Pre-registrations

By Jonas Kristoffer Lindeløv, April 8th, 2019.

Four student groups are conducting experiments as part of their 4th semester on Psychology at Aalborg University. They have self-selected one among several project ideas proposed by their supervisor, Jonas Kristoffer Lindeløv.

Jonas has quickly written this document to pre-register the planned analyses. The student projects will use averaging, etc. to make accessible for simple statistics a la frequentist t-tests and correlation. This pre-registration plans to do the analyses on raw data using Bayesian mixed models.

As of writing this document, data collection on all projects will commence in a few days, but only feasibility-piloting has been done at this point.

The validity of a 5-way CMC design in the study of individual differences in CMCs

The design is highly similar to the pre-registered protocol for the study “Individual Differences in Cross-Modal Correspondences” by Hazel Anderson, Jonas Kristoffer Lindeløv, and others. Indeed, most stimuli and response procedures are identical, as a several of the analyses. The present protocol can be seen as a update to that protocol, so I will point to *deviations* to the protocol for the previous study.

The main purpose of this project is to study whether the simultaneous alteration of multiple stimulus features reproduces the CMC effect sizes and directions obtained via “classical” one-inducer designs. This would many-fold the efficiency with which many CMCs could be studied (not unlike what the optimum1 paradigm did for the study of mismatch negativity), thus paving the way for more studies on individual differences.

We plan to recruit 50 participants.

# Design changes

The main changes in the experimental design are the following:

* **Hue is out.** Due to experimental time constraints and lack of prior literature, we “only” study shape, lightness, complexity, pitch, and timbre.
* **2-way on pitch-lightness and pitch-complexity.** These are studied using the same design as previously. To keep it short, response keys are not counterbalanced within-subject, but they are random between subjects. They were selected from the following two criteria. (1) pitch-lightness has a large body of literature and a large degree of consensus between participants as per the rating-data from our previous study. (2) pitch-complexity has a relatively small body of literature and a low degree of consensus between participants.
* **5-way.** We have a new condition, where there are four inducers instead of one, at any given trial. So while responding to pitch, you could experience a pointy (shape) bright (lightness) and complex (seven protrusions) figure while the sound is square-wave (timbre). 128 trials will be collected for each CMC for each participant (64 congruent and 64 incongruent), e.g. the effect of complexity on pitch.

# Analysis changes

The main outcome is whether the CMC effects from the 5-way design replicate the CMC effects from the 2-way design. Specifically:

* **Two-way hypothesis: effect sizes are identical.**  
  rt ~ design \* congruency \* task + (random…)  
    
  … where *design* is [2-way, 5-way]. For this analysis, we will subset the 5-way data so that only CMCs studied in the 2-way design is included, i.e., when participants are responding to pitch, lightness or complexity. *Congruency* will then be defined as in the 2-way design, irrespective of the other features.
* **Directional hypothesis: effect sizes have the same sign (same direction):** Compare effect directions. Studying the posteriors for each task separately, what is the joint probability that each CMC is in the same direction? I.e., in the upper right quadrant or the lower left quadrant. This will be done on a CMC-by-CMC basis.
* **Two-way hypothesis: we can replicate earlier 2-way results:** Some of the CMCs, we studied in our previous experiment. Others have been studied in the literature. See whether we can replicate them.

Other analyses made possible by the 5-way design:

* **Directional hypothesis: the more congruency with the task, the lower the RT:** Is the response time inversely proportional to the number of congruent features in a trial? In it’s most general form, this will be a linear or an ordinal regression across all conditions:  
    
  rt ~ n\_congruent + (0 + task|id)  
    
  … where n\_congruent is 0 to 4. One down-side to this is that some congruency effects are likely stronger than others and thus the model generalizes too much (large residuals). We may do this analysis for single tasks as well. A coarser analysis would be to compare RTs when n\_congruent = 0 and n\_congruent = 4. But there will be relatively few trials in each of these conditions, so we do not plan to do this.
* **Directional hypothesis: congruency among inducers decreases RT.** Very much akin to the analysis above, is RT lower when the inducers are congruent with each other, irrespective of whether they are congruent with the task?

