Test and Item Analysis Report

CZmaturaS
COMPS

Report created on **October 21, 2023**

Contents

Introduction

This report was created by R version 4.3.1 and its package ShinyItemAnalysis version 1.5.0. ShinyItemAnalysis provides test and item analysis and it is available on CRAN and also online.

To cite ShinyItemAnalysis application in publications, please, use:

[1] Martinková P., & Drabinová A. (2018) ShinyItemAnalysis for teaching psychometrics and to enforce routine analysis of educational tests. *The R Journal, 10*(2), 503-515. https://doi.org/10.32614/RJ-2018-074

Summary

Total scores

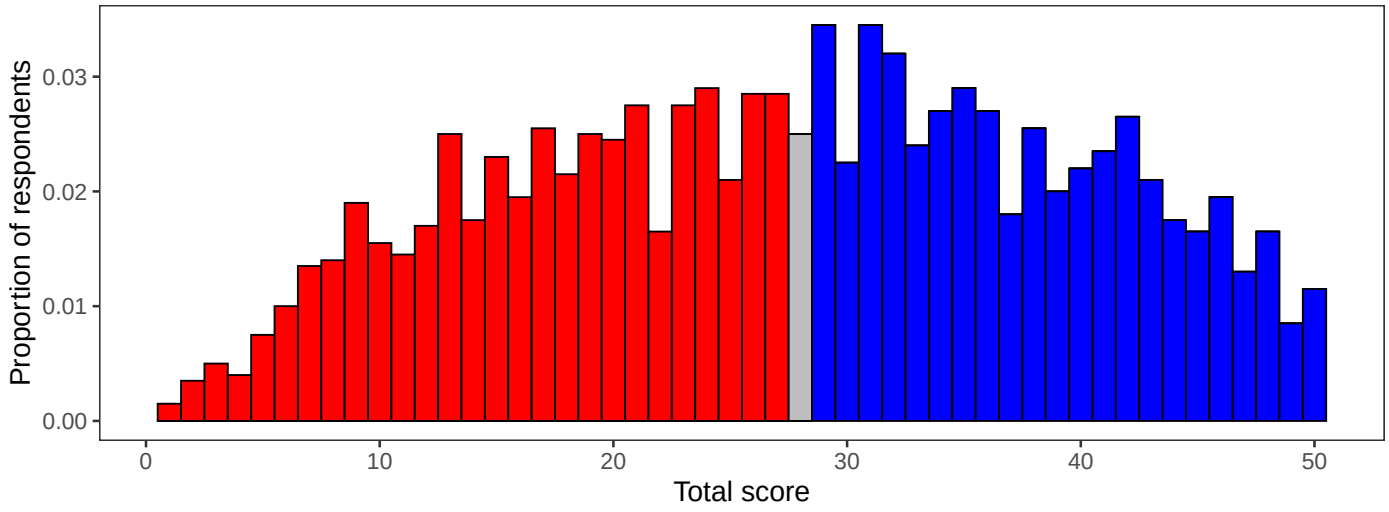
Summary table of total scores

The table below summarizes basic characteristics of total scores including number of respondents, minimum and maximum values, mean, median, standard deviation, skewness, and kurtosis. The skewness for normally distributed scores is near the value of 0 and the kurtosis is near the value of 3.

n	n_c	Min	Max	Mean	Median	SD	Skewness	Kurtosis
2000	2000	1	50	27.59	28.00	11.95	-0.09	2.08

Histogram of total scores

For selected cut-score 28, the blue part of a histogram shows respondents with a total score above the cut-score, the grey column shows respondents with a total score equal to the cut-score and the red part of a histogram shows respondents below the cut-score.



Scoring

Summary table of standard scores

The total score, also known as the raw score, is the total number of correct answers. It can be used to compare an individual score to a norm group, e.g. if the mean is 12, then an individual score can be compared to see if it is below or above this average. The percentile indicates the value below which a percentage of observations falls, e.g. an individual score at the 80th percentile means that the individual score is the same or higher than the scores of 80% of all respondents. The success rate is the percentage of correct answers, e.g. if the maximum points of a test is equal to 20 and an individual score is 12 then the success rate is 12/20 = 0.6, i.e. 60%. The Z-score, or the standardized score, is a linear transformation of the total score with a mean of 0 and with a variance of 1. If X is the total score, M is its mean and SD is its standard deviation then Z-score = (X - M) / SD. The T-score is the transformed Z-score with a mean of 50 and a standard deviation of 10. If Z is Z-score then T-score = (Z * 10) + 50.

Total score	Percentile	Success rate	Z-score	T-score
1	0.00	2	-2.22	27.75
2	0.00	4	-2.14	28.59
3	0.01	6	-2.06	29.43
4	0.01	8	-1.97	30.26
5	0.02	10	-1.89	31.10
6	0.03	12	-1.81	31.94
7	0.04	14	-1.72	32.77
8	0.06	16	-1.64	33.61
9	0.08	18	-1.56	34.45
10	0.09	20	-1.47	35.28
11	0.11	22	-1.39	36.12
12	0.12	24	-1.30	36.96
13	0.15	26	-1.22	37.79
14	0.17	28	-1.14	38.63
15	0.19	30	-1.05	39.46
16	0.21	32	-0.97	40.30
17	0.24	34	-0.89	41.14
18	0.26	36	-0.80	41.97
19	0.28	38	-0.72	42.81
20	0.31	40	-0.64	43.65
21	0.33	42	-0.55	44.48
22	0.35	44	-0.47	45.32
23	0.38	46	-0.38	46.16
24	0.41	48	-0.30	46.99
25	0.43	50	-0.22	47.83
26	0.46	52	-0.13	48.67
27	0.48	54	-0.05	49.50
28	0.51	56	0.03	50.34

Continued on next page

Total score	Percentile	Success rate	Z-score	T-score
29	0.54	58	0.12	51.18
30	0.57	60	0.20	52.01
31	0.60	62	0.28	52.85
32	0.63	64	0.37	53.69
33	0.66	66	0.45	54.52
34	0.68	68	0.54	55.36
35	0.71	70	0.62	56.19
36	0.74	72	0.70	57.03
37	0.76	74	0.79	57.87
38	0.78	76	0.87	58.70
39	0.80	78	0.95	59.54
40	0.83	80	1.04	60.38
41	0.85	82	1.12	61.21
42	0.88	84	1.21	62.05
43	0.90	86	1.29	62.89
44	0.91	88	1.37	63.72
45	0.93	90	1.46	64.56
46	0.95	92	1.54	65.40
47	0.96	94	1.62	66.23
48	0.98	96	1.71	67.07
49	0.99	98	1.79	67.91
50	1.00	100	1.87	68.74

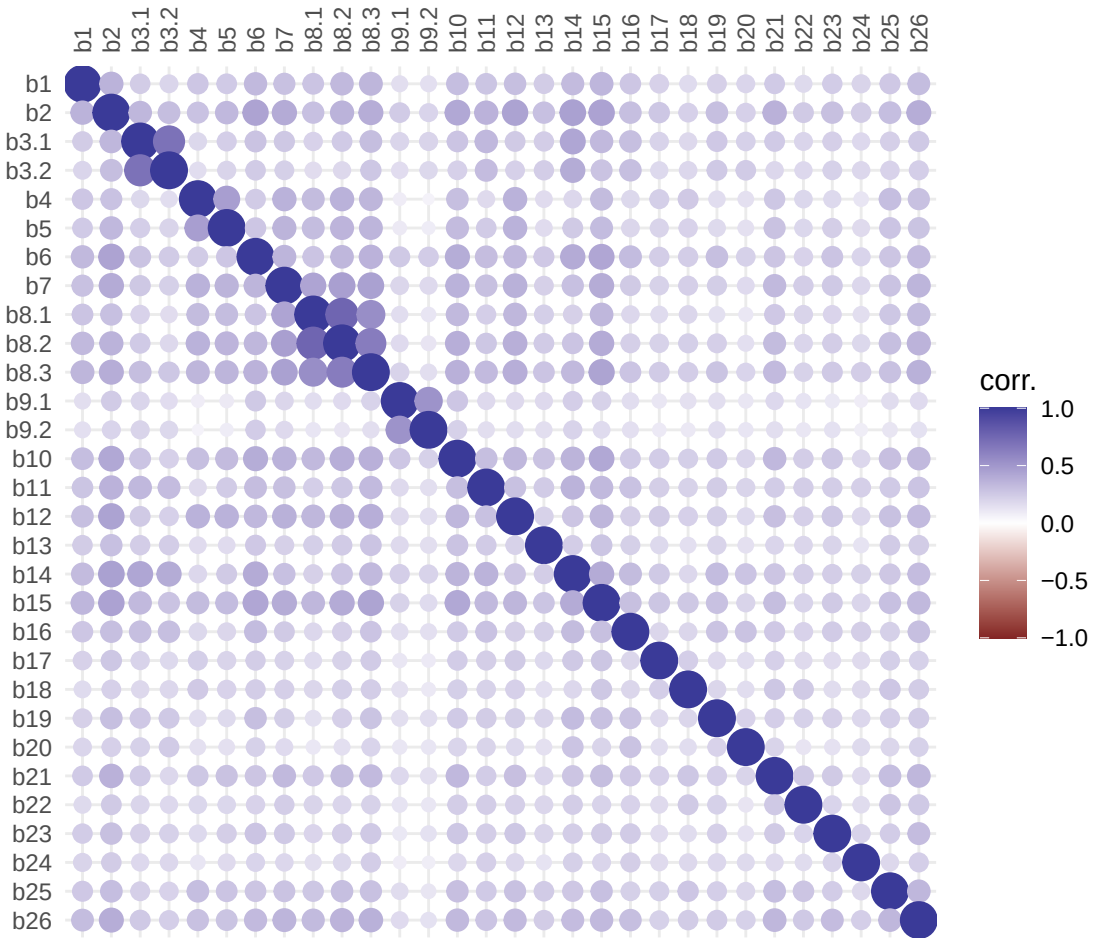
More complex estimates of ability are provided in the IRT section.

Validity

Correlation structure

Correlation heat map

A correlation heat map displays Pearson correlations of items. Pearson correlation coefficient describes linear correlation between two random variables X and Y.

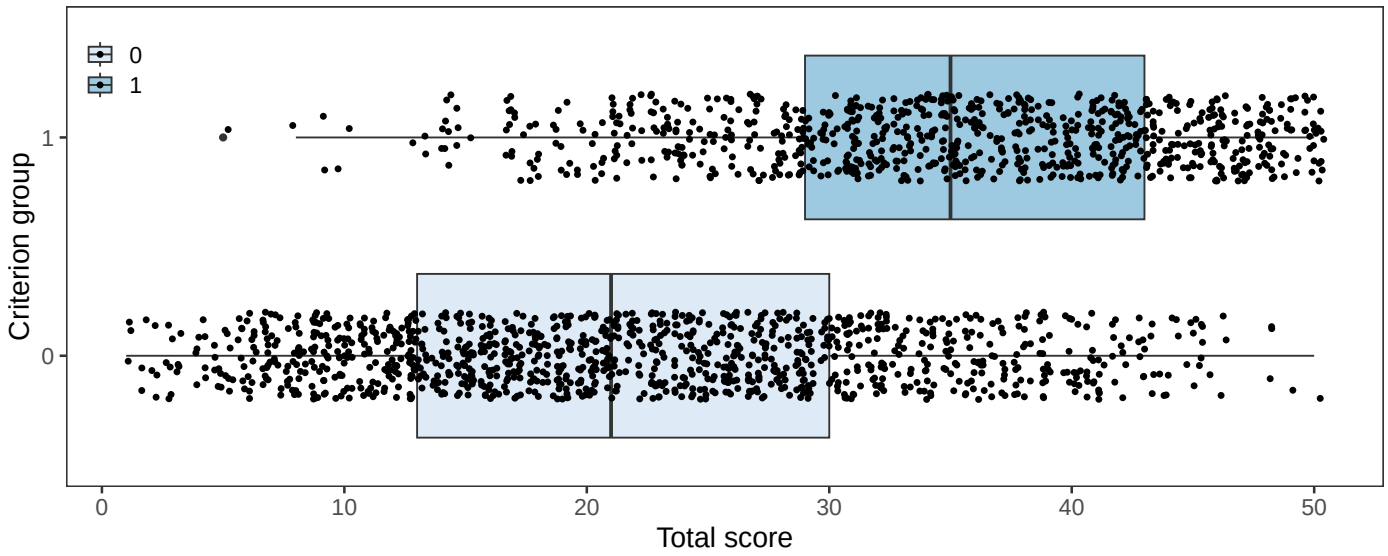


Predictive validity

An association between the total score and the criterion variable can be estimated using Pearson product-moment correlation coefficient *r*. The null hypothesis being tested states that correlation is exactly 0.

$r(1998) = .54$, $p = <.001$, 95% CI [.50, .57]

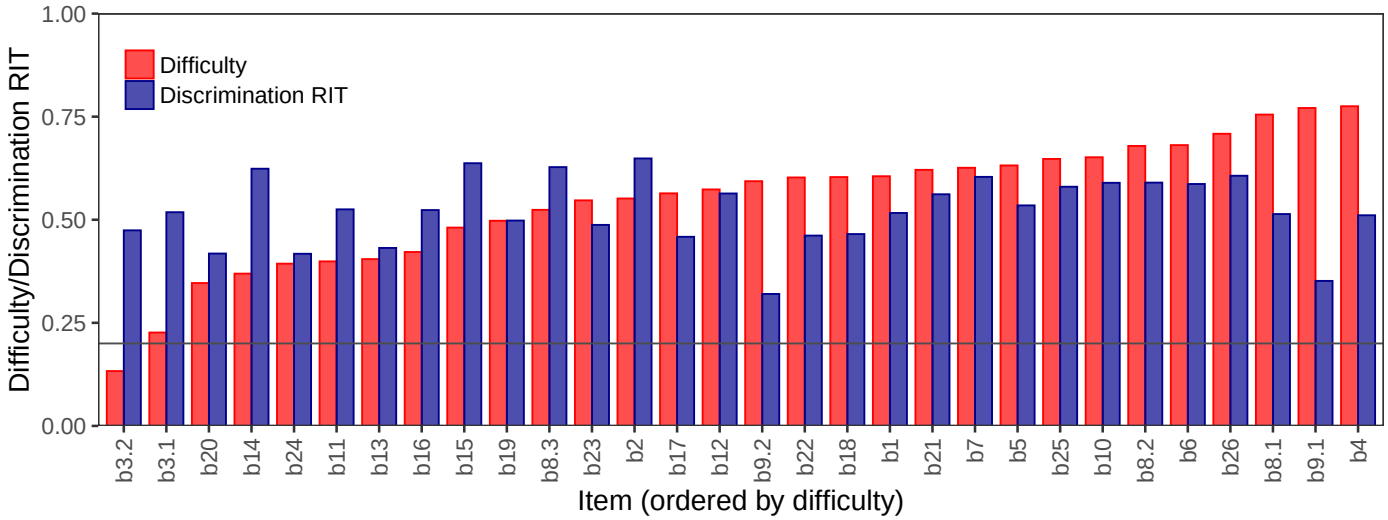
Interpretation:
The *p*-value is less than .05, thus we reject the null hypotheses. The total score and criterion variable are positively correlated.



Traditional item analysis

Item analysis

Difficulty/Discrimination plot



Cronbach's alpha

Cronbach's alpha is an estimate of internal consistency of a psychometric test. It is a function of the number of items in a test, the average covariance between item-pairs, and the variance of the total score (Cronbach, 1951).

Estimate	Confidence interval
0.890	(0.883, 0.897)

Traditional item analysis table

Item	Diff.	Avg. score	SD	ULI	RIT	RIR	Alpha Drop
b1	0.61	0.61	0.49	0.59	0.52	0.49	0.89
b2	0.55	0.55	0.50	0.76	0.65	0.62	0.88
b3.1	0.23	0.23	0.42	0.47	0.52	0.49	0.89
b3.2	0.13	0.13	0.34	0.33	0.47	0.45	0.89
b4	0.78	1.55	0.81	0.47	0.51	0.46	0.89
b5	0.63	1.26	0.92	0.56	0.53	0.48	0.89
b6	0.68	0.68	0.47	0.64	0.59	0.56	0.89
b7	0.63	1.25	0.91	0.64	0.60	0.55	0.88
b8.1	0.76	0.76	0.43	0.51	0.51	0.49	0.89
b8.2	0.68	0.68	0.47	0.65	0.59	0.56	0.89
b8.3	0.52	0.52	0.50	0.74	0.63	0.60	0.89
b9.1	0.77	0.77	0.42	0.34	0.35	0.32	0.89
b9.2	0.59	0.59	0.49	0.35	0.32	0.28	0.89
b10	0.65	0.65	0.48	0.67	0.59	0.56	0.89
b11	0.40	0.40	0.49	0.59	0.53	0.49	0.89
b12	0.57	0.57	0.49	0.66	0.56	0.53	0.89
b13	0.40	0.40	0.49	0.48	0.43	0.40	0.89
b14	0.37	1.11	1.42	0.68	0.62	0.54	0.89
b15	0.48	0.96	0.96	0.72	0.64	0.59	0.88
b16	0.42	0.84	0.88	0.53	0.52	0.47	0.89
b17	0.56	1.13	0.99	0.52	0.46	0.39	0.89
b18	0.60	1.21	0.98	0.53	0.47	0.40	0.89
b19	0.50	1.00	1.00	0.58	0.50	0.43	0.89
b20	0.35	0.69	0.95	0.45	0.42	0.35	0.89
b21	0.62	1.24	0.97	0.65	0.56	0.50	0.89
b22	0.60	1.21	0.98	0.51	0.46	0.39	0.89
b23	0.55	1.09	1.00	0.57	0.49	0.42	0.89
b24	0.39	0.79	0.98	0.46	0.42	0.35	0.89
b25	0.65	2.59	1.48	0.50	0.58	0.49	0.89
b26	0.71	2.13	1.01	0.47	0.61	0.55	0.88

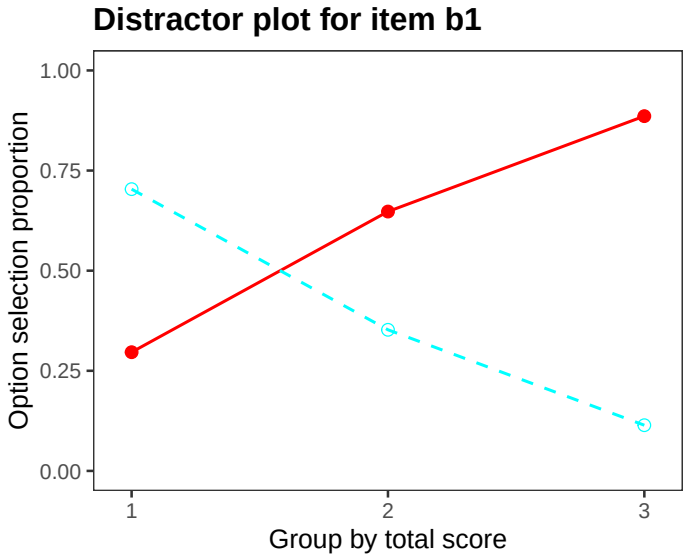
Distractor analysis

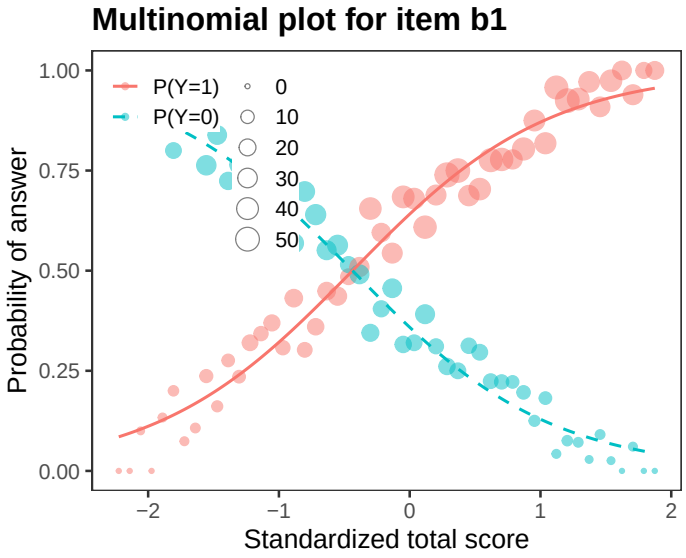
Distractor plot

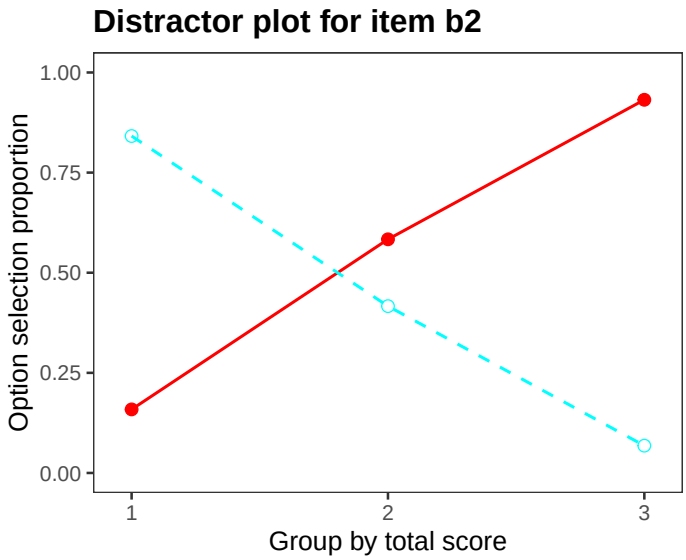
Respondents are divided into a selected number of groups by their total score. Subsequently, the percentage of respondents in each group who selected a given answer (correct answer or distractor) is displayed. The correct answer should be selected more often by the respondents with a higher total score than by those with a lower total score, i.e. the solid line should be increasing. The distractor should work in the opposite direction, i.e. the dotted lines should be decreasing.

