# D3mirt: Descriptive Three-Dimensional Multidimensional Item Response Theory for R

Erik Forsberg ⓘ [1*]

1 Division of Personality, Social and Developmental Psychology, Stockholm University, Sweden * These authors contributed equally.

Summary

The `D3mirt` package for R (R Core Team, 2021) offers functions for analyzing questionnaire items used in psychological research in a three-dimensional latent space. The application is based on descriptive multidimensional item response theory (DMIRT) (Reckase, 1985, 2009; Reckase & McKinley, 1991), a statistical framework incorporating vector geometry to describe item characteristics. The method is foremost visual, and the latent model can be plotted as an interactive graphical device with the help of a dedicated plot function based on the RGL 3D visualization device system for R (Adler & Murdoch, 2023). Alongside the plot function, the package also includes a model identification function that helps the user identify the DMIRT model and a model estimation function for extracting the necessary vector estimates. New additions to the DMIRT framework introduced in the `D3mirt` package include studying constructs (explained below) and individual scores plotted in the three-dimensional latent model.

Statement of need

Common to most item response theory (IRT) models is the assumption of *unidimensionality*, i.e., that a test or item measures simple structures (Hambleton & Rogers, 1991). There are, however, many occasions where this may be improper. Consider a mathematical word problem (Reckase, 1985, 2009; Reckase & McKinley, 1991). To solve a mathematical word problem, one must often have verbal and mathematical skills, referred to as abilities (denoted θ) in the literature on IRT. In other words, one's resulting answer would be a function based on one's ability to read, on the one hand, and one's ability to perform numerical manipulations, on the other. Accordingly, instead of a person's location on a unidimensional latent variable, the mathematical word problem illustrates a situation where it seems more reasonable to assume that a correct response is due to the respondent's location in a multidimensional latent variable space.

Descriptive multidimensional item response theory (DMIRT) (Reckase, 1985, 2009; Reckase & McKinley, 1991) was developed to handle the just mentioned situation. The method is based on using a *compensatory model*, i.e., a type of measurement model in multidimensional IRT that uses linear combinations of θ-values for ability assessment. This model assumes that the same probability score for a correct response can be reached by using different combinations of θ-values, as opposed to assuming that the relation is one-to-one. In turn, this implies that the compensatory model allows items to be unidimensional, i.e., that they measure a single ability, or *within-multidimensional*, i.e., that the items can measure more than one ability in the model space.

Note that the `D3mirt` approach is limited to two types of compensatory models, depending on

item type. If dichotomous items are used, the analysis is based on the multidimensional extension of the two-parameter logistic model (McKinley & Reckase, 1983) as the compensatory model. If polytomous items are used, the analysis is based on the two-parameter multidimensional graded response model (Muraki & Carlson, 1995) as the compensatory model.

Compared to other software, D3mirt is unique in implementing DMIRT methodology explicitly and comprehensively in a three-dimensional interactive environment. For instance, a compensatory model can be fitted with software dedicated to multidimensional IRT that supports dichotomous or polytomous data and allows the user to specify the necessary factor structure. This includes software such as Mplus (Muthén & Muthén, 1998-2017) or IRTPRO 2.1 (Thissen & du Toit, 2011) and flexMIRT (Houts & Cai, 2020) for the Windows environment. Using these software programs, the DMIRT estimates can then be derived manually with the help of general mathematical software, such as MATLAB (MATLAB, 2022) or Mathematica (Mathematica, 2023), and plotted with built-in options for creating vector plots in two or three dimensions. This is, however, often time-consuming, and the plotting methods are not optimized for DMIRT analysis and test development. Regarding R, the mirt package (Chalmers, 2012) can be used to fit the compensatory multidimensional model and derive the basic DMIRT item and person parameters while the vector plot options are limited. There is also the R package plink (Weeks, 2010) that offers two-dimensional vector plots suitable for DMIRT analysis but only for dichotomous items. Another more general limitation is that none of the formerly mentioned software applications provides a function to help identify the DMIRT model.

The D3mirt package was designed to counter many of the just mentioned shortages by implementing specialized functions for identifying the DMIRT model, calculation of the necessary DMIRT estimates, and plotting the results in an interactive three-dimensional latent environment (see Figure 1). An example of the utility of using the package in an empirical context for item and scale analysis has been presented in Forsberg & Sjöberg (2024).
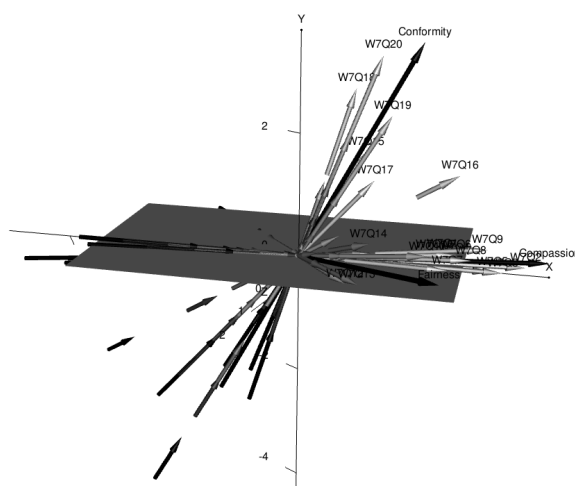

