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

Question 1: Hva er prediksjonsfeilens rolle I kognisjon? What is the role of prediction error in cognition?

Sensorveiledning: 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 onditioning, 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.

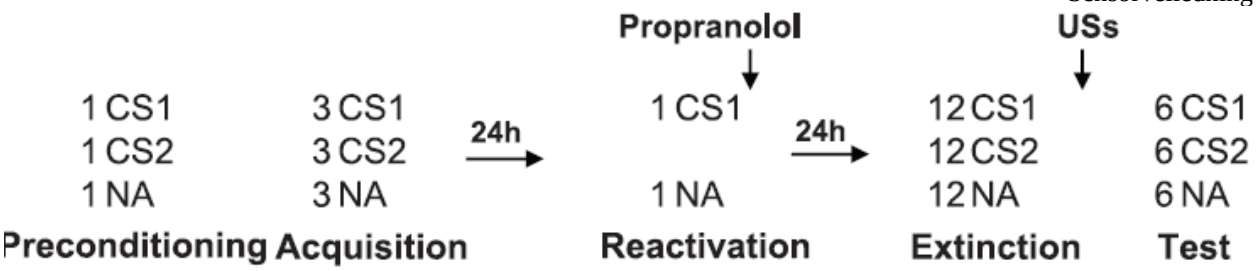
A third role that prediction error plays, in instrumental learning, is to help distinguish the effects of one's own actions from the effects of other causes. A possible solution is to predict the outcomes of one's own actions and to attribute any large enough deviation to external causes. For the purpose 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.

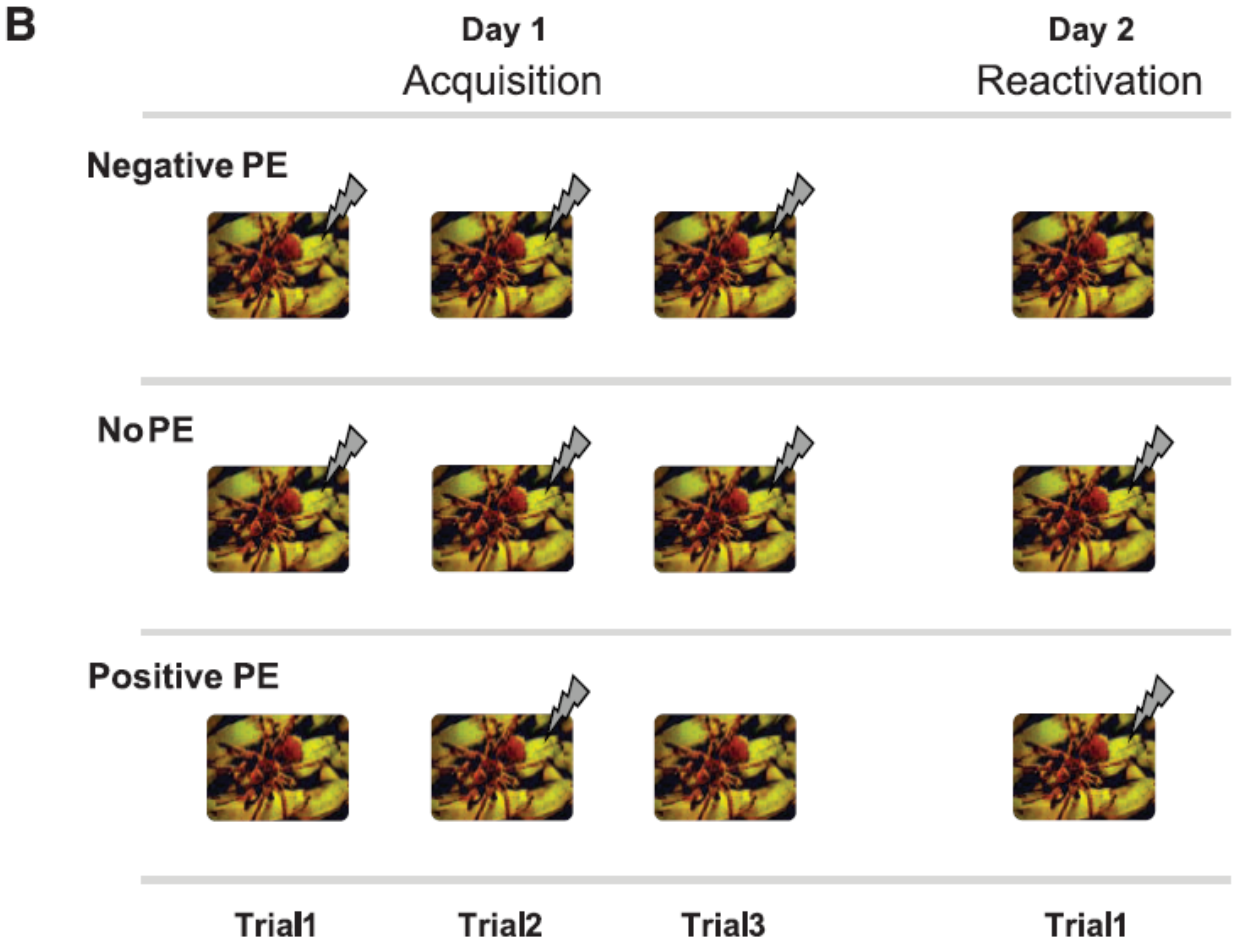
For example, say I practice archery. I expect my arrows to be normally distributed around the bull's eye, with a standard deviation of 10 cm. If I find a standard deviation of 40 cm, I keep being surprised by where my arrows go. I could recalibrate my estimate of my ability, assuming that I am having 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.