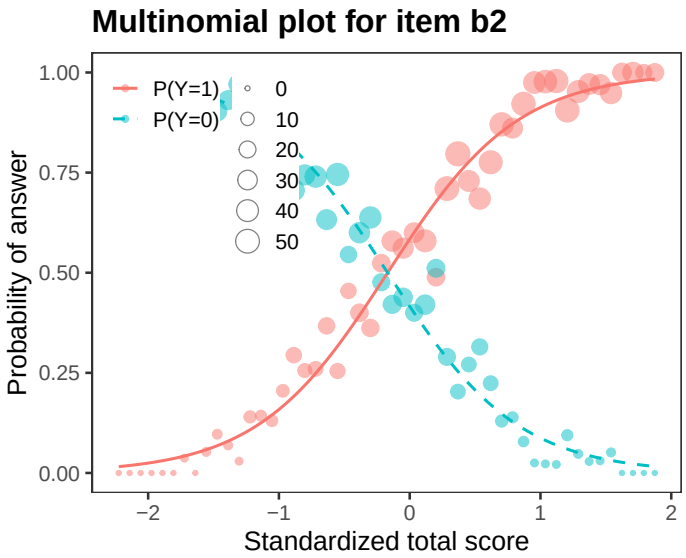
Multionomial plot

In the multinomial plot, points represent proportion of selected option with respect to total score. Their size is determined by count of respondents who achieved given level of total score and who selected given option.

