Item analysis

1. Data cleaning: by facet. No reverse coding or anything when nothing is sure.
2. Obtain in SPSS response frequencies and the histograms of the frequencies in order to detect any abnormal distribution. Normal distribution should have low frequencies for extreme response categories, but higher frequencies for more neutral ones.
3. After obtaining the frequencies distributions, put all items in GGUM and obtain the item parameters (using GGUM instead of MULTILOG because in GGUM, we don’t need to recode any responses, but in to use MULTILOG, the dominance model in other words, you need to first reversely code the items that require reverse coding).
4. After obtaining the item parameters, put them as well as item response into the excel MACRO called MODFIT, and with the fit plots and stats obtained, try to detect any abnormal items.
5. Mean chi^2/df – rule of thumb: 3
6. Plot: if discriminating, then item is fine; otherwise, not fine.
7. In DIF study, delete items based on plot that are problematic. If one item is not good in one group, delete it in both groups even though it works fine in the other group.
8. ~~Use the remaining items that exist in both groups to run a DIF analysis under GGUM, using Wei’s R code.~~

============================================================================

1. Before running DIF analysis for dominance model, items need to be all reversely coded.
2. To decide which items to be reversely coded, refer to GGUM model fit plots, FLs came along with the scale, and single item fit statistic.
3. If item performs well in one sample but not the other, then almost definitely this items has DIF, so don’t delete the item, but instead reverse (or not) the item based on the sample where it works fine.
4. For intermediate items, plots may indicate that the item does not perform poorly, but is not clear if the item should be reversed. In this situation, just keep the item without reversing it.

=============================================================================

1. Use all common items (reverse code all that need to be reversely coded) of both samples surviving the MODFIT plot check, and throw the item responses in MULTILOG, and obtain items pars
2. In MULTILOG, the a parameter is 1.7\*a, so after we obtain item pars from 13, select items again. This time, delete items that have a parameters smaller than 0.51 in both groups 🡪this how we got the final items used for dominance model DIF analysis (log likelihood ratio and effect sizes before summer break) 🡪 see folder 0420\_free&constrain\_2 under dropbox for pars
3. After obtaining the item parameters, do constrained baseline model first using R and data (item responses; reverse coding done) 🡪 constrained baseline model has inflated Type I error rate, meaning if an item is indicated as having no DIF using a constrained baseline model, then it really DOES NOT HAVE DIF 🡪 more conservative way to find non DIF items for linking purpose
4. After obtaining chi-square statistics for each item using the constrained baseline model, figure out item with no DIF 🡪 in our case, all items have DIF 🡪use items having chi-square statistics that are the closest to chi-square critical value as non DIF items and to serve as linking items for free baseline model (this conduct is valid because we have a large sample, which mean our tests have large power, and therefore, any diff across samples may be shown as significant)
5. With linking items identified for each facet respectively, we use the free baseline model for a more accurate DIF analysis 🡪chi-square values are negative, indicating null hypothesis fit better than H1 🡪 again, all items are shown to have DIF 🡪 large sample, high power, sigh….
6. Therefore, we switched to Nye’s effect size indices (MATLAB code can be found in dropbox, so can the revised version in R) 🡪rule of thumb: effect size <0.2 🡪 no DIF 🡪 we actually have a lot more items that have not DIF 🡪 yay!
7. Go back to looking at each items that have been shown to have DIF in 18, and figure out conceptually why they have DIF 🡪 based on the content and wording of the ites
8. Run DIF analysis using GGUM, and compare differences between GGUM and dominance model

=============================================================================

**1020\_2015**

1. GGUM gave unsatisfying item parameters; to be more specific, some estimates are extremely large, as what we encountered with the dominance model
2. Fritz suggested that we use MCMC GGUM to compute item parameters, because MCMC GGUM uses Bayesian estimation, which hopefully will return less crazy estimates of parameters.
3. Tried MCMC with default priors coming along with the example syntax, but didn’t work 🡪 not that the est. are too wild, but they are just too good to be true 🡪 everything is similar (small s.d.) and similarly good 🡪 might be because the priors are so strong that they overwhelmed the estimation.
4. Therefore, attempts were made to loosen (weaken) the priors a little bit 🡪 changing them to non-informative priors (AKA most of the time it’s the uniform distribution that is flat) 🡪 first only the prior for alpha was altered, but still results are similarly good; 🡪then both priors of alpha and delta were changed 🡪 still not much improvement 🡪then priors for all 5 pars (alpha, delta, tau 1, tau 2, and tau 3) were altered to uniform distribution (with the default support area) 🡪again, no improvement. Not sure what to do next
5. Thinking of using GGUM2004 to run the constrained and free baseline model as under the dominance model using MULTILOG, and then run effect size analysis?