

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

Mushfiqul Anwar Siraji¹, Rafael Robert Lazar², & Manuel Spitschan³

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

1. Describe EFA and CFA sample separately.
2. Sampling technique: Convince sampling (non-probability sample)
3. Method: cross-sectional survey
4. How many missing data?
5. How incomplete data were addressed.
6. Why such sample was chosen?

EFA: For exploring initial factor structure, a sample of 250-300 is recommended (Comrey & Lee, 1992; Schönbrodt & Perugini, 2013)

CFA: For estimating the sample size for the confirmatory factor analysis we followed the N:q rule (Bentler & Chou, 1987; Jackson, 2003; Kline, 2015; Worthington & Whittaker, 2006) where 10 participants per parameter is required to earn trustworthiness of the result. Our sample size exceeds the requirement.

Procedure

Development of the Scale.

1. How the items were generated

2. How the literature was reviewed to identify construct adequacy of the items.

3. Discuss the expert panel review process to assess content validity

Procedure

Our study had four objectives. First, to develop an instrument to assess an individual's light exposure behavior. Second, to conduct an exploratory factor analysis(EFA) to understand the latent structure. The third one is 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.

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 a polychoric correlation matrix (C. Desjardins & Bulut, 2018) for the EFA. We employed principal axis (pa) a factor extraction method with varimax rotation. PA is robust to the normality assumption violations (Watkins, 2020). The obtained latent structure was confirmed by the minimum residuals extraction method as well. We used a combination factor identification 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 determine 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 and Finney (2018);

Results

Sampling adequacy was checked using Kaiser-Meyer-Olkin (KMO) measures of sampling adequacy (Kaiser, 1974). The overall KMO value for 48 items was 0.63 which was above the cutoff value (.50) indicating a mediocre sample (Hutcheson, 1999).

Table 1 summarizes the univariate descriptive statistics for the 48 items. Some of the items were skewed with high Kurtosis values. Our data violated both univariate normality (Shapiro-Wilk statistics; (Shapiro & Wilk, 1965)) and multivariate normality assumptions (Marida's test; (Mardia, 1970)). Multivariate skew was = 583.80 ($p < 0.001$) and multivariate kurtosis was = 2,749.15 ($p < 0.001$). 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), $\chi^2 (1128) = 5042.86$, $p < .001$ indicated the correlations between items are adequate for the EFA. However only 4.96% of the inter-item correlation coefficients were greater than .30. The inter item correlation ranged between .44 to .91. And the corrected item-total correlations ranged between .10 to .44.

Scree plot (Figure 3) suggested a six-factor solution. Horn's parallel analysis (Horn, 1965) with 500 iterations also indicated a six-factor solution.

However, the 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.

Three rounds of EFA starting with all 48 items were conducted and problematic items were gradually discarded (cross-loading items and poor factor loading ($<.30$) items). Finally, a five-factor EFA solution with 25 items was accepted with low RMSR = 0.08 (Brown, 2015), all factor-loading higher than .30 and no cross-loading greater than .30. We confirmed this five-factor latent structure using varimax rotation with a minimum residual extraction method (see the supplementary). Table2 displays the factor-loading (structural coefficients) and communality of the items. The absolute value of the factor-loading ranged from -.49 to .99 indicating strong coefficients. The communalities ranged between .11 to .99. However, the histogram of the absolute values of non-redundant residual-correlations (Fig5 showed 26.00% correlations greater than the absolute value of .05, indicating a possible under-factoring. (C. D. Desjardins, 2018). Subsequently, we fitted a six-factor solution. However, a factor emerged with only one salient variable loading in the six-factor solution, thus disqualifying the six-factor solution.