Reconsolidation occurs when being reminded of some information in memory makes that information easier to change during a reconsolidation time window. It occurs in multiple different memory systems. In associative learning, there is evidence that a reminder alone is not enough to trigger reconsolidation, that the association only becomes labile if the conditioned stimulus is not followed by its usual consequence during the reminder. Agren's review provides a verbal description of Sevenster et al.'s (2012) experimental design. In fear conditioning, administration of propranolol combined with a reminder was previously found to selectively reduce the emotional response to a CS. In Sevenster et al.'s study, this was true only for participants who were connected to the electrodes that had previously delivered electric shock. People who were not connected, and who had no reason to expect any shock, and who therefore experienced no prediction error when shock was absent during the reminder, behaved like the placebo group. Their fear potentiated start did not decrease from the end of acquisition to the beginning of extinction. More importantly, the group without shock expectation and thus no prediction error showed reinstatement after being shocked again, which indicates that the original association still existed, and had only been masked by an inhibitory association created during extinction training.

A second study by Sevenster et al. (2013) is not described in Agren's review, but students may remember it from the lecture. Initial fear conditioning involved three presentations of a spider picture, with either one or three shocks in those three trials. A no prediction error group received a shock during the reminder trial, which was expected given the 100% reinforcement during acquisition. Negative prediction error came from the omission of an expected shock, positive prediction error came from a shock when past reinforcement history (spider picture followed by shock only one time out of three) indicated that no shock was more likely.





Although explicit expectation of shock depended on its past frequency, fear potentiated startle, the measure of emotional response, depended only on whether there was prediction error, not whether it was positive or negative. In the presence of propranolol, either prediction error led to an immediate decrease of fear potentiated startle after the reminder trial at the beginning of extinction, and after extinction training no reinstatement after being exposed to shock.

There is good evidence for prediction error being necessary for a reminder to make an association labile. It is not clear what prediction error even means in episodic memory. There, a reminder of a previously learned list can make the memory of that list vulnerable to intrusion errors from a newly learned list, consistent with the idea that the remembered list is being re-encoded while the new list is being learned, and so new list items are added to the memory of the old list. But what would be the prediction error in this situation? So the relevance of prediction error has been demonstrated when reconsolidating associative learning, so far specifically fear

conditioning, but it is unclear whether prediction error is relevant to reconsolidation in all memory systems.

Question 2: Hva slags impulsivitet er det? Har skillene virkninger på den virkelige verden?

What kinds of impulsivity are there? Do the distinctions have real world implications?

Sensorveiledning: Wiecki et al. list four kinds of impulsivity. The better students may notice something Wiecki et al did not mention: temporal or delay discounting, reward sensitivity and speed accuracy trade-offs apply to decisions with outcomes that are delayed over a wide range of time intervals, from less than a second to centuries. Failures of inhibition apply to responses typically less than a second delayed.

