

Software for IRT Analyses: Descriptions and Features¹

Yue Zhao and Ronald Hambleton

University of Massachusetts Amherst

Please forward information to either of the authors (yzhao@ets.org or rkh@educ.umass.edu) that can extend our summary, update the information, and/or correct any errors. Thanks in advance for helping us to keep the summary up-to-date.

Name, Authors, and URL	Capabilities (Model, Estimation Procedures)	Strengths		Weaknesses
		+	-	
BIGSTEPS John Linacre & Benjamin Wright http://www.winsteps.com	A DOS-based software package that runs most of the Rasch models.	+ Free software and manual	- Maximum of 3,000 items and 20,000 examinees	
BILOG Robert Mislevy & Darrell Bock http://www.ssicentral.com	Estimates IRT parameters for the one-, two-, and three- parameter item response models using marginal maximum likelihood estimation	+ Provides optional Bayes' estimates and priors to prevent extreme estimates + Handles omits and not presented data + Provides residual analysis	- Does not apply to polytomous data - Does not allow additional analysis, including DIF, Drift analyses, two-stage test and equating options	
BILOG-MG Michele Zimowski, Eiji Muraki, Robert Mislevy &	Estimates IRT parameters for multiple groups, allowing detection of differential item functioning, item	+ Has all the functions of BILOG, and handles multiple groups, multiple subtests, and multiple test	- Cannot handle polytomous response data	

¹ **Center for Educational Assessment Research Report No. 652.** Amherst, MA: University of Massachusetts, Center for Educational Assessment. (Updated on 1.6.09)

Darrell Bock http://www.ssicentral.com	parameter drift, equivalent and non-equivalent groups equating, vertical equating, two-stage testing, estimation of group means, standard deviations, and latent distributions.	<ul style="list-style-type: none"> + forms in one analysis + Easy to use with help function, features a dialog-box user interface to assist first-time users or occasional users with writing a command file + Handles omits and not-presented data + Provides graphics and residual analysis 	
ConQuest-2.0 Margaret Wu, Raymond Adams, & Mark Wilson http://www.assess.com	Fits Rasch logistic model, rating scale model, partial credit model, ordered partition model, linear logistic test model, multifaceted models, generalized unidimensional models, multidimensional item response models, and latent regression models	<ul style="list-style-type: none"> + Available with both a graphical user interface and a simple command line, or console, interface + Detects differential item functioning, explores rater effects, estimates latent correlations and test dimensionality and draw plausible values + According to the publicity, the 2007 release contains considerably more graphics and fit statistics, and a generally more user-friendly interface. + 200 + page manual is available. 	<ul style="list-style-type: none"> - Limits on maximum number of items and number of examinees
ConTEST E. Timminga, W. J. van der Linden & D.A. Schweizer	Uses pre-existing item banks and linear programming (LP) algorithms to create tests to meet specifications.	<ul style="list-style-type: none"> + The specifications can be based on IRT parameters, item and test information functions, test 	<ul style="list-style-type: none"> - Lacks a Windows interface

<http://www.assess.com>

		characteristic curves, item <i>p</i> -values and discrimination indices from classical test theory, test reliability, validity coefficients, and item format attributes	
DIMENSION John Hattie& Krzysztof Krakowski	Generates item response data according to compensatory and noncompensatory multidimensional item response models	+ Input includes the range of difficulty, the correlation between dimensions, the amount of guessing and various write and print controls + Output includes the product-moment correlation matrix among the dimensions and number-correct and/or examinee parameter data	- Handles number of dimensions up to 5, number of items up to 60 and sample size up to 1,000
DOMAIN Gary Phillips & Sandra Gedeik	Estimates domain scores in criterion-referenced tests using the Rasch model	+ Provides the test characteristic curve (TCC) + Estimates ability, standard scores and domain scores	- Limits up to 500 Rasch calibrated items with up to 10 item sets as subsets in the input file
DRAWICC Christine DeMars http://www.jmu.edu/assessment/icc/icc.htm	Consists of several SAS macros which read item parameter files created by PARSCALE or BILOG (for DOS) and graph the item characteristic curves, and optionally, the information function, with SAS GPLOT	+ The plots can be created with one, two, four, or six items per page + Free software	
Facets John Linacre	Handles applications of unidimensional Rasch measurement	+ Multiple different measurement models included in the same	