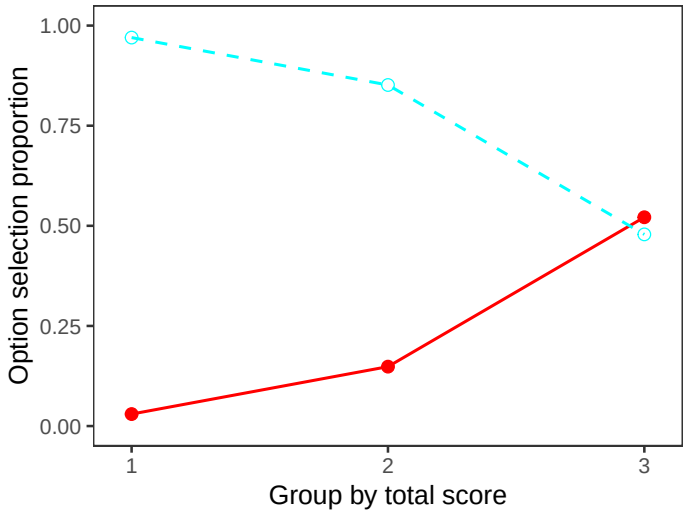




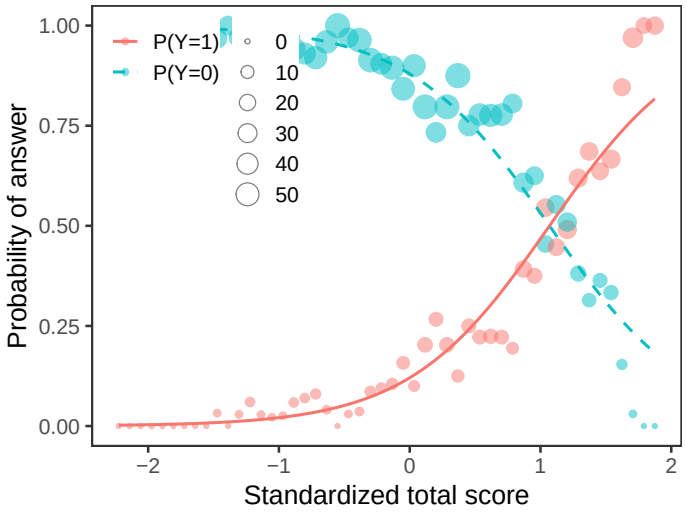




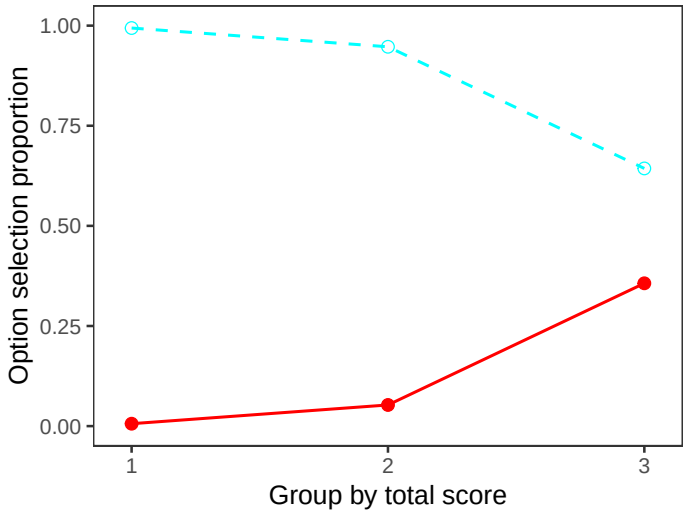
Distractor plot for item b3.1



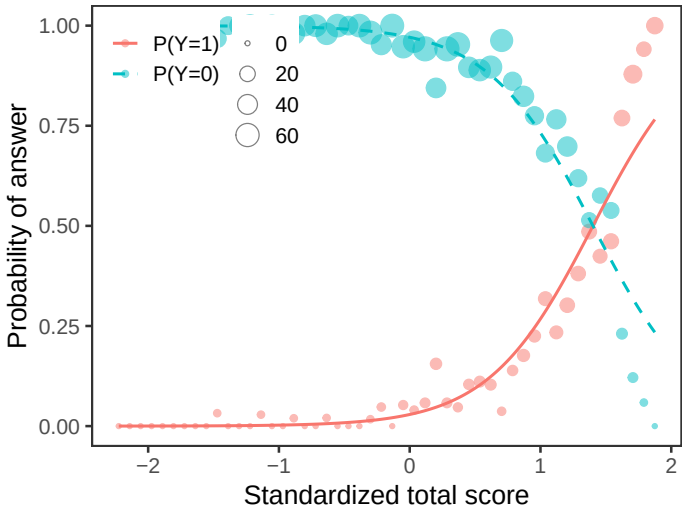
Multinomial plot for item b3.1



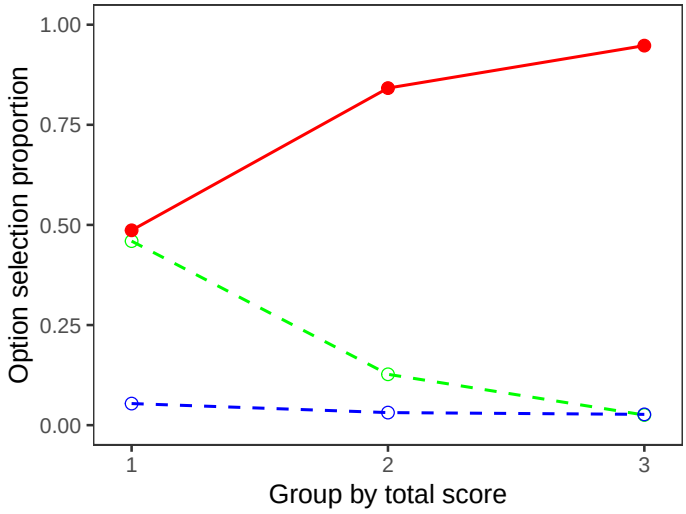
Distractor plot for item b3.2



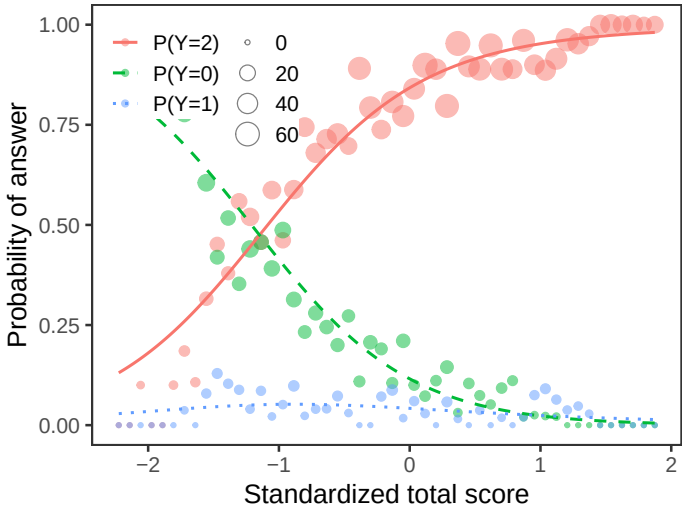
Multinomial plot for item b3.2



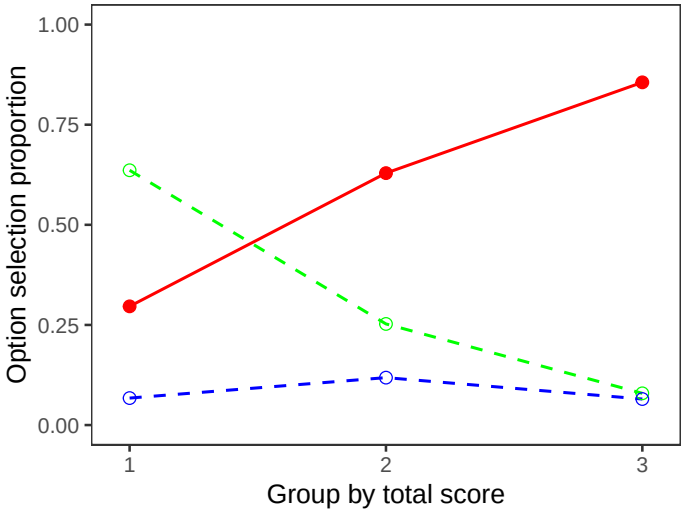
Distractor plot for item b4



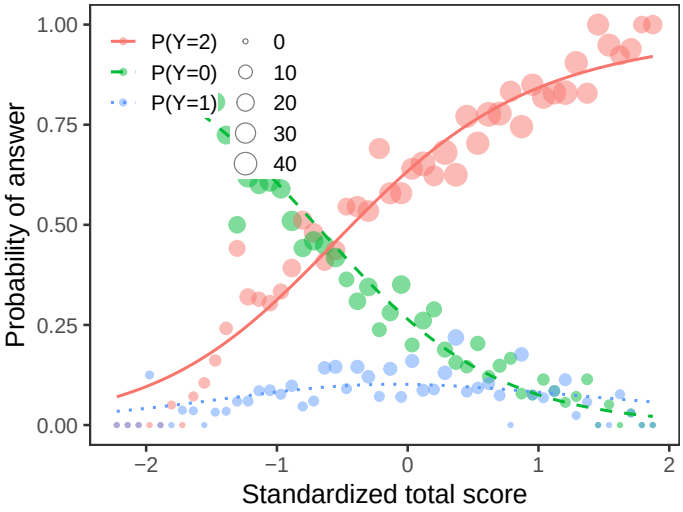
Multinomial plot for item b4



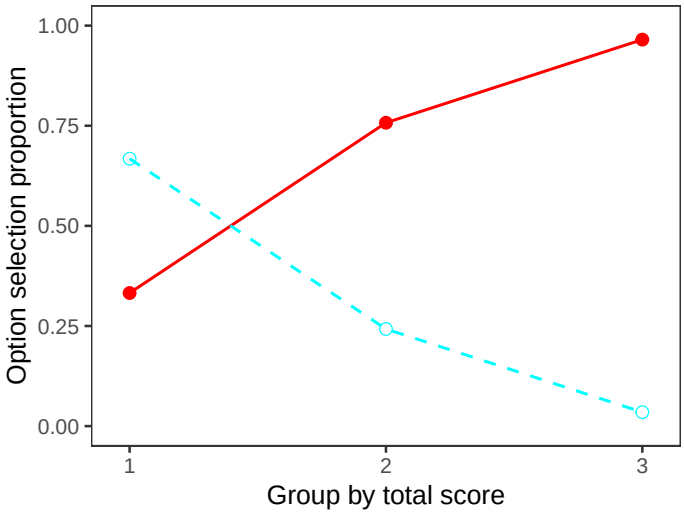
Distractor plot for item b5



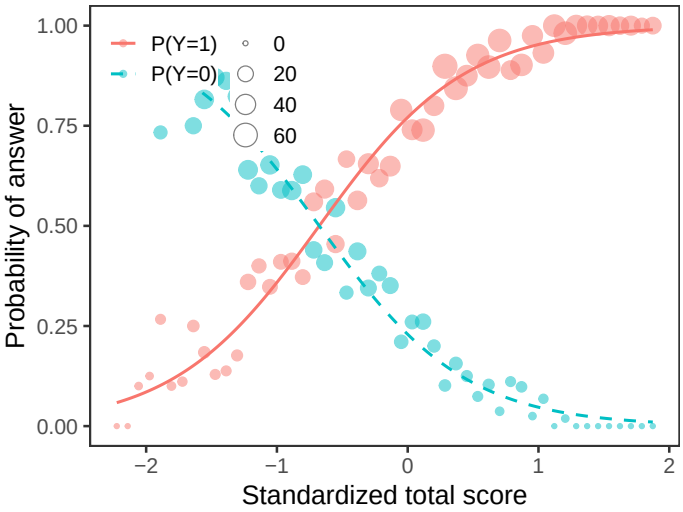
Multinomial plot for item b5



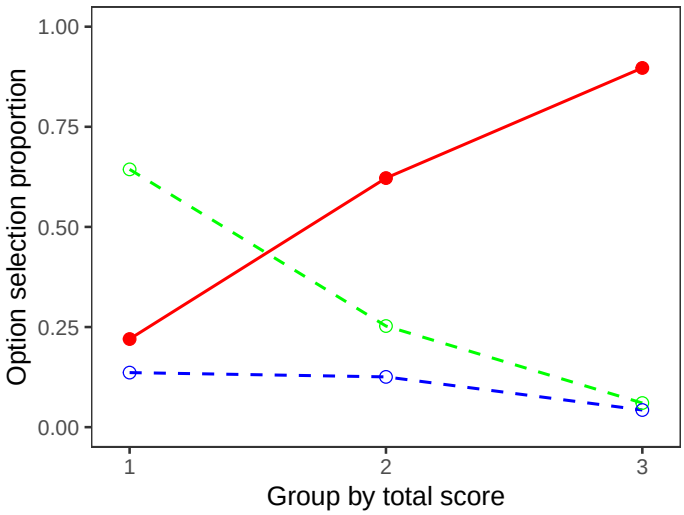
Distractor plot for item b6



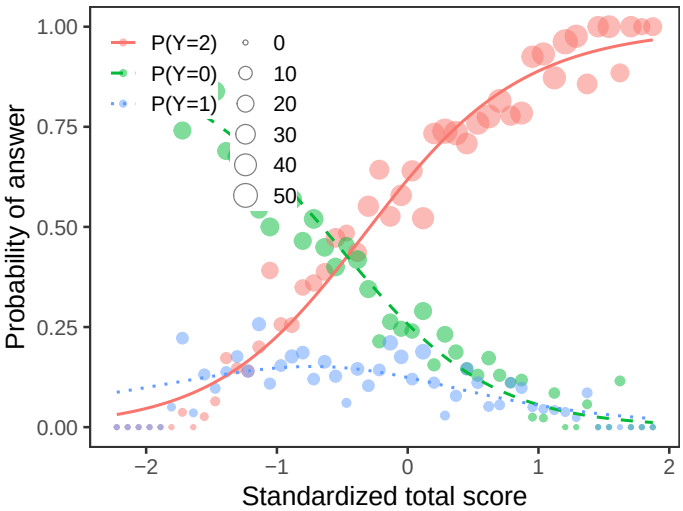
Multinomial plot for item b6



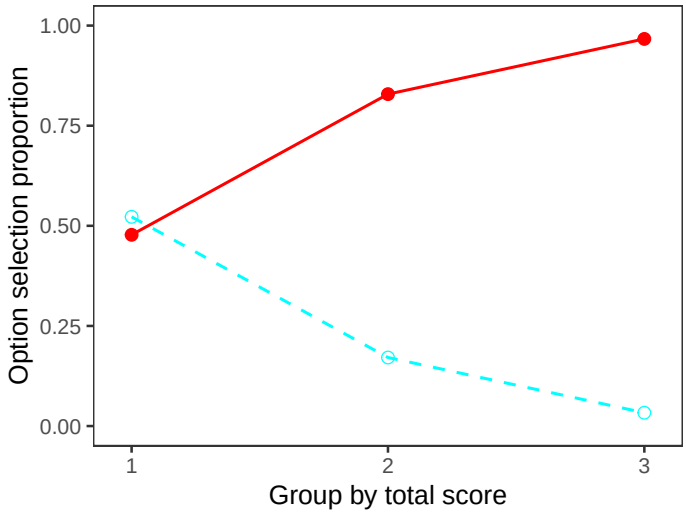
Distractor plot for item b7



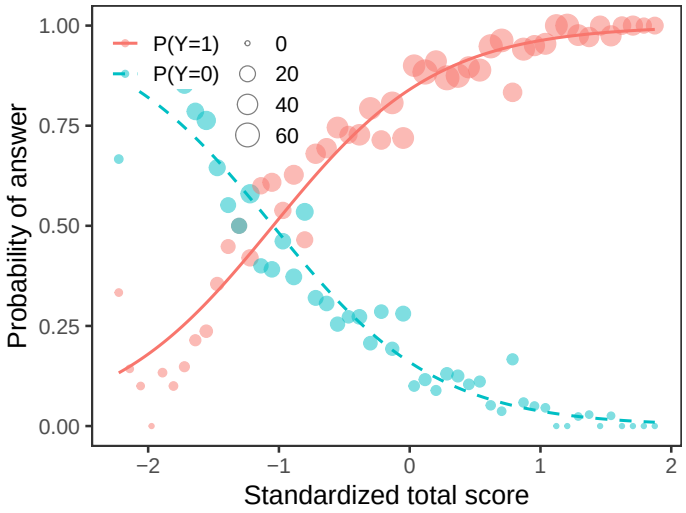
Multinomial plot for item b7



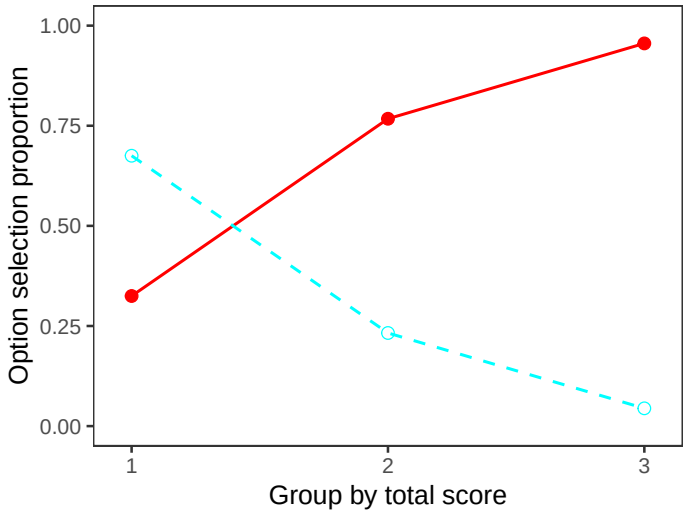
Distractor plot for item b8.1



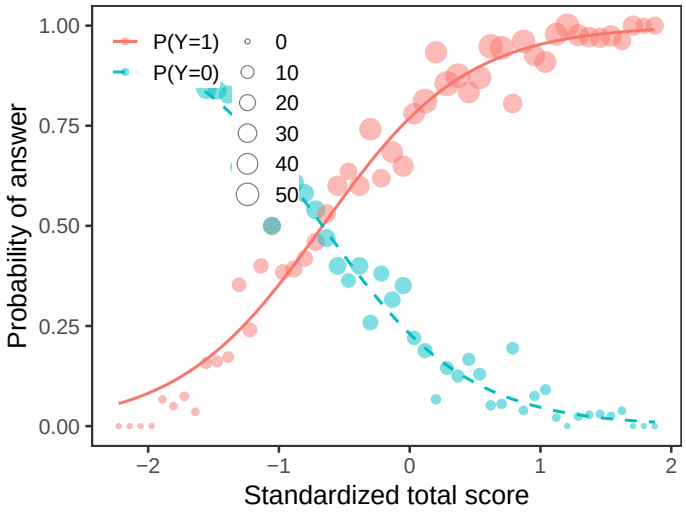
Multinomial plot for item b8.1



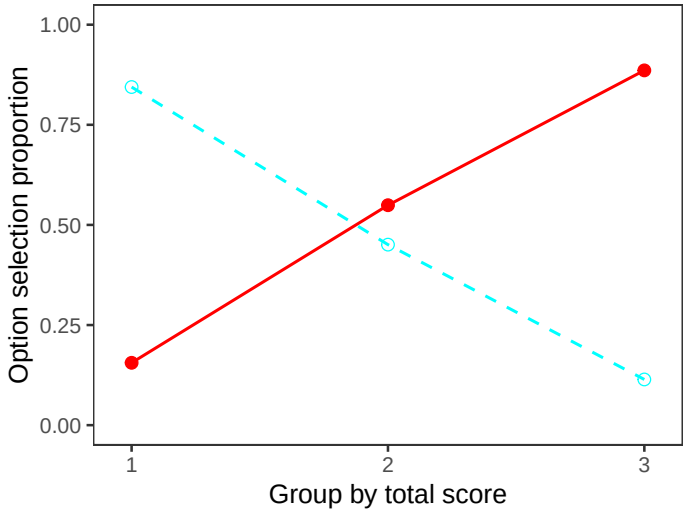
Distractor plot for item b8.2



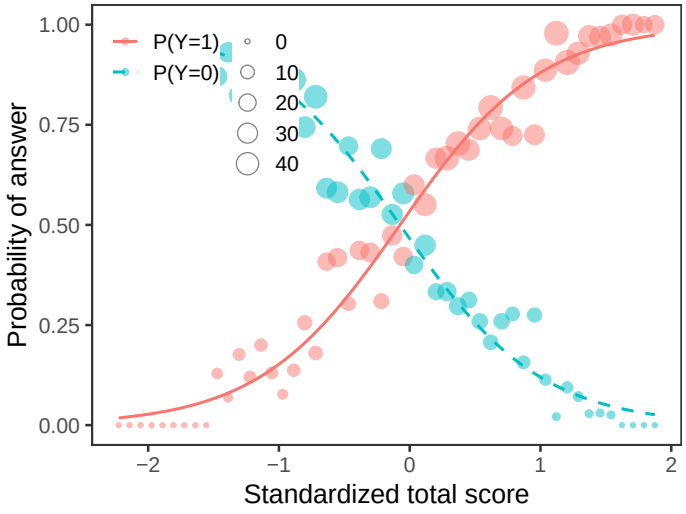
Multinomial plot for item b8.2



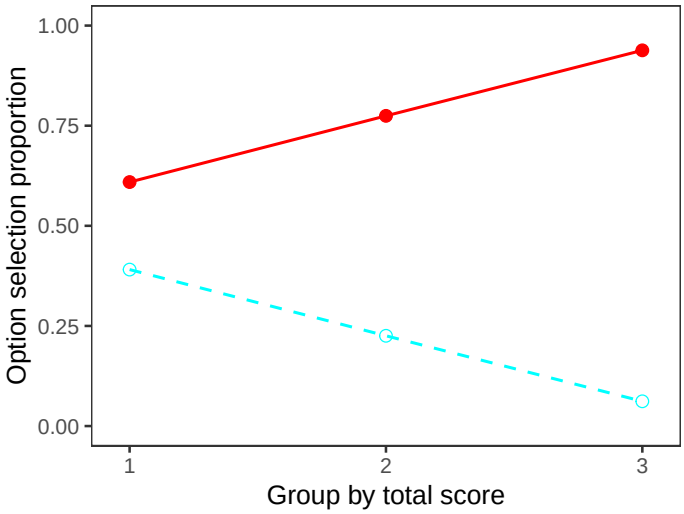
Distractor plot for item b8.3



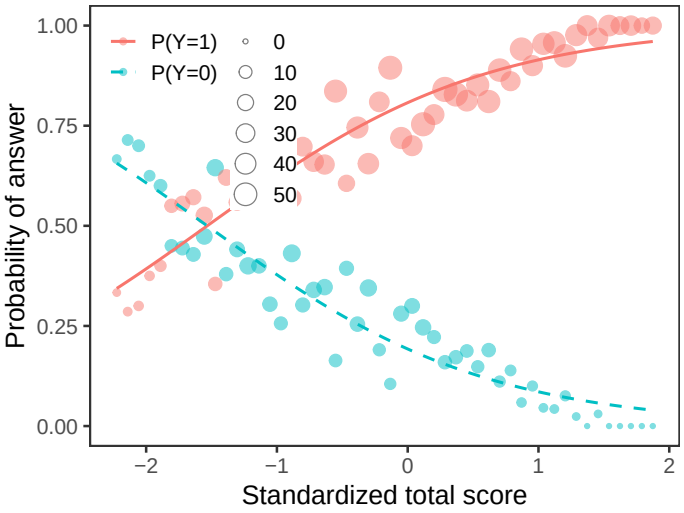
Multinomial plot for item b8.3



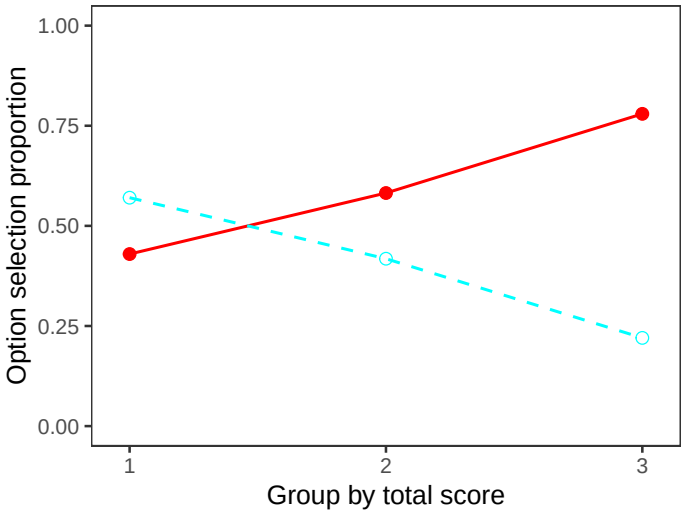
Distractor plot for item b9.1



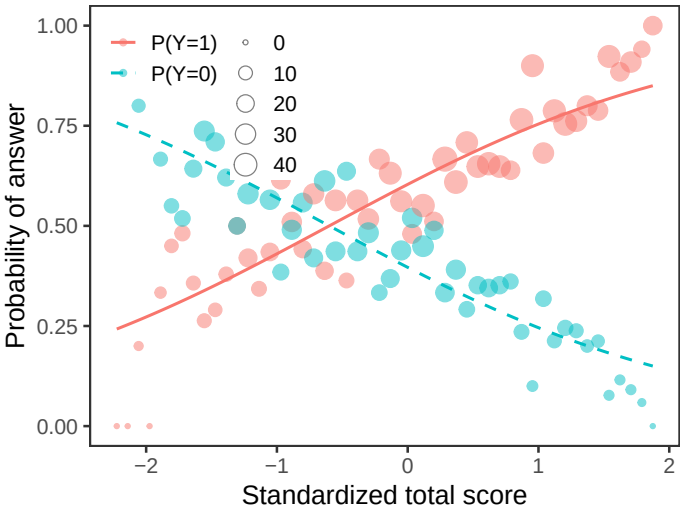
Multinomial plot for item b9.1



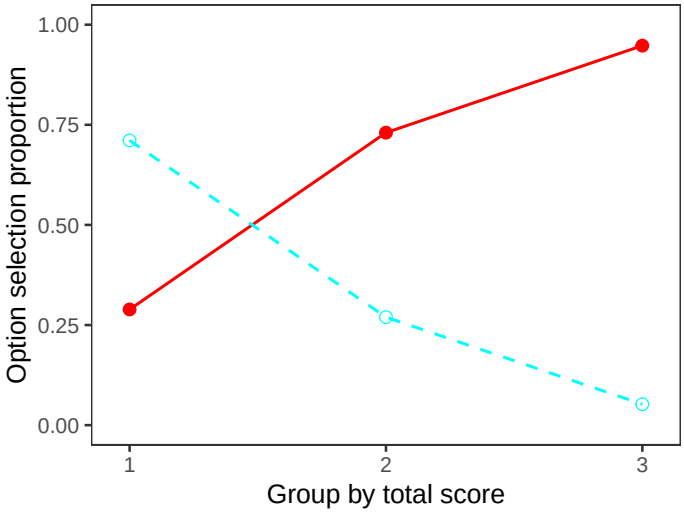
Distractor plot for item b9.2



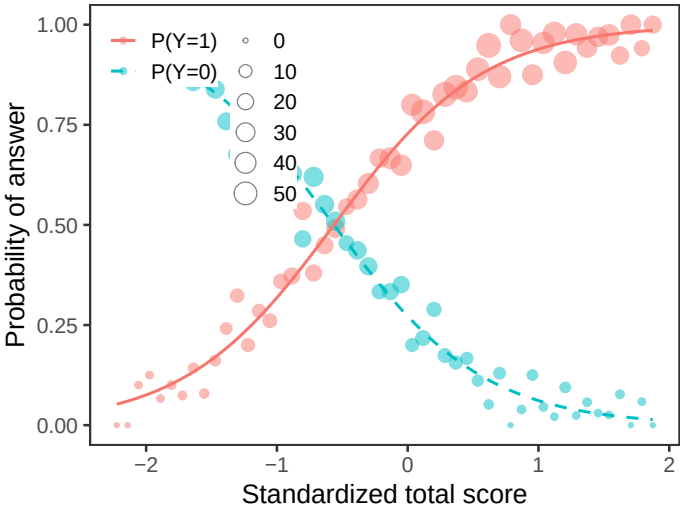
Multinomial plot for item b9.2



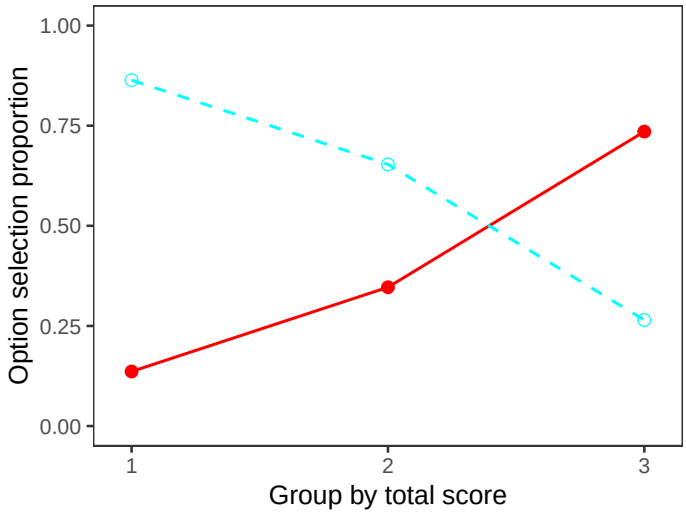
Distractor plot for item b10



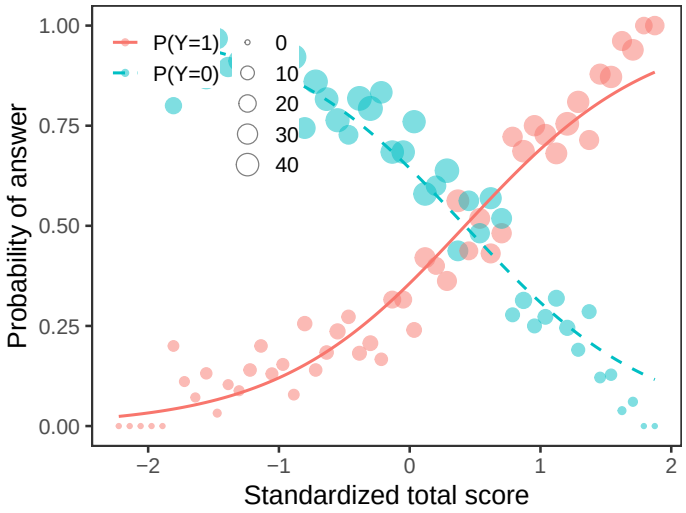
Multinomial plot for item b10



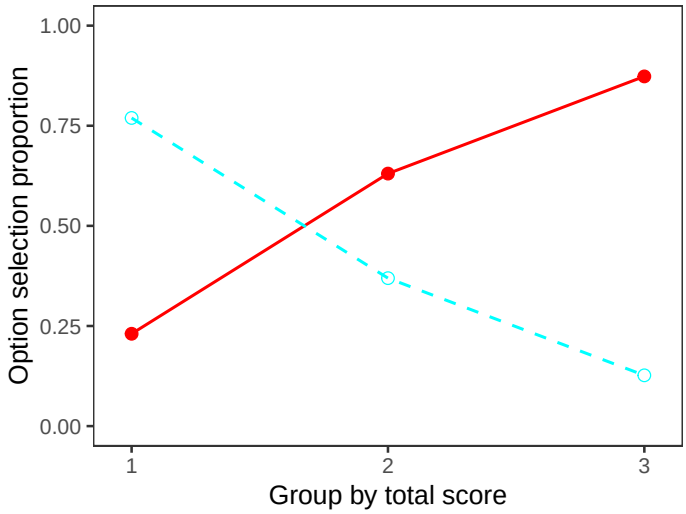
Distractor plot for item b11



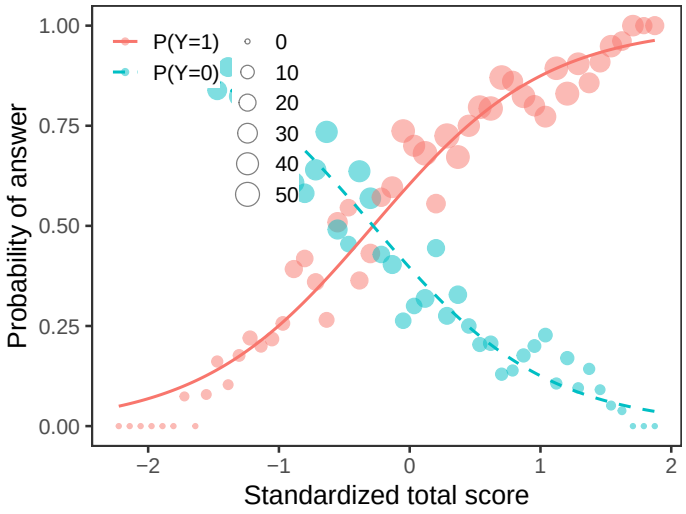
Multinomial plot for item b11



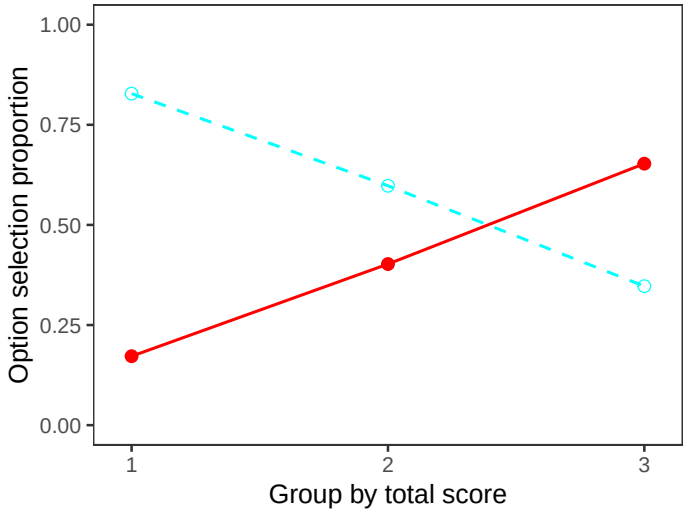
Distractor plot for item b12



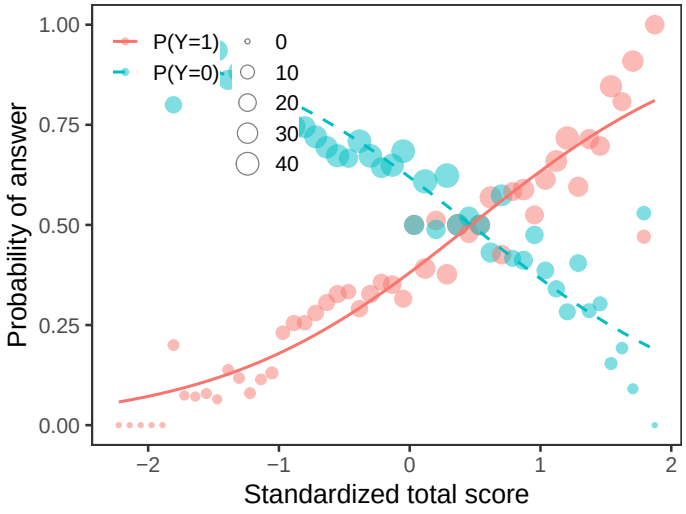
Multinomial plot for item b12



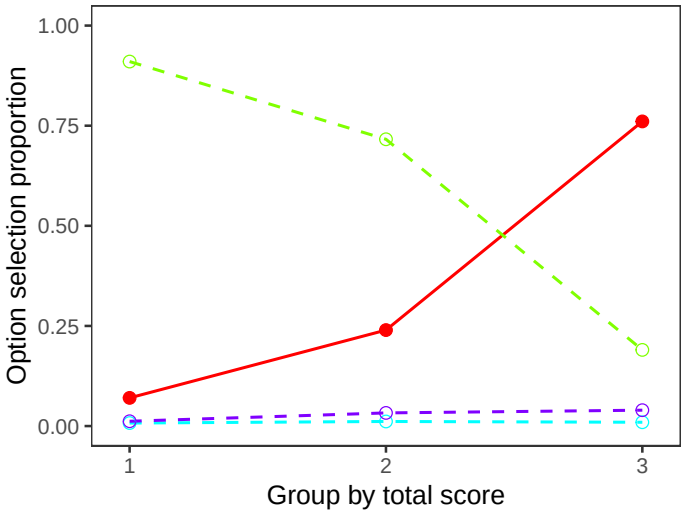
Distractor plot for item b13



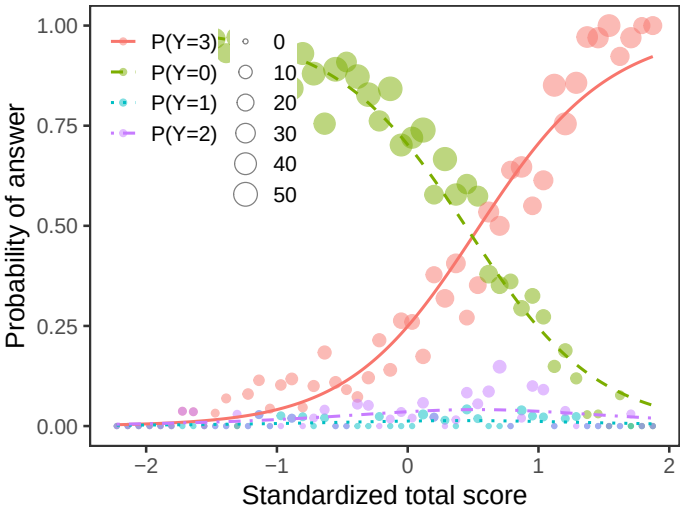
Multinomial plot for item b13



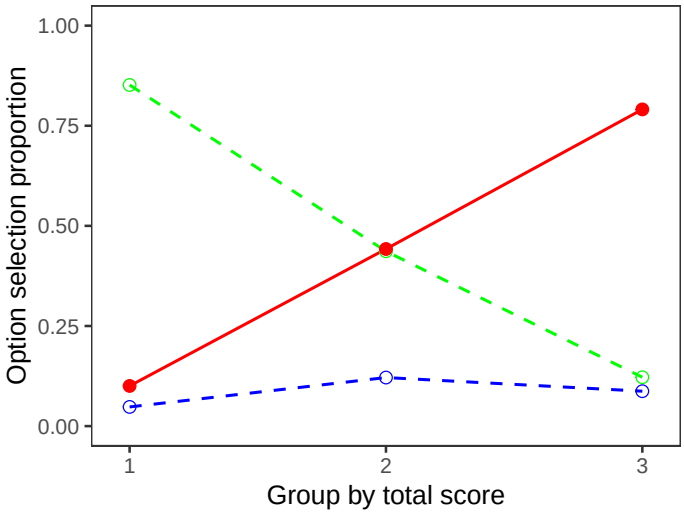
Distractor plot for item b14



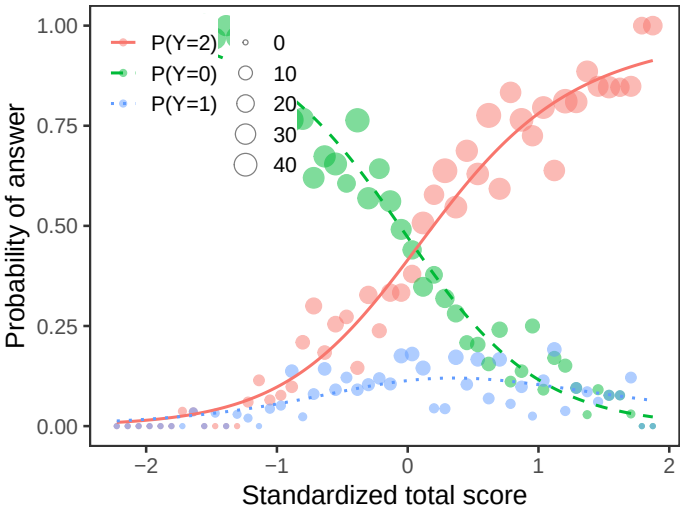
Multinomial plot for item b14



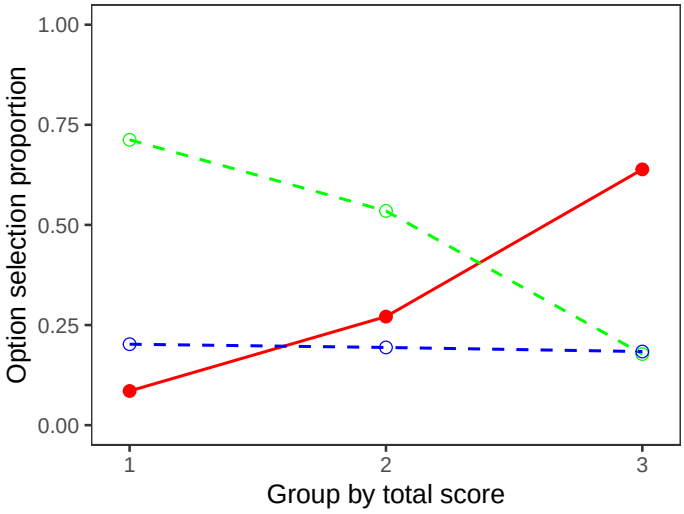
Distractor plot for item b15



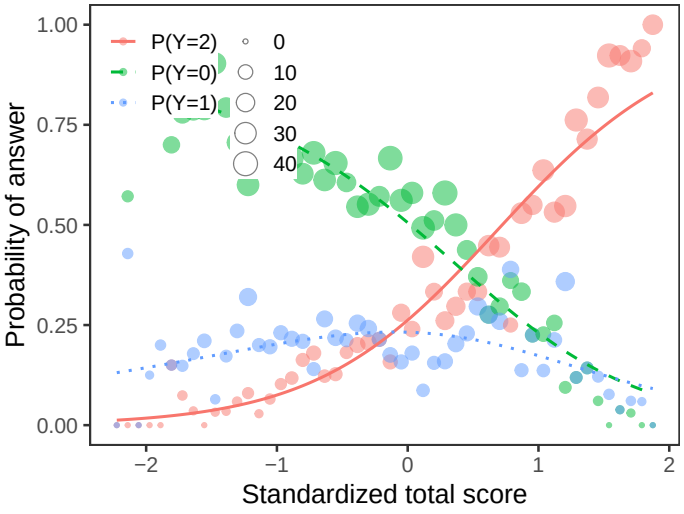
Multinomial plot for item b15



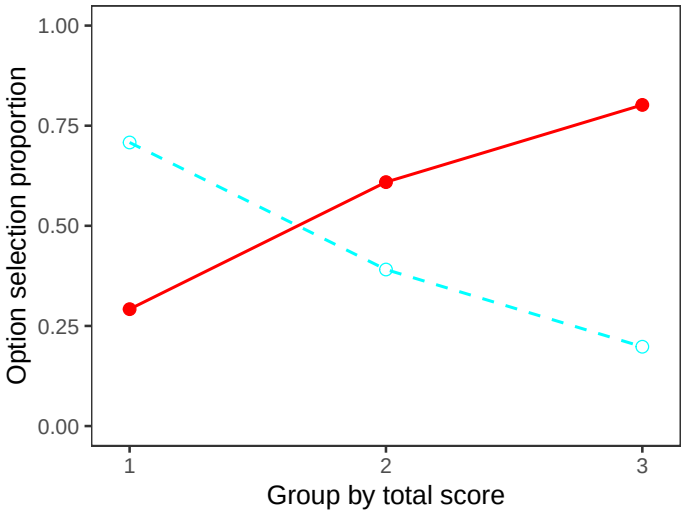
Distractor plot for item b16



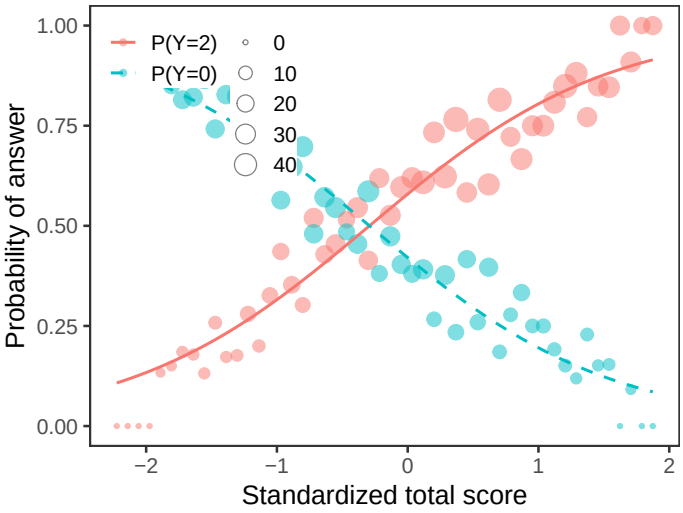
Multinomial plot for item b16



