

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

Erik Forsberg ¹

¹ Division of Personality, Social and Developmental Psychology, Stockholm University, Sweden

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: Øystein Sørensen 

Submitted: 13 March 2024

Published: unpublished

License

Authors of papers retain copyright
and release the work under a
Creative Commons Attribution 4.0
International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

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 that incorporates 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 RGL 3D visualization device system for R ([Adler & Murdoch, 2023](#)). Alongside the plot function, the package also includes a dedicated 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 the option of 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 et al., 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 both verbal and mathematical skills, or abilities (θ) as it is often called in the literature on item response theory. In other words, one's result would be a function of 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](#)) has been proposed as a data reduction technique for the just mentioned situation. The method is based on using a *compensatory model*, i.e., a type of measurement model that uses linear combinations of θ -values for ability assessment. This type of model assumes that the same sum can be reached by adding different combinations of θ -values. In turn, this implies that items can be unidimensional or *within-multidimensional*, i.e., that the item measures more than one ability.

The D3mirt approach is limited to two types of item models, dependent 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. The results of the

analysis are presented in tables and interactive three-dimensional devices with vector arrows representing item response functions (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 (2024).

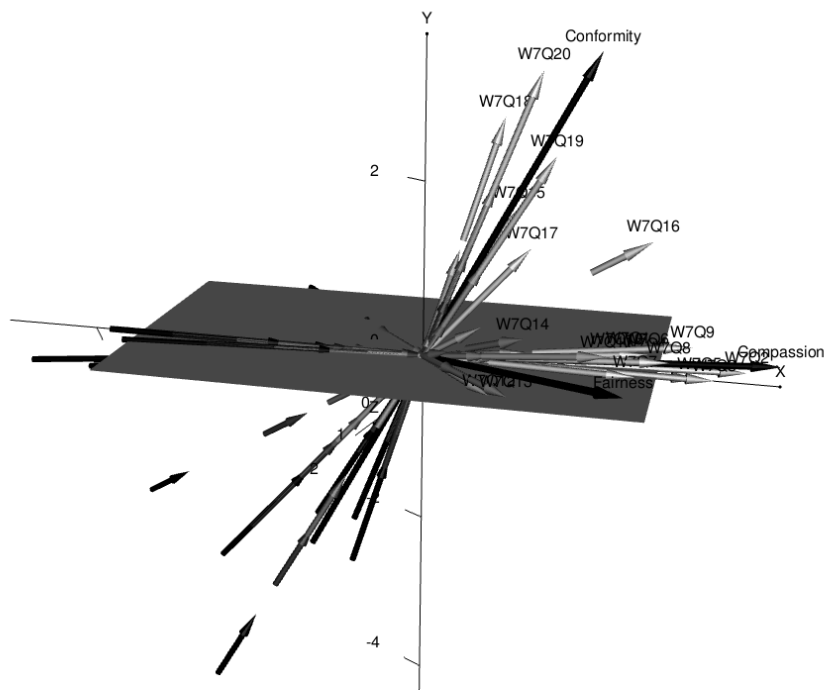


Figure 1: A still shot of the graphical output from D3mirt. The Figure illustrates a three-dimensional vector plot for items in the anes0809offwaves data set included in the package. The output also shows three construct vector arrows: Compassion, Fairness, and Conformity (solid black arrows).