Figure 1: A still shot of the graphical output from D3mirt. The Figure illustrates a three-dimensional vector plot for items in the anes0809offwaves data set included in the package. The output also shows three construct vector arrows: Compassion, Fairness, and Conformity (solid black arrows).

Multidimensional item parameters

The theoretical framework for DMIRT rests foremost on three assumptions (Reckase, 1985). Firstly, ability maps the probability monotonically, such that a higher level of ability implies a higher probability of answering an item correctly. Second, we wish to locate an item at a singular point at which it is possible to derive item characteristics for the multidimensional case conceptually similar to the unidimensional case. Thirdly, an item's maximum level of discrimination, i.e., its highest possible capacity to separate respondents on level of ability, is the best option for the singular point estimation. The most important parameter equations regarding the assumptions just mentioned will be briefly presented below.

Firstly, by using the discrimination score a_i on item i from the compensatory model, we can define the multidimensional discrimination index (MDISC) as follows.

$$MDISC := \sqrt{\sum_{k=1}^{m} a_{ik}^2}, \tag{1}$$

on m dimensions with the slope constant $\frac{1}{4}$ omitted (Reckase, 2009; Reckase & McKinley, 1991). The MDISC is sometimes denoted A_i to highlight the connection to the unidimensional a_i parameter (Reckase, 2009; Reckase & McKinley, 1991). Importantly, the MDISC sets the orientation of the item vectors in the multidimensional space (Reckase, 2009; Reckase & McKinley, 1991).

$$\omega_{il} = cos^{-1} \left(\frac{a_{il}}{\sqrt{\sum_{k=1}^{m} a_{ik}^2}} \right), \tag{2}$$

on latent axis l in the model. Note, the ω_{il} is in this solution a characteristic of the item i that tells in what direction i has its highest level of discrimination, assuming a multidimensional latent space (Reckase, 2009; Reckase & McKinley, 1991). This gives us the following criteria to use as a rule of thumb. Assume a two-dimensional space, an orientation of $0°$ with respect to any of the model axes indicates that the item is unidimensional. Such an item describes a singular trait only. In contrast, an orientation of $45°$ indicated that the item is within-multidimensional. Such an item describes both traits in the two-dimensional model equally well. The same criteria are extended to the three-dimensional case. The MDISC is also used in the graphical output to scale the length of the vector arrows representing the item response functions, e.g., so that longer vector arrows indicate higher discrimination, shorter arrows lower discrimination in the model, and so on.

Next, to assess multidimensional difficulty, the distance from the origin is calculated using the multidimensional difficulty (MDIFF), denoted B_i to highlight the connection to the unidimensional b_i parameter (Reckase, 1985).

$$MDIFF := \frac{-d_i}{\sqrt{\sum_{k=1}^{m} a_{ik}^2}}, \tag{3}$$

in which d_i is the d_i-parameter from the compensatory model. The MDIFF is, therefore, a difficulty characteristic of item i, such that higher MDIFF values indicate that higher levels of ability are necessary for a correct response (Reckase, 2009; Reckase & McKinley, 1991). Observe that the denominator in Equation 3 is the same expression as Equation 1.

Importantly, in DMIRT analysis, the MDISC and MDIFF only apply in the direction set by ω_{il} and Equation 2 (Reckase, 2009; Reckase & McKinley, 1991). Thus, we cannot compare these estimates directly across items, as would be the case in the unidimensional model.

This is because DMIRT seeks to maximize item discrimination as a global characteristic in a multidimensional environment. To estimate item discrimination as a local characteristic in the multidimensional space, it is, however, possible to select a common direction for the items and then recalculate the discrimination, i.e., to estimate the directional discrimination (DDISC), as follows.

$$DDISC := \sum_{k=1}^{m} a_{ik} \cos \omega_{ik}. \qquad (4)$$

Since the DDISC is a local characteristic in the model, it is always the case that $DDISC \leq MDISC$. In D3mirt, the DDISC is optional and implemented in D3mirt as optional *construct vectors* indicated by a subset of items or using spherical coordinates.

