PDR Clothing and Item Response Theory

Preregistration

Robert Chapman

11/28/2021

*Below are 10 questions to be included with the pre-registration form for secondary data analyses as recommended by* [*Mertens & Krypotos 2019*](https://psycnet.apa.org/doi/10.1037/abn0000424)

## What is the hypothesis that will be investigated?

**Research Question:**  
For the purpose of further evaluating and comparing the scale/item set of Painful, Distracting and Restrictive (PDR) clothing in [Engeln & Zola 2021](https://doi.org/10.1007/s11199-021-01230-9): what is the dimensionality and appropriateness of fitting of an Item Response Theory model to PDR clothing items representing ‘discomfort for appearance’

**Hypotheses:**

1. There is either a unidimensional or multidimensional construct of ‘discomfort for appearance’ as measured by the PDR clothing items?
2. An Item Response Theory model may be fit to the PDR clothing item data, and will allow further evaluation and comparison of PDR clothing items in relation to the central construct of ‘discomfort/distraction for appearance’.

## How will the crucial variables by operationalized?

1. Variables representing clothing choices, as presented and tested in the article (<https://osf.io/ajv5z/>)
2. Demographic variables:
   * Gender

## What is the source of data included on the analyses?

**Two samples:**

1. Online Qualtrics survey of undergraduate men and women through social media platforms (e.g., Facebook, Reddit)
2. Online survey via Mechanical Turk

**Source of Data**  
Openly accessible data, published by the authors, here: (<https://osf.io/ajv5z/>)

## How will this data be obtained?

Data will be obtained from the OSF repository, (<https://osf.io/ajv5z/>)

## Are there any exclusion criteria for the data?

In sample 1, persons were included in the analysis set if they were undergraduates who self-reported cisgender identity of male or female. Persons were excluded if they did not self-report cisgender identity of male or female.

In sample 2, the same inclusion/exclusion criteria were used, with the addition of exclusions from the analysis set for inattentiveness (via in-survey attention checks).

**Data will be excluded from analyses if they**

* cause a violation of assumptions in the statistical analysis
  + e.g., a single case exerts enough leverage/influence to cause a violation in normality
* represent an outlier (defined as > 3.29 Z-score)
* can be reasonably justified to be not a part of the population
* If they are missing a problematic amount of data
  + e.g., case is missing >50% of observations or statistical modelling requires complete data

## What are the planned statistical analyses?

**Set-up & Hypothesis Building**

1. Additional literature review (part of prereg)
2. Additional Hypothesis building (part of prereg)
3. Initial coding and scripting with simulated data from codebook & sample size

**Data Prep and Description**

1. Data Processing and Cleaning
   * Visual inspection of data for clear errors or empty rows/columns
     + *Base R: View()*
   * Checking that variables are correctly formatted
     + *Base R: class()*
   * Cross-check data variables with data dictionary
     + *Base R: names(Data) %in% names(Codebook)*
   * Checking that variables meet expected range/values
     + *Base R: unique(Data[,var]) %in% Codebook[Var]$Values*
   * Visualization
     + *Base R: plot(), pairs()*
2. Missing Data
   * Evaluate missingness patterns/Littles MNAR test/T test comparisons
     + *‘Missing Data’ script*
3. Normality
   * Skew/Kurt/QQplot
     + *‘Assumption Checking’ script*
   * Outlier detection
     + Boxplots/zRange/plausible meanSD’s
       - *‘Outlier Detection’ script*
4. Descriptives
   * Mean, Median, Standard Deviation, Frequency, Range
     + *Base R: mean(), median(), sd(), table(), range())*
   * Correlation matrix
     + *psych package: tetrachoric()*
   * Standardized Cronbach’s Alpha
     + *psych package: alpha()*

## Replication of paper analyses

1. Dimensionality
   * Scree plots
     + *psych package: fa.parallel()*
   * ICLUST
     + *psych package: ICLUST()*
   * Bi-factor modeling
     + *psych package: omega()*

*review of dimensionality analyses results alongside items and literature review to qualitative guide decision on number of factors to included or modelled*

1. IRT Assumption Checks  
   *Will need to be further fleshed out for which model is selected*
   * Monotonicity - *assumed for dichotomous data*
     + Unidimensionality
       - Omega hierarchical/total & Explained Common Variance
         * *psych package: omega()*
     + Local dependence
       - *Rasch: Q3*
       - *2pl: jackknife slope index – Q3 for rasch*
     + Item invariance (DIF)
       - *Rasch: Mantel-Haenszel*
       - *2PL: logistic ordinal regression for 2pl?*
2. IRT Modeling
   * Models to be evaluated:
     + Rasch
       - *psych package: rasch()*
     + Rasch w/guessing
       - *???psych package: rasch()*
     + 2PL
       - *psych package: ltm()*
     + Multidimensional IRT
       - *mirt package: mirt()*
     + IRTree Modeling
       - *???? package: ???()*
     + Explanatory IRT Modeling
       - *???? package: ???()*
   * Model fit
     + Goodness of fit/deviance
       - *Package: AIC,BIC, Chi-Sq?*
   * Item fit
     + Infit/outfit/signed chisq
       - *Package: TBD*

## What are the criteria for confirming and disconfirming the hypotheses?

### Hypothesis 1

There is either a unidimensional or multidimensional construct of ‘discomfort/distraction for appearance’ as measured by the PDR clothing items?

**Confirming**  
1. There is additional literature supporting a construct of ‘discomfort/distraction for appearance’ as indicated by wearing PDR clothing.  
2. Scree plots, ICLUST and Bifactor analyses indicate a reasonable number of factors (1 to 5) in the PDR clothing set

**Disconfirming**  
1. There is additional literature contradicting a construct of ‘discomfort/distraction for appearance’ as indicated by wearing PDR clothing.  
2. Scree plots, ICLUST and Bifactor analyses indicate an unreasonable or untestable number of factors (e.g., 6+) in the PDR clothing set

### Hypothesis 2

An Item Response Theory model can be fit to the dimensions in the PDR clothing item data, and will allow further evaluation and comparison of PDR clothing items in relation to the central construct(s) of ‘discomfort/distraction for appearance’.

**Confirming**  
1. Model assumptions will be met for each IRT model to be fit, including monotonicity, local dependence and item invariance  
2. Goodness of fit indices for the model is appropriate (e.g., CFI/TLI >0.95, RMSEA <0.08) and for item fit (e.g., non-significant signed chi-square)

**Disconfirming**  
1. Model Assumptions will are not met for each IRT model to be fit, including monotonicity, local dependence and item invariance  
2. Goodness of fit indices for the model are not appropriate (e.g., CFI/TLI <0.95, RMSEA >0.08) and item fit (e.g., significant signed chi-square)

## Have the analyses been validated on a subset of the data?

The analyses have not been validated on a subset of the data

## What is known about the data that could be relevant for the tested hypotheses?

Original publication introducing the PDR item displayed high internal consistency (cronbach’s alpha >0.9) and group-level differences between gender.

## Please provide a brief timeline for the different steps in the preregistration?