In the five-factor solution, the first factor contained three items and explained 10.25% of the total variance with a satisfactory internal reliability coefficient ($\alpha = .86$). All the items in this factor stemmed from the individual's preference to use blue light filters in different light environments. The second factor contained six items and explained 9.93% of the total variance with a satisfactory internal reliability coefficient ($\alpha = .71$). Items under this factor commonly investigate an individual's hours spent outdoor. The third factor contained five items and explained 8.83% of the total variance. Items under this factor dealt with the specific behaviors pertaining to sleep. However, the

internal consistency reliability coefficient was not satisfactory ($\alpha = .68$). The fourth factor contained five items and explained 8.44% of the total variance with an internal consistency coefficient, $\alpha = .62$. These five items stemmed from the behavior related to an individual's cellphone usage during the sleep-wakeup time. Lastly, the fifth factor contained six items and explained 6.14% of the total variance. This factor tried to measure an individual's behavior lead by the awareness of light's influence on health. However, this factor showed unsatisfactory internal consistency reliability ($\alpha = .53$). It is essential to attain a balance between psychometric properties and the interpretability of the common themes when exploring the latent structure. As all of the emerged factors are highly interpretable, regardless of the apparent low reliability of the two factors, we retain the five-factor solution with 23 items for our confirmatory factor analysis (CFA). Two items showed negative factor-loading (items 44 and 21). Upon inspection, it was understood that these items are negatively correlated to the common theme, and thus in the CFA analysis, we reversed the response code for these two items.

Confirmatory Factor Analysis

IRT

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.
- Bentler, P. M., & Chou, C.-P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, 16(1), 78–117. <https://doi.org/10.1177/0049124187016001004>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY, US: The Guilford Press.
- 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.

- 176 https://doi.org/10.1207/s15327906mbr0102_10
- 177 Chang, W., Cheng, J., Allaire, J., Sievert, C., Schloerke, B., Xie, Y., . . .
- 178 Borges, B. (2021). *Shiny: Web application framework for r*. Retrieved
- 179 from <https://CRAN.R-project.org/package=shiny>
- 180 Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis, 2nd*
- 181 *ed.* Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- 182 Desjardins, C., & Bulut, O. (2018). *Handbook of Educational Measurement*
- 183 *and Psychometrics Using R*. <https://doi.org/10.1201/b20498>
- 184 Desjardins, C. D. (2018). *Handbook of educational measurement and*
- 185 *psychometrics using R* (O. Bulut & ProQuest (Firm), Eds.). Boca Raton,
- 186 FL : CRC Press.
- 187 Dinno, A. (2018). *Paran: Horn's test of principal components/factors*.
- 188 Retrieved from <https://CRAN.R-project.org/package=paran>
- 189 Epskamp, S. (2019). *semPlot: Path diagrams and visual analysis of*
- 190 *various SEM packages' output*. Retrieved from
- 191 <https://CRAN.R-project.org/package=semPlot>
- 192 Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., &
- 193 Borsboom, D. (2012). qgraph: Network visualizations of relationships
- 194 in psychometric data. *Journal of Statistical Software*, 48(4), 1–18.
- 195 Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*
- 196 (Third). Thousand Oaks CA: Sage. Retrieved from
- 197 <https://socialsciences.mcmaster.ca/jfox/Books/Companion/>
- 198 Fox, J., Weisberg, S., & Price, B. (2020). *carData: Companion to applied*
- 199 *regression data sets*. Retrieved from
- 200 <https://CRAN.R-project.org/package=carData>

Harrell Jr, F. E., Charles Dupont, with contributions from, & others., many.
(2021). *Hmisc: Harrell miscellaneous*. Retrieved from
<https://CRAN.R-project.org/package=Hmisc>

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>

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter
Estimates: Some Support for the N:q Hypothesis. *Structural Equation
Modeling*, 10(1), 128–141.
https://doi.org/10.1207/S15328007SEM1001_6

