Light Exposure Behavior Assessment (LEBA): Develop of a novel instrument to capture light exposure-related behaviours

Mushfiqul Anwar Siraji[1], Rafael Robert Lazar[2], & Manuel Spitschan[3]

[1] Department of Psychology, Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Malaysia

[2] University of Basel

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Mushfiqul Anwar Siraji: Data Analysis, Writing - Original Draft Preparation, Data Visualization; Rafael Robert Lazar: Data Analysis, Writing - Original Draft Preparation, Data Visualization; Manuel Spitschan: Data Analysis, Writing - Original Draft Preparation, Data Visualization.

Correspondence concerning this article should be addressed to Manuel Spitschan, . E-mail:

Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

Keywords: keywords

Word count: X

Light Exposure Behavior Assessment (LEBA): Develop of a novel instrument to capture light exposure-related behaviours

Introduction

Methods

Participants

This line is just a test for pushing in the github repo.

Material

Procedure

Our study had four objectives. First, to develop an instrument to assess individual's light exposure behavior . Second, to conduct an exploratory factor analysis(EFA) to understand the latent structure. Third to gather structural validity evidence for the latent structure obtained in EFA. Lastly, we gathered item information using Item response theory (IRT)(Baker, 2017)

Data Collection. Timeline of data collection, ethical approval, mode of data collection, how consent was recorded.

Item generation and Content Validity: Expert Panel Review. How we developed the 48 items?

Analytic Strategies

We used R (version 4.1.0), including several R-packages for our analyses. Necessary assumptions of EFA, including sample adequacy, normality assumptions, quality of correlation matrix were assessed. Our data violated both the univariate and multivariate normality assumptions. Due to these violations and the ordinal nature of our response data we used polychoric correlation matrix (C. Desjardins & Bulut, 2018) for the EFA. We employed principal axis (pa) a factor extraction method with varimax rotation. PA is apparently robust to the normality assumption violations (Watkins, 2020). The obtained latent structure was confirmed by minimum residuals extraction method as well. We used a combination factor indentification method including scree plot(Cattell, 1966), Horn's parallel analysis (Horn, 1965), minimum average partials method(Velicer, 1976), and hull method (Lorenzo-Seva, Timmerman, & Kiers, 2011) to identify factor numbers. Additionally, to identify the simple structure we followed the following guidelines recommended by psychometricians (i) no factors with fewer than three items (ii) no factors with a factor loading <0.3 (iii) no items with cross-loading greater than .3 across factors (Bandalos & Finney, 2018; Child, 2006; Mulaik, 2009; Watkins, 2020)

Results

Sampling adequacy was investigated by Kaiser-Meyer-Olkin (KMO) measures of sampling adequacy(Kaiser, 1974) . The overall KMO vale for 23 items was 0.63 which was above the cutoff value of .50 indicating a mediocre sample (Hutcheson, 1999).

Table1 summarizes the univariate descriptive statistics for the 48 items. some of the items were skewed with high Kurtosis values. The Shapiro-Wilk test of normality (Shapiro & Wilk, 1965) indicated all the items violated normality assumptions. Multivariate normality assumptions were investigated by Marida's test (Mardia, 1970). Multivariate skew = 583.80 ($p < 0.001$) and multivariate kurtosis = 2,749.15 (p <0.001) indicated multivariate normality assumptions violation. Due to these violations and ordinal nature of the response data polychoric correlations over Pearson's correlations was chosen (C. Desjardins & Bulut, 2018). Bartlett's test of sphericity (Bartlett, 1954), χ^2 (1128) = 5042.86, p < .001] indicated the correlations between items are adequate for the EFA. However only 4.96% of the inter-item correlation coefficients were greater than .30 in the obtained matrix. The inter item correlation ranged between .44 to .91. The corrected item-total correlations ranged between .10 to .44.

Scree plot (Fig3) suggested a six-factor solution.Horn's parallel analysis (Horn, 1965), like the Monte Carlo study, draws several sets of random data with the same number of participants as the original data set and compares the mean eigenvalues among the simulated and original data sets to retain optimal factors.This extraction method also supported a five-factor model. In our data set parallel analysis with 500 iterations indicated six-factor solution. However, In MAP method (Velicer, 1976) and Hull method (Lorenzo-Seva, Timmerman, & Kiers, 2011) suggested a five-factor solution. As a result, we tested both five factor and six factor solutions.