Caswell et al. initially describe three kinds of impulsivity, and their factor analysis indicates there are four. Two of these are also in Wiecki et al.'s list, another two are different, but not precisely specified. The pensum thus describes six kinds of impulsivity, of which four are clearly explained:

1) Temporal discounting (what Caswell et al. call temporal impulsivity) is treating events as less important the more they are delayed. The standard procedure to estimate temporal discounting is to offer a choice between one fixed and once variable amounts, at one fixed and one variable delay. For example, offered a choice between two new 100 kr bank notes right now, one should be indifferent between the two. Next, offer 100kr tomorrow, and vary the alternative amount offered right now until finding the indifference point, say 90 kr. Then offer 100 kr in two days, and again find what amount is subjectively worth the same if paid immediately, say 82 kr. That variable amount indicates how much the 100 kr are worth at the delay that has been specified. The steeper that temporal discounting is (the more subjective value declines with delay), the greater the preference for an immediate over a delayed reward. Steep temporal discounting is one of the kinds of impulsivity associated with ADHD, and with nicotine use.
2) Reward sensitivity refers to how much people are motivated by positive as opposed to negative outcomes. It can be measured by learning tasks involving gains and losses, and checking to what extent choices are explained by memory for either gains or losses. Greater sensitivity for losses is associated with the negative symptoms of schizophrenia, greater sensitivity for gains with pathological gambling, and with ADHD. Reward sensitivity is not one of the four factors identified by Caswell et al.
3) Speed accuracy trade-offs occur in sequential sampling tasks in which people can choose how much data they sample before making a decision. Because the rate at which data become available is not, or not entirely, available under one's own control, gathering more data takes more time. That forces a trade-off between speed, a fast decision relying on les data, and accuracy. Examples of sequential sampling tasks in the laboratory would be the random dot motion task, the beads task, or the box task. Examples of real life sequential sampling problems would be deciding where to go on holiday, or whether to propose marriage. People with ADHD adjust speed, and consequently the amount of data

gathered, less than controls in response to task demands. Impaired speed accuracy trade-off is associated with greater risk of being a perpetrator or victim of violence

4) Motor impulsivity corresponds to failure of inhibition, as measured by the stop signal task or the Stroop task. Real life examples would be inability to stop oneself from making inappropriate remarks. This is factor 1 of Caswell et al.'s analysis. Motor impulsivity does occur in ADHD. Chamberlain et al. found reduced motor inhibition, as measured by the stop signal task, in both OCD and trichotillomania (compulsive hair pulling). The conditions differ in that OCD sufferers also show less cognitive flexibility.
5) Caswell et al. describe reflection impulsivity as involving a general preference for speed. The task they use to measure it does have speed and an accuracy version, but their analysis seems to be not sensitive to speed accuracy trade-offs. Reflection impulsivity seems to correspond to the jumping to conclusions bias in the beads task shown by schizophrenia patients, not an impaired speed accuracy trade-off. However, the link between reflection impulsivity and jumping to conclusions is not confirmed by empirical data, and therefore students are not expected to go into detail.
6) Caswell et al.'s third factor is measured by the immediate memory task, which does not seem to measure and it is not even clear why it is considered a measure of impulsivity. Therefore students are again not expected to go into detail.

Dalley et al. found that drug taking in humans is associated with a questionnaire-based measure of impulsivity, and with one aspect of sensation seeking. Rats bred for impulsivity consume more cocaine than rats bred to be less impulsive, and impulsive rats worked harder for nicotine and sugar.

