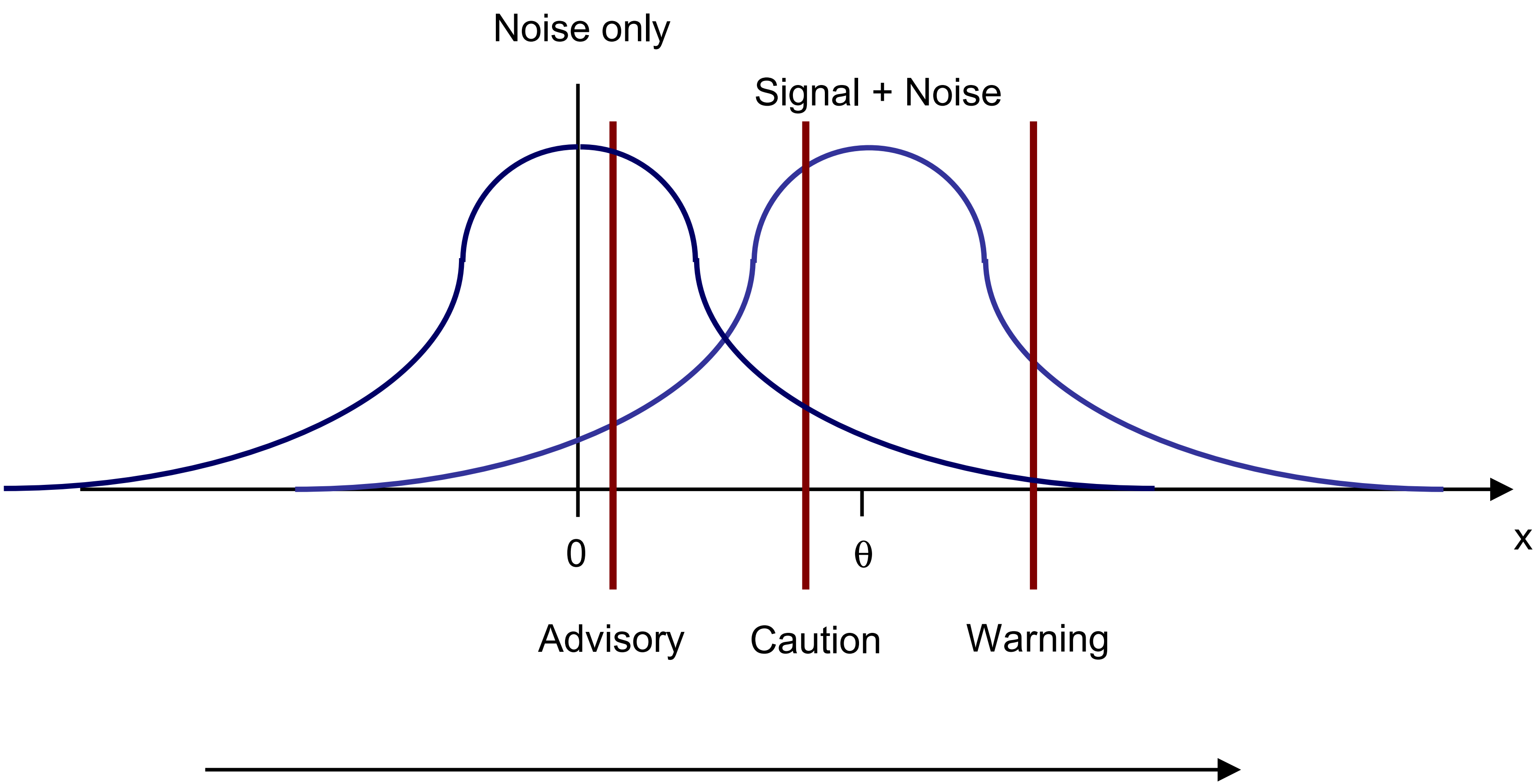
- Separate
  - Attention-getting display is different from that used to inform of the nature and urgency of the threat
  - Requires redirection of attention and interpretation
  - Able to cover many types of threats

- Aviation:
  - Master Warning / Master Caution
  - Detailed Information Display

Providing Information to Aid in Resolution

- Implicit
  - Trained response to stimulus (e.g., fire alarm)
- Explicit
  - Command is provided (e.g., “Climb!”)
- Guidance
  - Give safe target state and feedback on progress

Multiple Alert Stages



Increasing certainty that action is required
Increasing annoyance to false alarms
Progressing through stages also helps human get ready to take action

Dark Cockpit Philosophy

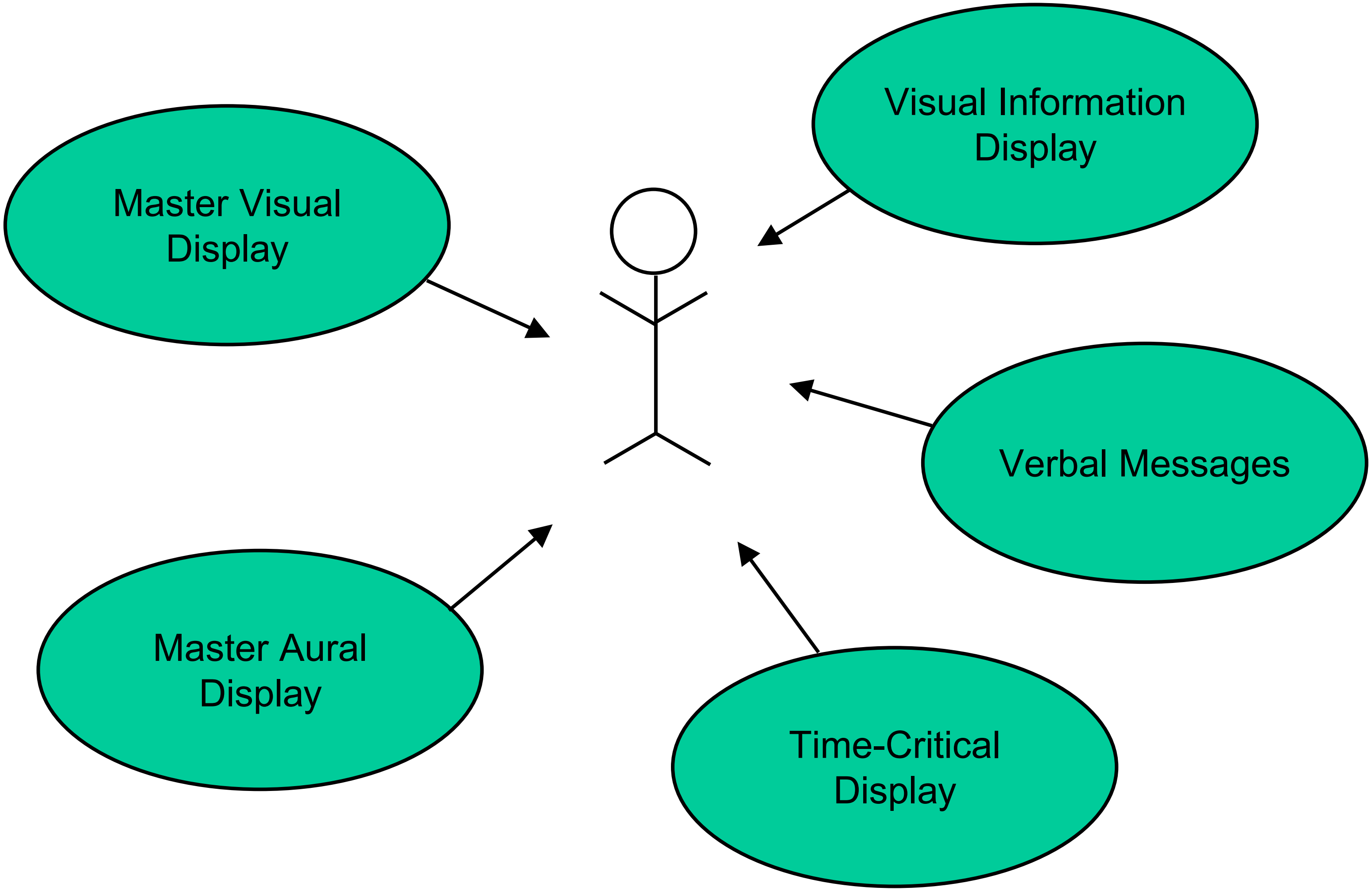
- If all is well, cockpit should be dark and quiet
  - Reduces nominal sensory and cognitive load
  - Aids in responding to new information rapidly

