**Overview of 2020 honours projects**

**Project 1: Taylor Kirkwood, working with Gavin.**

* **Research question:** Do people have implicit or explicit access to their point of subjective equivalence (PSE) for consumer like stimuli? How consistently do people make choices based on their PSE?
  + **Hypothesis:** Compared to the explicit method, the implicit method for identifying thresholds is more closely related to the choices people ultimately make between options. Therefore, we expect the implicit method to produce more consistent choices for options than the explicit method.
* Investigate preferential (rather than veridical) choice in the double factorial design. The aim is to identify each participant's 'threshold', which was explicitly assigned in the JMP paper, and then generate the cells of the double factorial paradigm around that threshold (H, L, D levels for each attribute).
* **Experiment:** This will use the task from the JMP paper - hotels varying on price and review rating - in the numeric presentation mode. The new experiment will have two phases:
  + **Phase 1.** Identify each participant's 'point of subjective equality' (PSE) / 'certainty equivalent' (CE). This will be investigated in two different ways in separate groups:
    - Implicit. Titrate to the PSE by presenting pairs of options and finding the point at which people are indifferent between the two. The options will vary on the two attributes and they will be adjusted by some adaptive algorithm (e.g., staircase, quest). This is like the PSE approach in psychophysics.
    - Explicit. Ask for ratings on two slider scales - one for price, one for quality. They should be negatively correlated to enforce the attribute tradeoff (so participants can't give $0 and 100% quality…). This will give us the 'threshold'. It's like the CE method in applied choice.
  + **Phase 2.** Just like the JMP paper: one stimulus is shown and the participant is asked if they would accept it or not. Each stimulus will come from one of the 9 cells of the double factorial design, with balanced number of trials per cell to facilitate modelling. The threshold identified in phase 1 will be used to generate H, L and D cells of the design.
* **Things to consider in the experimental design:**
  + The algorithm for the implicit condition in phase 1. There are a few possibilities. I suggest we draw from the literature to avoid recreating the wheel.
  + The way to generate the levels (e.g., means and SDs) for the H, L and D distributions. We will need to generate some rule-of-thumb that will remain valid across different participant thresholds.
* **Analysis approach (for honours):** 
  + Compare thresholds identified in phase 1: are they equal, on average, across the two methods? Is one more variable across participants? Also consider recording how long (in minutes) phase 1 took to run, which we could use to get an approximation to the efficiency of the two methods in identifying thresholds.
  + Compare choices and RTs across cells of the double factorial paradigm in phase 2. E.g., is the proportion of rejections for double distractor cells the same for implicit vs explicit conditions? This will tell us something about how well the explicitly and implicitly identified thresholds align with how people make choices for single products. The same comparison can be made across other focal cells of the design (HH, HL, LH, LL, and so on).