http://www.winsteps.com	and constructs measures from complex data involving heterogeneous combinations of examinees, items, tasks, judges along with further measurement and structural facets	+ analysis + Measures differential item functioning + Maximum of 1 million examinees + Measures can also be fixed individually or by group mean, facilitating equating and linking across test sessions	
GGUM2000 James Roberts http://www.psychology.gatech.edu/Unfolding/FreeSoftware.html	DOS-based system of FORTRAN programs that estimate parameters in the Generalized Graded Unfolding Model (Roberts, Donoghue & Laughlin, 2000)	+ Estimates parameters in seven other constrained versions of the model + Free software	- Lack a Windows interface - Allows for 100 items, with up to 10 response categories per item, and up to 2,000 respondents
GGUMLINK James Roberts http://www.psychology.gatech.edu/Unfolding/FreeSoftware.html	Windows-based FORTRAN program that equates parameter estimates from the Generalized Graded Unfolding Model using a variety of methods	+ Secures a common metric by using one of five methods that have recently been generalized to the GGUM + Free software	
GR-GRAFH David Gudanowski, Dawn Vreven, Lynda King & Daniel King	Set of Quattro Pro functions and macros located in spreadsheet templates that generate values and produce graphs and tables to accompany Samejima's graded response model analysis	+ Rely on a popular and widely available spreadsheet program- Quattro Pro + Provides information in both graphical and tabular form	- The items must use a 5-point Likert-type rating format - The templates can be used for up to 100 items
GUMJML James Roberts	DOS-based system of FORTRAN programs that estimates parameters in	+ Offers several diagnostic indexes of item and examinee fit	- Lack a Windows interface

http://www.psychology.gatech.edu/Unfolding/FreeSoftware.html	the Graded Unfolding Model using a joint maximum likelihood technique	+ Free software	- Allows for 200 items, with 2 to 9 response categories, and up to 2,000 respondents
IPLINK Kevin Lee & T.C. Oshima http://www.gsu.edu/~epstco/	Windows-based program that estimates linking coefficients that place item parameter estimates from separate calibrations onto a common trait metric for m-dimensional test data and the linking methods include the direct method, the equated function method, the test characteristic function method and the item characteristic function method (Oshima, Davey, & Lee, 1996).	+ Free software	- Linking methods are limited
IRT-Lab Randall Penfield	A program that provides graphical examination of item response functions, simulates item response and offers instructional windows used to facilitate the understanding of IRT concepts and it can be used with the one-, two-, three- parameter logistic, generalized partial credit and graded response models	+ Helps to improve students' understanding of IRT concepts using a semi-interactive, "point-and -click" Windows-based application	- Simulates up to 100 items
IRTDAT	An interactive or batch PASCALE	+ Permits flexible modeling	- Handles up to 3,000

George A. Johannson	program for generating one-, two-, and three-parameter logistic item response data	+ situations + Simulates common equating designs	simulated examinees and 200 test items
IRTDIF Seock-Ho Kim & Allan Cohen	Uses one-, two-, and three-parameter logistic item response models to provide measures of differential item functioning (DIF) including Lord's (1980) χ^2 statistics, the exact area measures and the closed-interval area measures	+ Significance levels are provided for Lord's (1980) χ^2 statistics, the exact area measures + Free software and manual	- Requires a numerical coprocessor for execution of the program
IRTFIT-RESAMPLE Clement Stone	An SAS program used to compute the goodness-of-fit statistics based on posterior probabilities for dichotomous and polytomous response models, such as the graded logistic response model, one-, two-, and three-parameter dichotomous logistic response models and generalized partial-credit models	+ Evaluates the fit of item response theory (IRT) models when ability is estimated imprecisely + Displays graphs of empirical- and model-based item category response functions	
IRTGEN Tiffany Whittaker, Steven Fitzpatrick, Natasha Williams & Barbara Dodd	A collection of SAS macros that can generate known ability scores according to the random normal or random uniform distribution and item responses for simulees based on the graded response, partial credit,	+ Free manual + Generates responses to items with differing numbers of categories when one of the polytomous IRT models is used	- Input must be created by a SAS set containing item parameters for the desired IRT model