- Information must earn its way into the cockpit

- Operator should only have necessary information
  - Ability to obtain other information if desired
  - Requires determining functional requirements for all tasks

- Information that is presented should adhere to strict conventions
  - Color (Yellow = Caution; Red = Warning)
  - Brightness
  - Size
  - Sound level and waveforms
  - Phraseology

Aviation Warning System Architecture




GEAR DOWN
GEAR UNSAFE
LEFT FUEL
RIGHT FUEL
SPEED BRAKE
ALT AIR
PITOT HEAT
HI/LO VAC
L.ALT VOLTS
R.ALT VOLTS
START POWER
STBY VAC
PROP DEICE
BOOST PUMP
PUSH ON ROT BCN
PUSH ON TAXI LITE
PUSH ON STRB LITE
PUSH ON RECOG LITE
PUSH ON LNDG LITE
PUSH ON NAV LITE
EXTERIOR LIGHTS
AIRSPEED KNOTS
TEMP.C
ENCODING
FUEL FLOW GPH
DEST RES
ENT TST
LEFT FUEL QTY
RIGHT FUEL QTY
AMPS
PUSH VOLTS
GEAR UP
DOWN
PSI FUEL PRESS
PSI OIL PRESS
°C OAT
°F TIT
OIL TEMP
°F CHT
NAV 1
NAV 2
AP
FD
WARN
SPEED BRAKE
D.C. ELEC.
TURN COORDINATOR
2 MIN.
NO PITCH INFORMATION
NAV
GPS
VERTICAL SPEED
100 FEET PER MIN
MAN PRESS IN HG
RPM X100
ELEV TRIM
WING FLAP
RUDDER TRIM
RUD DER TR IM
OFF R L BOTH START
PUSH ON ALT BAT
PUSH ON AVIONICS
FUEL
PUSH ON STBY VAC
PUSH ON PROP DE-ICE
PUSH ON PITOT HEAT
PUSH ON BOOST PUMP
PUSH ON ELEV TRIM
CLSD
OPEN
COWL FLAP
COWL FLAP OPEN
FLAPS UP
FLAPS DOWN
THROTTLE
PROP
MIX
PARKING BRAKES - Press PERIOD (.) to release.

Master Visual Display Guidelines

- Function: attract attention and convey urgency level
- Locate within 15 degrees of normal line of sight
- Subtend at least 1 degree of visual angle
- Remain on until cancelled or resolved
- Flashing better if many competing background lights, otherwise keep steady (less distracting)
- Faster responses to black text on colored background than vice versa

Master Aural Display Guidelines

- Function: attract attention and convey urgency level

- Aural generally has faster response than visual
  - but may compete with radio, other sounds
  - Combined aural + visual is most effective

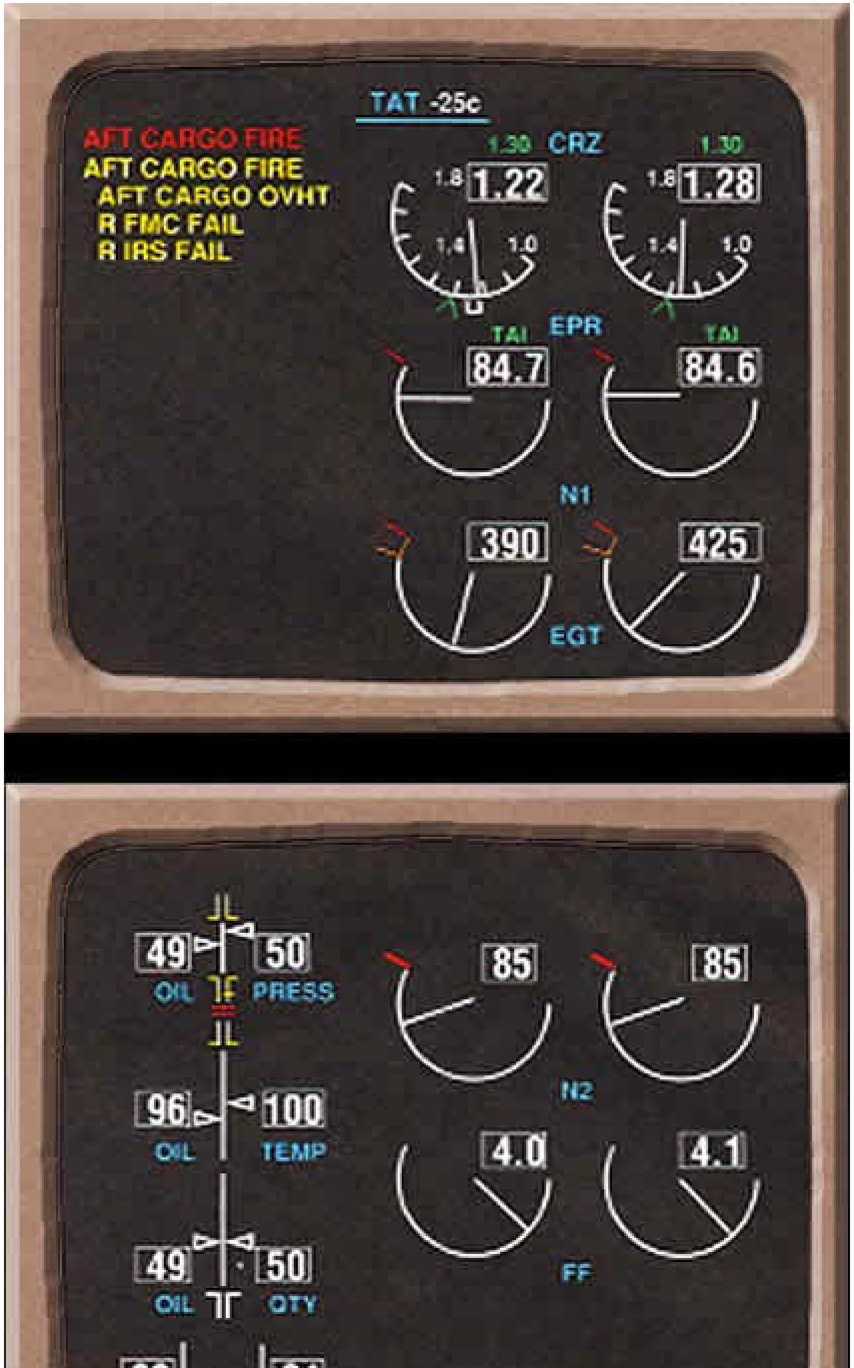
- Advisory: short single stroke sound (chime)
- Warning: alternating high & low frequency, bell, etc.