**Project 2: Jessica Grimmond, working with Gavin.**

* **Research question:** How are consumer-like options mentally represented? Does this differ when options have correlated vs uncorrelated features? How does that mental representation affect choices and RTs between those options?
  + **Hypothesis:** Choice is harder when stimuli are more similar, translating to slower RTs and more variable choices. Mental representations are more complex (lower similarity) for options with uncorrelated features.
* **Experiment:** Trials in this experiment will mostly look like a DCE with two options. The options will be mobile phones each described by 5 attributes, drawn from the database collected by Daniel O'Brien (Summer internship student) from the JB-HiFi website. The correlated and uncorrelated options is a between subjects manipulation.
  + **Phase 1.** Each trial will present a pair of options and ask the participant to rate the similarity between the two options on a rating scale from "highly dissimilar" to "highly similar", or some wording like that. The ratings scale should have 7 or 9 points. Participants are not given rules or guidance on what 'similar' means - they decide on that. However, it's important that they are instructed to use the full range of the scale across trials. Each possible pair of options is presented once.
  + **Phase 2.** Each of the pairs of options from phase 1 are re-presented (random order) though the rating scale is now removed. Instead, participants are instructed to choose their preferred option. Each possible pair of options is presented at least once. Ideally we'd get more, but this will depend on the total number of options we consider in the experiment.
* **Things to consider in the experimental design:**
  + How many options? This will be driven by how many trials we can fit in the time limit of the session. If N is the number of options then the number of unique pairs of options is N\*(N-1) / 2. This ignores order (left vs right, right vs left), which is fine as we will randomise option position on the display. This blows up quickly! (some examples below) I suggest that N=20 is already quite high since the total number of trials in the experiment will be number of pairs x2 (i.e., phases 1 and 2).
    - N=5 -> 10 pairs
    - N=10 -> 45 pairs
    - N=20 -> 190 pairs
    - N=30 -> 435 pairs
  + Composition of the uncorrelated options. There are at least two options for this:
    - Generate a fixed set of options that is constant across participants. Pros: matches the correlation condition in that all participants see the same set of options. Cons: What if we happen to sample a strange set of uncorrelated options? One way around this could be to deliberately generate (rather than randomly sample) the uncorrelated options.
    - Randomly generate a set of options that differs for every participant. Pros: avoid the possible con of a fixed set across participants. Cons: Some participants will, by chance, see correlated options in the uncorrelated condition (due to randomly sampling a small-ish set of options). Also more difficult to aggregate data across participants…
    - At this point, I probably lean toward #1. Happy to discuss further.
  + Coding item pairs. It will help in the analysis phase if we have a coding system recorded in the data file where each unique pair is given a unique (arbitrary) ID. This can be anything - eg, the digits 1 through to K, where K is the number of unique pairs. This will make analysis a lot easier because we can easily link the similarity ratings from phase 1 to the choice/RT from phase 2 for each unique pair (rather than the more painful process of matching attributes of the options in each pair across phases).
* **Analysis approach (for honours):**
  + Simple: Mean similarity for each pair of options in phase 1 compared to mean RT for choices between the pairs in phase 2 (this is where the ID for each pair will be helpful!). See whether this relationship differs between the correlated and uncorrelated conditions. Also test mean similarity ratings for correlated and uncorrelated options… This one might still need some thought; I expect lower overall similarity ratings for uncorrelated attributes *however* this is manipulated between subjects and participants are instructed to use the full range of the scale… which may 'wash out' any differences in mean similarity across the (between subject) conditions. There's a more straight forward test of the latter idea in the 'complex' analysis
  + Complex: This one depends on Jess (honours student) to some extent, and what she feels capable of doing. Instead of analysing mean similarity ratings, use them to get an MDS representation, and then relate the MDS representation to RTs in phase 2. We can also look at whether the MDS representation differs for correlated and uncorrelated options. This is interesting though also a bit tricky for honours as there is no simple hypothesis testing approach to compare MDS representations, other than something like Quentin Gronau's approach from his AMPC 2020 talk… but this should be fine for honours.