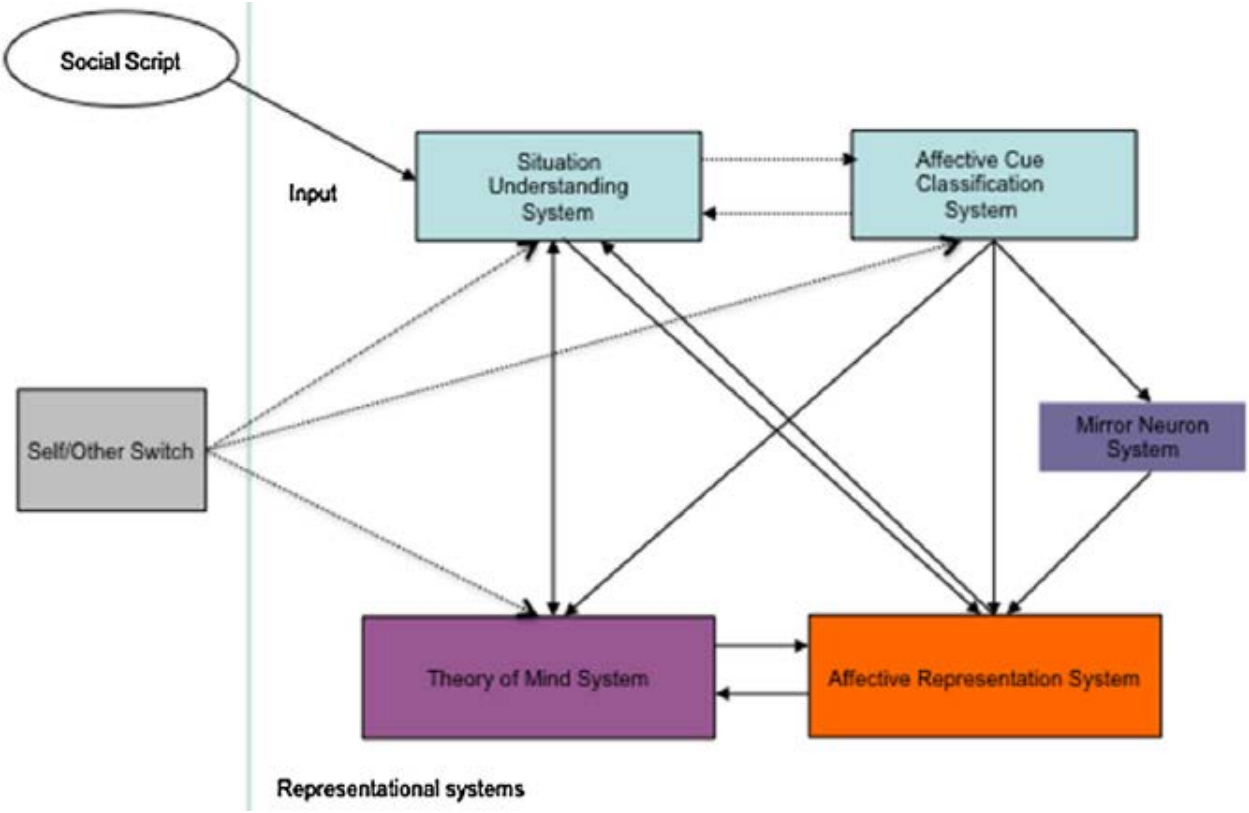
There is some evidence that different kinds of impulsivity are selectively associated with different real world problems. It is less clear, from the material in the pensum, how impulsivity is associated with mental health conditions. Wiecki et al. mention some kind of impulsivity being associated with ADHD, OCD, Tourette's syndrome, substance abuse, gambling, and eating disorders. It is less clear how specific those associations are, because patients with a particular condition may not have been tested for all kinds of impulsivity. ADHD is associated with all four kinds of impulsivity mentioned by Wiecki et al. Gambling, OCD and trichotillomania are mentioned in connection with only one kind of impulsivity each, but either patients were only tested for one kind, or it is not clear whether there were other tests.

Question 3: Forklar Bird og Viding sin empatimodell. Hva er komponentene, hvordan redegjøres deres interaksjon for svekkelser i sosial erkjennelse? Hva er forskjellen mellom empati og emosjonell smitte? Kan modellen redegjøre for modulering av empati, som ble funnet da menns smertenettverk var aktivt da de visste at noen som hadde behandlet dem rettferdig fikk elektrisk støt, men ikke når noen som hadde behandlet dem urettferdig fikk det samme elektriske støtet? Hvis modellen kan redegjøre for det, hvordan? Hvis ikke, hvorfor ikke?

Explain Bird and Viding's model of empathy. What are the components, how does their interaction explain impairments in social cognition? What is the difference between empathy and emotional contagion? Can the model account for modulation of empathy, as found when

men's pain network was active when they knew that someone who had treated them fairly was receiving electric shock, but not when someone who had treated them unfairly was getting the same electric shock? If the model can account for that, how? If not, why not?

Sensorveiledning: The components of Bird and Viding's model are summarised in this figure:



Social scripts are stories with interpretations of the motives of the protagonists, as supplied by culture and personal experience. They provide data for the situation understanding system, that attempts to match current experience to the social scripts.