	generalized partial credit, successive intervals and three-parameter logistic models	
IRTGRAPH Ruth Childs	A program that processes item parameter estimate files for one-, two-, three- parameter logistic, graded response models and generalized partial credit models from MULTILOG and PARSCALE .	<ul style="list-style-type: none"> + Would be particularly useful for producing large numbers of uniformly formatted graphics + Includes very simple SAS/GRAFH formatting options
IRTLRDIF David Thissen	Computes likelihood ratio tests of DIF for the three- parameter logistic and graded IRT models	<ul style="list-style-type: none"> + Free download
http://www.unc.edu/~dthiss/en/dl.html		
IRTnew David Thissen	Computes indices of local dependence---diagnostic statistics that may be used to supplement parameter estimates and other indices of goodness of fit produced in the analysis of test data using item response theory	<ul style="list-style-type: none"> + Free download - Lacks a Windows interface
http://www.unc.edu/~dthiss/en/dl.html		
IRT Painter Ning Han	Plots IRT graphs such as item characteristic curves (ICCs), item information functioning (IIFs), test characteristic curves (TCCs), and test information functioning (TIFs)	<ul style="list-style-type: none"> + Reads directly the output from commonly used commercial software + Input data file can be created by the user + The output plots are editable + Displays multiple item plots on - Not commercially available

one screen simultaneously

IRTScore David Thissen http://www.unc.edu/~dthissen/dl.html	Computes summed-score to EAP(theta) translation tables, and the values and weights used in linear IRT response-pattern scoring, given parameters from Multilog output files or space- or tab-delimited files		
LINKDIF Niels Waller http://peabody.vanderbilt.edu/depts/psych_and_hd/faculty/wallern/	An S-Plus function for linking IRT item parameters and calculating several measures of differential test (DTF) and item functioning (DIF)	+ Computes Lord's chi square, associated significance levels, the signed (ESA), unsigned (EUA) area measures between two item response functions, z-values for the ESA and EUA measures, Raju's compensatory (CDIF) and noncompensatory (NCDIF) DIF indexes, Raju's index of DTF with the associated chi square and significance level	
LOGIMO H. Kelderman & R. Steen http://www.assess.com	Estimates and tests ordinary loglinear models and loglinear IRT models	+ Can analyze various types of logistic IRT models: partial credit models, multidimensional partial credit models and multidimensional polytomous latent trait models + Computes maximum likelihood estimates and standard errors of the model parameters, observed	- Lack a Windows interface

		and expected sufficient statistics, the kernel of the likelihood, likelihood ratio goodness-of-fit statistics, and Pearson goodness-of-fit statistics	
LOGIST M. Wingersky, P. Barton, & F. Lord http://www.ets.org	Estimates IRT parameters for the one-, two-, or three- parameter logistic item-response models using unconditional maximum likelihood	<ul style="list-style-type: none"> + LOGIST V provides standard errors + Gives many options and is very flexible + Allows omits and not reached data 	<ul style="list-style-type: none"> - Input specifications are complex - Places many constraints on the parameters to obtain convergence
LPCM-Win Gerhard Fischer http://www.assess.com	Applies linear partial credit models (LPCM) in item analysis and the measurement of change; allows estimation of the Rasch model (RM), the multifactorial (multifacet) RM, the linear logistic test model (LLTM), the rating scale model (RSM), the partial-credit model (PCM), and <u>extensions of these models</u>	<ul style="list-style-type: none"> + The data may be dichotomous or polytomous items, ratings, or symptoms + Applies to multidimensional items in the measurement of change and the assessment of treatment effects 	<ul style="list-style-type: none"> - High cost
LRSM Gerhard Fischer & Peter Parzer	Computes the conditional maximum likelihood estimates of the basic model parameters for the linear rating scale model (Fischer & Parzer)	<ul style="list-style-type: none"> + Two versions available depending on whether each examinee responses to the same items or not and whether weights of the basic parameters can be different for each examinee and each item or not 	<ul style="list-style-type: none"> - Runs only under MS-DOS and CMS
LTDOMAIN Yuan Hwang Li	A look-up table for the corresponding estimated one-parameter logistic	<ul style="list-style-type: none"> + Prompts the user to enter the value of a scaling factor D 	<ul style="list-style-type: none"> - Runs only under DOS and MS-DOS

