The Space between Pre-Registration and Publication: Deviation Documentation

Issues:

* How to set-up a pre-registration to be evaluated by the research team, and others, after data is collected
* How and where to communicate deviations from the pre-registration?
* What to report?
  + Study design: were changes made to the study design during the research process?
  + Data analyses: were changes made to the planned analyses?
  + Additional analyses: were new analyses conducted during the research process?
* **Suggestion:** create reporting table in the pre-registration document that can be updated throughout the research process
* Perhaps add relevant questions to the Transparency Checklist (see the Shiny App linked in the “relevant resources” section)

Relevant Resources

* Developing a lab culture for handling making mistakes in the lab: <https://psyarxiv.com/gxcy5>
  + AE: My thinking about the above was that making a deviation from a pre-registration is not a “mistake” per se, but more about taking inspiration from Jeff’s paper about how to plan for deviations and make robust technology for adjusting to and documenting them.
* The replication recipe has some items regarding deviations, original paper versus replications. Might be able to adjust things.
* Relevant OSF page re Pre-registration: <https://osf.io/4acje/>

Improvements on pre-registration templates:

1. Moving from free-text to components that can be clearly evaluated.
2. Adding evaluation tables in the pre-registration to make it easier to evaluate later.
   1. Promotes thinking in advance about any deviations that may occur or be required

Potential Table for Documenting Deviations

* Summary table at end of pre-registration document (?)
* Columns?
  + Essentially, what was planned?
    - Specific plans (e.g., confirmatory hypothesis); open plans (e.g., run an EFA with no specific expectations for the outcome)
      * Was the plan followed? Yes, no, yes but some changes made
  + What was not planned but done?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Components** | **Location of the preregistered decision/plan** | **Location of the rationale for the decision/plan (if any)** | **Were there deviations? (no, minor,  major)** | **If yes - describe the deviation(s) or link to location in manuscript/supplementary** | **Rationale for deviation or link to location in manuscript/supplementary** | **How might the results be different if you had not deviated** |
| 01\_Data Import |  |  | Minor | Imported data for each wave based on real data | Else import would not have worked for real data | No difference |
| 01\_Data Import |  |  | Minor | Removed first check for Rows and CASEID | Was repeated after connection of whole dataframe, thus unnecessary repetition | No difference |
| 02\_Datawrangling  ... |  |  | Minor | Changed name of df according to 01\_Data Import | Code would not work else | No difference |
| 02\_Datawrangling |  |  | Minor | Added: table(cccu\_complete1$EC\_recode)  crosstabs( ~ is.na(EC) + WAVE, data = cccu\_complete1)#item is missing at t1 | To look for high number of NA | No difference |
| 02\_Datawrangling |  |  | Minor | Added:  crosstabs( ~ variable + WAVE, data = cccu\_complete1) | To look if proportions are the same | No difference |
| 02\_Datawrangling |  |  | Minor | Selected further necessary variables to describe sample: HHINC, REGION,  STATE, EMPSTAT,  RACETH, MARSTAT | to describe sample | No difference |
| 02\_Datawrangling |  |  | Minor | Recoded HHINC, REGION,  STATE, EMPSTAT,  RACETH | To describe sample |  |
| ...03\_Exclusion |  |  | Minor | Changed name of df according to 01\_Data Import | Code would not work else | No difference |
| 03\_Exclusion |  |  | Minor | Changed recoding of cccu\_complete1$exclusion\_R3 | Code was wrong | No difference |
| 03\_Exclusion |  |  | Minor | Added code to see number of participants after each exclusion step | To describe exclusion | No difference |
| 05\_Descriptives |  |  | Minor | Described Participants at t1, compared included vs excluded sample, described variables at t1 and made correlation matrix, visual checks for normality | To describe sample and variables as code was missing | No difference |
| 06\_MainAnalysis |  |  | Minor | Changed name of df according to 03\_Exclusion | Code would not work else | No difference |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Minor | Added optimizer to models | For model convergence | No difference |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Minor | Defined method “wald” for confidence intervals | Computational power problems with other methods | Difference in exactness of confidence intervals |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Minor | Plotted all effects via plot(alleffects) | Was missing | No difference |
| 02\_Datawrangling |  |  | Minor | Combined categories “no sex” and “once” for sex\_freq | As categories contained only few participants 🡪 to combine into one categorie | Minor difference as reference category for sex\_freq is now containing no sex or once |
| 02\_Datawrangling |  |  | Minor | Defined factor levels for sex\_freq and hc | To define contrasts in models | Different reference contrast |
| 03\_Exclusion |  |  | No | Excluded steps to deal with high missingness in the simulated data | As I was now working with the real dataset | No difference |
| 07\_Subanalysis |  |  | Minor | Changed name of dataframe according to 03\_Exclusion | Code would not work else | No difference |
| 08\_Robustnessanalyses |  |  | Minor | Changed name of dataframe according to 03\_Exclusion | Code would not work else | No difference |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Minor | Computed average marginal effects as effect size | To compute effect size | No difference |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Medium | Spitted models into hormonal / non-hormonal users | to compute average marginal effects for the interaction terms if effects were significant | No difference |
| 06\_MainAnalysis, 07\_Subanalysis, 08\_Robustnessanalyses |  |  | Medium | Computed model 4 without multilevel structure as glm model for the first possible measurement occasion (only one wave), repeated model 3 the same way | As model 4 did not converge otherwise (and to control for dependence of longitudinal data) and to see if effects in model 3 change because of different modelling | Minor difference because of different modelling |
| 06\_MainAnalysis |  |  | Minor | Computed sensitivity analysis: E-values | Computation was still missing | No difference |
| 04\_Codebook |  |  | Minor | Added code for Codebook | Was missing | No difference |
| 09\_Plots |  |  | Minor | Added code for plots | Was missing | No difference |

**Locations should include page number and paragraph or line number (as specific as possible)**

Do you think it affects validity of statistical inference - are you less sure about the result because you had to deviate?  (Note: self-report/introspection is only worth so much, but still good to know what author thinks)

Prioritizing the things that are most important? (if such a thing is possible)