The affective cue classification system infers emotion from such information as facial expressions, body language, or tone of voice.

Both the situation understanding system and the affective cue classification system contribute to the affective representation system being influenced by the emotions of others. If the affective representation system lacks information on whose affect it represents, then the influence from another's affect is emotional contagion. If the influence is noticed, and the emotion is tagged as belonging to someone else, that is instead empathy.

The model resolves disagreements between situation understanding system and the affective cue classification system by recruiting the theory of mind system, which is portrayed as a reasoning system with apparently little impact on the intensity of one's own affect. Theory of mind resolves conflicts over what affect should be inferred if affective cues don't match the situation. That input to the affective representation system is itself affectively neutral, but may secondarily provoke affect.

The mirror neuron system is described as providing a parallel path for information from the affective cue classification system to reach the affective representation system, but bird and Viding do not explain what additional function mirror neurons contribute.

The self/other switch has the dual function of directing attention either to oneself or to others, and of tagging affect as belonging to someone else. It has no direct access to the affective representation system, though. The tagging can only occur through the theory of mind system.

Autism is explained as an impaired theory of mind system, resulting in reduced ability to tag affect as originating from someone else, and a reduced ability to resolve disagreements between the situation understanding system and the affective cue classification system. Bird and Viding also propose that atypical social scripts, caused by reduced attention to social situations, increase false inferences from the situation. Bird and Viding also claim that the probability of conflict between the situation understanding system and affective cue classification system is reduced, even though increasing either noise or systematic error in one of the systems should increase the probability of conflict. They do state that the affective cue classification system is most likely intact. They also speculate that the self/other switch may be biased towards the self in autism.

Bird and Viding describe alexithymia as an impairment in the affective representation system, resulting in impaired representation of affective states of both self and others. They further argue that many impairments commonly ascribed to autism are instead due to comorbid alexithymia.

Psychopathy is explained at a level not directly captured by the model, in two ways: first, there is a reduced ability to recognize and feel fear and sadness; second, that the impairment reduces attention to these emotional states throughout development, and so the affective representation system never learns to represent them properly.

The model does not have an obvious mechanism for implementing selective modulation of empathy, as occurred in the experiment when men lost empathy for unfair players. Possibly the self/other switch might be used, but the mechanism that would bias the self/other switch is not included.

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

Question 1: Hva er konsekvensene av at ulik læring er kontekstspesifikk i forskjellige grader? Hvordan definerer du grad av kontekstspesifisitet?

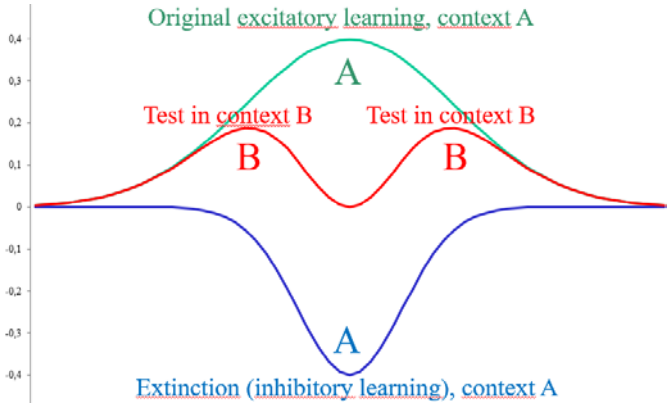
What are the consequences of different learning being context specific to different degrees? How do you define degree of context specificity?

Sensorveiledning: Learning is context specific to the extent that behavior differs between the context in which behavior was learned and the context in which it occurs, or fails to occur, later. For example, if an association between a stimulus and outcome is learned in context A, then context specificity means that the relationship between stimulus and outcome is treated as valid in context A, but as uncertain in other contexts.

Extinction training (reducing the probability of the stimulus being followed by the outcome) can lead to the learning of new inhibitory associations. That means a new association between the stimulus and the *absence* of a relevant outcome. When the new inhibitory association is combined with the old excitatory association, the net outcome is reduced or zero expectation of the outcome.