	model scale score and the unbiased domain score for each number-correct score		
MIRTE Jim Carlson	Estimates IRT parameters for the one-, two-, and three- parameter logistic item-response models using unconditional maximum likelihood	+ Fits a multidimensional model + Gives standard errors + Includes residual analysis	- the guessing (<i>c</i>) parameter is fixed
MOKSCAL Johannes Kingma & Terry Taerum	A program for the Mokken scale analysis based on a nonparametric item response model that makes no assumption about the functional form of the item trace lines	+ Two versions available: one version can be used as an SPSS-X user procedure and the other is a stand-alone program + Provides search procedure	- Output printed in up to 80 columns
MSP W. Molenaar, P. Debets, K. Sijtsma, & B. Hemker http://www.assess.com	Mokken scale analysis for polytomous items, scales test item responses using nonparametric cumulative item response theory	+ Can analyze either dichotomous or polytomous data + Can test scalability of a given scale or construct one or more unidimensional scales from an item pool	- Does not allow missing data - Handles up to 100 polytomous items with a maximum of 10 ordered categories within a given range (900 item steps) - Can analyze up to 32,000 examinees
MULTILOG David Thissen, Wen-Hung Chen, & Darrell Bock http://www.assess.com http://www.ssicentral.com http://www.unc.edu/~dthissen/dl.html	Handles one, two, and three parameter logistic, multiple nominal categories, graded rating scale, partial credit, multiple-choice, and constrained parameter models	+ Features a dialog-box user interface to assist first-time users or occasional users with writing such a command file + Presents quality IRT graphics which can be imported in other formats	- Some bugs reported in 6.0 that were corrected in 6.2/6.3

		<ul style="list-style-type: none"> + Analyze mixtures of items types + Handles data from several populations simultaneously + Can analyze DIF and standardized residuals + Handles unlimited number of items and unlimited number of examinees 	
MULTISIM Educational Measurement Lab at the University of Illinois at Urbana-Champaign http://www.stat.uiuc.edu/psychometrics/software.html	Simulates dichotomous multidimensional test responses using a multidimensional compensatory logistic IRT model	<ul style="list-style-type: none"> + Unlimited number of examinees simulated + A user-specified multivariate normal distribution of the underlying latent ability distribution 	<ul style="list-style-type: none"> - Generate test response data having up to four latent dimensions for up to 120 items
MultiRa Claus Carstensen & Jürgen Rost http://www.multira.de/	Implements an algorithm for the multidimensional item component Rasch model and other Rasch models; Functions includes: Examinee and item parameter estimates according to the joint maximum likelihood and conditional maximum likelihood estimation, model fit measures and MKAT Algorithm by J. Rost for the Rasch multidimensional Model (1961)	<ul style="list-style-type: none"> + English version software is available 	<ul style="list-style-type: none"> - Manual is only available in German version
NOHARM	Fits both unidimensional and	<ul style="list-style-type: none"> + Free software 	