- Use frequencies between 250 - 4000 Hz
- Use 2+ frequencies simultaneously
  - masking, aging effects

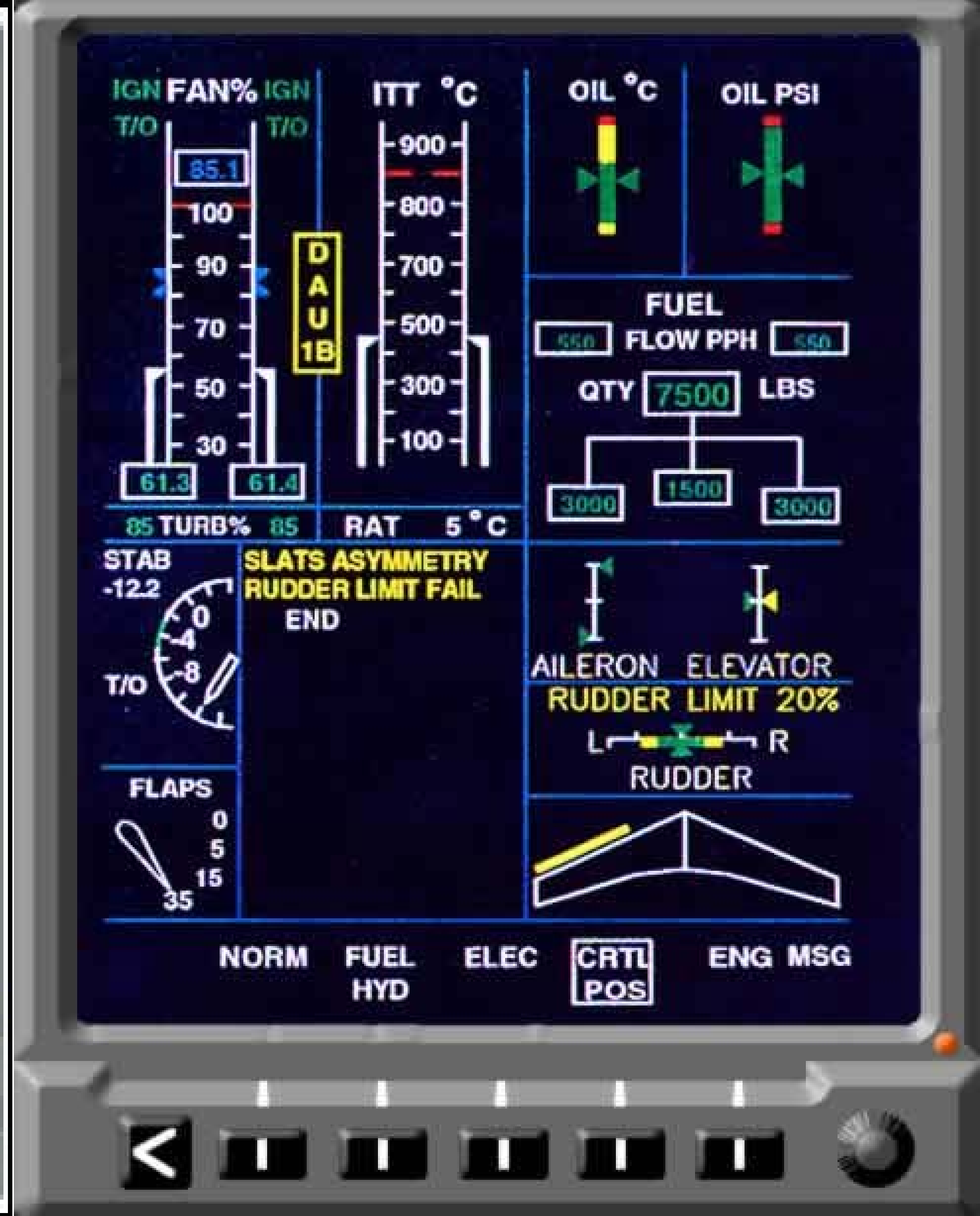
- Separate by 90 degrees from other sound sources if possible

Visual Information Display Guidelines

- Function:
  - Convey type and urgency of threat
  - Provide assistance in resolving threat
  - Provide feedback when faults are corrected
- Best if located within 30 degrees of line of sight
- Text or graphics
- Group by urgency level and chronology
- Cue new information (flashing, box outline)


TAT -25c
AFT CARGO FIRE
AFT CARGO FIRE
AFT CARGO OVHT
R FMC FAIL
R IRS FAIL
1.30 CRZ 1.30
1.22
1.28
EPR
TAI
TAI
84.7
84.6
N1
390
425
EGT
49
50
OIL PRESS
85
85
N2
96
100
OIL TEMP
4.0
4.1
49
50
OIL QTY
FF


CANC
RCL
EVENT
RCD
TAT +6c
TO 2
1.60
1.60
1.60
1.60
EPR
0.0
0.0
0.0
0.0
N1
EGT
HYD PRESS SYS 3
HYD PRESS SYS 2
ELEC AC BUS 4
ELEC AC BUS 3
ELEC AC BUS 2
ELEC AC BUS 1
ELEC UTIL BUS R
ELEC UTIL BUS L
PASS SIGNS ON
AUTOBRAKES RTO
APU RUNNING
PG 1
DOWN
GEAR
TOTAL FUEL
113.0
LBS x1000
UP
OFF
DN


N1
97.2
59.3
ITT
758
677
N2
OXY QTY
GEN AMPS HIGH
PIT TRIM MSCMP
CAB PRESS FAIL
CKLST MISMATCH
ELEVATOR DISC
END
OIL PSI
OIL °C
FF PPH
SAT °C
FUEL
LBS
PITCH TRIM
FLAPS
ELEC
HYD/ECS
FLT
VOLTS
EMER-V
AMPS
TEMP°C
MAIN
B-ACUM
CAB ALT
OXY QTY
SPLR
PIT
AIL
RUD
SUMRY
ELEC
HYD
ECS
FLT
FUEL
CAS