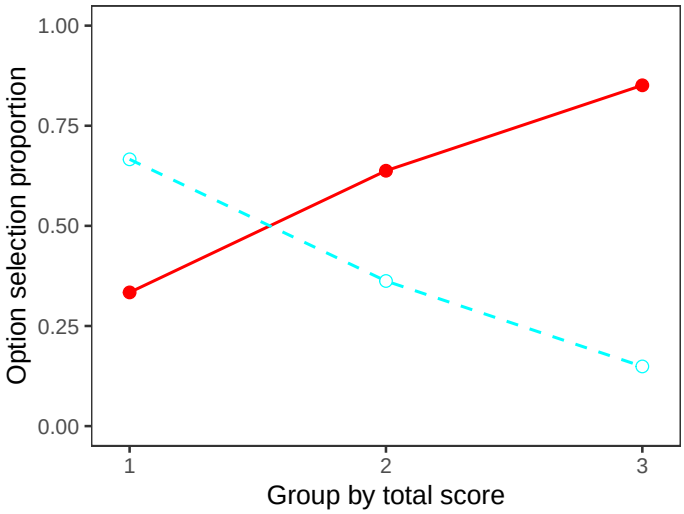
Distractor plot for item b17



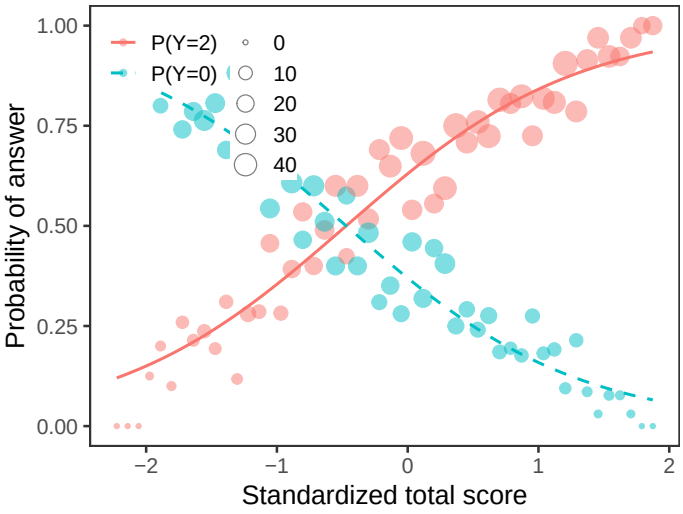
Multinomial plot for item b17



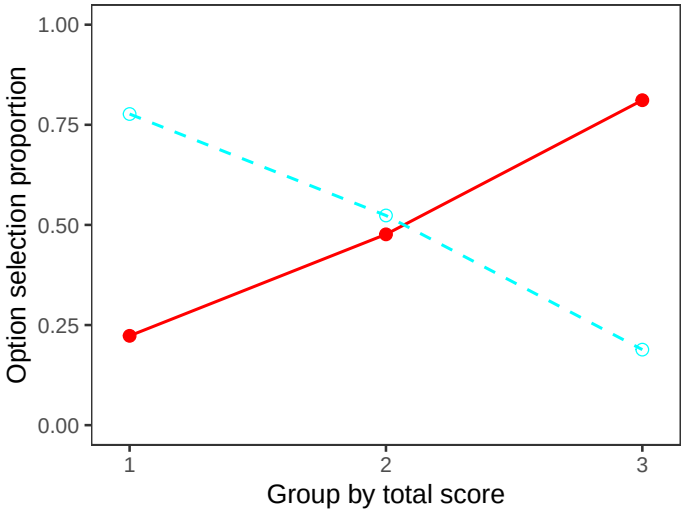
Distractor plot for item b18



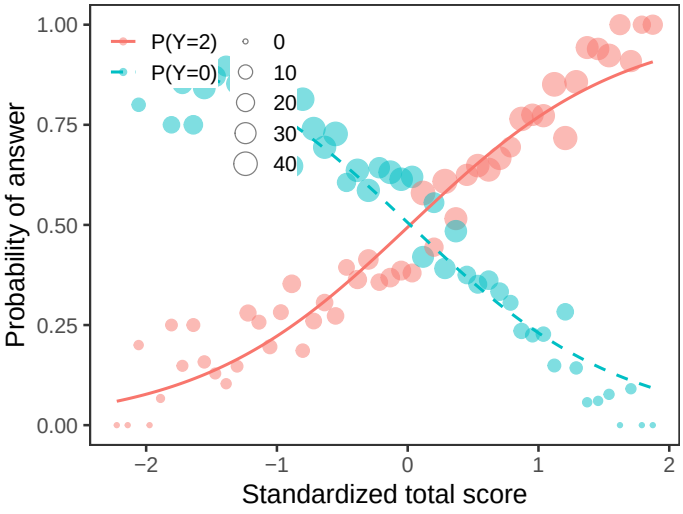
Multinomial plot for item b18



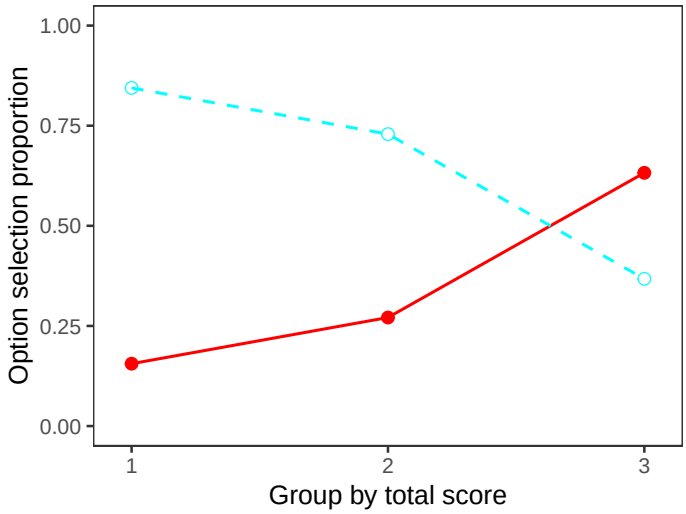
Distractor plot for item b19



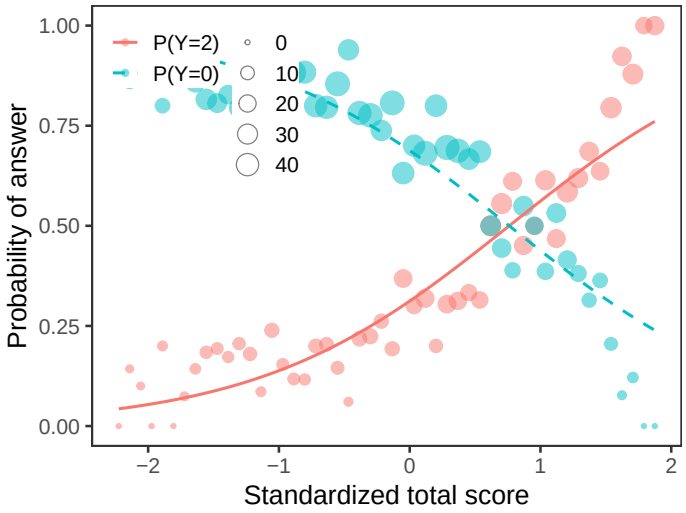
Multinomial plot for item b19



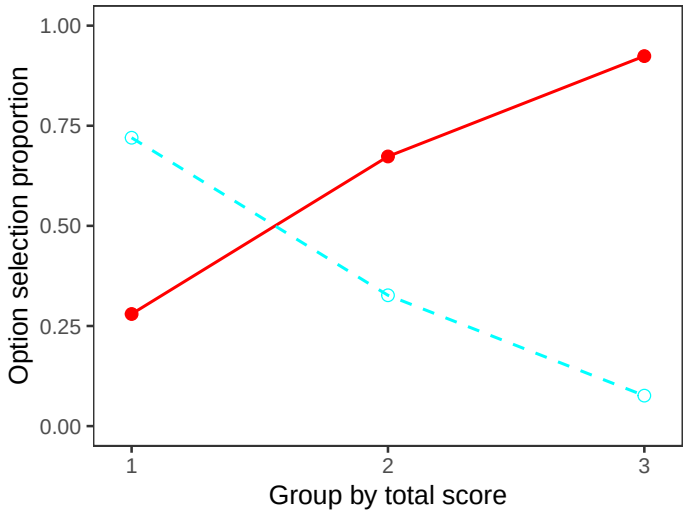
Distractor plot for item b20



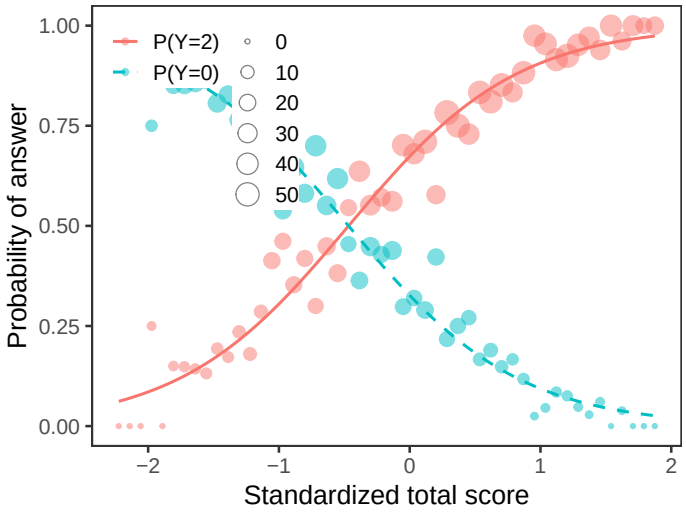
Multinomial plot for item b20



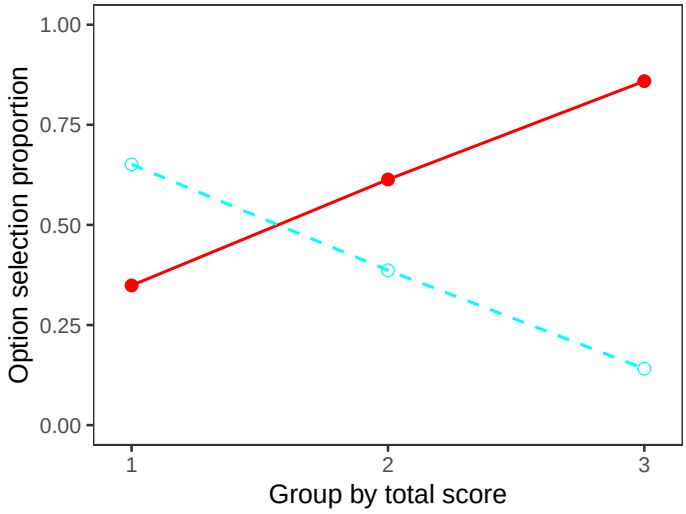
Distractor plot for item b21



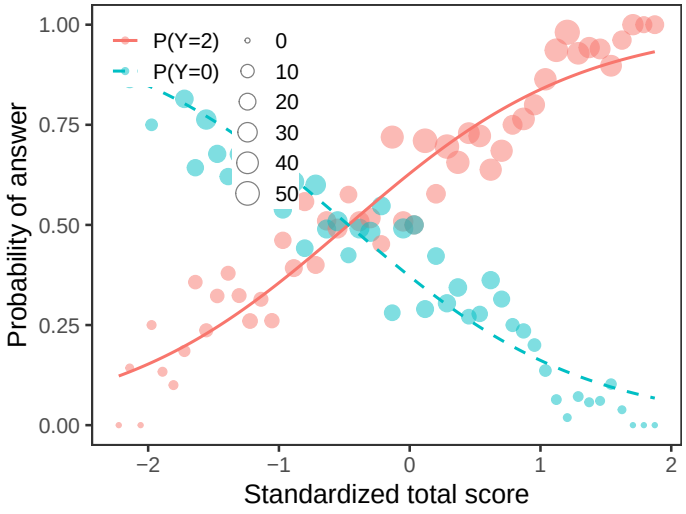
Multinomial plot for item b21



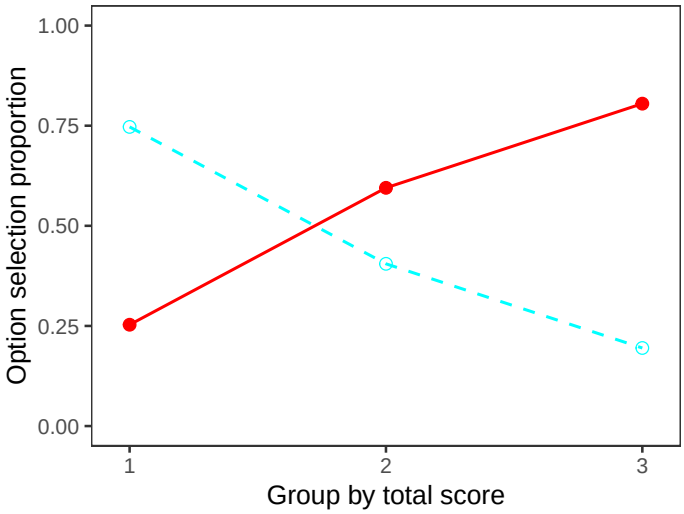
Distractor plot for item b22



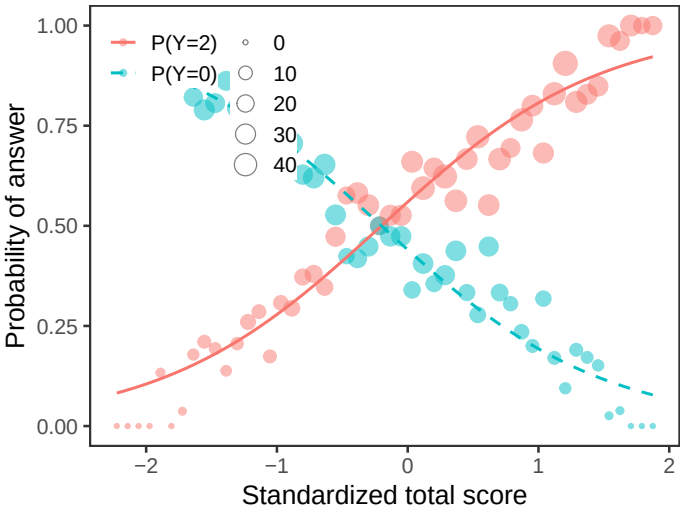
Multinomial plot for item b22



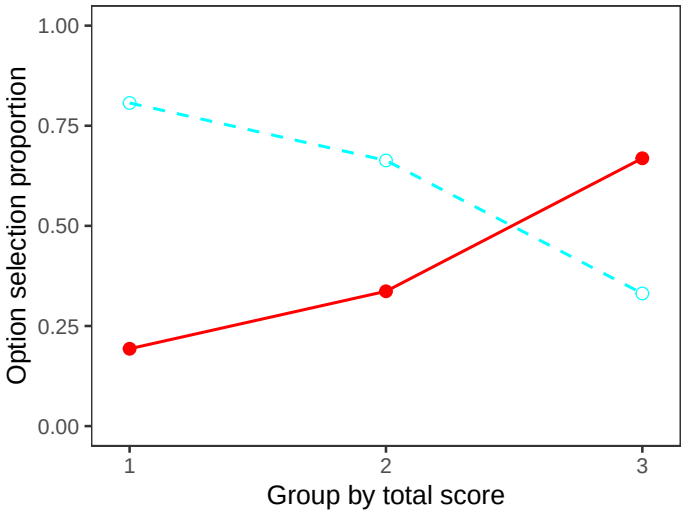
Distractor plot for item b23



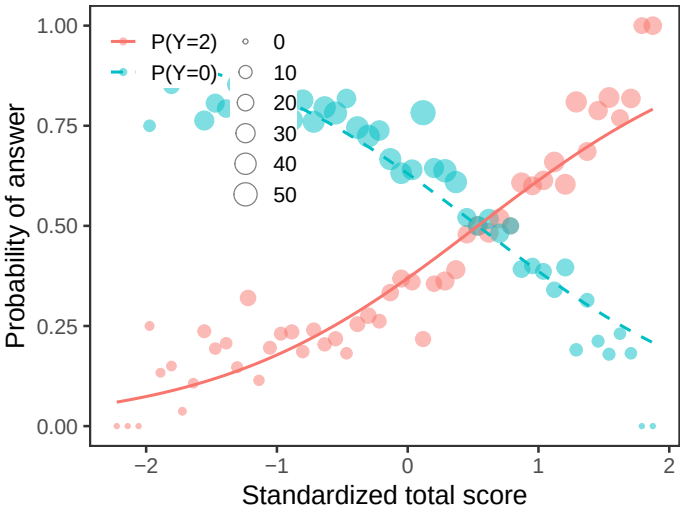
Multinomial plot for item b23



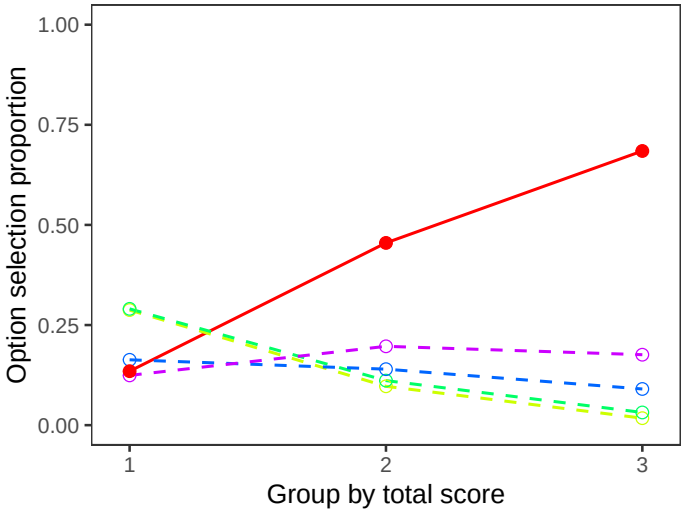
Distractor plot for item b24



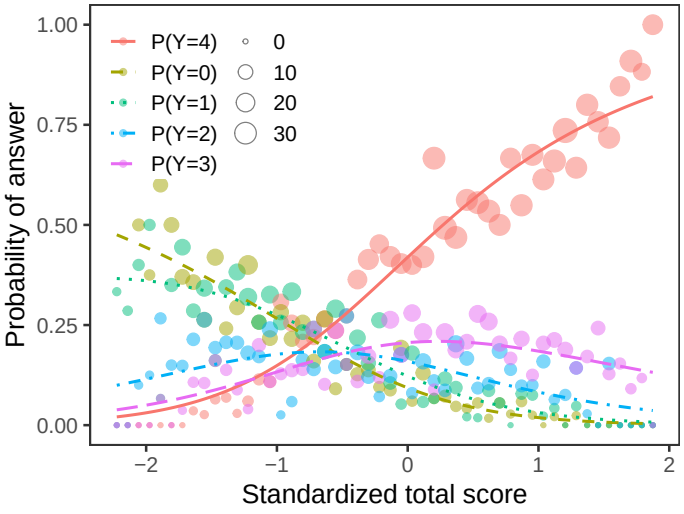
Multinomial plot for item b24



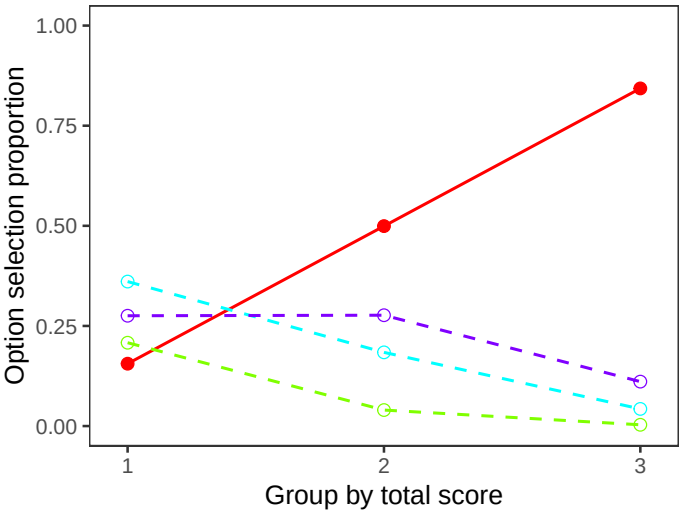
Distractor plot for item b25



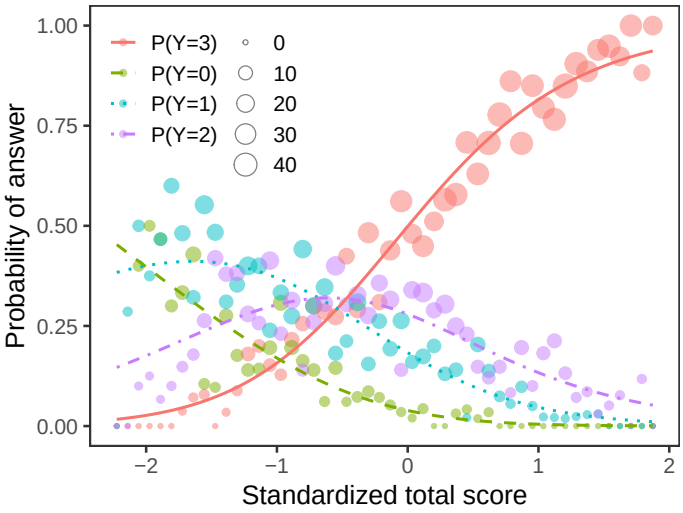
Multinomial plot for item b25



Distractor plot for item b26



Multinomial plot for item b26

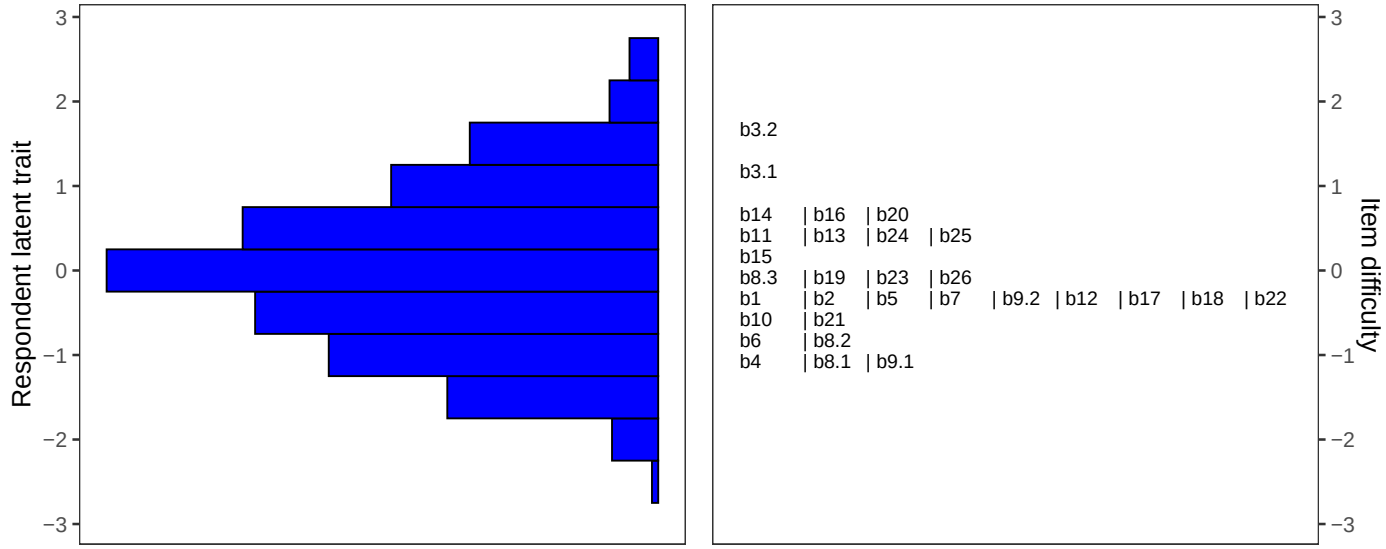


IRT models

Item Response Theory (IRT) models are mixed-effect regression models in which the respondent's ability θ is assumed to be a latent and is estimated together with item parameters.

Wright (item-person) map using 1PL IRT model

A Wright map, also called an item-person map, is a graphical tool to display person estimates and item parameters. The person side (left) represents a histogram of estimated knowledge of the respondents. The item side (right) displays estimates of the difficulty of particular items.



Equation

All subsequent analyses are based on the selected 1PL IRT model:

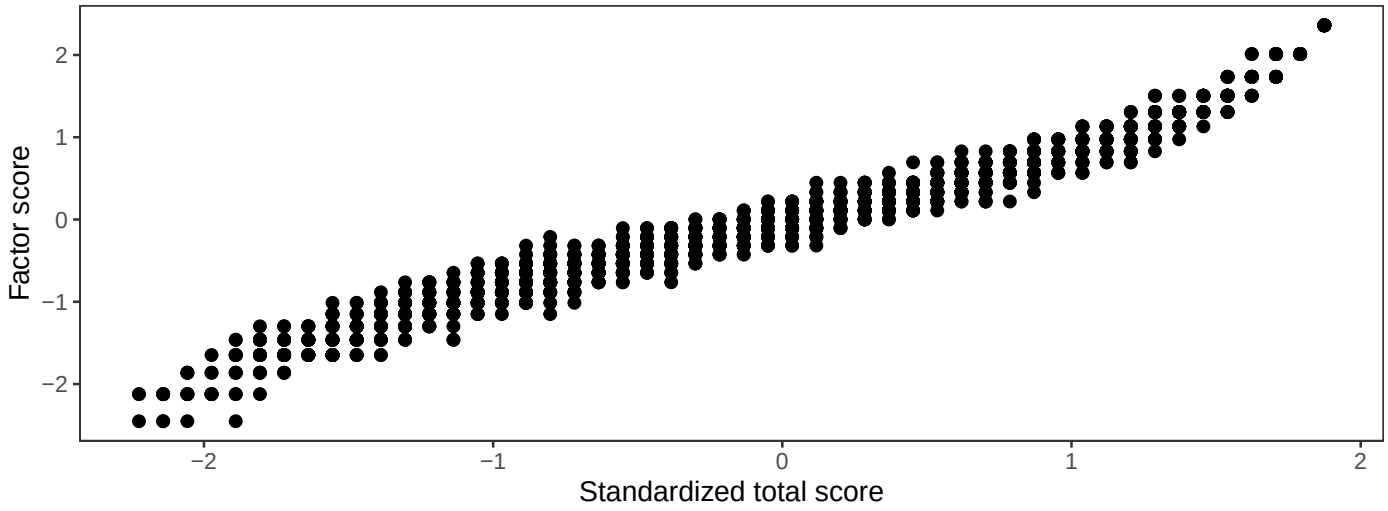
$$\mathrm{P}(Y_{pi} = 1|\theta_p) = \pi_{pi} = \frac{e^{a(\theta_p - b_i)}}{1 + e^{a(\theta_p - b_i)}}$$

Model parameters are estimated using a marginal maximum likelihood method. Ability θ is assumed to follow standard normal distribution.

Ability estimates

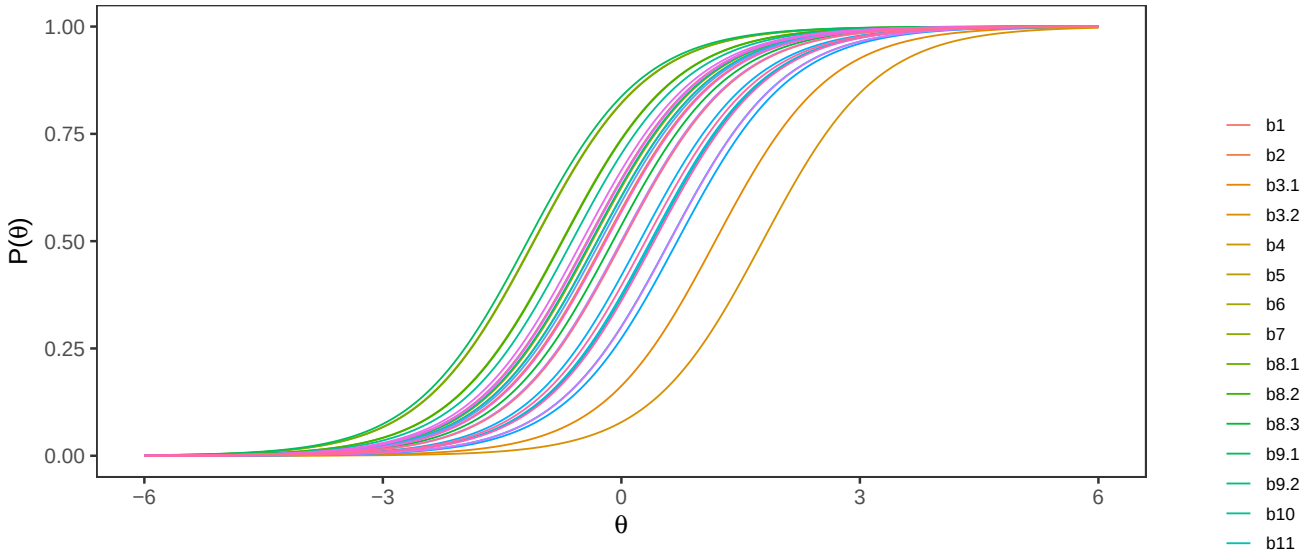
Ability is estimated using an EAP algorithm and a 1PL IRT model. The relationship between ability estimates (factor scores, F-scores) and standardized total test scores (Z-scores) can be seen on the scatter plot below. A table with ability estimates for all respondents can be downloaded from the application.

	Min	Max	Mean	Median	SD	Skewness	Kurtosis
Total Scores	1.00	50.00	27.59	28.00	11.95	-0.09	2.08