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*

- 226 *using 'ggplot2'*. Retrieved from
227 <https://CRAN.R-project.org/package=ggcorrplot>
- 228 Kline, R. B. (2015). *Principles and practice of structural equation*
229 *modeling*. The Guilford Press.
- 230 Lorenzo-Seva, U., Timmerman, M., & Kiers, H. (2011). The Hull Method
231 for Selecting the Number of Common Factors. *Multivariate Behavioral*
232 *Research*, 46, 340–364.
233 <https://doi.org/10.1080/00273171.2011.564527>
- 234 Makowski, D., Ben-Shachar, M. S., Patil, I., & Lüdtke, D. (2020). Methods
235 and algorithms for correlation analysis in r. *Journal of Open Source*
236 *Software*, 5(51), 2306. <https://doi.org/10.21105/joss.02306>
- 237 Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with
238 applications. *Biometrika*, 57(3), 519–530.
239 <https://doi.org/10.1093/biomet/57.3.519>
- 240 Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved
241 from <https://CRAN.R-project.org/package=tibble>
- 242 Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2021). *EFA.MRFA:*
243 *Dimensionality assessment using minimum rank factor analysis*.
244 Retrieved from <https://CRAN.R-project.org/package=EFA.MRFA>
- 245 Ooms, J. (2021). *Rsvg: Render SVG images into PDF, PNG, PostScript, or*
246 *bitmap arrays*. Retrieved from
247 <https://CRAN.R-project.org/package=rsvg>
- 248 R Core Team. (2021). *R: A language and environment for statistical*
249 *computing*. Vienna, Austria: R Foundation for Statistical Computing.
250 Retrieved from <https://www.R-project.org/>

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

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

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

260 Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New
261 York: Springer. Retrieved from <http://lmdvr.r-forge.r-project.org>

262 Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do
263 correlations stabilize? *Journal of Research in Personality*, 47(5),
264 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>

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

268 Terry M. Therneau, & Patricia M. Grambsch. (2000). *Modeling survival*
269 *data: Extending the Cox model*. New York: Springer.

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

273 Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s*
274 (Fourth). New York: Springer. Retrieved from
275 <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>

Worthington, R. L., & Whittaker, T. A. (2006). Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist*, 34(6), 806–838.

301 <https://doi.org/10.1177/0011000006288127>

302 Zeileis, A., & Croissant, Y. (2010). Extended model formulas in R: Multiple

303 parts and multiple responses. *Journal of Statistical Software*, 34(1),

304 1–13. <https://doi.org/10.18637/jss.v034.i01>

305 Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and*

306 *pipe syntax*. Retrieved from

307 <https://CRAN.R-project.org/package=kableExtra>

Table 1

Descriptive Statistics

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item1	2.27	1.39	0.74	-0.81	0.81*	.25
Item2	2.87	1.59	0.08	-1.60	0.81*	.19
Item3	3.36	1.38	-0.48	-1.03	0.87*	.16
Item4	1.47	1.18	2.38	4.00	0.43*	.28
Item5	4.01	1.40	-1.22	0.07	0.70*	.13
Item6	2.79	1.55	0.19	-1.48	0.85*	.20
Item7	2.26	1.25	0.70	-0.60	0.85*	.19
Item8	2.97	1.20	-0.06	-0.94	0.91*	-.10
Item9	2.94	1.03	-0.12	-0.40	0.91*	.10
Item10	2.74	1.04	0.09	-0.74	0.91*	.28
Item11	2.18	0.90	0.60	0.12	0.86*	.26
Item12	2.36	1.22	0.59	-0.62	0.87*	.25
Item13	2.73	1.46	0.20	-1.36	0.87*	.33
Item14	2.14	1.31	0.77	-0.78	0.80*	.26
Item15	3.26	1.09	-0.26	-0.45	0.80*	.14
Item16	1.56	1.23	2.00	2.45	0.50*	.32
Item17	1.54	1.21	2.07	2.75	0.49*	.31
Item18	1.12	0.49	5.02	27.80	0.25*	.16
Item19	1.05	0.36	7.23	52.98	0.13*	.18
Item20	1.04	0.33	8.99	85.28	0.10*	.16
Item21	1.14	0.59	4.79	24.05	0.25*	.16
Item22	3.57	1.07	-0.65	-0.17	0.88*	.21
Item23	2.56	1.27	0.33	-1.00	0.89*	.11