IGN FAN% IGN
T/O
T/O
85.1
100
90
70
50
30
61.3
61.4
85 TURB% 85
DAU 1B
ITT °C
900
800
700
500
300
100
RAT 5 °C
OIL °C
OIL PSI
FUEL
FLOW PPH
QTY 7500 LBS
3000
1500
3000
STAB
-12.2
0
-4
-8
T/O
SLATS ASYMMETRY
RUDDER LIMIT FAIL
END
AILERON
ELEVATOR
RUDDER LIMIT 20%
L
R
RUDDER
FLAPS
0
5
15
35
NORM
FUEL HYD
ELEC
CRTL POS
ENG MSG

Voice Message Guidelines

- Use when rapid action required
- Use when attention can’t be diverted from visual tasks
- Use when aural signal may not be memorized

- Tone-voice-visual best if no other talking
- Tone-visual best if other concurrent talking

- Multiple words better than single words
  - Keep short
  - Repeat once
  - Don’t use “Don’t”

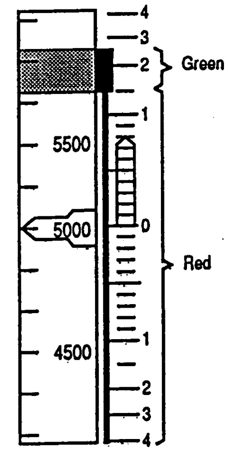
- Caution: State the type of problem (e.g., “Terrain”, “Traffic”)
- Warning: Provide action command (e.g., “Pull Up!”, “Descend!”)

Time Critical Display Guidelines

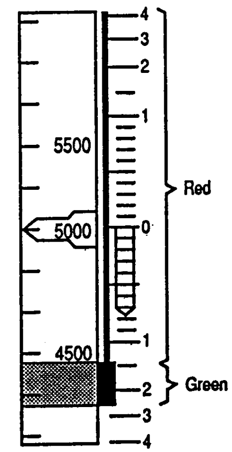
- Function: time-critical warnings
- Operator needs to get warning guidance and feedback from same display as used to correct the problem
- Locate within 15 degrees of primary field of view
- Subtend at least 2 degrees of visual angle
- Separate displays for each pilot
- Provide guidance rather than status (e.g., “Climb” vs. “too low”)
- Graphics preferred to text

TCAS CONTROLS AND INDICATORS (CONT'D)

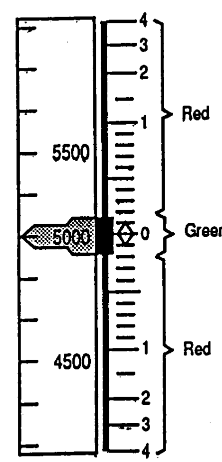
Primary Flight Display TCAS Resolution Advisories



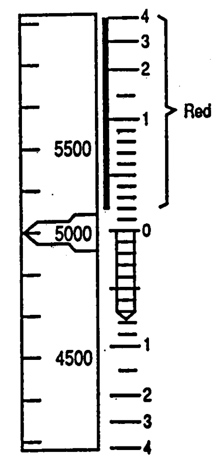
Corrective RA
Up Advisory
Climb > 1500 fpm
Voice warning:
climb, climb, climb.



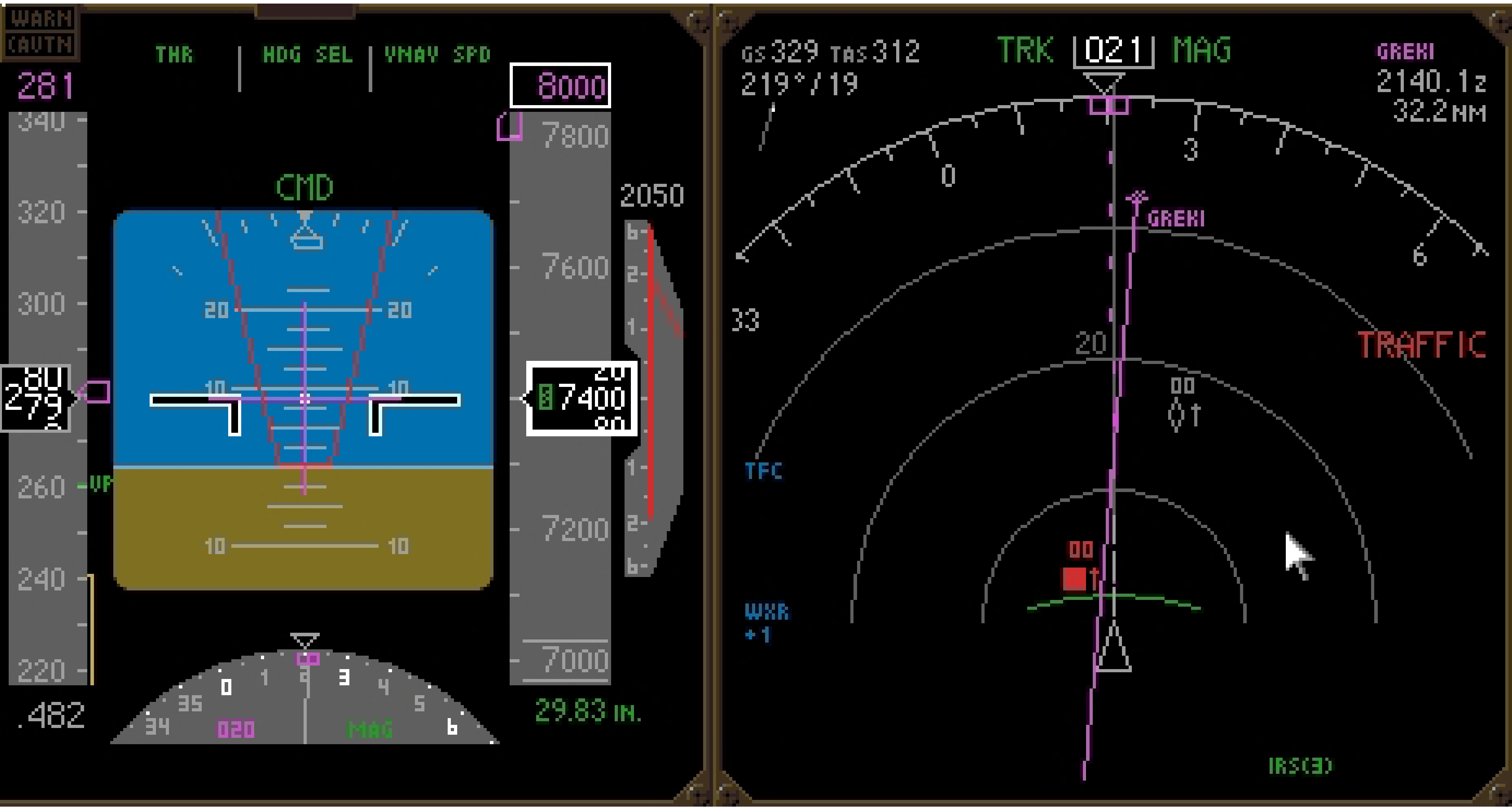
Corrective RA
Down Advisory
Descend > 1500 fpm
Voice warning:
descend, descend, descend



Preventive RA
Don't Climb
Don't Descend
Voice warning: monitor
vertical speed.



