ADDENDUM

We hereby note departures from the PAP made while we were still using scrambled data:

Choices not pre-specified

- for the variable `num_calm_methods` , we pre-registered that it would be coded as a sum of the available choices of calming methods. Yet, in coding the data, we discovered a great number of unanticipated options provided in the `calm_method_oth` category. We decided to score these with a count, and add it to the sum used for `num_calm_methods` using the following decision rules:
  - If the respondent suggested crying, fighting, arguing, or drinking alcohol 0 was added to `num_calm_methods`
  - If the respondent suggested dancing, praying, visiting friends / family, talking calmly, singing, working, playing with children, or other non-confrontational / non-destructive reactions, we added 1-3 depending on the number counted

Field Issues

- for `emo_reg_*` variables we adjusted the choice options after piloting the questionnaire in the field. They are now

```
  1      Almost never
  2         Sometimes
  3 Most of the time
  4    Almost always
-99        Don't know
-88 Refuse to answer
```

- for `fear_partner` we added a category that is `5 IN THE PAST (NO LONGER AFRAID NOW)` thus we need to re-write coding in PAP so `fear_partner_i` is not just based on `fear_partner > 1` but rather should have following explicit coding:

```
1      1                              NEVER = 0
2      2                          SOMETIMES = 1
3      3                         MANY TIMES = 1
4      4                 MOST/ALL OF THE TIMES = 1
5      5 IN THE PAST (NO LONGER AFRAID NOW) = 0
```

- for `pair_*` questions we added another response category `Often` after piloting

```
1   Never
2   Rarely
3   Sometimes
4   Often
5   Always
```

Divergences from pre-specification

- Departure from pre-specified lasso procedure. In the pre-analysis plan, we state "Since this regularization parameter cannot be optimally chosen in advance, we will select it using 10-fold cross-validation. Specifically, for each outcome, we will choose the lambda that minimizes the 10-fold cross-validation error averaged over 10 runs (since the folds are chosen at random). Only the covariates retained by the lasso will be included in the specification." Upon further inspection of the lasso's performance, we determined that the procedure tends to result in quite different values of lambda, which in turn result in very different model selections. This primarily happens because 10-fold CV doesn't give a very good estimate of the optimal lambda. By contrast, using 20 or 30-fold cross-validation produces very stable estimates of lambda. Given the size of our sample, we are able to use a much greater number of folds and thus get better estimates of lambda. Hence, we are now using 30-fold validation, and taking the average lambda across simulations. We now note: "Since this regularization parameter cannot be optimally chosen in advance, we will select it using 30-fold cross-validation. Specifically, for each outcome, we will choose the lambda that minimizes the 30-fold cross-validation error, averaged over 10 runs (since the folds are chosen at random). Only the covariates retained by the lasso will be included in the specification."

- Departure from pre-specified definition of compliance types and non-compliance analysis. In the pre-analysis plan we defined the following compliance types for individuals:

Strata	Definition
Compliers	Attend at least one session during the first cohort when assigned to treatment, attend no sessions in first or second cohort when assigned to control.
Always-Takers	Attend at least one session during the first or second cohort whether assigned to control or to treatment.
Never-Takers	Never attend at least one session, irrespective of the cohort to which they were assigned.
Defiers	Attend at least one session during the first or second cohort when assigned to control, and no sessions when assigned to treatment.

Based on the observed rates of compliance we decided to make the following (blind) modifications to these definitions.

Strata	Definition
Compliers	Attend **at least 50% of sessions** during the first cohort when assigned to treatment, attend no sessions in first or second cohort when assigned to control.
Always-Takers	Attend at least one session during the first or second cohort when assigned to control **and attend at least 50% of sessions during first cohort when assigned to treatment**.
Dropouts (Never-Takers)	Attend **less than 50% of sessions** during the first cohort when assigned to treatment, attend no sessions in the first or second cohort when assigned to control.
Defiers	Attend at least one session during the first or second cohort when assigned to control, and **less than 50% of sessions** when assigned to treatment.

- In the description of the BART compliance model in the PAP, we mention "*We will train the models using 10-fold cross validation, and then obtain predictions of always-takers in the treatment and never-takers in the control.*" We were only able to do **5-fold CV** due to resource constraints.

Notational discrepancies

- In the pre-analysis plan, we denoted variables that were recoded to point in a direction consistent with the meaning of an index using the suffix `_i` . We denoted whether variables at the couple level were referring to men or women using the suffixes `_m` and `_w` , respectively. We still do so. However, in the PAP we imagined that index recoding would be done after the couple-level transformation, so that variables would have the suffix `_m_i` , for example. However, we are recoding before reshaping, so that the suffix is instead of the form `_i_m` . We do not anticipate this will change results in any substantive way.

- the `disc_sex` variable was changed to `disc_sex_alt` in the midline survey (this is not reflected in the PAP)

- For "meta-beliefs", we pre-registered the following notation to describe the difference in the respondent's belief about whether their partner trusts them (`P_trusts_R_r`) and what their partner says about trusting the respondent (`R_trusts_P_p`): `(P_trusts_R_r – R_trusts_P_p)^2` . The problem with this notation is that the first term is from the perspective of the respondent, whereas the second is from the perspective of their partner, which is confusing. Instead, we are now adopting the following convention:

  - Everything is from the perspective of the respondent
  - The final suffix `_r` or `_p` denotes whether the response was provided by the respondent or by the partner, respectively.
  - E.g. `income_r_sep_r` is asking respondent about money they keep separate, while `income_p_sep_r` is asking respondent about money their partner keeps separate, `income_r_sep_p` is asking respondent's partner about their beliefs about money respondent keeps separate, and `income_p_sep_p` is asking respondent's partner about money partner keeps separate.

- two of the sexual violence questions had their variable names shortened in the midline (not reflected in the PAP). Namely `sexual_forced_intercourse_5mo_w` became `sex_forced_intercourse_5mo_w` and `sexual_forced_other_acts_5mo_w` became `sex_forced_other_acts_5mo_w` .

- `perc_att_abuse` was changed to `perc_att_unfaithful` in midline.

Errors in PAP

- `relation_quality_divergence` should be defined as `(vignette_3_w – vignette_3_m)^2` **not** `(vignette_3_m – vignette_3_m)^2`