Table 1 continued

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item24	4.14	0.99	-1.23	1.14	0.79*	.19
Item25	2.59	1.41	0.27	-1.27	0.86*	.19
Item26	2.25	1.27	0.69	-0.64	0.84*	.18
Item27	3.80	1.29	-0.87	-0.42	0.82*	.17
Item28	3.76	1.14	-0.68	-0.45	0.86*	.00
Item29	2.44	1.31	0.38	-1.14	0.86*	.11
Item30	1.48	1.11	2.18	3.35	0.48*	.24
Item31	3.00	1.62	-0.08	-1.61	0.83*	.44
Item32	3.55	1.65	-0.60	-1.34	0.76*	.43
Item33	3.62	1.64	-0.68	-1.25	0.74*	.32
Item34	3.42	1.83	-0.45	-1.69	0.69*	.33
Item35	3.86	1.67	-0.99	-0.85	0.65*	.23
Item36	1.54	1.25	2.13	2.86	0.46*	.36
Item37	1.33	0.91	3.03	8.43	0.41*	.01
Item38	4.30	1.08	-1.79	2.53	0.67*	.22
Item39	1.96	0.98	1.02	0.69	0.82*	.05
Item40	2.16	1.19	0.71	-0.54	0.84*	.14
Item41	1.31	0.81	2.75	6.92	0.43*	.21
Item42	3.93	1.48	-1.06	-0.44	0.71*	.18
Item43	1.64	1.18	1.79	2.02	0.60*	.15
Item44	3.51	1.30	-0.70	-0.59	0.85*	.39
Item45	2.22	1.48	0.71	-1.02	0.76*	.30
Item46	1.76	1.23	1.35	0.44	0.66*	.38
Item47	2.11	1.17	0.77	-0.39	0.83*	.32

Table 1 continued

	Mean	SD	Skew	Kurtosis	Shapiro-Wilk Statistics	Item-Total Correlation
Item48	2.60	1.25	0.29	-0.86	0.89*	.35

Note. *p<.001

Table 2

	F1	F2	F3	F4	F5	Communality
item16	0.99	-	-	-	-	0.99
item36	0.94	-	-	-	-	0.9
item17	0.8	-	-	-	-	0.66
item11	-	0.8	-	-	-	0.64
item10	-	0.76	-	-	-	0.59
item12	-	0.65	-	-	-	0.47
item7	-	0.5	-	-	-	0.27
item8	-	-0.5	-	-	-	0.25
item9	-	0.32	-	-	-	0.11
item27	-	-	0.8	-	-	0.66
item3	-	-	0.8	-	-	0.68
item40	-	-	0.65	-	-	0.46
item30	-	-	0.45	-	-	0.35
item41	-	-	-0.36	-	-	0.33
item33	-	-	-	0.74	-	0.56
item32	-	-	-	0.73	-	0.62
item35	-	-	-	0.66	-	0.45
item37	-	-	-	-0.39	-	0.17
item38	-	-	-	0.36	-	0.18
item46	-	-	-	-	0.6	0.42
item45	-	-	-	-	0.59	0.37
item25	-	-	-	-	0.41	0.19
item4	-	-	-	-	0.41	0.22
item1	-	-	-	-	0.4	0.17
item26	-	-	-	-	0.39	0.17
internal Consistency	.86	.71	.68	.62	.53	-
% of variance	10.25	9.93	8.83	8.44	6.14	-