The initial five-factor solution with all 48 items showed the presence of cross-loading items (item 42, 16, & 1) and poor factor loading (<.30) items (item 20,3 ,15, 17, 40, 4, 11, 39, 18, 45, 29, 25, 8 , & 46). At first we discarded the items with poor factor loading and ran another EFA on the remaining 34 items. This iteration of EFA also appeared as a misfit in terms of poor factor-loading (Item 12, 22, 38, 6) and cross-loading (items 23, 31, 37, 48) . Another two rounds of EFA were conducted with gradually identifying problematic items and discarding them from the model. Finally, a five-factor EFA solution with 23 items was accepted with low RMSR = 0.04, no loading smaller than .30 and no cross-loading greater than .30. The obtained latent construct was also confirmed by using minimum residual extraction method (see the supplementary). Table**??** displays the factor loadin (structural coefficients) and commonality of the items. The absolute value of the factor-loading ranged from .47 to .99 indicating strong coefficients. The commonalities ranged between .10 to .99. However, the histogram of the absolute values of non-redundant residual-correlations (Fig4 showed 26.09% correlations greater than the absolute value of .05, indicating under-factoring. (C. D. Desjardins, 2018). Subsequently, we fitted a six-factor solution. However, in the six factor solution a factor emerged with only two salient variable loading thus disqualifying the six-factor solution.

Confirmatory Factor Analysis

Discussion

References

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. Retrieved from https://github.com/crsh/papaja

Baker, F. B. (2017). *The Basics of Item Response Theory Using R* (1st ed. 2017.). Springer.

Bandalos, D. L., & Finney, S. J. (2018). Factor analysis: Exploratory and confirmatory. In *The reviewer's guide to quantitative methods in the social sciences* (pp. 98–122). Routledge.

Barth, M. (2021). *tinylabels: Lightweight variable labels*. Retrieved from https://github.com/mariusbarth/tinylabels

Bartlett, M. (1954). A Note on the Multiplying Factors for Various Chi-square Approximations. *Journal of the Royal Statistical Society. Series B, Methodological*, *16*(2), 296–298.

Buchanan, E. M., Gillenwaters, A., Scofield, J. E., & Valentine, K. D. (2019). *MOTE: Measure of the Effect: Package to assist in effect size calculations and their confidence intervals*. Retrieved from http://github.com/doomlab/MOTE

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., . . . Borges, B. (2021). *Shiny: Web application framework for r*. Retrieved from https://CRAN.R-project.org/package=shiny

Child, D. (2006). *Essentials of factor analysis* (3rd ed.). New York:

Continuum.

Desjardins, C., & Bulut, O. (2018). *Handbook of Educational Measurement and Psychometrics Using R*. https://doi.org/10.1201/b20498

Desjardins, C. D. (2018). *Handbook of educational measurement and psychometrics using R* (O. Bulut & ProQuest (Firm), Eds.). Boca Raton, FL : CRC Press.

Dinno, A. (2018). *Paran: Horn's test of principal components/factors*. Retrieved from https://CRAN.R-project.org/package=paran

Epskamp, S. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output*. Retrieved from https://CRAN.R-project.org/package=semPlot

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18.

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Hutcheson, G. D. (1999). *The multivariate social scientist : Introductory statistics using generalized linear models*. London : SAGE.

Iannone, R. (2016). *DiagrammeRsvg: Export DiagrammeR graphviz graphs as SVG*. Retrieved from https://CRAN.R-project.org/package=DiagrammeRsvg

Iannone, R. (2021). *DiagrammeR: Graph/network visualization*. Retrieved from https://github.com/rich-iannone/DiagrammeR

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling*. Retrieved from https://CRAN.R-project.org/package=semTools

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/bf02291575

Kassambara, A. (2019). *Ggcorrplot: Visualization of a correlation matrix using 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggcorrplot

Lorenzo-Seva, U., Timmerman, M., & Kiers, H. (2011). The Hull Method for Selecting the Number of Common Factors. *Multivariate Behavioral Research*, *46*, 340–364. https://doi.org/10.1080/00273171.2011.564527

Makowski, D., Ben-Shachar, M. S., Patil, I., & Lüdecke, D. (2020). Methods and algorithms for correlation analysis in r. *Journal of Open Source Software*, *5*(51), 2306. https://doi.org/10.21105/joss.02306

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.1093/biomet/57.3.519

Mulaik, S. A. (2009). *Foundations of Factor Analysis* (Vol. 7). London: London: Chapman and Hall/CRC. https://doi.org/10.1201/b15851

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2021). *EFA.MRFA: Dimensionality assessment using minimum rank factor analysis*.