Z-Scores	-2.22	1.87	-0.00	0.03	1.00	-0.09	2.08
F-scores	-2.45	2.36	-0.00	0.00	0.95	0.11	2.67

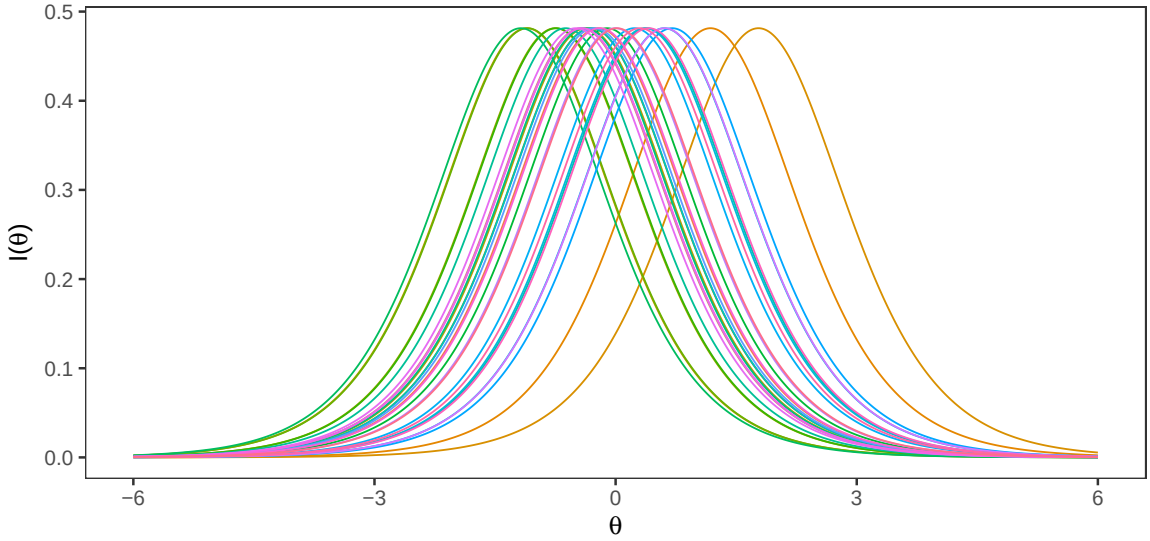


Item characteristic and information curves

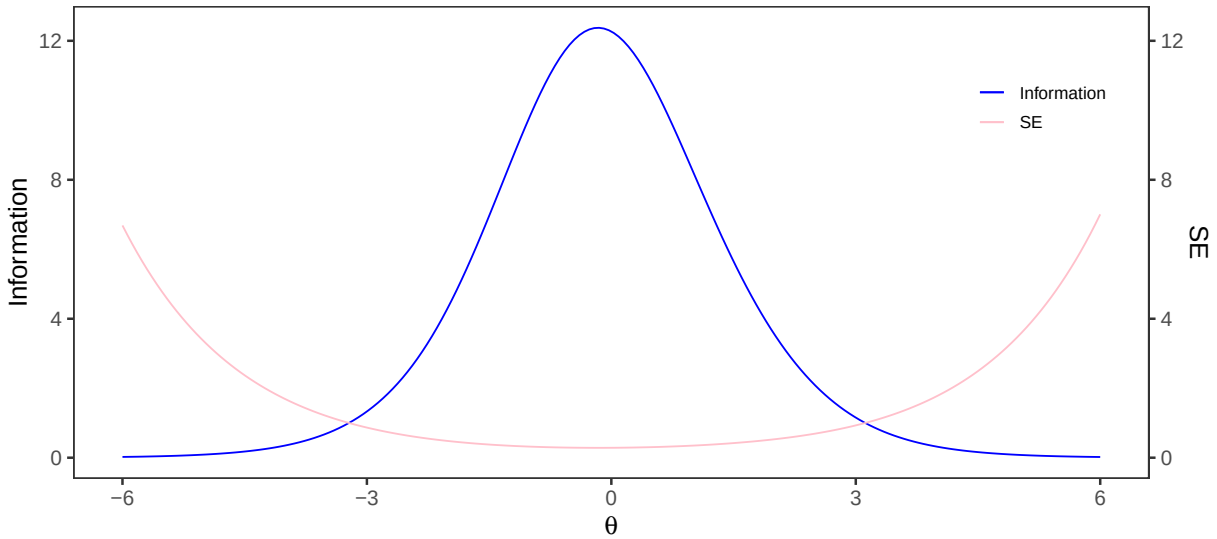
Item characteristic curves



Item information curves



Test information function



Parameter estimates and item fit

Estimates of parameters are completed by their standard errors (SE) and by signed Chi-squared statistics S-X2 (see Orlando and Thissen, 2000). P-values lower than 0.05 indicate suspicious items not fitting the selected IRT model. S-X2 is computed only when no missing data are present.

	a	SE(a)	b	SE(b)	c	SE(c)	d	SE(d)	S-X2	df	p-value
b1	1.39	0.03	-0.43	0.05	0.00		1.00		15.49	25	0.93
b2	1.39	0.03	-0.21	0.04	0.00		1.00		89.72	25	0.00
b3.1	1.39	0.03	1.18	0.05	0.00		1.00		50.81	25	0.00
b3.2	1.39	0.03	1.78	0.07	0.00		1.00		70.12	22	0.00
b4	1.39	0.03	-1.10	0.05	0.00		1.00		20.79	25	0.70
b5	1.39	0.03	-0.37	0.04	0.00		1.00		33.35	25	0.12
b6	1.39	0.03	-0.75	0.05	0.00		1.00		50.96	26	0.00
b7	1.39	0.03	-0.31	0.04	0.00		1.00		43.63	25	0.01
b8.1	1.39	0.03	-1.10	0.05	0.00		1.00		49.70	25	0.00
b8.2	1.39	0.03	-0.74	0.05	0.00		1.00		75.72	26	0.00
b8.3	1.39	0.03	-0.11	0.04	0.00		1.00		81.80	25	0.00
b9.1	1.39	0.03	-1.18	0.05	0.00		1.00		51.79	25	0.00
b9.2	1.39	0.03	-0.38	0.04	0.00		1.00		168.75	25	0.00
b10	1.39	0.03	-0.62	0.05	0.00		1.00		70.50	25	0.00
b11	1.39	0.03	0.39	0.05	0.00		1.00		23.20	25	0.57
b12	1.39	0.03	-0.30	0.04	0.00		1.00		37.33	25	0.05
b13	1.39	0.03	0.37	0.04	0.00		1.00		99.80	25	0.00
b14	1.39	0.03	0.61	0.05	0.00		1.00		41.86	25	0.02
b15	1.39	0.03	0.23	0.04	0.00		1.00		65.77	25	0.00
b16	1.39	0.03	0.70	0.05	0.00		1.00		33.36	25	0.12
b17	1.39	0.03	-0.26	0.04	0.00		1.00		71.53	25	0.00
b18	1.39	0.03	-0.42	0.04	0.00		1.00		63.28	25	0.00
b19	1.39	0.03	-0.00	0.04	0.00		1.00		65.10	25	0.00
b20	1.39	0.03	0.61	0.05	0.00		1.00		130.21	25	0.00
b21	1.39	0.03	-0.49	0.05	0.00		1.00		23.95	25	0.52
b22	1.39	0.03	-0.42	0.04	0.00		1.00		93.98	25	0.00
b23	1.39	0.03	-0.20	0.04	0.00		1.00		39.51	25	0.03
b24	1.39	0.03	0.41	0.05	0.00		1.00		143.75	25	0.00
b25	1.39	0.03	0.30	0.04	0.00		1.00		75.31	25	0.00
b26	1.39	0.03	0.02	0.04	0.00		1.00		20.68	25	0.71

DIF/Fairness analysis

Delta plot method

A delta plot compares the proportions of correct answers per item in the two groups. It displays non-linear transformation of these proportions using quantiles of standard normal distributions (so called delta scores) of each item for the two groups in a scatterplot called diagonal plot or delta plot. An item is under suspicion of DIF if the delta point departs considerably from the major axis.

Summary table

Item purification was not applied. The detection threshold is 1.5.

No DIF item was detected.

