Long questions (Choose <u>two</u> of three, each counts as 40% of the final grade)

Langspørsmål (Velg <u>to</u> av de tre oppgavene, oppgavene teller 40 % hver av den endelige karakteren)

Question 1: Hva er forholdet mellom deteksjonsrate (detection rate), falsk-alarm-rate (false alarm rate) og d'? I tabellen under finner du deteksjonsrate (detection rate), falsk-alarm-rate (false alarm rate) og d'. Regn ut de manglende verdiene i hver rad, og vis dine utregninger. For ***<u>en</u>*** av de manglende verdiene; vis hvor kriterium er, og forklar hva som skjer med deteksjonsraten (detection rate), falsk-alarm-raten (false alarm rate) og d' når kriterium flyttes et halvt standardavvik opp ***<u>eller</u>*** ned. Hvilke implikasjoner har dette for medisinsk diagnose? Hva er kildene til variasjonen standardavviket representerer i fordelingen?

Detection rate	False alarm rate	d'
94%	5 %	
68 %	1 %	
87 %	44 %	
59 %	1 %	
97 %		2,00
88 %		1,20
89 %		4,84
99,9 %		1,77
	0,08 %	2,64
	2 %	2,64
	6%	1,50
	6%	3,00

Question 2: Hvordan kan falske minner bli opprettet?

Question 3: Hva er prediksjonsfeilens rolle i kognisjon?

Short questions (Choose <u>one</u> of two, each answer counts as 20% of the final grade)

Kortspørsmål (Velg en av de to oppgavene, oppgaven teller 20 % av den endelige karakteren)

Question 1: Hva er ″attribution of agency″, og hvorfor er det relevant for vrangforestillinger av kontroll?

Question 2: Forklar hva som skal skje, i henhold til driftdiffusjonsmodellen, til fordelingen av responstider for riktige og feil beslutninger i en oppgave som "dot motion" når du gjør hver av de følgende endringene:

1) Flytter avgjørelsesgrensene nærmere eller lenger fra hverandre.

2) Endrer forutinntatthet (bias).

3) Endrer drivhastighet (drift rate).

4) Endrer ikke-beslutningstid non-decision time.

Hvilke av disse parameterne har adferdskorrelater som er klinisk interessante?

Long questions (Choose two of three, each counts as 40% of the final grade)

Question 1: Hva er forholdet mellom deteksjonsrate (detection rate), falsk-alarm-rate (false alarm rate) og d'? I tabellen under finner du deteksjonsrate (detection rate), falsk-alarm-rate (false alarm rate) og d'. Regn ut de manglende verdiene i hver rad, og vis dine utregninger. For ***en*** av de manglende verdiene; vis hvor kriterium er, og forklar hva som skjer med deteksjonsraten (detection rate), falsk-alarm-raten (false alarm rate) og d' når kriterium flyttes et halvt standardavvik opp ***eller*** ned. Hvilke implikasjoner har dette for medisinsk diagnose? Hva er kildene til variasjonen standardavviket representerer i fordelingen?

What is the relationship between detection rate, false alarm rate, and d'? In the following table of detection rates, false alarm rates and d', calculate the number missing in each row, including the steps of your calculation. For one of the cases, show where the criterion is, and explain what happens to detection rate, false alarm rate and d' when you move the criterion half a standard deviation up or half a standard deviation down. What are the implications of your findings for medical diagnosis? What are the sources of the variation represented by the standard deviation of the distributions?