Preventive RA
Don't Climb
Voice warning: monitor
vertical speed.

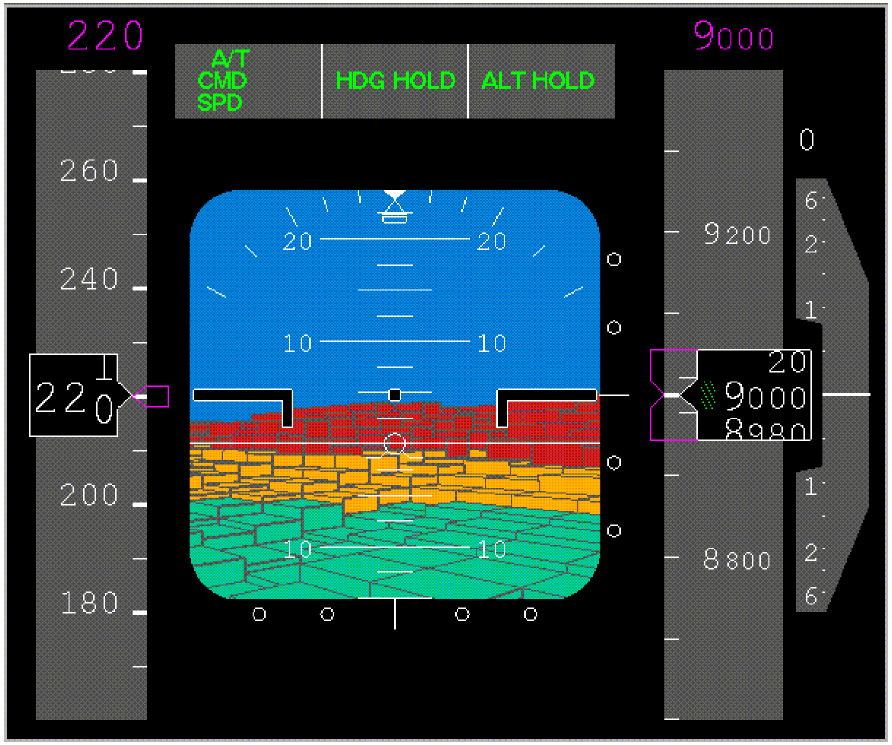

WARN
CAUTN
THR
HDG SEL
VNAV SPD
281
8000
CMD
2050
7400
29.83 IN.
.482
020
MAG
GS 329 TAS 312
219°/19
TRK 021 MAG
GREKI
2140.1z
32.2 NM
TRAFFIC
TFC
WXR
+1
IRS(3)

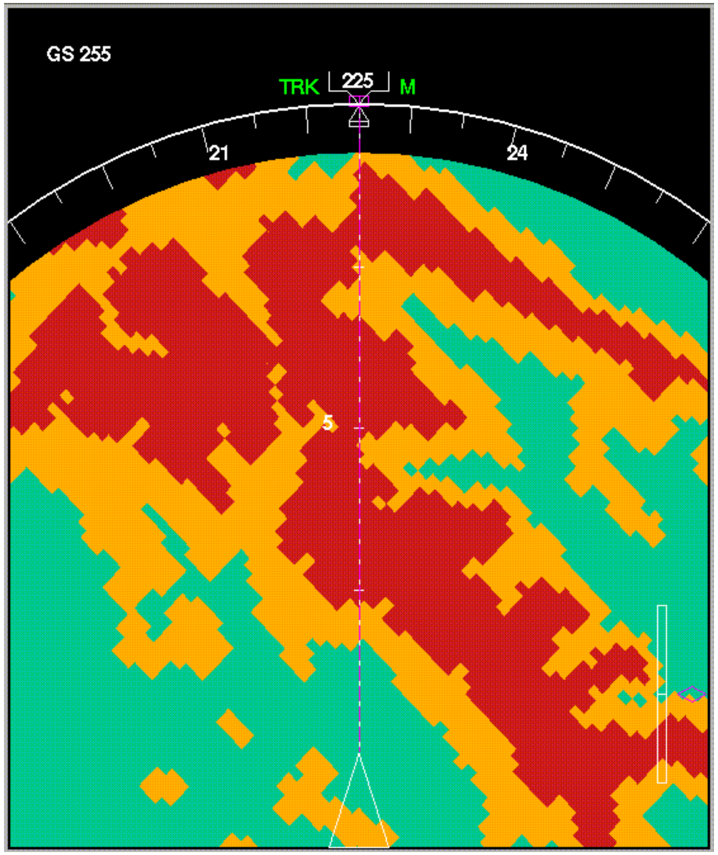
NASA Ames Basic CDTI Cockpit Display

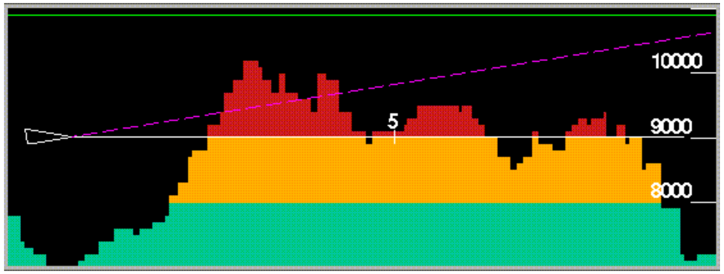
- **4D intent and traffic information**
  - 3D Flight Plans
  - Individual aircraft ID blocks
  - Static and dynamic predictors
  - Three levels of relative altitude color coding (co-altitude, above, and below)
- **Situational awareness information**
  - Multistage strategic conflict alerts
  - Traffic relevance coding (i.e. temporal proximity, “free flight” status) using intensity levels and symbol shape (nose)
- **Anti-clutter features:**
  - Full and partial data blocks (Tail tags)
  - Individually controllable data blocks
  - Smart Tags
  - Global ID and Route Declutter
- **Captain/First Officer display sharing**
- **Touchpad and panel-mounted controls**

Effect of Display on Response

- Terrain display induced turning rate
  - Plan view: 80%
  - Profile / side view: 5%
  - Perspective view: 30%