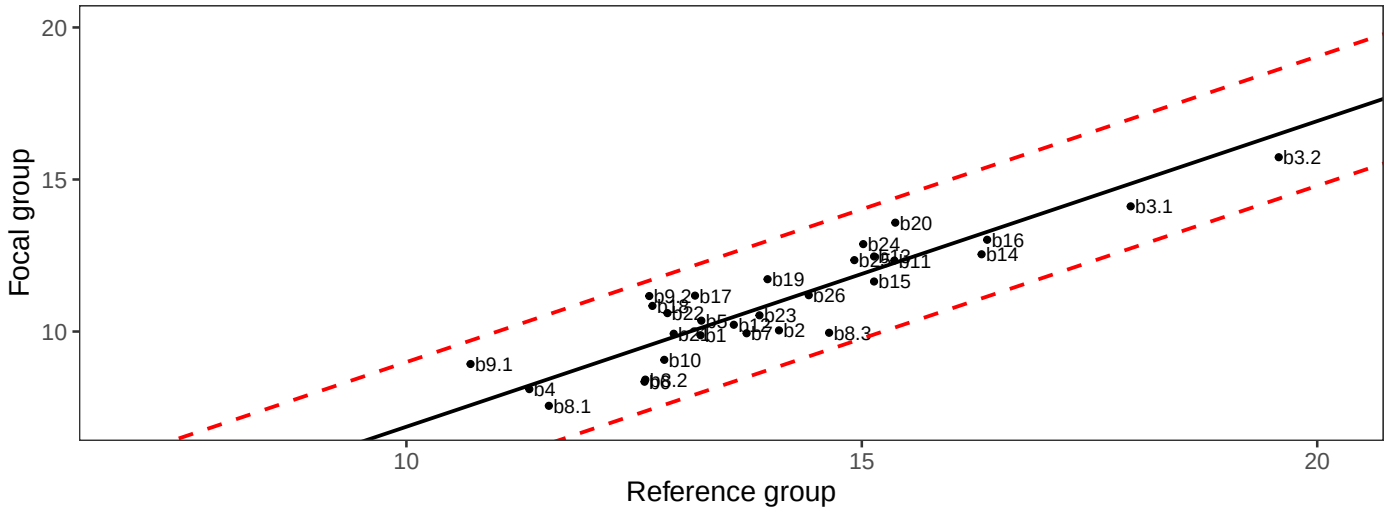
	Prop. Ref	Prop. Foc	Delta Ref	Delta Foc	Dist.
b1	0.477	0.782	13.229	9.879	0.167
b2	0.393	0.771	14.090	10.038	0.665
b3.1	0.108	0.390	17.952	14.117	0.525
b3.2	0.050	0.247	19.578	15.732	0.539
b4	0.660	0.889	11.350	8.106	0.084
b5	0.476	0.746	13.238	10.358	-0.165
b6	0.538	0.878	12.614	8.349	0.810
b7	0.427	0.778	13.735	9.943	0.480
b8.1	0.640	0.913	11.564	7.557	0.623
b8.2	0.538	0.874	12.623	8.419	0.767
b8.3	0.341	0.776	14.641	9.959	1.111
b9.1	0.717	0.845	10.704	8.932	-0.956
b9.2	0.533	0.677	12.667	11.167	-1.141
b10	0.517	0.837	12.831	9.070	0.455
b11	0.278	0.566	15.357	12.335	-0.057
b12	0.441	0.756	13.595	10.223	0.183
b13	0.297	0.553	15.134	12.468	-0.309
b14	0.204	0.546	16.315	12.540	0.477
b15	0.297	0.633	15.134	11.645	0.271
b16	0.199	0.498	16.376	13.018	0.184
b17	0.483	0.675	13.169	11.181	-0.795
b18	0.530	0.705	12.701	10.843	-0.888
b19	0.405	0.625	13.965	11.721	-0.611
b20	0.277	0.442	15.368	13.580	-0.928
b21	0.506	0.779	12.935	9.927	-0.076
b22	0.513	0.725	12.866	10.605	-0.603
b23	0.413	0.731	13.876	10.533	0.164
b24	0.307	0.512	15.016	12.875	-0.680
b25	0.316	0.565	14.918	12.347	-0.377

Continued on next page

	Prop. Ref	Prop. Foc	Delta Ref	Delta Foc	Dist.
b26	0.362	0.674	14.418	11.194	0.082

Delta plot

Points represents delta scores. Black solid is a major axis. Red dashed lines stand for detection threshold.



Mantel-Haenszel test

Summary table

Item purification was not applied. No p-value adjustment for multiple comparisons was used. The detection threshold is 3.84.

Items detected as DIF: b6, b8.3, b9.1, b9.2, b16, b17, b18, b19, b20, b23.

	Stat.	P-value	α_{MH}	δ_{MH}
b1	1.005	0.316	0.874	0.317 A
b2	0.436	0.509	0.907	0.230 A
b3.1	2.737	0.098.	0.771	0.610 A
b3.2	0.008	0.931	0.963	0.089 A
b4	0.724	0.395	0.865	0.341 A
b5	0.230	0.631	0.936	0.156 A
b6	4.974	0.026*	0.706	0.818 A
b7	1.897	0.168	0.830	0.438 A
b8.1	2.668	0.102	0.757	0.655 A
b8.2	2.457	0.117	0.781	0.581 A
b8.3	17.992	0.000***	0.571	1.317 B
b9.1	11.016	0.001***	1.633	-1.152 B
b9.2	7.982	0.005**	1.403	-0.796 A
b10	0.023	0.880	0.968	0.076 A
b11	0.592	0.442	1.110	-0.245 A
b12	0.091	0.763	1.048	-0.111 A
b13	0.381	0.537	0.923	0.189 A
b14	2.377	0.123	0.814	0.484 A
b15	0.082	0.774	0.957	0.103 A
b16	4.447	0.035*	0.759	0.647 A
b17	6.698	0.010**	1.377	-0.752 A
b18	11.359	0.001***	1.527	-0.995 A
b19	6.503	0.011*	1.371	-0.742 A
b20	10.067	0.002**	1.500	-0.953 A
b21	0.781	0.377	1.135	-0.297 A
b22	1.131	0.288	1.148	-0.324 A
b23	12.261	0.000***	0.661	0.974 A
b24	0.752	0.386	1.117	-0.261 A
b25	0.000	0.992	1.006	-0.013 A
b26	0.333	0.564	1.079	-0.179 A

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Effect size (ETS Delta scale):

'A': negligible effect

'B': moderate effect

'C': large effect

Effect size codes: 0 'A' 1.0 'B' 1.5 'C' (for absolute values of δ_{MH})

Logistic regression method

Logistic regression allows for detection of uniform and non-uniform DIF by adding a group specific intercept (uniform DIF) and group specific interaction (non-uniform DIF) into the model and by testing for their significance.

Summary table

Both types of DIF were tested. Item purification was not applied. No p-value adjustment for multiple comparisons was used. The detection threshold is 5.99.

Items detected as DIF: b3.1, b8.1, b8.3, b9.1, b9.2, b15, b16, b17, b18, b19, b20, b23, b24.

	Stat.	P-value	R^2	ZT	JG
b1	1.021	0.600	0.000	A	A
b2	1.875	0.392	0.001	A	A
b3.1	7.116	0.029*	0.002	A	A
b3.2	1.943	0.379	0.001	A	A
b4	3.340	0.188	0.001	A	A
b5	0.266	0.875	0.000	A	A
b6	5.306	0.070.	0.002	A	A
b7	1.961	0.375	0.001	A	A
b8.1	7.553	0.023*	0.003	A	A
b8.2	2.827	0.243	0.001	A	A
b8.3	20.339	0.000***	0.006	A	A
b9.1	14.565	0.001***	0.006	A	A
b9.2	15.720	0.000***	0.007	A	A
b10	0.179	0.915	0.000	A	A
b11	5.783	0.055.	0.002	A	A
b12	0.150	0.928	0.000	A	A
b13	0.199	0.905	0.000	A	A
b14	2.734	0.255	0.001	A	A
b15	10.989	0.004**	0.004	A	A
b16	15.904	0.000***	0.006	A	A
b17	9.542	0.008**	0.004	A	A
b18	13.779	0.001***	0.006	A	A
b19	10.865	0.004**	0.004	A	A
b20	28.519	0.000***	0.012	A	A
b21	2.646	0.266	0.001	A	A
b22	3.782	0.151	0.002	A	A
b23	11.734	0.003**	0.005	A	A
b24	20.804	0.000***	0.009	A	A
b25	0.191	0.909	0.000	A	A
b26	4.065	0.131	0.001	A	A

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Effect size is based on Nagelkerke's R^2.

'A' means negligible, 'B' moderate and 'C' large effect size

The thresholds are:

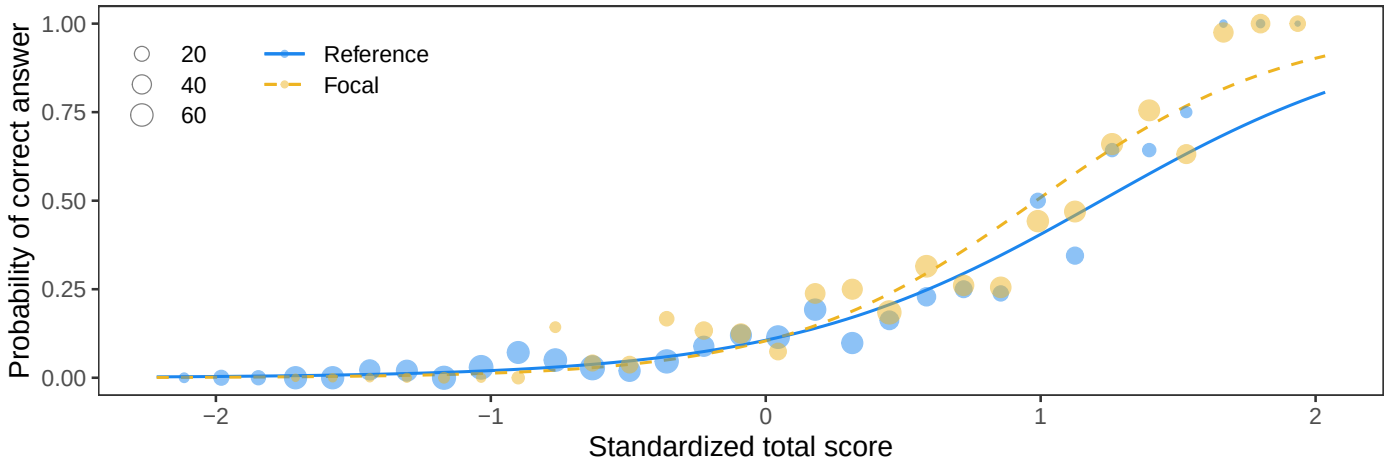
Zumbo & Thomas (ZT): 0 'A' 0.13 'B' 0.26 'C' 1

Jodoin & Gierl (JG): 0 'A' 0.035 'B' 0.07 'C' 1.

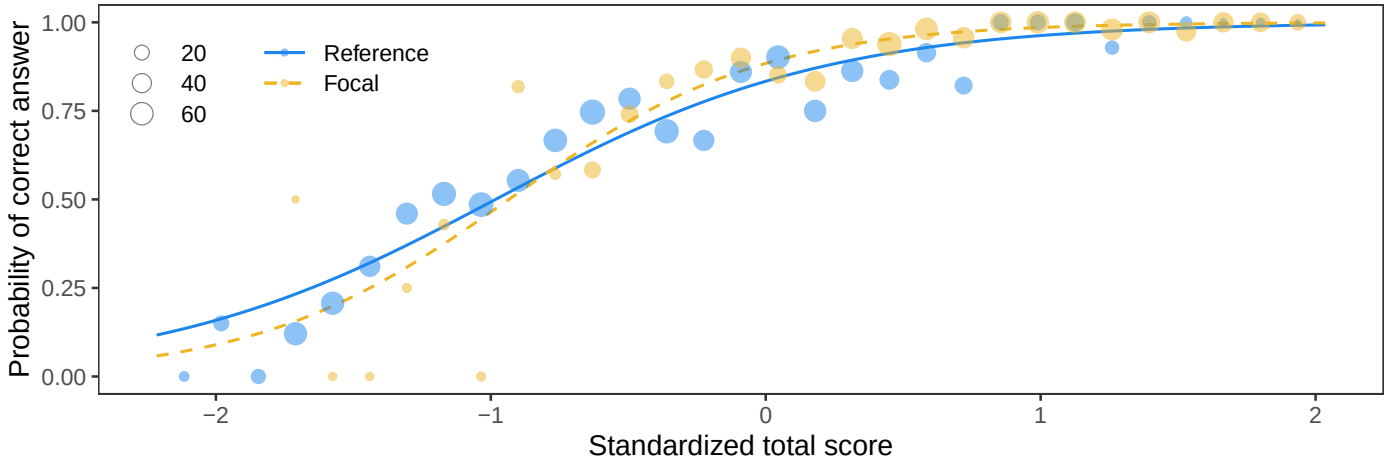
Characteristic curves of DIF items

Plots are based on DIF logistic procedure without any correction method. Points represent a proportion of a correct answer with respect to standardized total score. Their size is determined by count of the respondents who achieved a given level of standardized total score.