Detection rate	False alarm rate	d'
94%	5 %	
68 %	1 %	
87 %	44 %	
59 %	1 %	

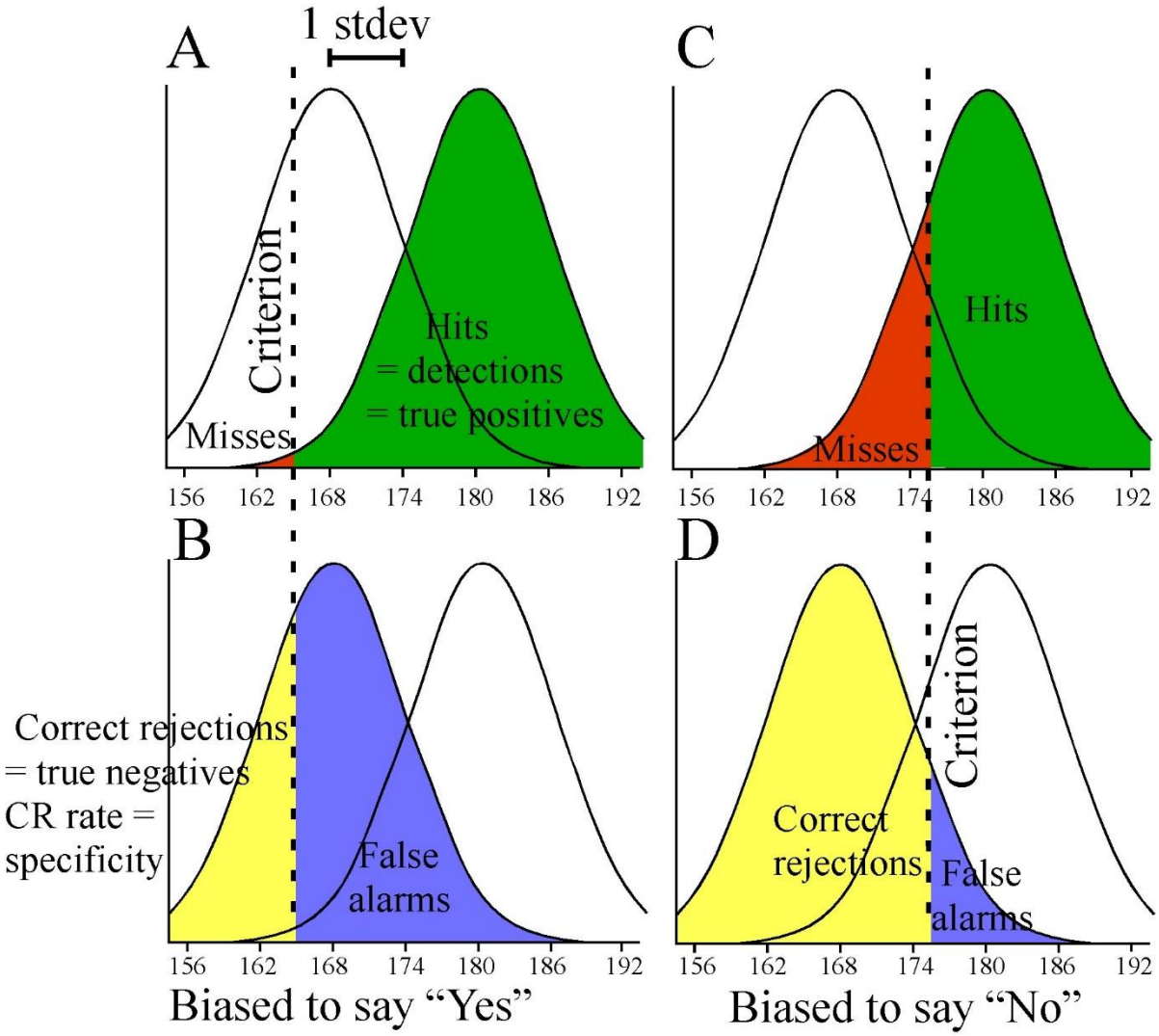
97 %		2,00
88 %		1,20
89 %		4,84
99,9 %		1,77
	0,08 %	2,64
	2 %	2,64
	6%	1,50
	6%	3,00

Sensorveiledning: A simple example that illustrates the relationship between detection, rate, false alarm rate and d' is this: you have a list of surnames and heights. You want to determine which of those people are male and which are female. Say you phrase the question as "is this person male?" On average, men are a little taller than women, in Norway 180 ± 6 cm versus 168 ± 6 cm. If you want to represent the underlying statistical reality, you draw one normal or Gaussian distribution with an average of 168 ± 6 cm and another with an average of 180 ± 6 cm. Then you must set a criterion.