Undesirable Behavioral Impacts

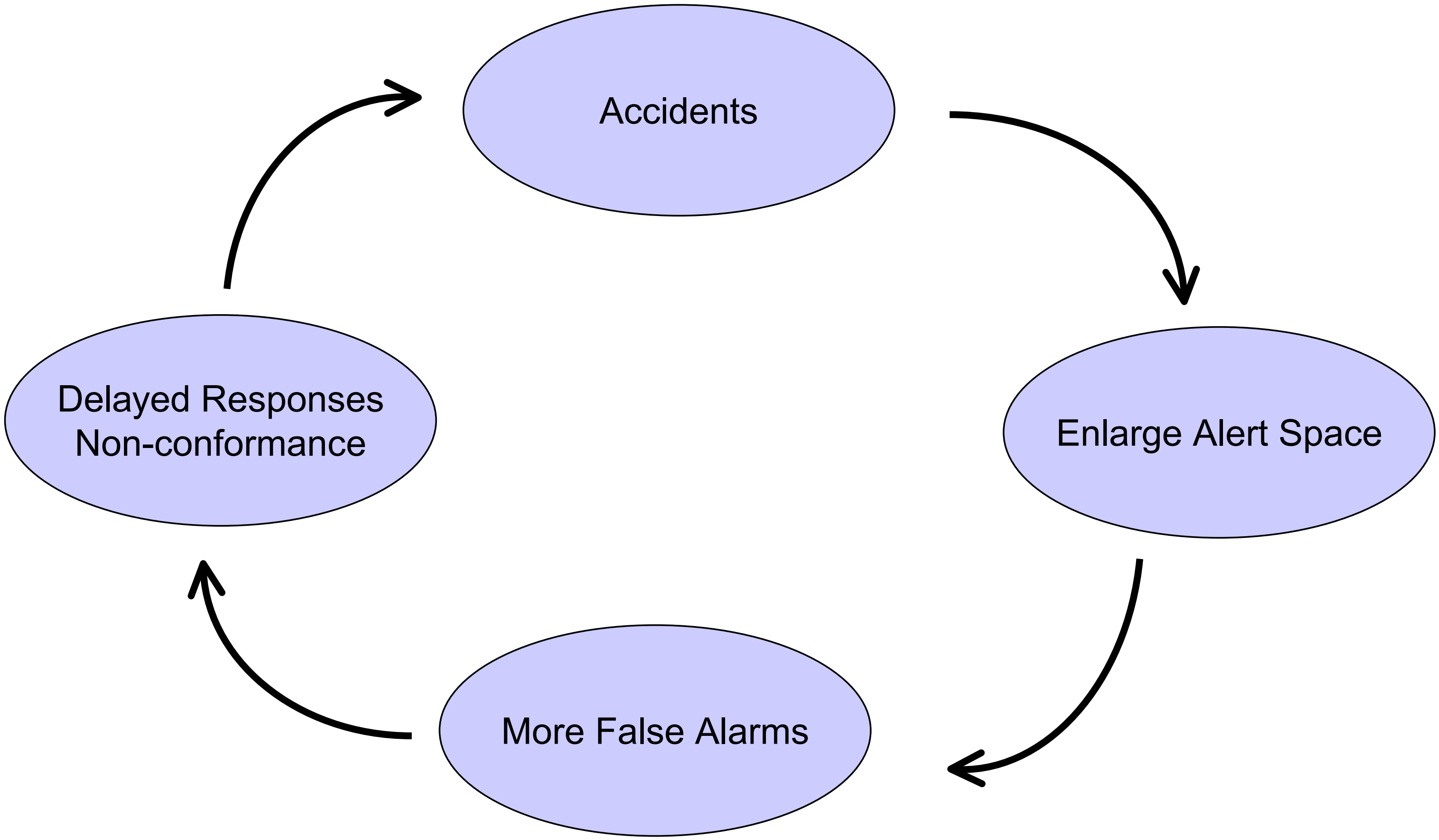
- Risk homeostasis
  - People adjust their behavior to maintain constant risk
  - e.g., introduce automatic braking system (ABS), people drive faster in wet weather

- Non-conformance to alerts
  - Pritchett: 40% nonconformance rate to collision warning
  - Mismatch of human’s internal mental model vs. automation
  - Disregard alerts
  - Take action contrary to alert command

- Viscious cycle of false alarm effects

- Need improved feedback on justification for alerts

Viscious Cycle of False Alarms and Accidents



Alert Inhibition and Prioritization

- Inhibit alerts when they would distract from primary task
  - 767: fire bell and master warning inhibited from nose gear extension until 20 seconds elapsed or reaching 400 ft

- When does human need to know information?

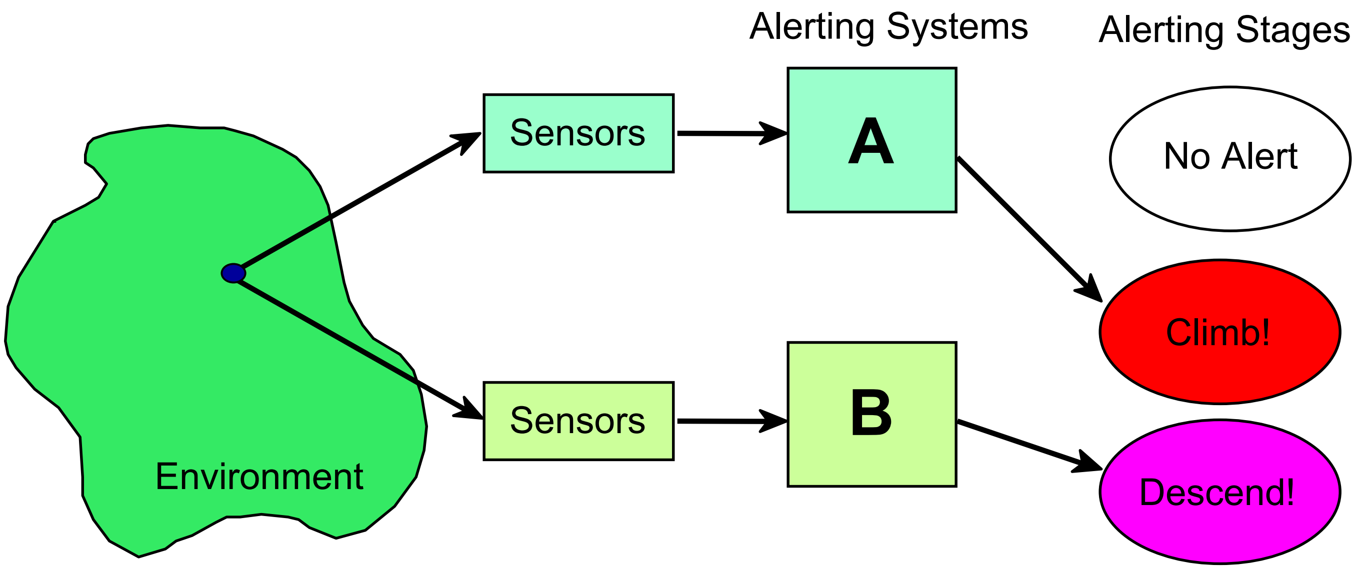
- Is it ok for automation to withhold safety-related information?