<p>(Normal Ogive Harmonic Analysis Robust Method)</p> <p>Colin Fraser & Roderick McDonald</p> <p>http://www.unt.edu/rss/clas/s/rich/5840/mcdonald/hohar/m/NOHARM%20Downloads.htm</p>	<p>multidimensional normal ogive models of IRT</p>	<p>+ Includes residual analysis</p>	
<p>OPLM (One Parameter Logistic model)</p> <p>Verhelst, Glas, & Verstralen</p> <p>http://www.citogroep.nl/ex/p/poc/OPLM.htm</p>	<p>Combines the attractive mathematical properties of the Rasch model with the flexibility of the two-parameter logistic model.</p>	<p>+ Difficulty parameters are estimated and discrimination indices are imputed as known constants</p>	<p>- Lacks a Windows interface</p>
<p>PARDSIM</p> <p>Michael Yoes</p> <p>http://www.assess.com</p>	<p>Generates IRT item and examinee parameters and/or item response (data) files for the Rasch, two-parameter, and three-parameter IRT models using Monte Carlo simulation methods</p>	<p>+ Item response files generated with PARDSIM are ready for analysis using XCALIBRE, RASCAL, or a similar IRT analysis software program</p> <p>+ User configurable options allow to specify the type of distribution for each parameter</p>	
<p>PARELLA</p> <p>H. Hoijtink, W. Molenaar, & W. Post</p>	<p>A parametric item response model that can be used for measurements of attitudes and preferences and the</p>	<p>+ Offers several goodness-of-fit tests and diagnostics</p>	<p>- The maximum input for the PARELLA program is 60 items, 10</p>

http://www.assess.com	locations of the items, and the (nonparametric) density function of the examinees' locations can be estimated using marginal maximum likelihood	subsamples, and 300 examinees - Only applies to dichotomous response - lacks a Windows interface
PARSCALE Eiji Muraki, Darrell Bock http://www.assess.com http://www.ssientral.com	Estimates IRT parameters for one, two, and three-parameter models, Samejima's graded responses model, Master's partial credit model and Generalized partial credit models	<ul style="list-style-type: none"> + Analyzes rating scale items, multiple-choice items, differential item functioning (DIF), mixtures of item types, rater-effect; applies to multiple-group polytomous item response models + Presents quality IRT graphics + Can be imported in Word, Access, etc. + Handles unlimited number of items and unlimited number of examinees
Plotlog David Thissen http://www.unc.edu/~dthissen/dl.html	Displays various graphics associated with item response theory, such as trace lines, information curves, test characteristic functions, and likelihoods (or posterior distributions)	<ul style="list-style-type: none"> + Uses item parameters that maybe supplied by the user at the keyboard, or from Mutilog's .SAV files - Lacks a Windows interface - Limited to number of items up to 200 items.
PRASCH John Grego	Fits latent class polytomous response Rasch models using conditional maximum likelihood estimates (CMLEs) and tests a moment	<ul style="list-style-type: none"> + Input file available in three possible formats: tabular, grouped with counts and ungrouped - Not widely used

	condition on a sequence of ratio statistics generated from the CMLEs		
PRASCH/ECIZ Randall Nelson & Steven Chatman	A program that uses the PROC MATRIX commands of SAS to score a test, compute item and examinee parameters for the Rasch model, and provide several measures of fit to the model for each examinee	<ul style="list-style-type: none"> + Reports examinee fit: mean square fit(Wright & Stone), two extended caution indices, EC12 and EC14 (Tatsuoka & Linn), two standardized extended caution indices and EC1Z2 and EC1Z4 (Tatsuoka) + Runs on any system that supports the SAS MATRIX languages 	<ul style="list-style-type: none"> - Analyses with large numbers of examinees and items requires a large memory region and may necessitate running the program in segments
RASCAL Assessment System Corp. http://www.assess.com	Estimates the item difficulty and examinee (ability) parameters based on the one-parameter logistic IRT model for dichotomous data	<ul style="list-style-type: none"> + Provides a graphical “items-by-examinees” map and graphical displays of the test characteristic curve and test information function 	<ul style="list-style-type: none"> - Handles up to 750 items with unlimited sample size for the extended and Windows Versions
RIDA Glas	Estimates item parameters based on the one-parameter logistic model using conditional or marginal maximum likelihood estimation	<ul style="list-style-type: none"> + Provides a complete analysis of examinees + Handles incomplete designs for test equating + Includes fit analysis 	
Resgen Eiji Muraki	Generates simulated latent trait distributions and dichotomous or polytomous item responses for One-, Two-, or Three-Parameter Model; Logistic or Normal Ogive Model; unidimensional or multidimensional Model; and Graded Response or	<ul style="list-style-type: none"> + The latent trait distributions can be univariate or multivariate normal, log-normal, uniform, or gamma. + Be capable of simulating realistic testing situations 	