Assume that you randomly choose to set the criterion at 164 cm. Anyone taller than that you classify as a man. When you meet all the people on your list of names, you should find that you correctly detected men in 99.38% of cases, and you missed only 0.62% of men. If you look up the z-score associated with that detection rate, you find that you get that detection rate if you set the criterion 2.5 standard deviations below the average height of Norwegian men. The z-score is -2.5, as shown in part A of the figure.

You might be dismayed to realise, though, how often you wrongly categorised women as men. Your false alarm rate on the question "is this person male?" turns out to be 69% (see part B of the figure). You should expect that if you set the criterion below the average height of women.

Deeply disappointed by random selection of the criterion, you decide to move it some place more sensible. You choose 175.5 cm. Now your detection rate falls to 77.4%, but you have reduced the false alarm rate to 10.6%.



You can work out how much overlap there is between the heights of men and women, which is expressed how many standard deviations there are between men's and women's average heights. When you used the liberal or lax criterion, with a strong bias to say "yes, this person is male", you found your criterion to be 0.5 standard deviations below the average female height and 2.5 standard deviations below the average male height. The difference is called sensitivity in signal detection theory and abbreviated as d'. The formula for calculating d' is:

$$d' = z(\text{detection rate}) - z(\text{false alarm rate})$$

The averages and standard deviations do not change when you move the criterion. Therefore the difference between the average heights, in multiples of standard deviations, should not change either. The z-score associated with a detection rate of 77.4% is -0.75 and the z-score associated with a 10.6% false alarm rate is 1.25. Insert into the equation and you get d' = 1.25 – (-0.75) = 1.25 + 0.75 = 2. Same result as before, as it should be.

d' then offers you a measure of performance that is independent of criterion. If you test memory by recognition, you have a yes/no decision, you have two distinct populations of items, those you have shown to your participants during training and those that are new, and you can apply signal detection theory to the result. d' is thus a far better measure than the proportion

correct. d' also offers you a measure of how well a test discriminates that is independent of the criterion you choose.

The question then is where you should put the criterion. For the moment, assume that correct choices make no difference to anyone's life, but that mistakes have a cost. You want to minimise the total cost. You *cannot* eliminate cost, because you cannot eliminate all mistakes. If you only shift your decision criterion, you trade one type of error against another; misses against false alarms. However, the rates at which the two type of mistake change are usually not equal.

Look again at A and B. The curve for hits and misses is both quite low and quite flat, so moving the criterion a fraction of a standard deviation will not make much difference there. But that same horizontal shift of the criterion will make a lot more difference to the false alarms and correct rejections, because that curve is both steep and at a higher level. The same movement of the criterion changes areas a lot more. So you can lose more false alarms than you gain misses if you move the criterion to the right.

If we let the areas under the curve represent not just probabilities, but the total number of cases, we have to adjust the heights of the curves to represent the base rate and 1 – base rate. You know from your calculations of Bayes' theorem that even if the false alarm rate is quite modest, the total number of false alarms can be quite high if the base rate is low. You must take that into account if you want to find the choice of criterion that will give you the smallest total number of mistakes.

Next, you may not be concerned about the total number of mistakes, but the total cost. If one type of mistake is much more expensive than the other, you can't afford so much of the expensive mistake and must shift the criterion accordingly. Likewise, if one of the correct choices gives you a greater benefit, you shift the criterion to increase the proportion of that choice. The optimum location of the criterion depends on the payoffs of all choices, and on the prior probabilities of the outcomes.

An important aspect of all this is that moving the decision criterion only trades different kinds of errors off against each other. The only way to reduce both false alarms *and* misses is to increase d' by reducing measurement error. In the example of the height of men and women, the average difference is 12cm is two standard deviations only because the standard deviation happens to be 6cm. Some of those 6cm represent real variation in height. And some of those 6cm reflect random errors in measurement. Eliminating all measurement error, if that were possible, would increase d'. However, if we are dealing with normally distributed data, there will always be overlap between the two normal distributions, and therefore the probability of errors can never be driven to 0.

In summary, signal detection theory can be used when categorically distinct states of the world (for example, infected with a virus or not) need to be inferred from noisy data. If you choose a decision criterion, signal detection theory can tell you how often you will make what kind of error, it lets you choose a decision criterion that minimises the average cost of errors, and it makes clear why even the best decision can still result in a mistake.