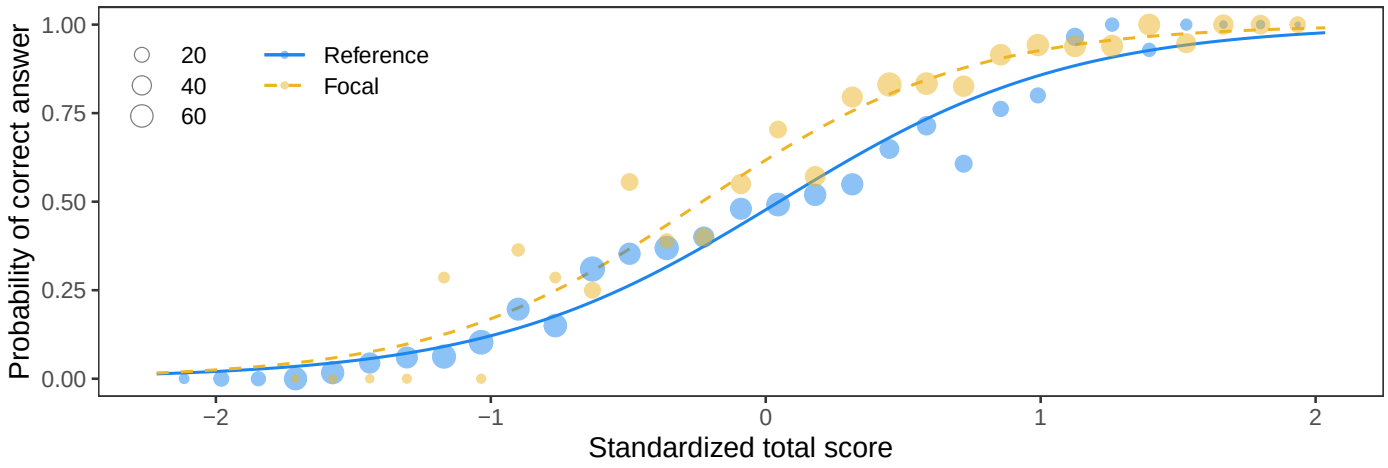
DIF logistic plot for b3.1



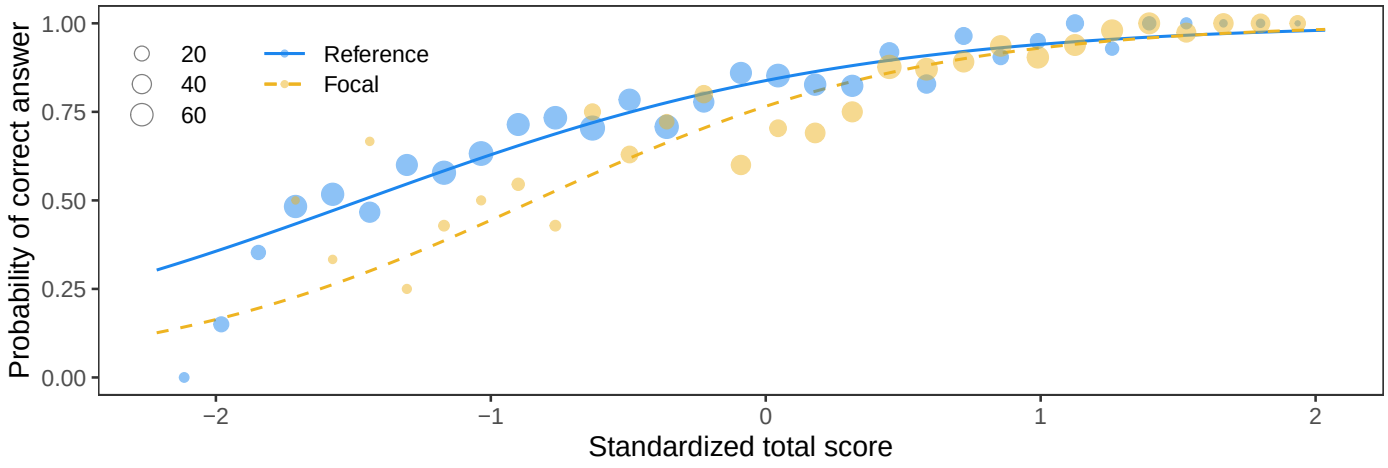
DIF logistic plot for b8.1



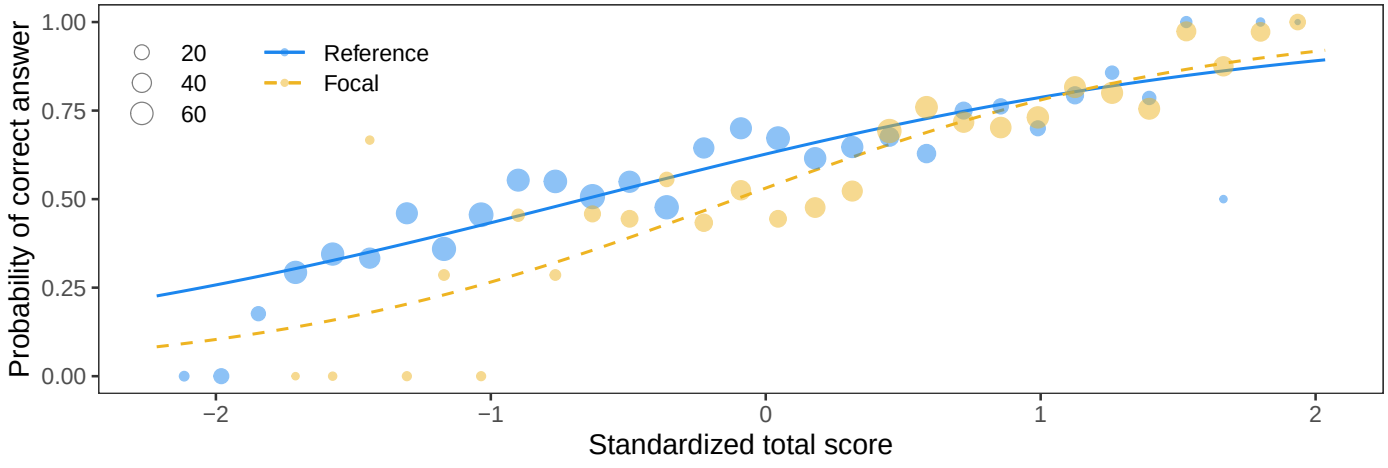
DIF logistic plot for b8.3



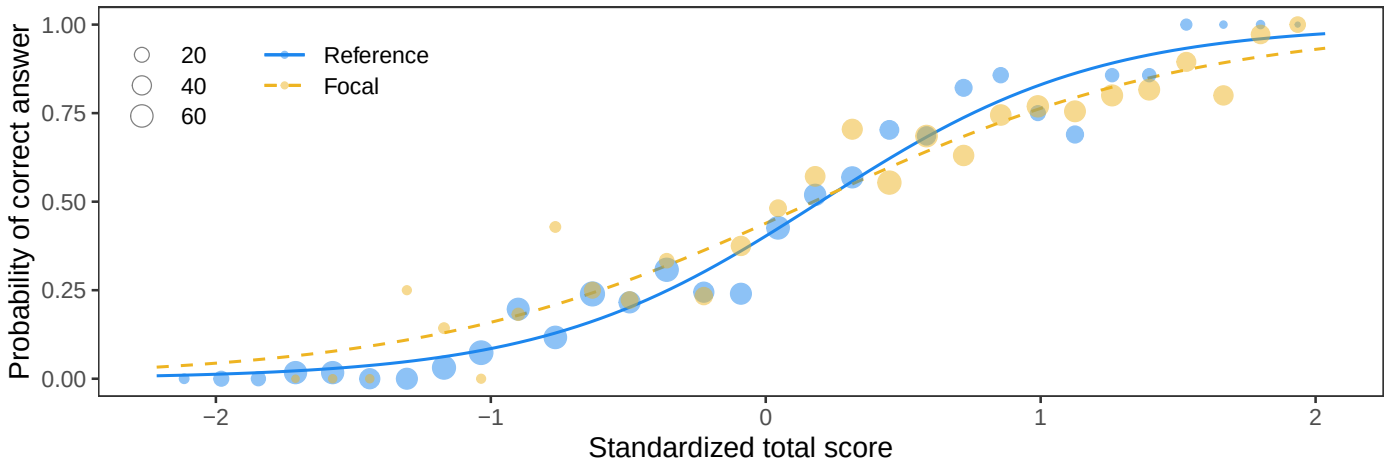
DIF logistic plot for b9.1



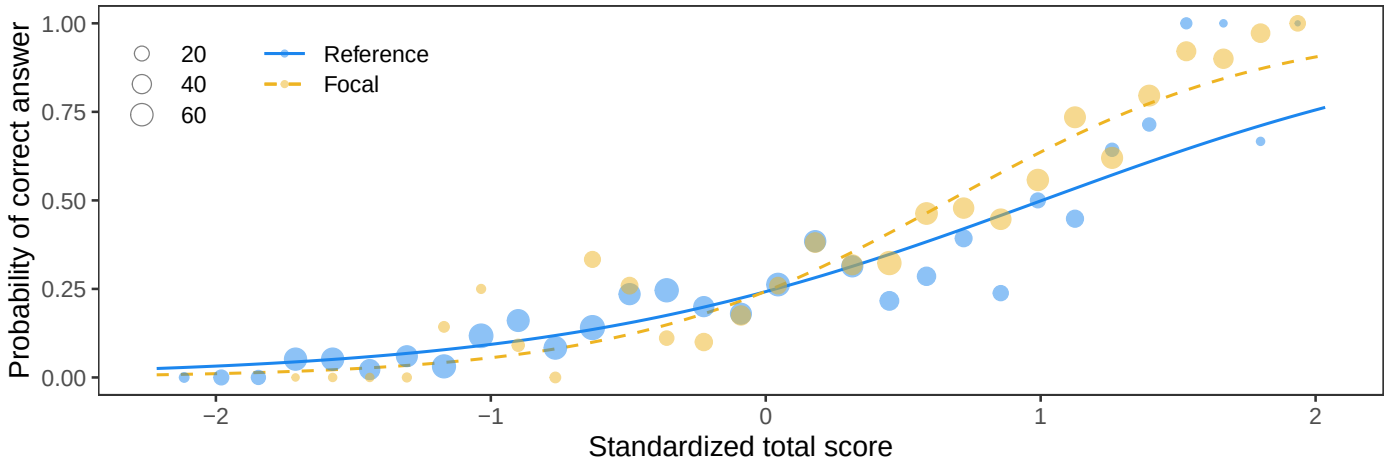
DIF logistic plot for b9.2



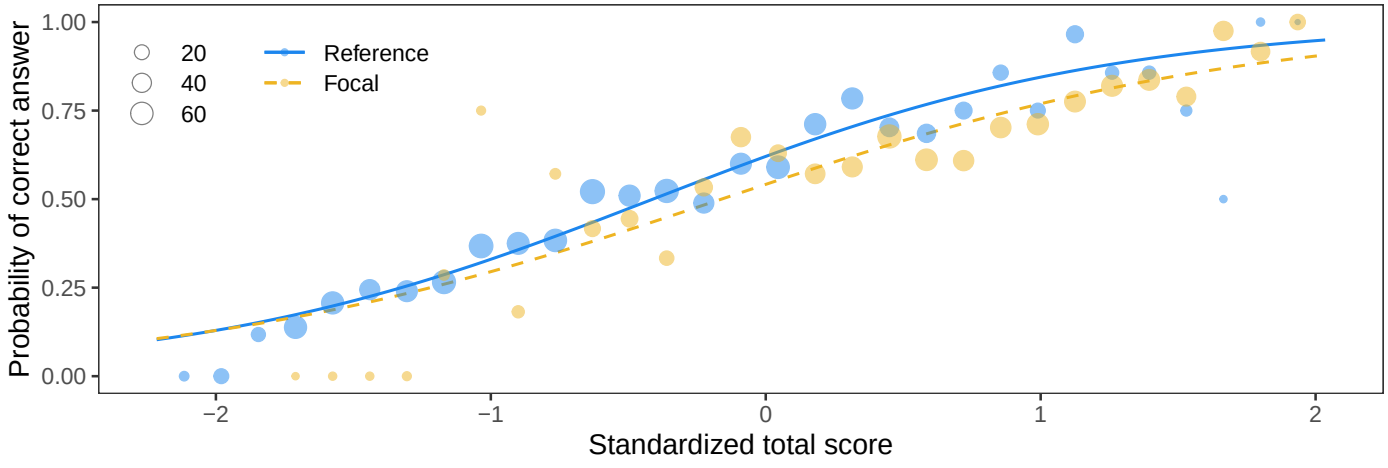
DIF logistic plot for b15



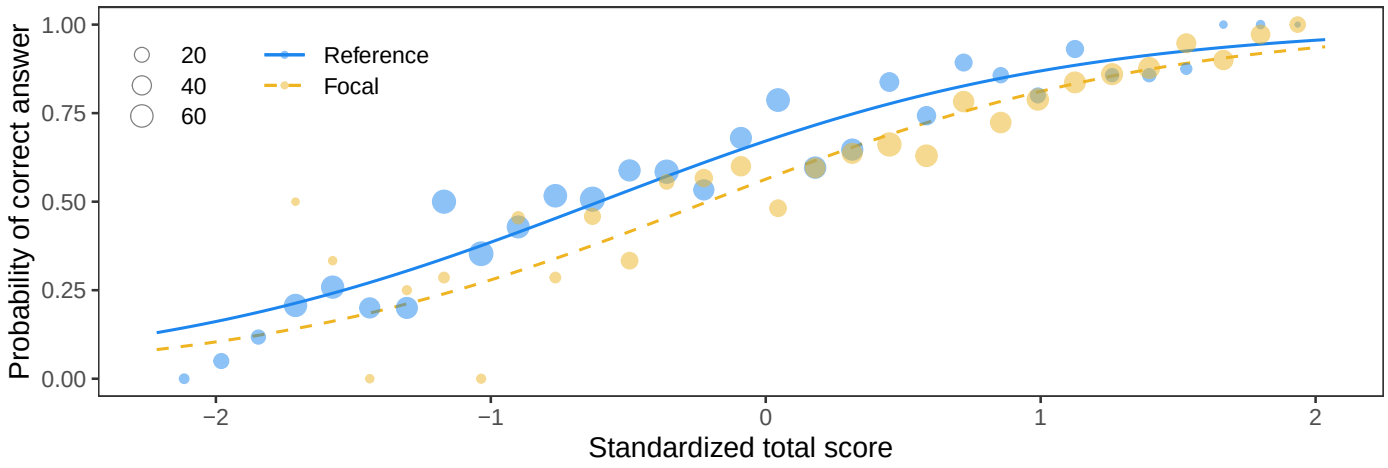
DIF logistic plot for b16



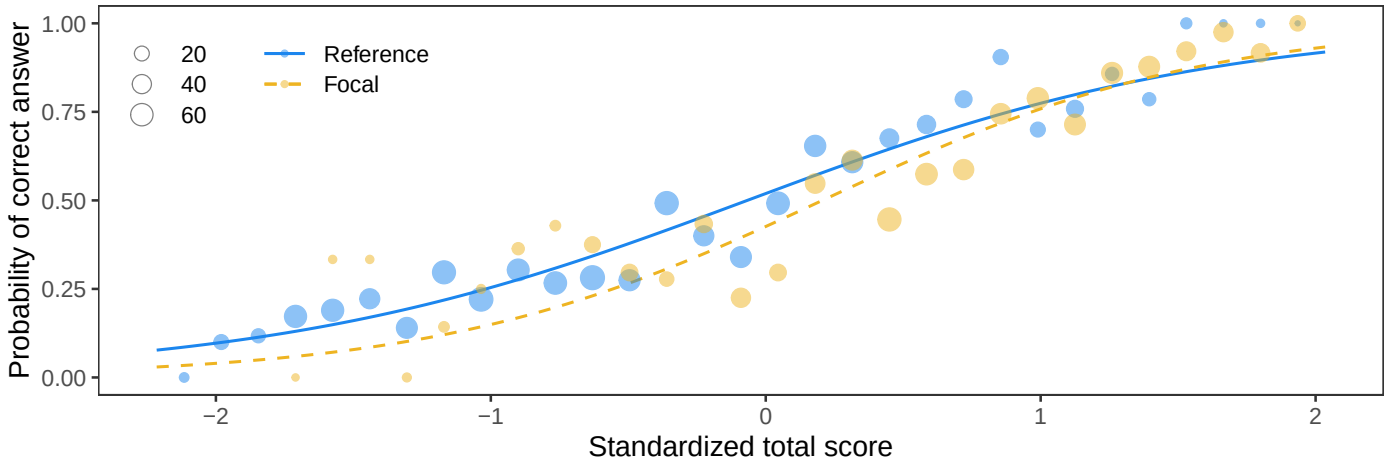
DIF logistic plot for b17



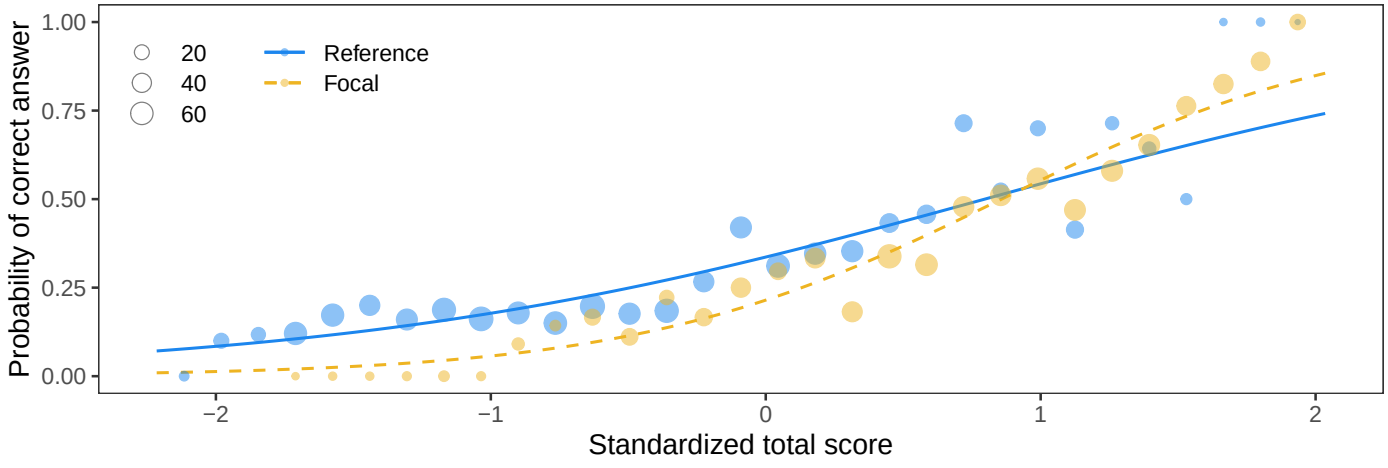
DIF logistic plot for b18



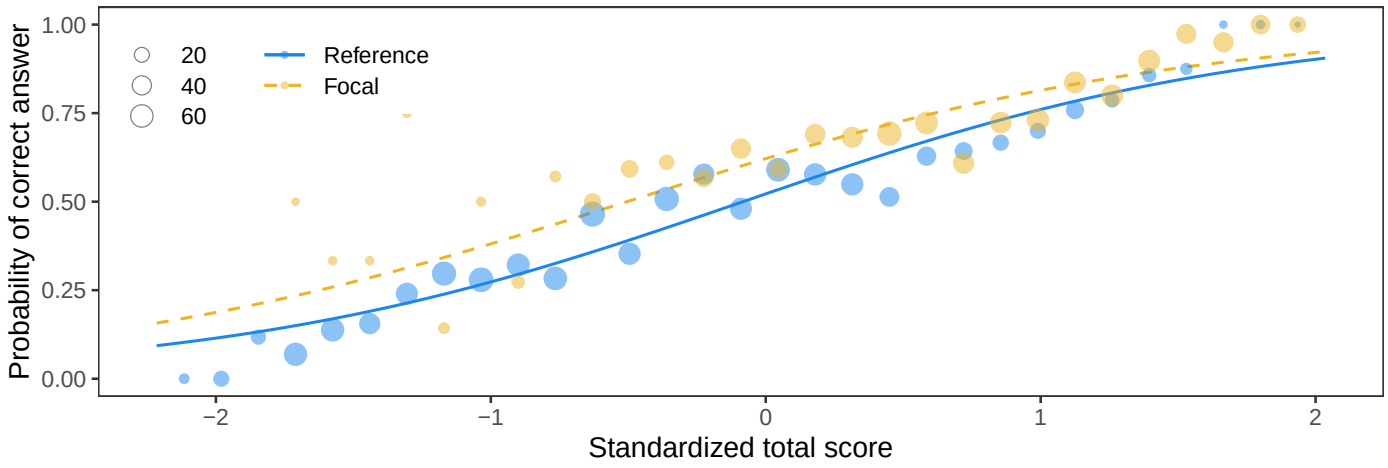
DIF logistic plot for b19



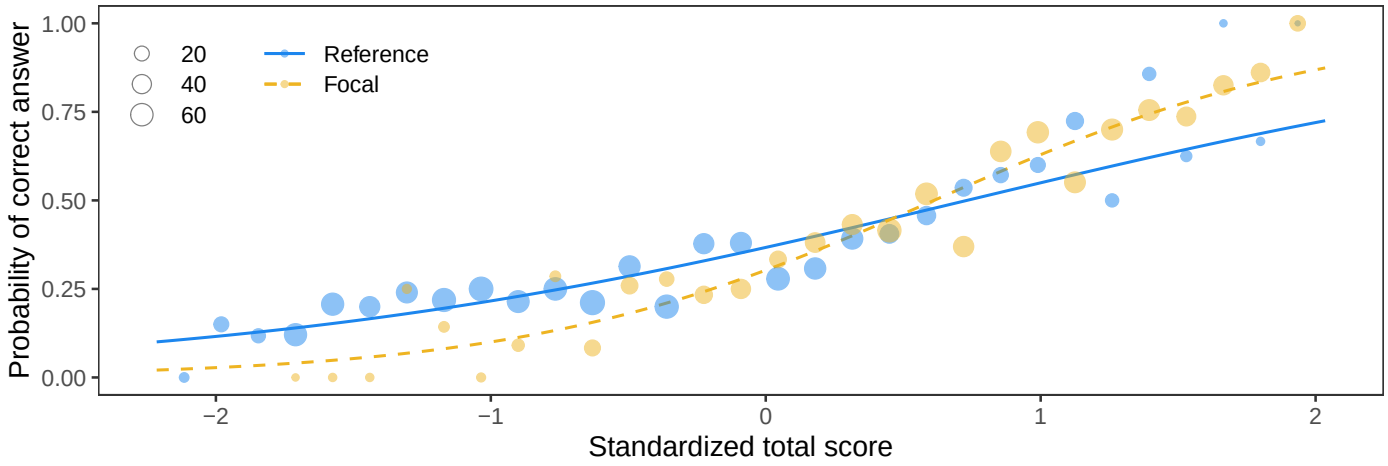
DIF logistic plot for b20



DIF logistic plot for b23



DIF logistic plot for b24



DDF detection using multinomial regression method

Differential Distractor Functioning (DDF) occurs when respondents from different groups but with the same knowledge have a different probability of selecting at least one distractor choice. DDF is examined here by a multinomial log-linear regression model with Z-score and group membership as covariates.

Summary table

Both types were of DDF tested. Item purification was not applied. No p-value adjustment for multiple comparisons was used.

Items detected as DDF: b3.1, b7, b8.1, b8.3, b9.1, b9.2, b15, b16, b17, b18, b19, b20, b23, b24.

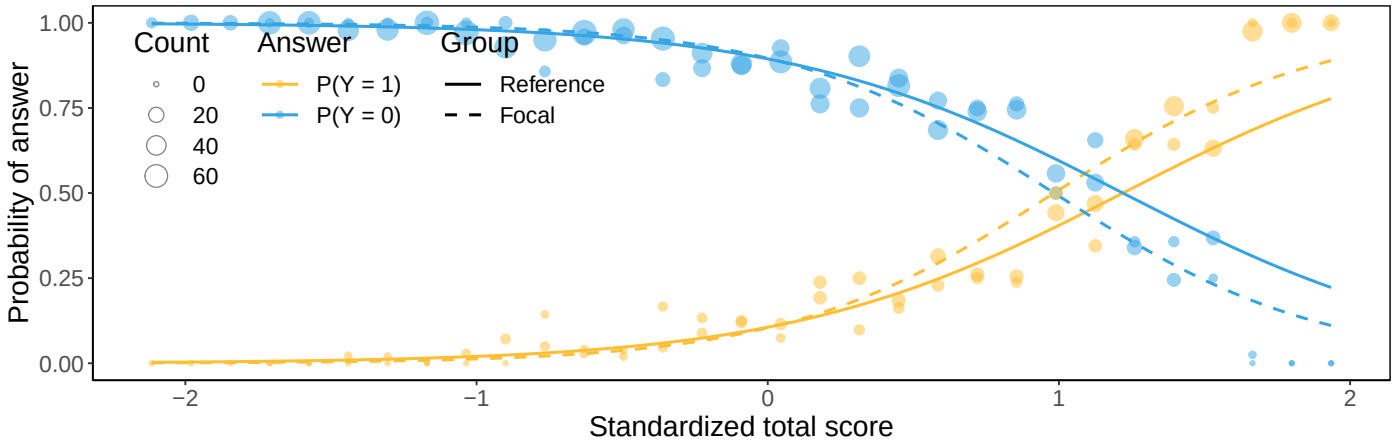
	Stat.	P-value
b1	1.021	0.600
b2	1.875	0.392
b3.1	7.116	0.029*
b3.2	1.943	0.379
b4	4.234	0.375
b5	6.022	0.198
b6	5.306	0.070.
b7	13.180	0.010*
b8.1	7.553	0.023*
b8.2	2.827	0.243
b8.3	20.339	0.000***
b9.1	14.565	0.001***
b9.2	15.720	0.000***
b10	0.179	0.915
b11	5.783	0.055.
b12	0.150	0.928
b13	0.199	0.905
b14	11.194	0.083.
b15	29.101	0.000***
b16	17.490	0.002**
b17	9.543	0.008**
b18	13.779	0.001**
b19	10.865	0.004**
b20	28.519	0.000***
b21	2.646	0.266
b22	3.782	0.151
b23	11.734	0.003**
b24	20.804	0.000***
b25	3.299	0.914
b26	10.199	0.117

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

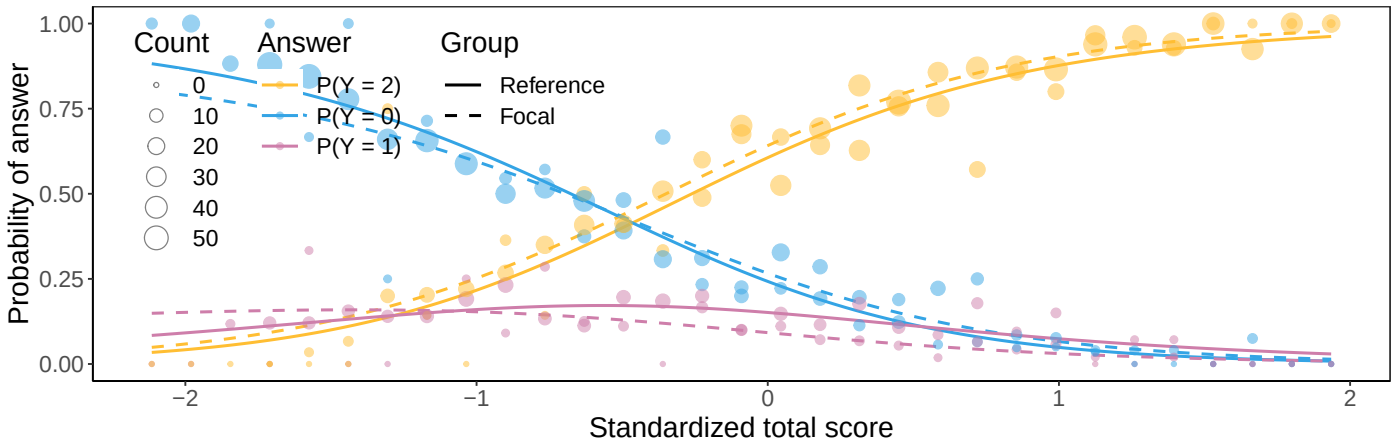
Characteristic curves of DDF items

Points represent a proportion of a selected answer with respect to standardized total score. Their size is determined by count of the respondents who achieved a given level of standardized total score and who selected a given option with respect to the group membership.