Partial Credit Model.			
RUMM2010 Sheridan, Andrich, & Luo To install after you download, double click on the file and follow the prompts accepting the default settings. It might complain about a DLL - just ignore that and keep pressing enter, continue or finish. The site is www.rummlab.com/demo/installrumm2030wksheval.exe	A comprehensive item analysis package which fits most IRT rating-scale models, including Samejima's graded response model generalized for rating scale and Masters' partial credit model with or without discriminating parameter	<ul style="list-style-type: none"> + User friendly interaction procedure + Can be imported in Word etc. + Provides easy to read Tables and Plots + Handles multiple subsets and weighted combinations of subtest scores + Includes adjustments for differences in rater severity , multi-group DIF analysis and Examinee Factor and Item Factor (or Facet) analyses 	
RUMMFOLD David Andrich http://www.assess.com/	A Windows program for scaling attitude and preference data by single-stimulus (direct-response) designs, estimating the examinee trait levels and item location parameters of the one-parameter logistic Rasch unfolding measurement model (RUMM)	<ul style="list-style-type: none"> + Easy-to-use mouse-driven Windows program + Output options allow for graphical display of the results for both examinee and item parameter estimates and related statistics 	<ul style="list-style-type: none"> - Maximum of 100 items and 5,000 examinees
RUMMFOLD David Andrich http://www.assess.com/	A Windows program for scaling attitude and preference data by a paired-comparison data collection design, estimating the examinee trait	<ul style="list-style-type: none"> + Easy-to-use mouse-driven Windows program + Output options allow for graphical display of the results for both 	<ul style="list-style-type: none"> - Maximum of 100 items and 5,000 examinees

	scores and item location parameters of the one-parameter logistic unfolding measurement model (RUMM)	examinee and item parameter estimates and related statistics	
SCOREALL http://www.assess.com/	Calculates the maximum-likelihood and Bayesian scores for tests administered by paper-and-pencil	<ul style="list-style-type: none"> + Output contains the number-correct score, simple formula score, Bayesian-modal and maximum-likelihood ability estimates, and EAP Bayesian ability estimate for each examinee 	<ul style="list-style-type: none"> - Examinee response data and a, b, and c item parameters as input are required
SScore David Thissen http://www.unc.edu/~dthissen/dl.html	An interim version of a stand-alone application, programmed in C++, that computes <i>expected a posteriori</i> (EAP) scores, the associated standard deviations, and scale scores (and associated standard errors) for summed scores, given 3PL item parameters	<ul style="list-style-type: none"> + Input file of 3PL item parameters may be in the .sav output file format produced by Multilog or a file with the a, b, and c parameters in columns in a space-delimited format + Free program 	<ul style="list-style-type: none"> - The executable lacks any facility for dialog-box input file selection
T-Rasch I. Ponocny & E. Ponocny-Seliger http://www.assess.com/	A tool for test construction and detailed item analysis using the Rasch model	<ul style="list-style-type: none"> + Implements non-parametric goodness-of-fit tests for the Rasch model + Calculates non-asymptotic p-values referring to most powerful tests for each of the scales separately + Can do item bias analysis + Especially useful with small samples of examinees 	<ul style="list-style-type: none"> - Number of items is limited up to 30 items
TESTINFO	A graphics-oriented program for	<ul style="list-style-type: none"> + Displays test information 	<ul style="list-style-type: none"> - Displays item functions