Detection rate	z(TP)	False alarm rate	z(FA)	d'
94 %	-1,55	5 %	1,64	3,20
68 %	-0,47	1 %	2,33	2,79
87 %	-1,13	44 %	0,15	1,28
59 %	-0,23	1 %	2,33	2,55
97 %	-1,88	45 %	0,12	2,00
88 %	-1,17	49 %	0,03	1,20
89 %	-1,23	12 %	1,17	2,40
99,9 %	-3,09	54 %	-0,09	3,00
30 %	0,52	0,08 %	3,16	2,64
72 %	-0,59	2,00 %	2,05	2,64
48 %	0,05	6,00 %	1,55	1,50
93 %	-1,45	6,00 %	1,55	3,00

Sensorveiledning: the pensum contains explanations of two mechanisms. Students who remember the introductory cognition course may add a third. First, reconsolidation can alter memories. This applies both to associative learning and to episodic memory. In associative learning, conditioned fear has been reduced through either a reminder with prediction error, then extinction during the reconsolidation time window, or through a reminder with prediction error while under the influence of propranolol. The same mechanism offers a plausible account for people retrospectively changing not just their interpretation, but their memory of events. In Linton's diary study of autobiographical memory (not in the pensum, included here only for illustration; students may offer other examples), one of the recorded events was Linton's first meeting with the man she later married. At the time of testing, she remembered him having made a strong impression straight away, which turned you not to match the diary entry. A plausible explanation for that bias is that Linton later remembered that meeting, putting that memory in a labile state, then re-encoded it with the emotion attached that she felt at the time of re-encoding, when she had fallen in love. Emotional associated with autobiographical memory may be changed by reconsolidation.

Reconsolidation can also change memories of events. Hupbach et al. had people encode one list of objects on a Monday. On Wednesday, the experimental group was asked to remember Monday's list before learning a new list of objects. That group encoded the Wednesday list at the same time they re-encoded the Monday list, and their memory of the Monday list had more intrusions from the Wednesday list than the other way round, and more intrusions than the control group, which had no reminder. Reconsolidation is also the most plausible explanation when episodic memories are implanted, as in Shaw and Porter's study of false memories of committing crime. Experimenters established their credibility by presenting with memories of real childhood events, but they also asked about criminal behavior that, according to the experimenters, the participants' parents had reported. After repeated sessions of guided imagery and being encouraged to treat the imagined events as real, the majority of participants accepted that they had committed crimes. Repeated retrieval of information, elaboration, and

encoding of the altered, elaborated memory is a good candidate for the underlying mechanism. The same applies to Geraerts et al.'s study of recovered memory. They found only memories of childhood abuse recovered outside therapy were supported by other evidence. Geraerts et al. did not report what kind of therapy patients had experienced. Recovered memory therapy uses many of the same techniques that Shaw and Porter used to establish false memories. There is no reason to believe that all therapy carries the same risk of creating false memories.

Second, alternative associations may be created. This occurs with standard extinction training, and the original memory still being present explains spontaneous recovery and various forms of renewal. So when people are presented with information that conflicts with their original experience, it is also possible that they create an alternative, conflicting memory. If that alternative memory is more accessible at the time of retrieval, that will result in a false memory being reported. This is only possible if there *is* a memory of an original event. Where episodic memory is concerned, there is no information in the pensum on what favours either the creation of alternative memories or modification of the original memory. For that, we have evidence only from associative learning, where it is not clear that a new association created in extinction, or in a different context, is *false*.

The third mechanism, which students may remember form a different course, is priming, as in the Deese-Roediger-McDermott memory illusion. A list is made from the strongest associates of a core word, for example sleep. Then only the associates (pillow, dream, bed, alarm clock, etc.) are presented, but *not* the core word itself. Retrieval of each list word conceptually primes words with related meaning. All list words are related to the core word, which gets primed each time and so becomes very accessible. That core word is then a frequent false memory, both in recognition and in recall.

Question 3: Hva er prediksjonsfeilens rolle i kognisjon?

What is the role of prediction error in cognition?