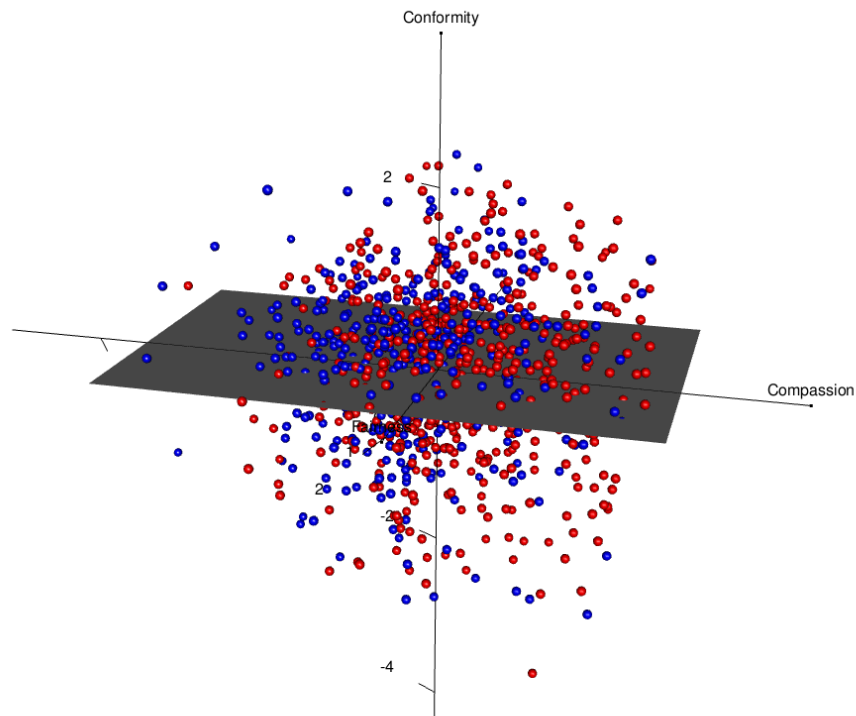


Figure 2: A still shot of the graphical output from D3mirt illustrating respondent scores in the latent space separated on sex (male in blue and female in red) from the anes0809offwaves data set included in the package.

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. Thirdly, an item's maximum level of discrimination, i.e., its highest possible sensitivity score for measuring ability, is the best option for the singular point estimation. The most important parameter equations capturing the just mentioned assumptions in DMIRT will be presented below.

Firstly, by using the discrimination score a_i from the compensatory model we can derive the multidimensional analog to the unidimensional discrimination parameter, i.e., the multidimensional discrimination index (MDISC), denoted A_i for item i , to highlight the connection to the unidimensional a_i parameter (Reckase, 2009; Reckase & McKinley, 1991),

$$MDISC = A_i = \sqrt{\sum_{k=1}^m a_{ik}^2}, \quad (1)$$

with the slope constant $\frac{1}{4}$ omitted (Reckase, 2009; Reckase & McKinley, 1991). Importantly, the MDISC sets the orientation of the item vectors in the multidimensional space (Reckase, 2009; Reckase & McKinley, 1991), as follows,

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

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) index (Reckase, 1985):

$$MDIFF = B_i = \frac{-d_i}{\sqrt{\sum_{k=1}^m a_{ik}^2}} \quad (3)$$

in which d is the d -parameter index from the compensatory model. The MDIFF is denoted B as the DMIRT counterpart to the b -parameter in the unidimensional IRT model. The MDIFF is, therefore, a 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),

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

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*.

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

93 to plot individual scores (i.e., *profile analysis*) in the three-dimensional latent space (see
94 [Figure 2](#)). This can be useful for studying respondents location conditioned on some external
95 variable, e.g., sex, age, political preference, and so on. Instructions on the method, such as
96 model identification, model estimation, plotting, and profile analysis, are given in the package
97 vignette.

98 Acknowledgements

99 I acknowledge support, advice, and suggestions for improvements from my supervisor Dr. Anders
100 Sjöberg, Stockholm University. I also would like to express gratitude to Professor Torun
101 Lindholm Öjmyr and Professor Mats Nilsson, Stockholm University, for their support and
102 professional advice.

103 References

- 104 Adler, D., & Murdoch, D. (2023). *Rgl: 3D visualization using OpenGL*. <https://dmurdoch.github.io/rgl/index.html>
- 105
- 106 Forsberg, A., E. & Sjöberg. (2024). A Validation Study of the Extended Relevance Scale
107 using the D3mirt Package for R. *Measurement: Interdisciplinary Research and Perspectives*,
108 1–23. <https://doi.org/10.1080/15366367.2023.2282418>
- 109 Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response*
110 *theory*. SAGE Publications.
- 111 McKinley, R. L., & Reckase, M. D. (1983). *An extension of the two-parameter logistic model*
112 *to the multidimensional latent space* (No. ONR83-2). American College Testing Program.
- 113 Muraki, E., & Carlson, J. E. (1995). Full-information factor analysis for polytomous item
114 responses. *Applied Psychological Measurement*, 1(19), 73–90. [https://doi.org/10.1177/](https://doi.org/10.1177/014662169501900109)
115 [014662169501900109](https://doi.org/10.1177/014662169501900109)
- 116 R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R
117 Foundation for Statistical Computing. <https://www.R-project.org/>
- 118 Reckase, M. D. (1985). The Difficulty of Test Items That Measure More Than One Ability.
119 *Applied Psychological Measurement*, 4(9), 401–412.
- 120 Reckase, M. D. (2009). *Multidimensional Item Response Theory*. Springer-Verlag. <https://doi.org/10.1007/978-0-387-89976-3>
121
- 122 Reckase, M. D., & McKinley, R. L. (1991). The Discriminating Power of Items That Measure
123 More Than One Dimension. *Applied Psychological Measurement*, 4(15), 361–373.