Retrieved from https://CRAN.R-project.org/package=EFA.MRFA

Ooms, J. (2021). *Rsvg: Render SVG images into PDF, PNG, PostScript, or bitmap arrays*. Retrieved from https://CRAN.R-project.org/package=rsvg

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. Retrieved from https://www.jstatsoft.org/v48/i02/

Ryu, C. (2021). *Dlookr: Tools for data diagnosis, exploration, transformation*. Retrieved from https://CRAN.R-project.org/package=dlookr

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Velicer, W. (1976). Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika*, *41*, 321–327. https://doi.org/10.1007/BF02293557

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth). New York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/

Watkins, M. (2020). *A Step-by-Step Guide to Exploratory Factor Analysis with R and RStudio*. https://doi.org/10.4324/9781003120001

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). *Readxl: Read excel files*. Retrieved from https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., & Hester, J. (2021). *Readr: Read rectangular text data*. Retrieved from https://CRAN.R-project.org/package=readr

Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*. Retrieved from https://CRAN.R-project.org/package=kableExtra

Table 1

Descriptive Statistics

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item1	1.12	0.49	5.02	27.80	0.25*	.16
Item2	2.16	1.19	0.71	-0.54	0.84*	.14
Item3	4.14	0.99	-1.23	1.14	0.79*	.19
Item4	2.87	1.59	0.08	-1.60	0.83*	.19
Item5	1.76	1.23	1.35	0.44	0.66*	.38
Item6	2.73	1.46	0.20	-1.36	0.87*	.33
Item7	3.86	1.67	-0.99	-0.85	0.65*	.23
Item8	3.76	1.14	-0.68	-0.45	0.86*	.00
Item9	3.42	1.83	-0.45	-1.69	0.69*	.33
Item10	2.74	1.04	0.09	-0.74	0.91*	.28
Item11	2.60	1.25	0.29	-0.86	0.89*	.35
Item12	2.11	1.17	0.77	-0.39	0.83*	.32
Item13	2.94	1.03	-0.12	-0.40	0.91*	.10
Item14	3.62	1.64	-0.68	-1.25	0.74*	.32
Item15	1.64	1.18	1.79	2.02	0.60*	.15
Item16	3.51	1.30	-0.70	-0.59	0.85*	.39
Item17	1.96	0.98	1.02	0.69	0.82*	.05
Item18	2.44	1.31	0.38	-1.14	0.86*	.11
Item19	3.80	1.29	-0.87	-0.42	0.82*	.17
Item20	4.01	1.40	-1.22	0.07	0.70*	.13
Item21	1.33	0.91	3.03	8.43	0.41*	.01
Item22	2.59	1.41	0.27	-1.27	0.86*	.19
Item23	1.31	0.81	2.75	6.92	0.43*	.21

Table 1 continued

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item24	1.47	1.18	2.38	4.00	0.43*	.28
Item25	2.56	1.27	0.33	-1.00	0.89*	.11
Item26	1.54	1.25	2.13	2.86	0.46*	.36
Item27	4.30	1.08	-1.79	2.53	0.67*	.22
Item28	2.27	1.39	0.74	-0.81	0.81*	.25
Item29	3.26	1.09	-0.26	-0.45	0.91*	.14
Item30	2.22	1.48	0.71	-1.02	0.76*	.30
Item31	1.05	0.36	7.23	52.98	0.13*	.18
Item32	1.54	1.21	2.07	2.75	0.49*	.31
Item33	1.04	0.33	8.99	85.28	0.10*	.16
Item34	3.36	1.38	-0.48	-1.03	0.87*	.16
Item35	2.26	1.25	0.70	-0.60	0.85*	.19
Item36	2.36	1.22	0.59	-0.62	0.87*	.25
Item37	1.14	0.59	4.79	24.05	0.25*	.16
Item38	2.25	1.27	0.69	-0.64	0.84*	.18
Item39	3.93	1.48	-1.06	-0.44	0.71*	.18
Item40	3.57	1.07	-0.65	-0.17	0.88*	.21
Item41	3.55	1.65	-0.60	-1.34	0.76*	.43
Item42	3.00	1.62	-0.08	-1.61	0.83*	.44
Item43	1.56	1.23	2.00	2.45	0.50*	.32
Item44	2.97	1.20	-0.06	-0.94	0.91*	-.10
Item45	2.79	1.55	0.19	-1.48	0.85*	.20
Item46	2.14	1.31	0.77	-0.78	0.80*	.26
Item47	2.18	0.90	0.60	0.12	0.86*	.26