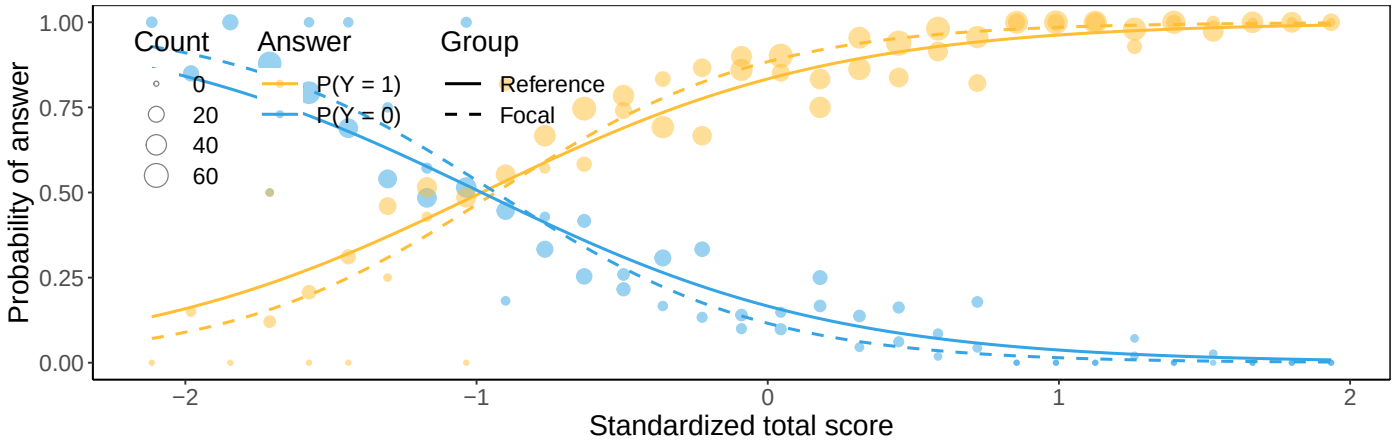
DDF multinomial plot for item b3.1



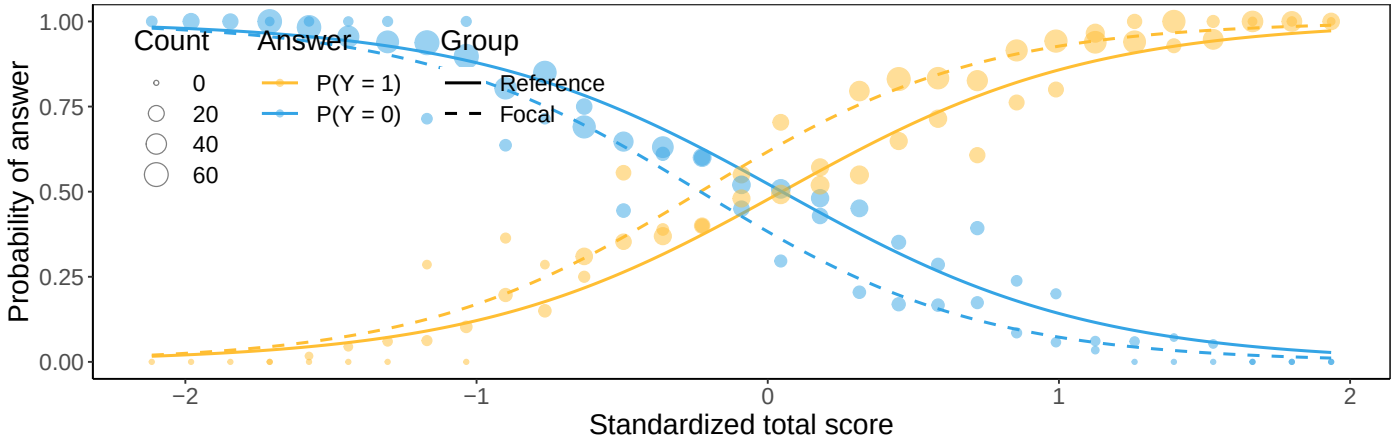
DDF multinomial plot for item b7



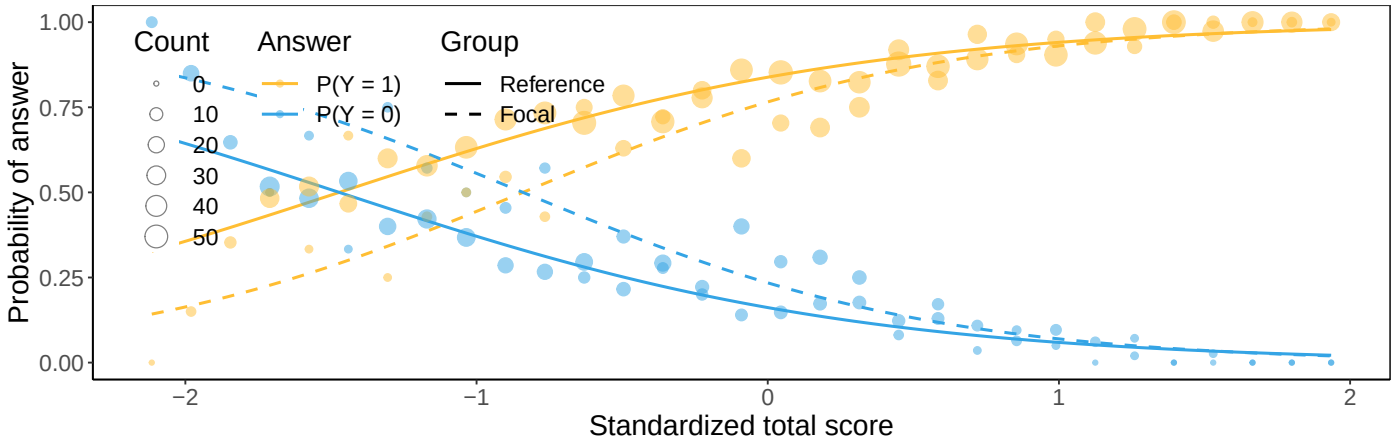
DDF multinomial plot for item b8.1



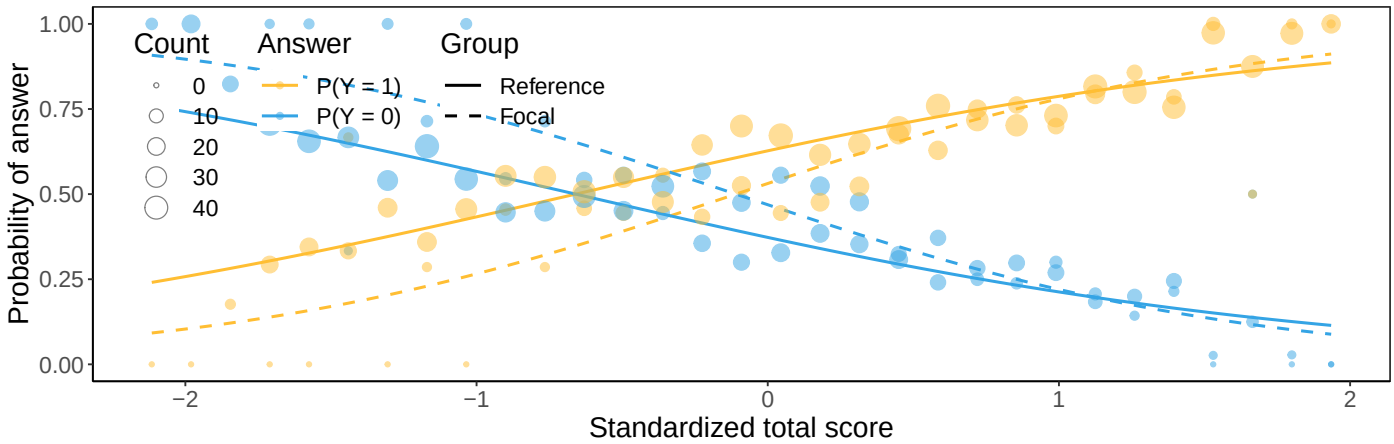
DDF multinomial plot for item b8.3



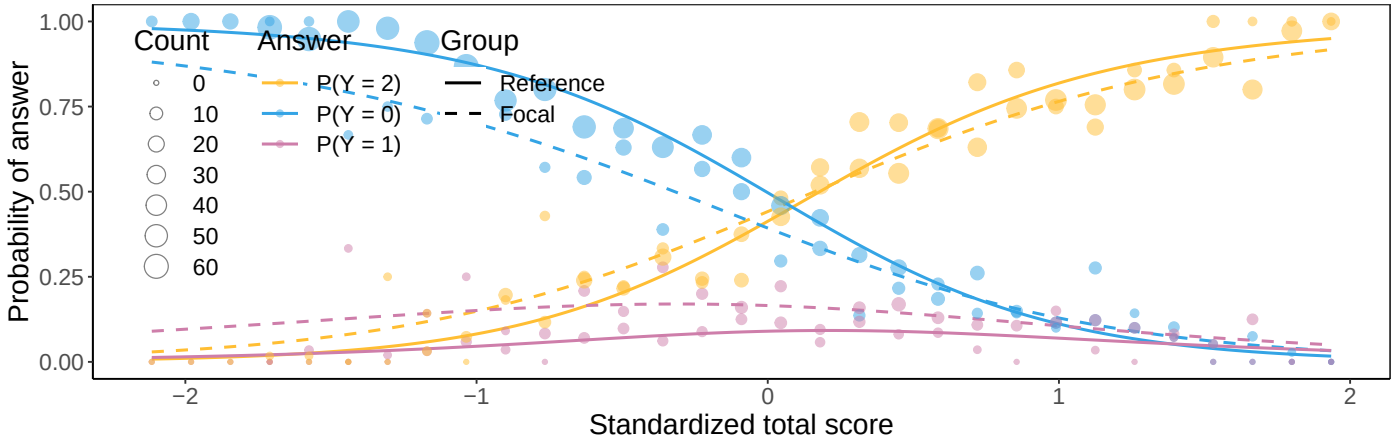
DDF multinomial plot for item b9.1



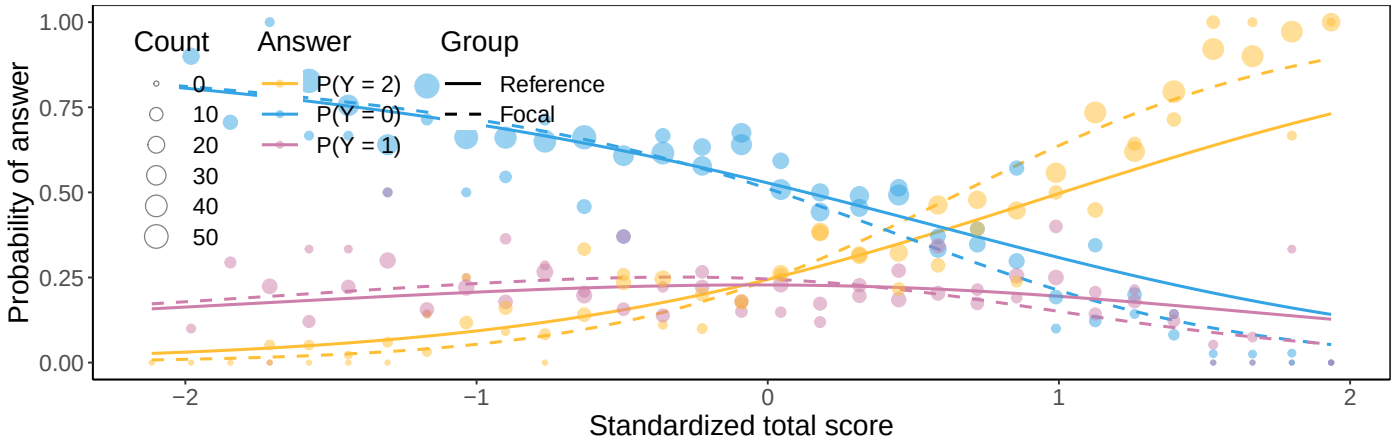
DDF multinomial plot for item b9.2



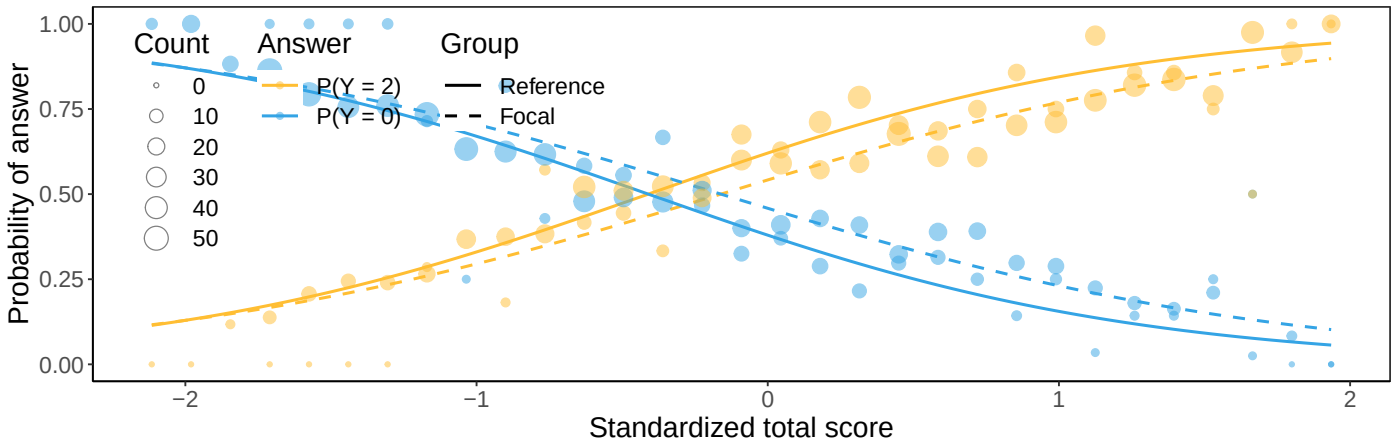
DDF multinomial plot for item b15



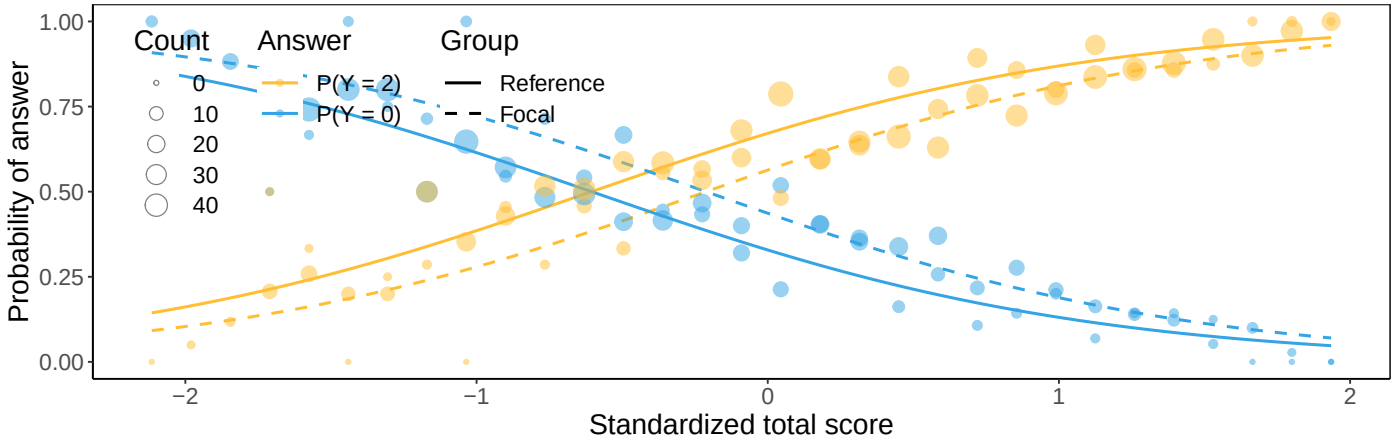
DDF multinomial plot for item b16



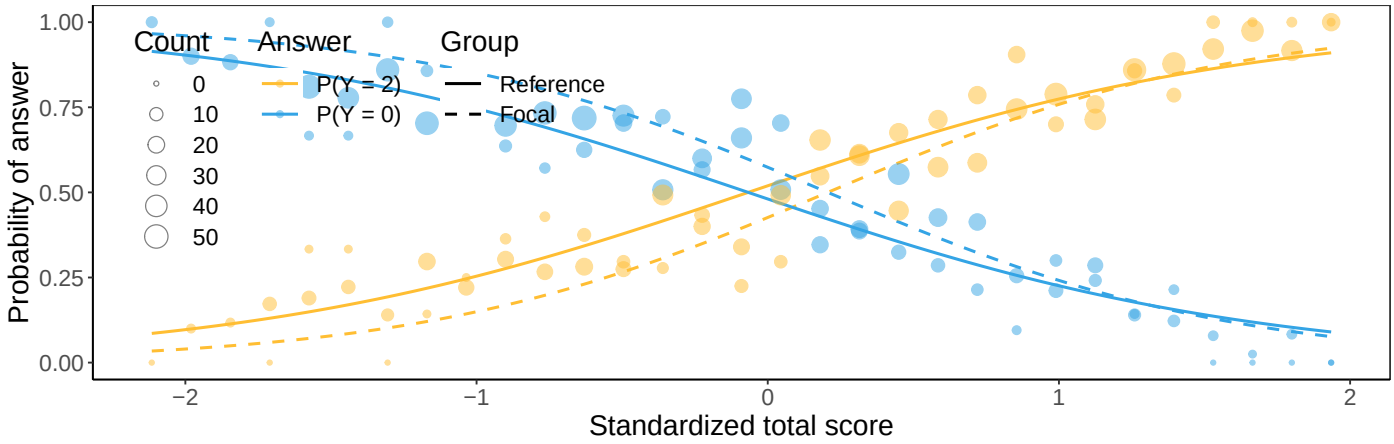
DDF multinomial plot for item b17



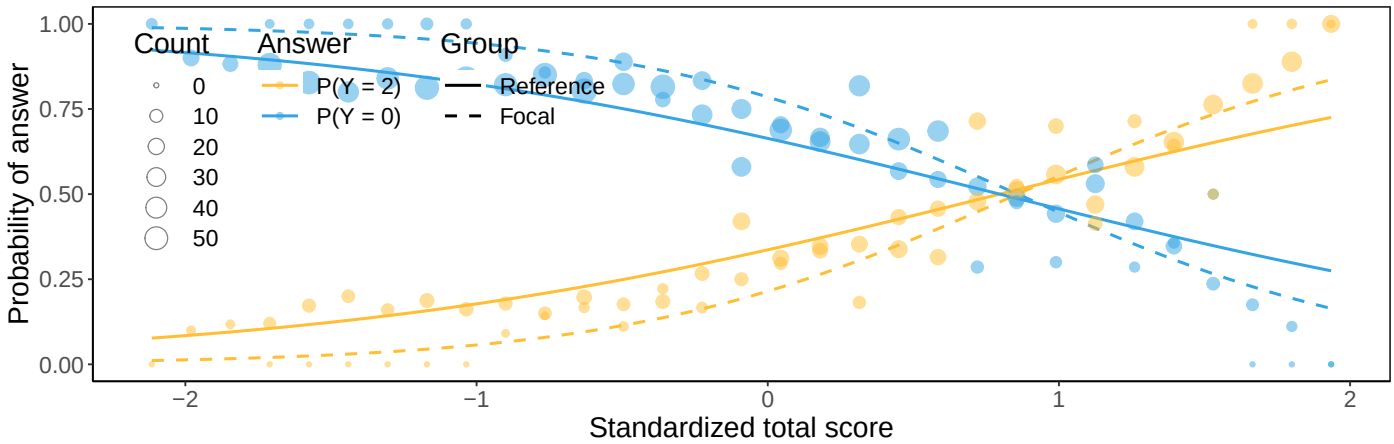
DDF multinomial plot for item b18



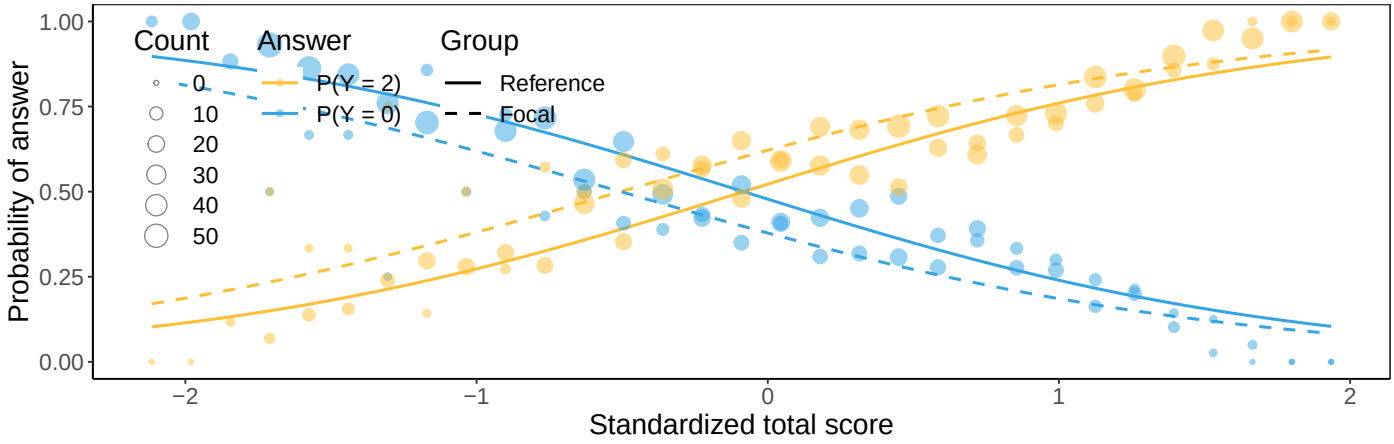
DDF multinomial plot for item b19



DDF multinomial plot for item b20



DDF multinomial plot for item b23



DDF multinomial plot for item b24