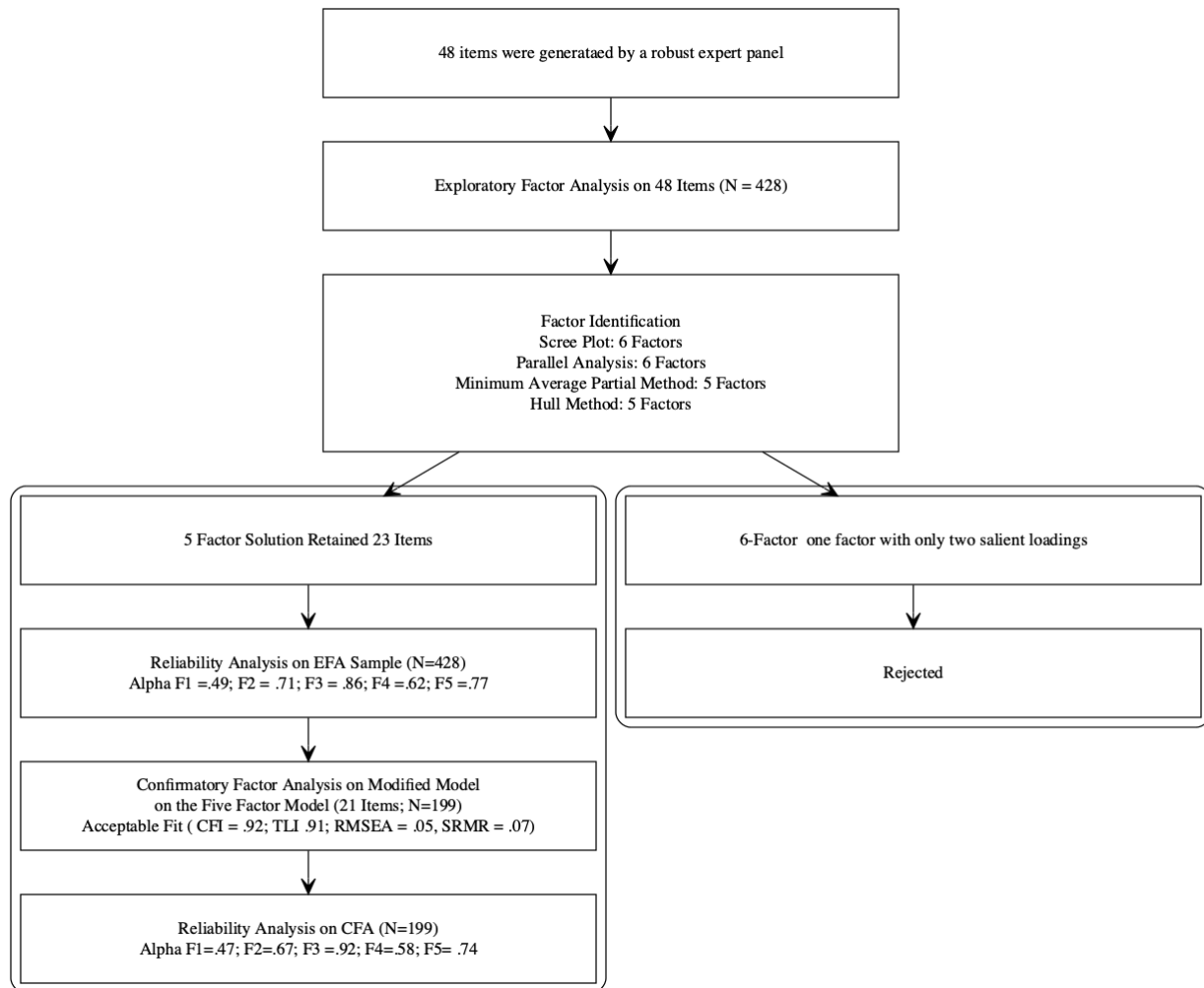


Figure 1. Development and psychometric properties of LEBA

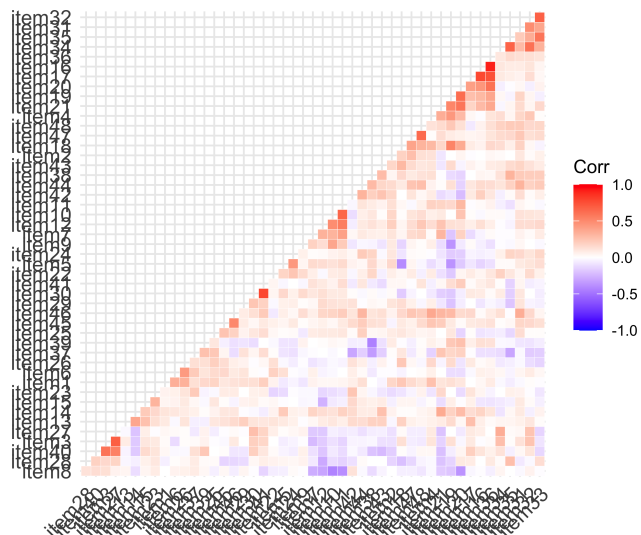