Table 1 continued

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item48	1.48	1.11	2.18	3.35	0.48*	.24

Note. *p<.001

Table 2

	F1	F2	F3	F4	F5	Communalities
item1	0.06	-0.03	0.01	0.03	0.35	0.13
item2	0.12	-0.10	-0.11	0.69	-0.03	0.51
item5	0.01	0.16	0.09	0.01	0.69	0.52
item7	0.06	-0.09	0.66	-0.01	-0.03	0.45
item10	-0.01	0.82	0.07	0.02	0.02	0.68
item13	-0.06	0.34	-0.03	0.10	0.00	0.13
item14	0.00	0.05	0.89	-0.08	-0.08	0.81
item16	0.10	0.05	0.29	-0.11	0.31	0.21
item19	0.02	-0.06	0.00	0.80	0.03	0.64
item21	-0.05	-0.02	-0.34	0.03	-0.06	0.12
item24	-0.03	0.10	0.10	0.11	0.54	0.33
item26	0.93	0.00	0.13	-0.01	0.13	0.90
item27	-0.01	0.07	0.38	-0.12	0.21	0.21
item28	0.02	0.00	-0.05	0.01	0.31	0.10
item30	0.06	0.01	0.11	-0.04	0.52	0.29
item32	0.80	0.00	0.05	0.13	0.10	0.67
item34	-0.01	-0.14	0.02	0.84	0.12	0.74
item35	-0.04	0.46	0.04	-0.17	0.04	0.25
item36	0.09	0.63	0.10	-0.15	0.11	0.45
item41	0.05	0.07	0.70	0.30	0.14	0.60
item43	0.99	0.00	0.06	0.01	0.03	0.99
item44	-0.03	-0.47	-0.01	0.10	0.01	0.24
item47	0.02	0.82	-0.05	-0.06	0.16	0.70