# Reproducing earlier analyses

I just want to re-iterate, that we also want to extend some analyses in the “old” pre-registration.

* **Directional hypothesis: Rating-defined congruency better predicts CMC than one-size-fits-all congruency?** If the 5-way design is successful in eliciting CMCs, we can study for 20 CMCs where we previously “only” studied eight. This should allow us to overcome a major issue with our previous study: we selected CMCs where participants had a large degree of consensus about their direction.
* We can now do a factor-analysis on many more CMCs. This will be vastly underpowered (to state it in frequentist terms), and this will be communicated clearly in the reporting.

The accuracy of subjective judgment of objective performance

There is an extensive literature, which suggests that we are relatively poor at judging our own performance.

# Design

In this study, participants are asked to do a number of tasks and to estimate their own performance. We only seek to improve a bit methodologically, and to study some potential covariates. We plan to recruit 120 participants to be able to detect medium-sized correlations.

The performance tasks are:

* WAIS-IV Working memory
  + WAIS-IV digit span
  + WAIS-IV arithmetic
* WAIS-IV Processing speed
  + WAIS-IV symbol search
  + WAIS-IV digit-symbol coding

After each task, they will be asked to

estimate their percentile score. The objective percentile score will be calculated using the norm WAIS-IV norm materials, so the two can be compared.

In addition, they will be asked to complete two brief questionnaires:

* Self-efficacy using the General Self-Efficacy scale (GSE)
* Sub-clinical narcissism traits using the Narcissistic Personality Inventory (NSI).

Planned analyses:

* **Directional hypothesis: there is a weak positive correlation between subjective and objective percentile scores/ratings.** This is a simple correlation. The “weak” hypothesis will be characterized by a ROPE in the interval *r* = 0.0 to 0.25.
* **Two-way hypothesis: Subjective percentile is predicted by narcissism and self-effiacy over and above objective performance.** Using multiple regression, we model:  
    
  perc\_subj

~ perc\_obj + GSE + NCI

* **Two-way hypothesis: the *precision* of subjective judgement is predicted by narcissism and self-efficacy.** The precision is simply the absolute difference between the objective and the subjective percentiles. So the test becomes:  
    
  perc\_diff ~ GSE + NCI

Does a transfer-motive improve transfer from n-back training?

This design is a mix of the setup of Thorndike (1901) and many of the N-back training studies that ensued the Jaeggi (2008) paper, which we believe is deeply flawed, as exposed by Shipstead, Redick, et al. The underlying hypothesis here is something akin to what Judd (1908) proposed, that we need to directly put the participants in a “transfer-forming state of mind” to facilitate transfer.

We stay within the n-back framework, so all effects are relatively near transfer as compared to the general literature. Check out the quadro N-back. It is fun!

# Design

We pre- and post-test participants on

* Nearer transfer: Dual 2-back
* Medium transfer: Triple 1-back
* Farther transfer: Quadro 1-back

The tested features are:

* Position in a 3x3 grid
* Shape of the figure: circle, square, triangle, star.
* Hue of the figure: around 7 alternatives.
* Sound: equally-sounding Danish letters: B, C, D, E, T, …

All participants get to do a small bit of practice on the pre-test before the experiment starts. They receive feedback on their responses and lacking responses. During pre-test and post-test, no feedback is given.

After the pre-test (~10 minutes), participants train a Dual 2-back on position and sound for around 20 minutes. They receive feedback during training, and intermittent status screens with their performance on the two modalities so far.

Participants are randomly divided into one of two groups for the training part.

* The “training-motive” group receives an instruction to focus exclusively on improving the training.
* The “transfer-motive” group receives an instruction to use the training to improve on the post-test.

We plan to recruit 60 participants – 30 for each group.

# Analyses

* **Directional hypothesis: the transfer-motive group has superior improvement on each of the post-test tasks.** A d-prime will score will be calculated for each task in the pretest and posttest. The difference in d from pretest to posttest is the dependent variable. An independent t-test on each outcome will answer the hypothesis.