John H. Neel	investigating the effect of item selection on test information and standard error of measurement for one-, two-, and three-parameter models	functions and standard error of measurement functions	for up to 56 items - Lacks a Windows interface
TestFact Douglas Wilson, Robert Wood, Stephen Schilling & Robert Gibbons http://www.assess.com/	Performs full information factor analysis using multi-dimensional IRT models	+ Handles up to 10 factors using numerical quadrature: up to 5 for non-adaptive, up to 10 for adaptive quadrature and up to 15 factors using Monte Carlo integration techniques + Simulates responses to items based on user specified parameters + Handles unlimited sample size	- Number of items is limited up to 1,000 items
TestGraf James O. Ramsay ftp://ego.psych.mcgill.ca/pub/ramsay/testgraf/ http://www.psych.mcgill.ca/faculty/ramsay/TestGraf.html	A MS-DOS machine that provides a graphical analysis of multiple-choice test items and/or rated responses using Ramsay's "kernel smoothing" approach to item response theory	+ Would be useful when the number of examinees is of the order of 100 or more and the number of questions or choices exceeds 20 or so + Enables better estimates of examinee proficiency or ability by making use of the information provided by which wrong options were chosen for incorrectly answered items + Displays graphically what range of proficiency is reasonably consistent with the set of choices	- Lacks a Windows interface

		made by an individual
WinGen Chris Han http://www.umass.edu/rem/p/software/wingen/	A software package for generating data for research purposes. Software can generate ability scores and item responses consistent with all popular IRT models for handling dichotomous and polytomous response models. Can generate multidimensional data too (up to five dimensions) for any correlational structure. Ability scores and item parameter values can be chosen from several distributions specified by the user. Parameter files from BILOG-MG or Parscale can be read in also.	<ul style="list-style-type: none"> + Free program + Highly flexible software for generating item response data to fit man + Produces lots of graphics + A terrific tool for instructing students who want to learn about IRT + No practical limit on the length of the test or the number of examinees + Has special features for conducting item bias research
WINMIRA 2001 Matthias von Davier http://www.assess.com/	Stand-alone software that estimates and tests a large number of discrete mixture models for categorical variables and it can be used for analyses with the latent class analysis (LCA), with the Rasch model (RM), with the mixed Rasch model (MRM) and with hybrid models (HYBRID) for dichotomous and polytomous data	<ul style="list-style-type: none"> + Reads and writes data directly in SPSS file format and EXCEL, and other spreadsheet data files can be imported and exported + Estimates the partial credit model, the rating scale model, the equidistance model and the dispersion model for polytomous data + Can be used for scale construction + Can be imported to EXCEL and SPSS etc.

WINSTEPS John M. Linacre & Benjamin D. Wright http://www.winsteps.com	A Windows-based software that runs most of the Rasch models and can handle dichotomous, multiple-choice, and multiple rating-scale and partial credit items	+ Straightforward to use in combination with other software + Can process up to 1,000,000 examinees, 30,000 items, and each item can have a rating scale of up to 255 categories	
XCALIBRE http://www.assess.com/	A marginal maximum-likelihood estimation (MMLE) program for computing item and trait parameters for the two-parameter logistic (2PLM) and the three-parameter logistic (3PLM) item response theory (IRT) models	+ Estimates the parameters from dichotomously scored test data using the expectation-maximization (EM) algorithm to implement MMLE + Logically formats the graphical user interface and easy to use + Handles unlimited sample size	- Number of items is limited at 750 items - No on-line help - Produces test-level rather than item-level graphics - Allows only one sample run
YeomanDG Ralph De Ayala	A program that generates data for the one-, two-, and three- parameter dichotomous models, the graded response, nominal response, and partial credit models, the multidimensional three-parameter dichotomous model and the multidimensional graded response models	+ Data can be generated according to a linear factor analytic approach, a latent class model or randomly + Ability scores may be randomly generated from normal, uniform, or beta distributions	- Maximum number of 150, six-option items and 32,000 examinees - Limited to four-dimensional problems for the multidimensional models and specifies up to five latent classes for latent class models

Version: January 6, 2009