Prediction error is the difference between the expected outcome and the experienced outcome. One fundamental role of prediction error is to drive learning. The more experience deviates from prediction, the more there is to learn, and the larger is the learning rate. That is the reason why Craske et al propose maximising prediction error in extinction training. Their main theoretical concern is that habituation, which reduces the effect of the US, reduces prediction error and therefore reduces the learning rate. They therefore advocate avoiding habituation by not gradually increasing the intensity of exposure, but instead randomly mixing the order. They also assume that the effects of extinction training will be context specific, and the way to work around that is to extinguish in a wide variety of contexts.

Prediction error is also relevant to solving a different problem in Pavlovian or classical conditioning, namely working out whether current experience results from an already known cause-effect relationship, or whether a new, unknown cause-effect relationship operates. Gershman refers to this as state discovery: is the world still in the same state, meaning familiar cause effect relationships, or in a new state, meaning new cause effect relationships? It appears that evolution has given us a rule of thumb that if the world is as we know it, prediction errors

should be relatively small, reflecting incomplete learning, random variation, or a gradual change. The appropriate response is to adjust the strengths of existing associations. If instead prediction errors are large, it is likely that some unknown factor is at work, and new learning is needed. Then it makes sense to create new associations. That reasoning is behind Gershman et al.'s recommendation that the creation of context-specific inhibitory associations in extinction can be avoided by changing prediction error gradually, instead of introducing a sudden, large change.

A third role that prediction error plays, in instrumental learning, is to help distinguish the effects of one's own actions from the effects of other causes. A possible solution is to predict the outcomes of one's own actions and to attribute any large enough deviation to external causes. For the purpose of solving this attribution problem, "large enough" must be defined in statistical terms: How unlikely is the observed event given how precise prediction and the observatyion of the result are believed to be? If causal attribution depends on this subjective estimate of precision, then miscalibration (i.e., a mismatch between perceived and actual precision) should result in erroneous attributions.

For example, say I practice archery. I expect my arrows to be normally distributed around the bull's eye, with a standard deviation of 10 cm. If I find a standard deviation of 40 cm, I keep being surprised by where my arrows go. I could recalibrate my estimate of my ability, assuming that I am having a bad day, or that I am just not as good as I thought. Or if I am, for some reason, unable to recalibrate, I can explain away my poor performance, perhaps choosing to believe that someone spiked my drink. An unspoken assumption of Frith's proposal is that schizophrenia interferes with recalibration.

Frith (2005) proposed that believing one's predictions and/or perception to be more precise than they really are contributes to hallucinations. If I keep dropping things and fumbling about, yet I remain convinced that I can precisely predict my movements, I may conclude that my movements are not under my control. The same principle applies to predicting sensations. Most people can't tickle themselves. If the sensations predicted based on the planned movements are close enough to the actual sensations, they attract little attention. However, there is always random variation in the predictions that are generated, and random variation in the movements carried out. If I underestimate that randomness, if I believe I can predict the sensations resulting from my movements very precisely, then the unappreciated randomness means that my experience is unlikely to be close enough to prediction to be ignored. It will feel vivid, because the less sensations match prediction, the more attention they get. Sensations that are well predicted, are suppressed. And patients with schizophrenia do find the sensations produced by their own actions equally as ticklish as the sensations caused by other people.

Short questions (Choose one of two, each answer counts as 20% of the final grade)

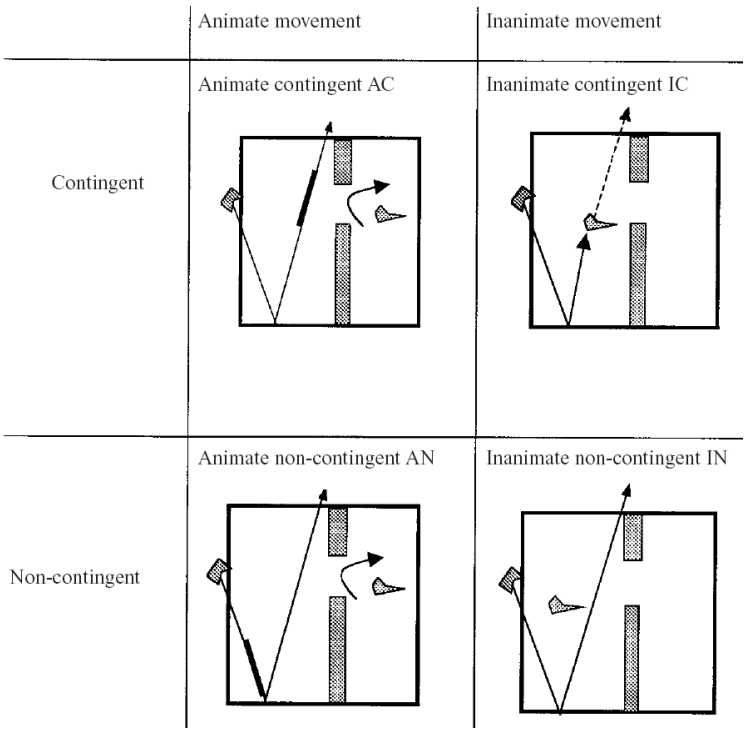
Question 1: Hva er ″attribution of agency″, og hvorfor er det relevant for kontrollvrangforestillinger?