- Simplified prioritization scheme
  - Windshear
  - Ground proximity
  - Engine failure
  - Gear / flaps
  - Traffic collision

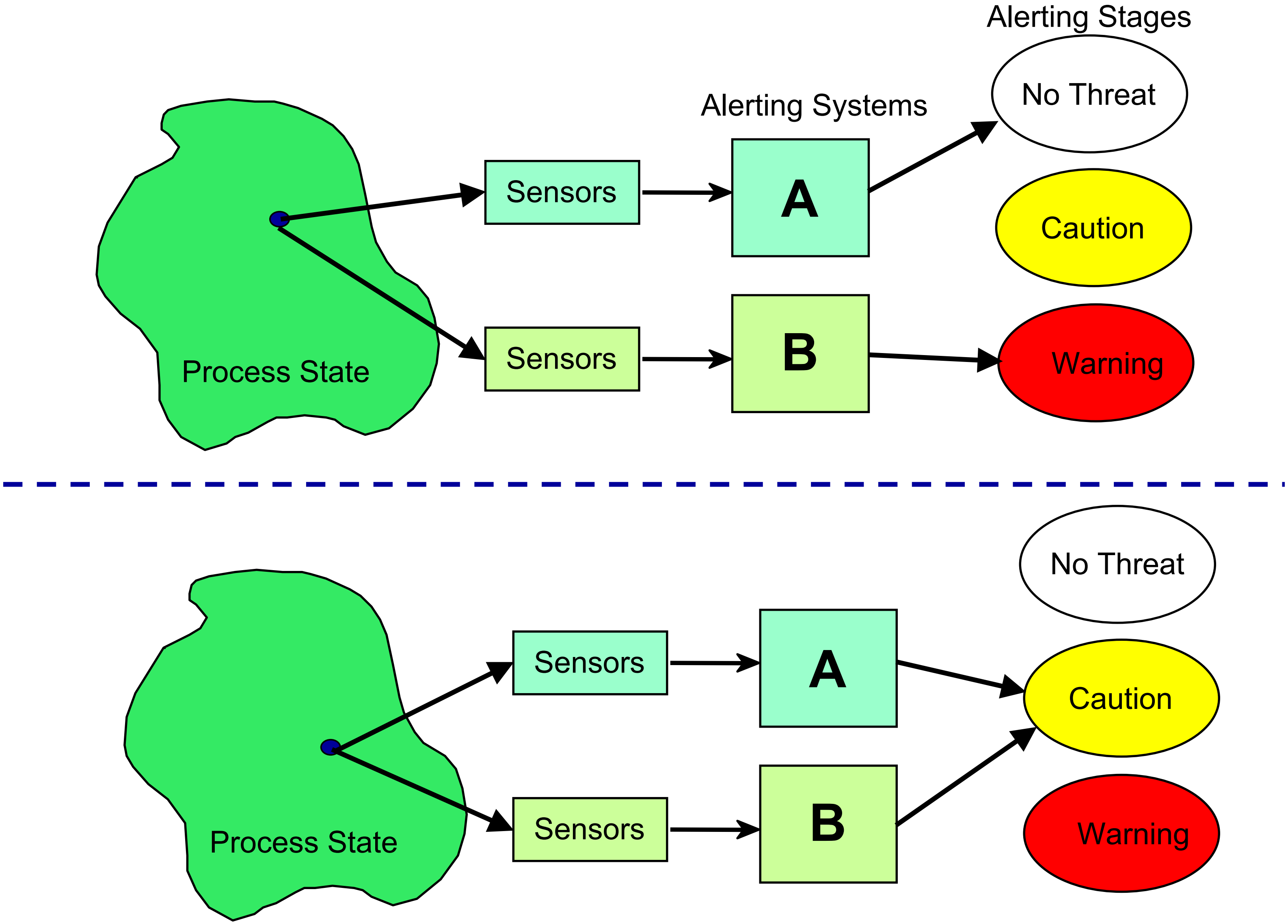
Multiple Alerting Dissonance

Proliferation of decision support functions

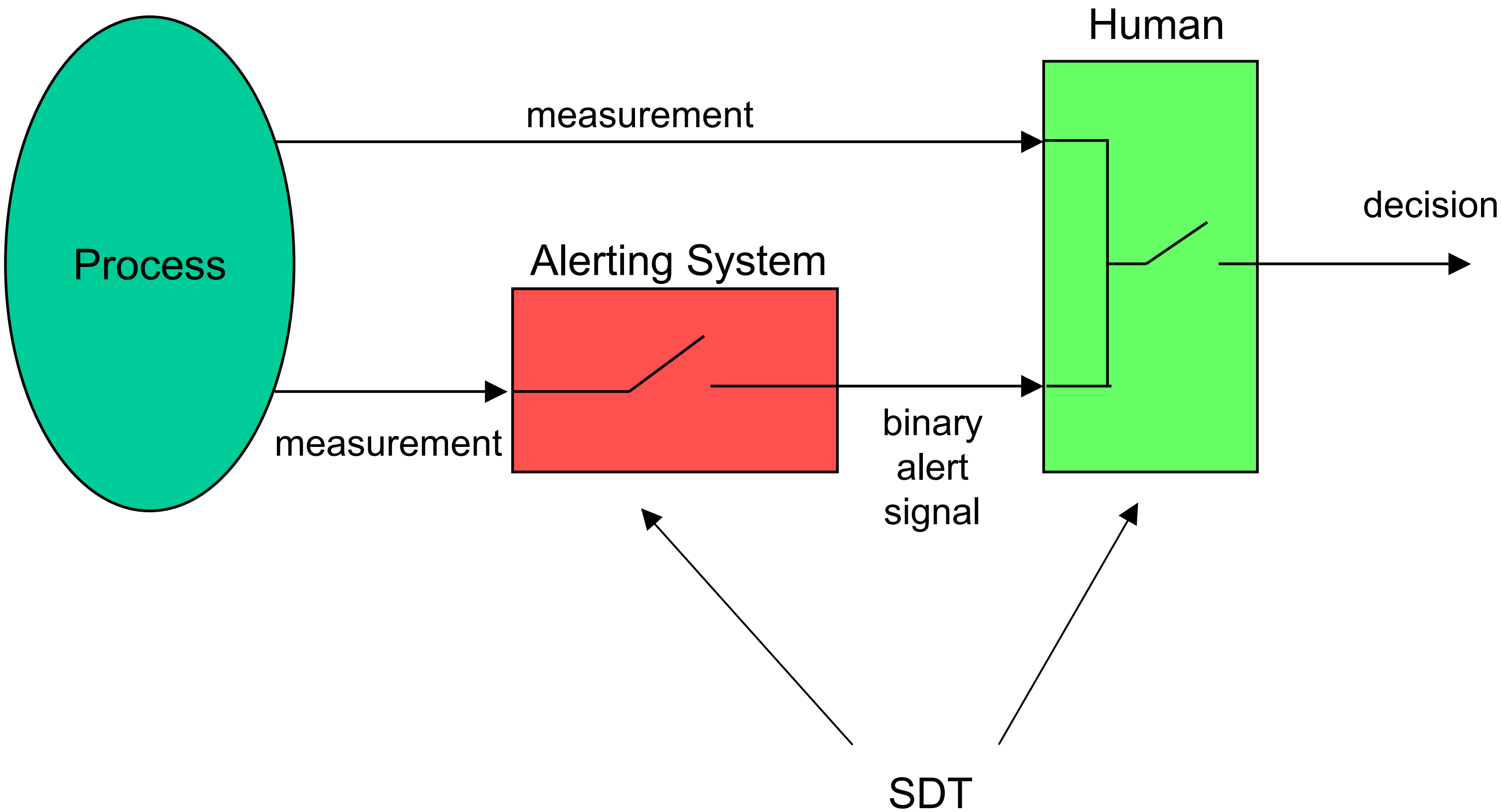
Formally identifying types of dissonance, impact, solutions