Figure 2. Correlation plot of the items

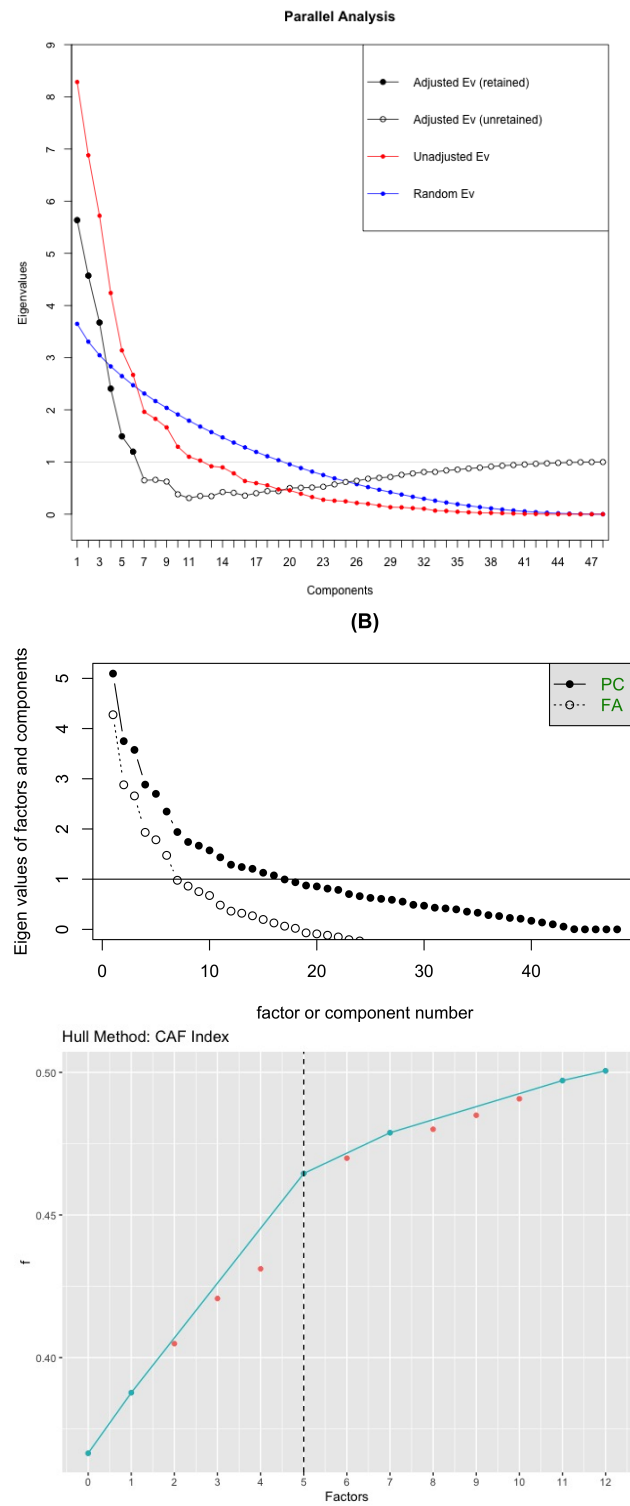


Figure 3. Factor Identification (A) Parallel analysis (B) Scree Plot, (C) Hull method