What is attribution of agency, and why is it relevant to delusions of control?

Sensorveiledning: Attribution of agency is explaining events as the consequences of the actions of an agent, that is someone who has intentions. Frith argues that excessive attribution of agency is needed to explain delusions of control (the belief that someone other than the patient controls the patients actions or thoughts) because problems with the forward model are not a sufficient explanation.

A forward model predicts the motoric and sensory outcomes of motor commands. A large enough deviation from that prediction can reasonably be interpreted as indicating an external causal factor that affects the outcome. If the forward model is faulty and produces larger prediction errors than justified, then deviations from prediction that actually are small and random can be interpreted as indicating an external influence. The anarchic hand phenomenon, caused by lesions to supplementary motor cortex or anterior corpus callosum, demonstrates that problems with the forward model can only explain the feeling of not controlling one's own actions, but not the feeling that someone else controls one's actions. Anarchic hand patients feel one of their hands has a will of its own, but not that someone else controls it.

Excessive attribution of agency means seeing evidence of intentional action where the general population sees less or none. The most intuitive data presented to the students comes from an experiment on moving geometric shapes. Subjects were shown all four combination of animate and inanimate shapes (they move under their own power or only passively, like billiard balls) and contingent (the shapes interact) versus non-contingent.



The question was whether there would be an interaction between animation, contingency and diagnosis. The difference between inanimate contingent and inanimate non-contingent provides a baseline, and was similar in delusional patients and controls. However, delusional patients perceived no difference between animate contingent and animate non-contingent conditions, attributing intentions to the animate-non-contingent movements that controls did not. When this exaggerated attribution of intention to some outside agency is combined with not feeling in control of one's own movements or thoughts, the result is to attribute control to someone else. That is a delusion of control.

Question 15: Forklar hva som skal skje, i henhold til driftdiffusjonsmodellen, til fordelingen av responstider for riktige og feil beslutninger i en oppgave som "dot motion" når du gjør hvert av følgende:

1) Flytt avgjørelsesgrensene nærmere eller lenger fra hverandre.

2) Endre forutintatthet (bias).

3) Endre drivhastighet (drift rate).

4) Endre ikke-beslutningstid non-decision time.

Hvilke av disse parametrene har adferdskorrelater som er klinisk interessante?