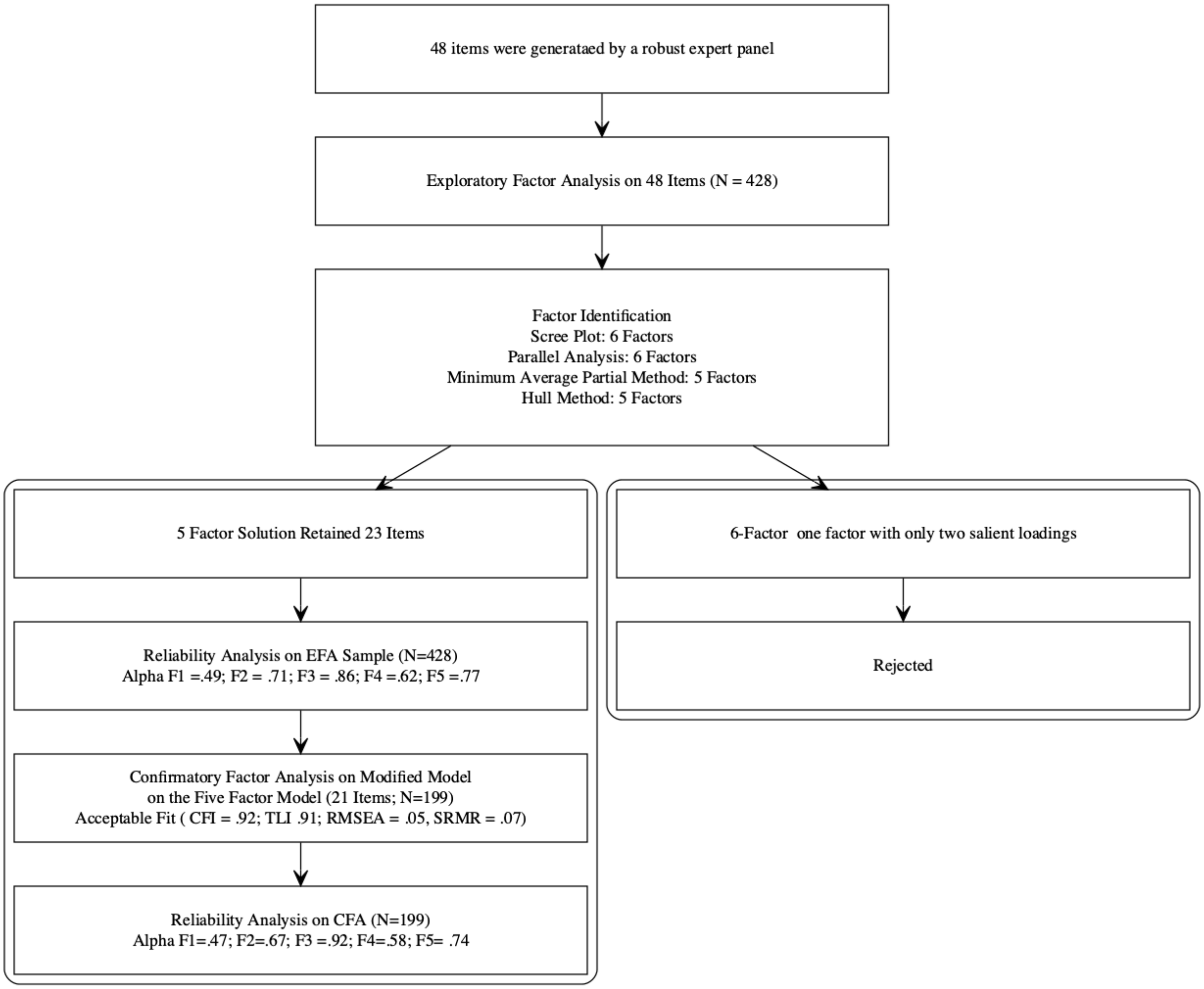
48 items were generataed by a robust expert panel
Exploratory Factor Analysis on 48 Items (N = 428)
Factor Identification
Scree Plot: 6 Factors
Parallel Analysis: 6 Factors
Minimum Average Partial Method: 5 Factors
Hull Method: 5 Factors
5 Factor Solution Retained 23 Items
Reliability Analysis on EFA Sample (N=428)
Alpha F1 =.49; F2 = .71; F3 = .86; F4 =.62; F5 =.77
Confirmatory Factor Analysis on Modified Model
on the Five Factor Model (21 Items; N=199)
Acceptable Fit (CFI = .92; TLI .91; RMSEA = .05, SRMR = .07)
Reliability Analysis on CFA (N=199)
Alpha F1=.47; F2=.67; F3 =.92; F4=.58; F5= .74
6-Factor one factor with only two salient loadings
Rejected