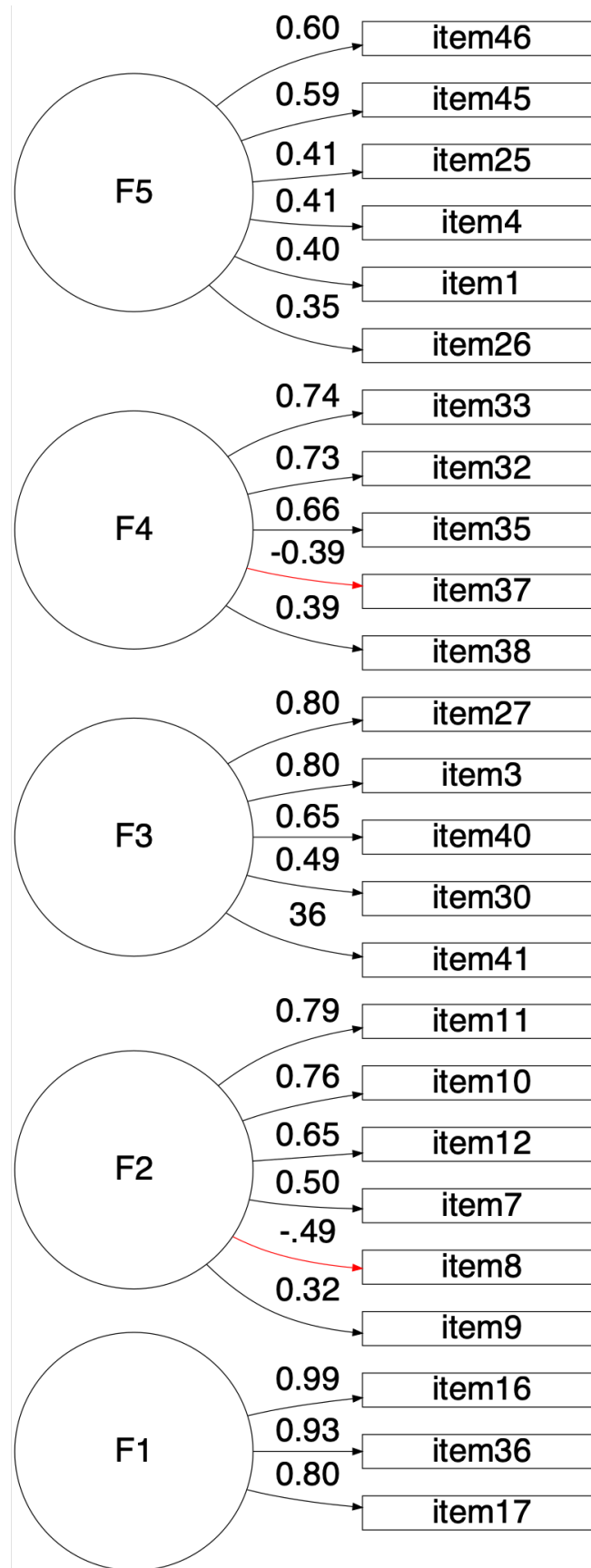


Figure 4. Five Factor Solution

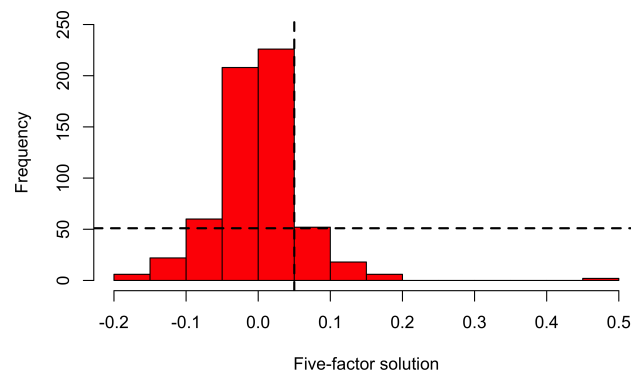


Figure 5. Histogram of residulas: five-factor solution