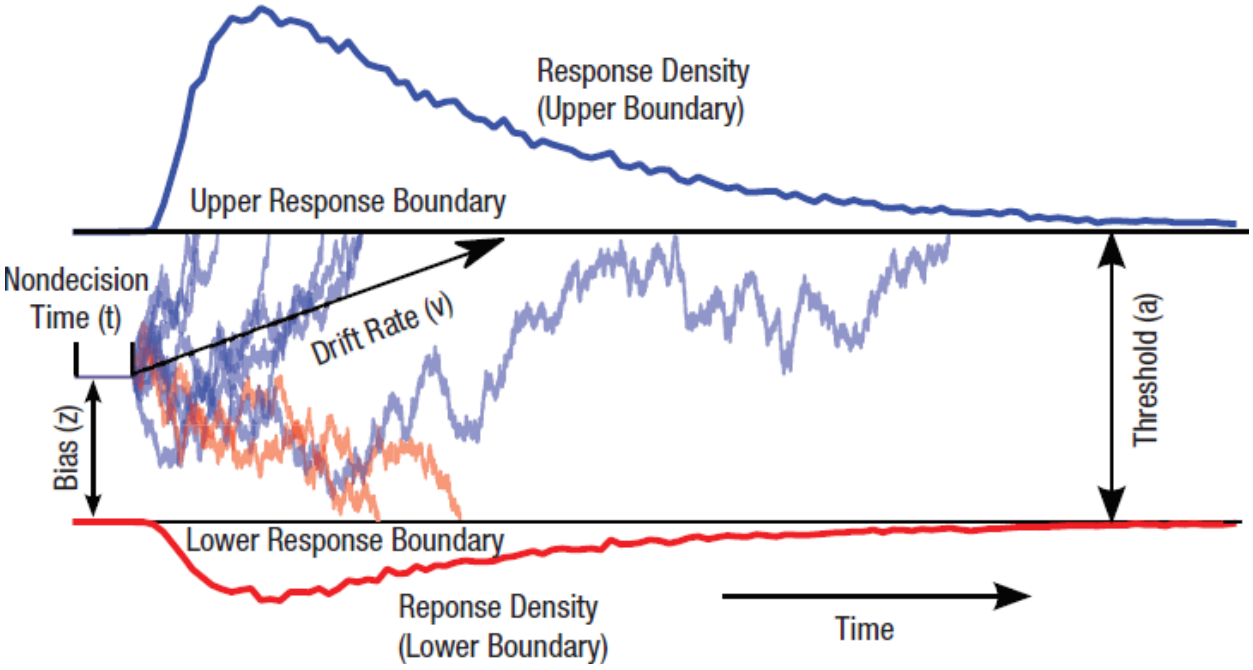
Explain what should happen, according to the drift-diffusion model, to the distributions of response times for correct and wrong decisions in a task such as dot motion when you do each of the following:

1) Move the decision boundaries closer together or further apart.
2) Change bias.

3) Change drift rate.
4) Change non-decision time.

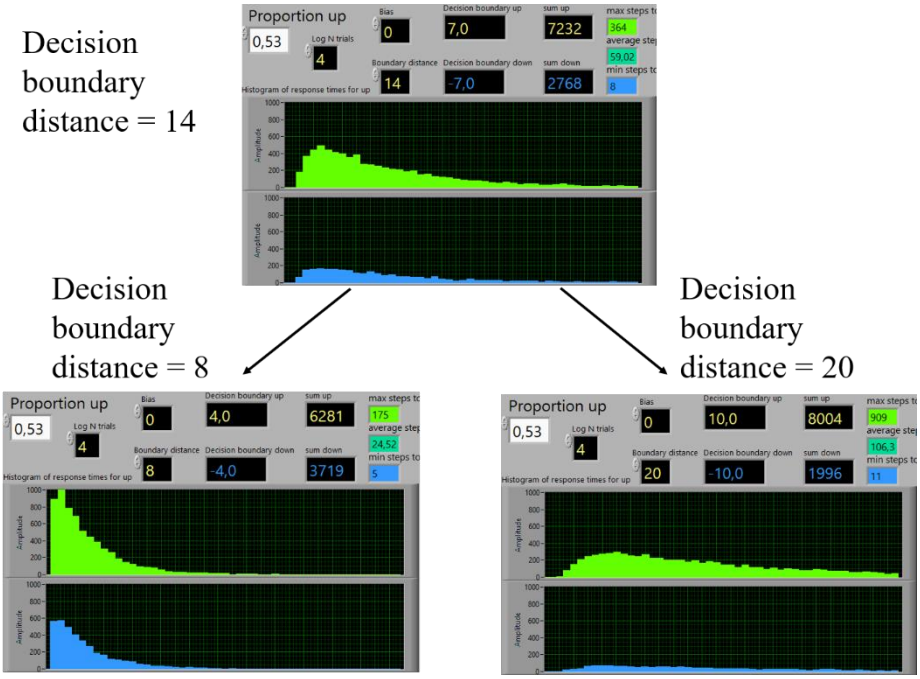
Which of these parameters have behavioural correlates that are clinically interesting?