Figure 1. ABC

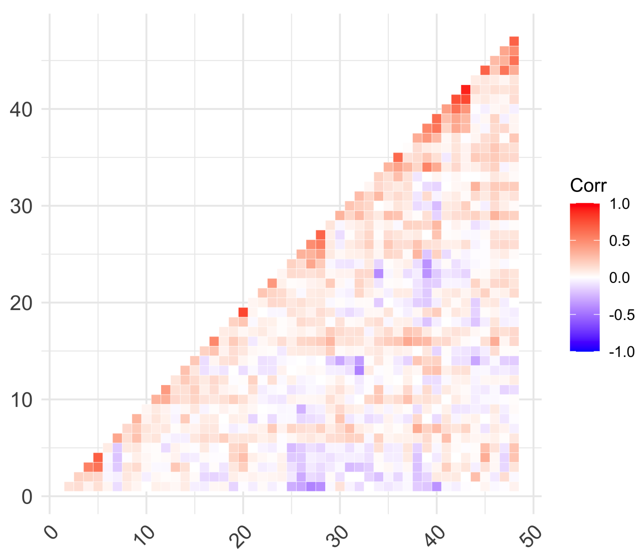

Figure 2. Iter-correlation of the items

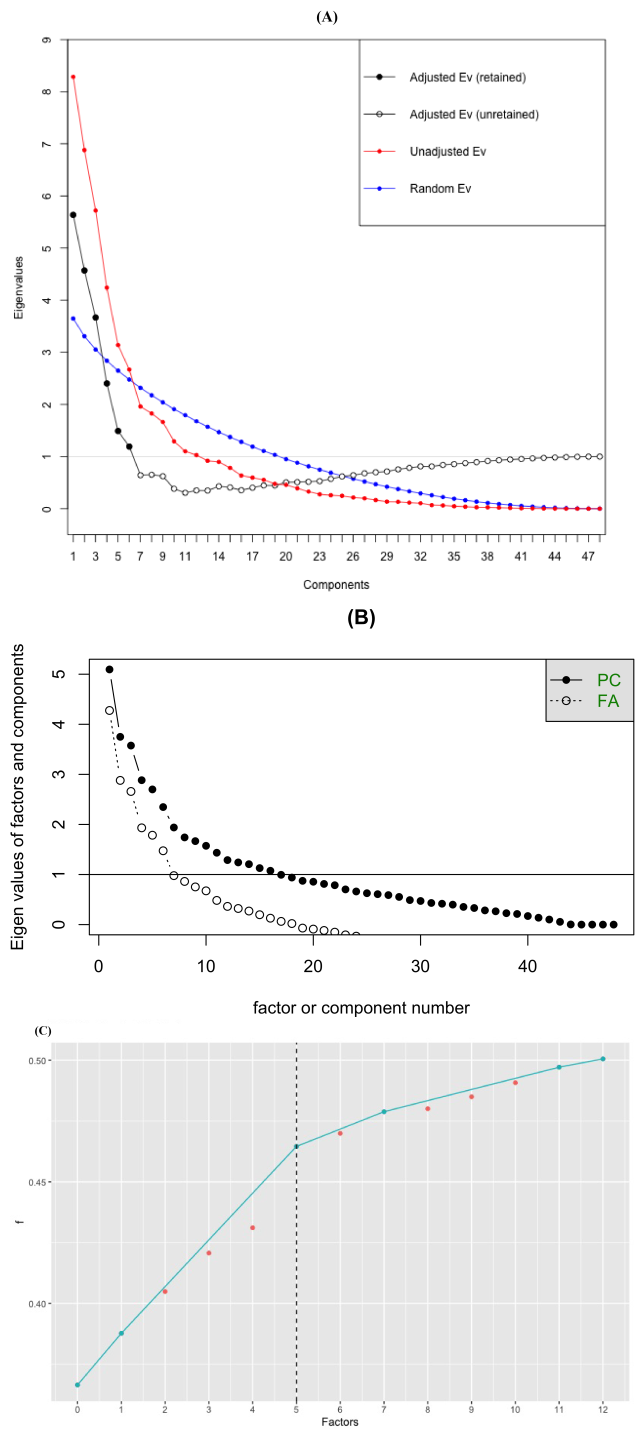

Figure 3. Factor Identification (A) Parallel analysis (B) Scree Plot, (C) Hull method

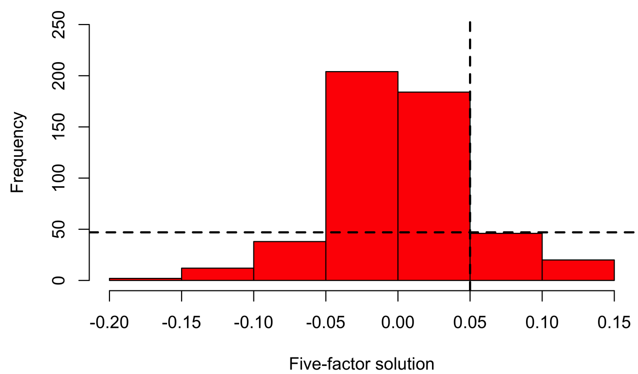

Figure 4. Histogram of residulas: five-factor solution