The results include tables for the MDISC and MDIFF estimates as well as spherical coordinates describing the location of the vector arrows. If construct vectors are used, the output also includes DDISC scores for all items showing the constrained discrimination. It is also possible to plot individual scores (i.e., *profile analysis*) in the three-dimensional latent space (see Figure 2). This can be useful for studying respondents' location conditioned on some external variable, e.g., sex, age, political preference, and so on. Instructions on the method, such as model identification, model estimation, plotting, and profile analysis, are given in the package vignette.
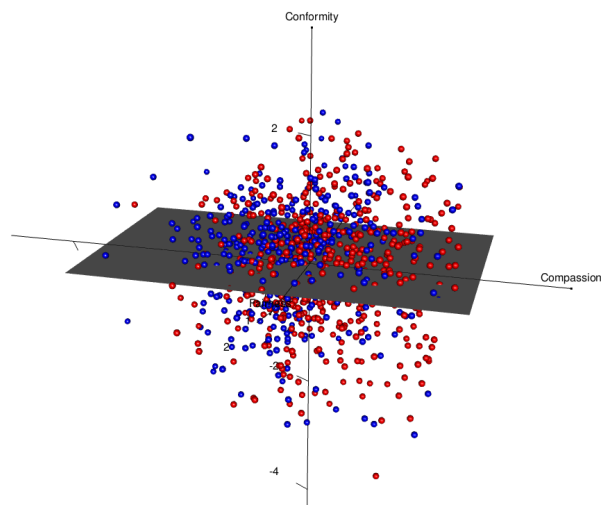

Figure 2: A still shot of the graphical output from D3mirt illustrating respondent scores in the latent space separated on sex (male in blue and female in red) from the anes0809offwaves data set included in the package.

To report issues, seek support, or for developers wishing to contribute to the software, contact the author via the dedicated GitHub page (https://github.com/ForsbergPyschometrics/D3mirt) or email (forsbergpsychometrics@gmail.com).

Forsberg. (2024). D3mirt: Descriptive Three-Dimensional Multidimensional Item Response Theory for R. *Journal of Open Source Software*, *9*(99), 46523. https://doi.org/10.21105/joss.06523

Acknowledgements

References

Adler, D., & Murdoch, D. (2023). *Rgl: 3D visualization using OpenGL*. https://doi.org/10.32614/CRAN.package.rgl

Chalmers, R. P. (2012). mirt: A Multidimensional Item Response Theory Package for the R Environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Forsberg, E., & Sjöberg, A. (2024). A Validation Study of the Extended Relevance Scale using the D3mirt Package for R. *Measurement: Interdisciplinary Research and Perspectives*, 1–23. https://doi.org/10.1080/15366367.2023.2282418

Hambleton, S., R. K., & Rogers, H. J. (1991). *Fundamentals of item response theory*. SAGE Publications.

Houts, C. R., & Cai, L. (2020). *flexMIRT user's manual version 3.52: Flexible multilevel multidimensional item analysis and test scoring*. Vector Psychometric Group.

Mathematica. (2023). *Mathematica, Version 13.3*. Wolfram Research Inc. https://www.wolfram.com/mathematica

MATLAB. (2022). *MATLAB Version: 9.13.0 (R2022b)*. The MathWorks Inc. https://www.mathworks.com

McKinley, R. L., & Reckase, M. D. (1983). *An extension of the two-parameter logistic model to the multidimensional latent space* (No. ONR83-2). American College Testing Program.

Muraki, E., & Carlson, J. E. (1995). Full-information factor analysis for polytomous item responses. *Applied Psychological Measurement*, *1*(19), 73–90. https://doi.org/10.1177/014662169501900109

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide. Eight Edition*. Muthén & Muthén. https://www.statmodel.com

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reckase, M. D. (1985). The Difficulty of Test Items That Measure More Than One Ability. *Applied Psychological Measurement*, *4*(9), 401–412. https://doi.org/10.1177/014662168500900409

Reckase, M. D. (2009). *Multidimensional Item Response Theory*. Springer-Verlag. https://doi.org/10.1007/978-0-387-89976-3

Reckase, M. D., & McKinley, R. L. (1991). The Discriminating Power of Items That Measure More Than One Dimension. *Applied Psychological Measurement*, *4*(15), 361–373. https://doi.org/10.1177/014662169101500407

Thissen, C. L., & du Toit, S. H. C. (2011). *IRTPRO for Windows*. Scientific Software International.

Weeks, J. P. (2010). plink: An R Package for Linking Mixed-Format Tests Using IRT-Based Methods. *Journal of Statistical Software*, *35*(12), 1–33. https://doi.org/10.18637/jss.

v035.i12

Forsberg. (2024). D3mirt: Descriptive Three-Dimensional Multidimensional Item Response Theory for R. *Journal of Open Source Software*, *9*(99), 6523. https://doi.org/10.21105/joss.06523