**Project 3: Mikaylee Baldwin, working with J-P.**

* **Research question:** Two possibilities. How does time pressure affect the perception of attribute correlations? Do attribute correlations facilitate decisions under time pressure? Or both.
  + **Hypothesis:** 
    1. The correlation between attribute utility parameters will be greater when the stimulus has attribute correlations than when it does not have attribute correlations.
    2. Correlations in the stimulus attributes will facilitate decisions under time pressure.
* **Experiment:** This is a 2x2 mixed design: correlated vs uncorrelated attributes (between subjects) and speed x accuracy instructions (within subjects). I suggest using the speed vs accuracy DCE as a template. Each trial will have 3 options, as in the previous DCE, where each option has 5 attributes, based on our mobile phone database.
  + SAT manipulation: I think this should happen in the same way asthe previous SAT instructions DCE - a cue for each trial, where speed vs accuracy is randomised across trials.
  + Correlated vs uncorrelated attributes manipulation:
    1. Correlated attributes: On each trial, randomly sample three rows from the phone database. This will preserve the correlations present in the real mobile phones from the marketplace. This should be random sampling *without replacement* to make sure we don't have any phone repeats on a trial.
    2. Uncorrelated attributes: On each trial and separately for each attribute, randomly sample a level from each attribute column. This will preserve the base rate frequency of each attribute level (e.g., 12MP phones appear more often than other MPs in the real phones, and sampling with this approach will conserve that base rate in the uncorrelated condition) and will greatly attenuate the attribute correlations. It won't completely remove the attribute correlations, owing to the different base rates for the levels of different attributes, though they will be much lower than the correlated attributes condition.
  + Number of trials per participant will need pilot testing. My guess is that it will be similar to the previous SAT experiments, perhaps slightly fewer trials since there are more attributes in the new experiment (5 vs 3/4 in the previous experiments).
* **Analysis approach (for honours):**
  + Different test for the two hypothesis.
  + Apply a mixed logit model to the choice data separately in the 4 cells of the design: correlated vs uncorrelated stimulus attributes crossed with speed vs accuracy instructions. For honours, then conduct a hypothesis tests on the utility correlations that are estimated in the mixed logit analysis.
  + Use the mixed logit analysis from step 1 to calculate the estimated R-squared for the analysis and compare across cells of the design. This tells us how well the model accounted for the observed choices. We then test if this is larger for correlated vs uncorrelated stimulus attributes for speed instructions. Overall, we expect choices to be less variable for accuracy than speed instructions, so R-squared should be greater for accuracy.

**Project 4: Kezia Beale, working with J-P.**

* **Research question:** What is the time course of information processing for multi-attribute consumer-like options?
  + **Hypothesis:** The time course of information processing will differ between the two attributes. This can be interpreted to reflect a processing preference/bias for multi-attribute stimuli.
* **Experiment:** Completely within-subjects experiment based on the response-signal (RS) task run in 2019. I suggest repeating that task with the following changes:
  + Adding a visual mask in the "respond now" interval. I suspect this will lead to lower choice proportions (closer to 50%) for the very fast display times, which is more sensible. The mask will replace the text shown in the cells of the DCE and should match the important stimulus properties: font type, size, alignment, etc. To maximise the effectiveness of the mask, I suggest that we randomly sample from the letters that make up both levels of both attributes (i.e., the letters "H", "i", "g", "h", "Q", "u", "a", …). This will match the base rate occurrence of letters across all stimuli with letters in the mask. The "respond now" text can be relocated (e.g., underneath the stimulus) or cued in some other way, like a coloured box around the stimulus (DRT style). I don't mind how the cue looks as long as it is very salient (e.g., a different colour to the stimuli, heavy/bolded lines, etc) and the participant is given clear instructions on what to do when the cue appears.
  + 9 stimulus display durations (ms): 0, 50, 100, 200, 400, 800, 1200, 2000, 4000. The display times from 100ms and above should match the 2019 task (please adjust the above if I've misremembered them). One new aspect to this experiment is the addition of the very fast durations: 0ms and 50ms.
  + The number of trials per block will be slightly higher. I think we decided that there are 6 possible pairs of options and 9 display times for a total of 54 trials per block. There will need to be a corresponding decrease in the number of blocks so that total experiment time remains similar to the 2019 RS task.
* **Analysis approach (for honours):**
  + 2 (attribute) x 9 (display time) within-subjects ANOVA. This will definitely give a main effect of display time and will likely also result in an interaction between the two factors (given the preference for quality over price in the 2019 RS task). If there is an interaction, do pairwise comparisons between adjacent levels of the stimulus display time (rather than all pairwise combinations). This will test whether each increment to display time led to statistically reliable change (increase) in choice proportions. These could be one-sided t-tests owing to the clear directional hypothesis.
  + An extended analysis could take the average SAT functions and fit separate shifted exponential functions to the quality and price choice data. This would allow some discussion of the shift, rate and asymptote parameters. Even if this is a qualitative comparison of the parameters it could still be an interesting discussion point for an honours thesis.