Example Dynamic Conflict



Alerting as a Tandem SDT Problem



Alerting System in SDT

- Alerting system has decision threshold

  $P(FA \mid N)$

  $P(CD \mid N)$

  Can set threshold using SDT principles

  $P(Warning) = P(FA \mid N)(1-P(S)) + P(CD \mid S)P(S)$

  $P(No\ Warning) = 1 - P(Warning)$

Human + Alerting System in SDT

- Human also has decision threshold
- Binary alert signal changes estimate of P(S) for human
- If there is a warning
  - $P(S \mid Warning) = P(CD \mid S)\, P(S) / P(Warning)$
- If there is no warning
  - $P(S \mid No\ Warning) = (1-P(CD \mid S))P(S) / P(No\ Warning)$

Example

- Reactor temperature experiment:
  - $P(S) = 0.5$, $V(FA) = 1$, $V(MD) = 2$
  - Human alone: $\beta^* = 0.5$

- Add warning system which has $x_c = 0.5$
  $P(FA \mid N) = 0.31$
  $P(CD \mid S) = 0.69$

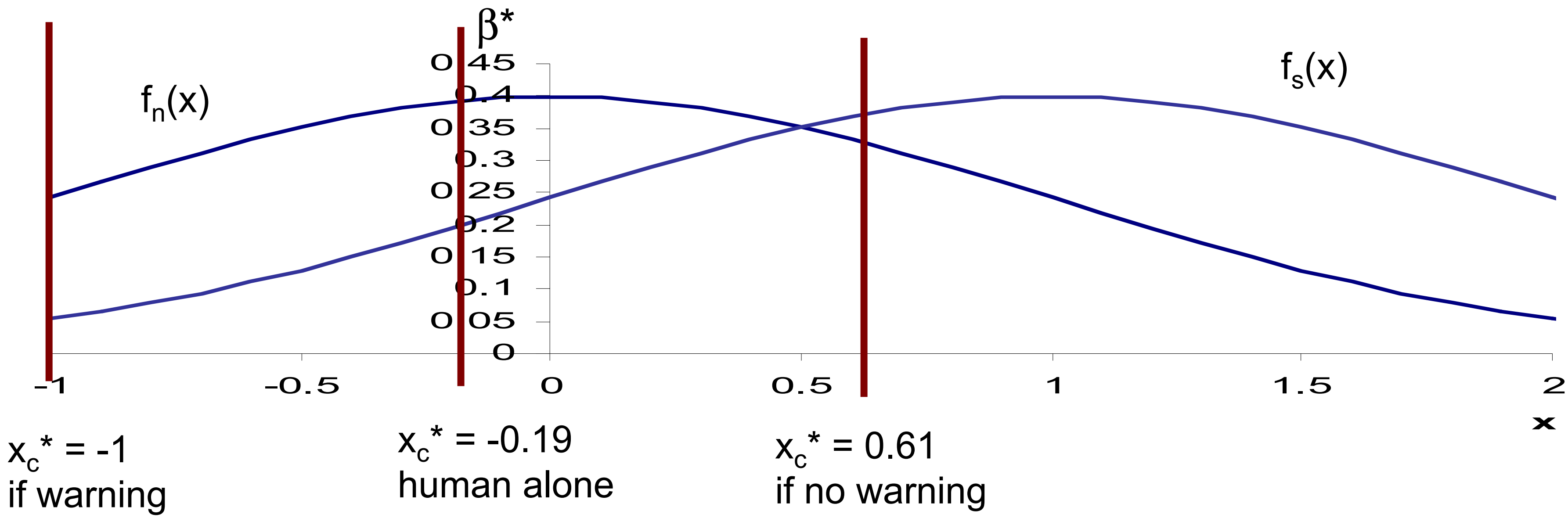
  $P(\text{Warning}) = P(FA \mid N)(1-P(S)) + P(CD \mid S)P(S) = 0.5$
  $P(\text{No Warning}) = 1 - P(\text{Warning}) = 0.5$

- If there is a warning
  $P(S \mid \text{Warning}) = P(CD \mid S)\, P(S) \,/\, P(\text{Warning}) = 0.69$

- If there is no warning
  $P(S \mid \text{No Warning}) = (1-P(CD|S))P(S) \,/\, P(\text{No Warning}) = 0.31$

Example (cont'd)

- Human alone, $\beta^* = 0.50$
- So, if there is a warning, $P(S) = 0.69$, and now
  $\beta^* = (0.31/0.69)*(1/2) = 0.23$, $x_c = -1$
- If there is no warning, $P(S) = 0.31$, and now
  $\beta^* = (0.69/0.31)*(1/2) = 1.11$, $x_c = 0.61$



Limiting Cases

- Perfect warning system, P(FA | N) = 0, P(CD | S) = 1
  - If signal is present, there is always a warning
  - P(S | Warning) = 1
  - P(S | No Warning) = 0
  - For human, if warning, $\beta^* = 0$, $x_c = -\infty$
    if no warning, $\beta^* = \infty$, $x_c = +\infty$
  - Human should just agree to whatever the warning system says

- Useless warning system, P(FA | N) = .5, P(CD | S) = .5
  - P(S | Warning) = 0.5
  - P(S | No Warning) = 0.5
  - For human, if warning, $\beta^* = 0.5$, $x_c = -0.19$
    if no warning, $\beta^* = 0.5$, $x_c = -0.19$
  - Human should ignore warning, keep original threshold

Summary

- Adding a warning system improves the estimate of P(S)
  - Moves closer to 1 if there is a warning
  - Moves closer to 0 if there is no warning

- Optimal threshold for human should move in response
  - If warning, move threshold left (smaller β)
    - need less evidence to decide signal is present

  - If no warning, move threshold right (larger β)
    - need even more evidence to decide signal is present

- Sluggish beta effect means that human will probably not shift the threshold as much as he/she should

- Can extend to multiple decision-making systems in series or parallel