Sensorveiledning: The underlying decision model is summarised in the figure:



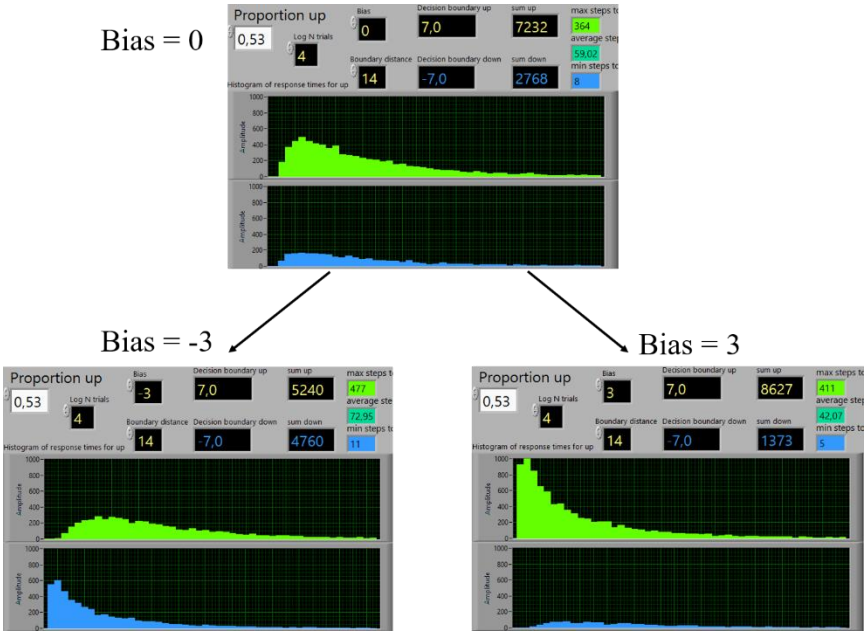
There is a decision variable, which can be thought of as each piece of evidence getting a vote that favours one of the two alternative interpretations of the evidence, or two candidates for election. The votes come in random order, and the election ends the moment one of the candidates has N votes more than the other. Each track in the figure represents the progress of one election, more specifically the *difference* in votes. The midpoint between the decision boundaries represents both candidates having an equal number of votes. The response densities show the results of many elections with the same voters casting their votes in random order.

The distance between the midpoint and each of the decision boundaries represents how large a majority one alternative needs to be chosen. (The distance from the midpoint to a decision boundaries is N, but students don't need to be that specific.) The larger the majority that one alternative needs, the further apart are the decision boundaries. If N = 50, then the earliest possible decision can come after 50 votes. If N = 10, the earliest decision comes after 10 votes. Also, the smaller N is, the easier it is for random fluctuations to create a majority. In the extreme case, if N = 1, the first vote leads to a decision, and the election stops. Therefore the closer the decision boundaries, the faster but less accurate the decisions, because it is easier for random fluctuations in the order of votes to give election victory to the less favoured candidate. The response density curves would be squished together to the left, and the total number of elections won by each candidate would be more evenly distributed, like here:



The clinically relevant application of that idea is that one form of impulsivity is problems with appropriately trading off speed against accuracy. If speed matters, faster, but less accurate decisions can be made by moving decisions boundaries closer together. If accuracy matters, move decision boundaries further apart. If someone keeps decision boundaries closer together than normal, they make faster, but less accurate decisions, which is one way of being impulsive.

Bias then corresponds to giving one candidate some extra votes to start with. That is sensible if one outcome is more common or associated with a greater payoff. The response density curve for the favoured alternative will be squished together to the left, and will have more area under the curve, representing a larger number of elections won. The response density curve for the less favoured alternative will be pulled apart to the right, and will have less area under the curve, representing a smaller number of elections won.



A change in drift rate corresponds to the votes coming in at a different rate. If votes come in faster, then both response density curves will be squished together to the left. The pensum contains no example of this. Attentional processes may increase drift rate.

A change in non-decision time moves both response density curves in the same direction without changing their shapes. Non-decision time comes from everything that takes time other than the decision process itself, such as the time it takes information to get into the brain, and then out again to the muscles. The pensum suggests no relevance to a clinical psychologist.