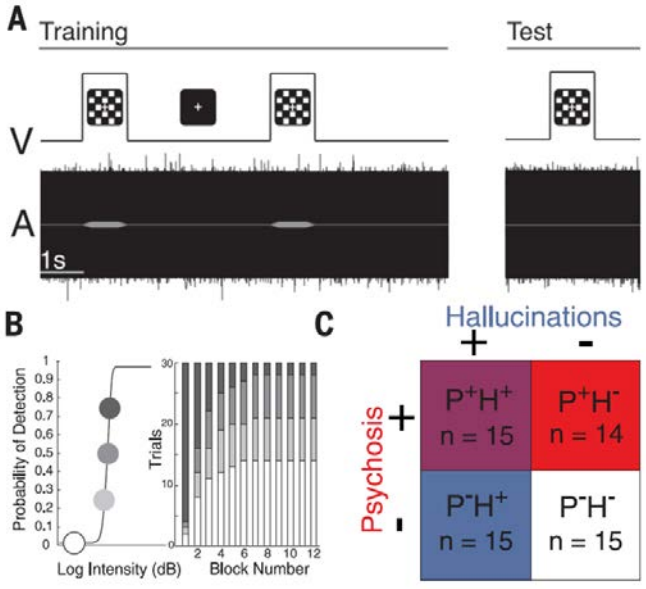
Renewal after extinction training is the recovery of a response after a change of context. ABA renewal can be explained purely qualitatively: there is less confidence that the outcome expected in the learning context A will also occur in B, and so even if extinction training creates an inhibitory association that there is no response (and presumably no expectation), the inhibitory association does not need to be as strong as the original excitatory association in context A. On return to the context A for testing, the already weaker inhibitory learning also generalizes only to a limited extent, so the original excitatory learning is stronger. A similar argument can be made for ABC renewal.

Explaining AAB renewal requires the additional assumption that inhibitory associations are more strictly context specific than excitatory associations. If context is treated as continuously variable, meaning it is possible to say *how* different two contexts are, then greater context specificity means a narrower generalisation gradient, that responding declines faster for inhibitory associations as context changes. Then the inhibitory learning needed to counter the original excitatory learning needs to be as strong as that excitatory learning, but because it generalises less than the excitatory learning, more of the excitatory is left in somewhat different contexts.



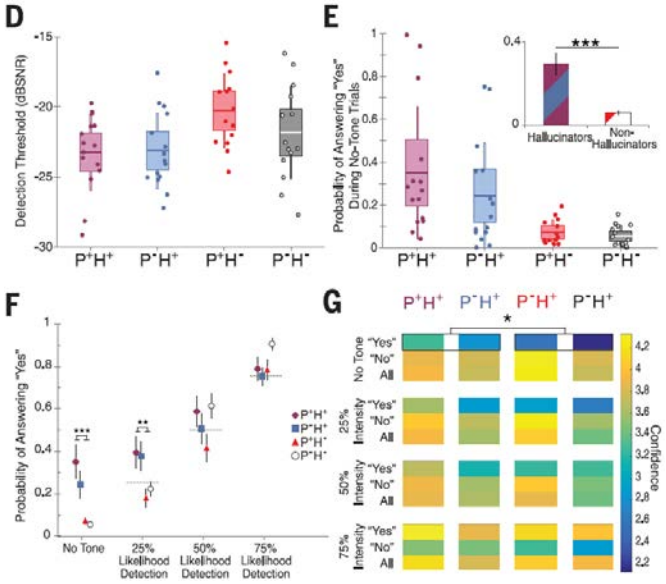
Question 2: Forklar Powers et al. Sin eksperimentelle prosedyre for å indusere hallusinasjoner gjennom pavlovsk kondisjonering. Hva er deres viktigste funn? (I dette korte spørsmålet kan du ignorere beregningsmodellen.)

Explain Powers et al.'s experimental procedure for inducing hallucinations through Pavlovian conditioning. What are their most important findings? (In this short question, you can ignore the computational model.)



Sensorveiledning: The procedure is

summarised in Powers et al.'s figure 1. A visual stimulus is paired with a tone at four different sound levels, partially masked by noise (A). The sound levels are individually calibrated to be audible 75%, 50%, 25% and 0% of the time (B). In the first block of training, the loudest sound is most common. In later blocks, the quietest sound becomes most common.



Participants are divided into four groups according to the presence or absence of psychosis and the presence or absence of hallucinations.

The results indicate that only the presence of hallucinations is associated with a higher probability of hearing a sound when there is none, and that people with hallucinations are more confident of hearing a sound when there is none.