Assessment Analysis

sample\_student\_response

Generated On: 2019-08-14 12:00:18

# Summary

This report identifies assessment items flagged for review based on either (A) general population (all students tested) item response analysis characteristics or (B) differential item functioning (DIF) for one or more of the following demographic subgroupings:

- Gender

- Ethnicity

- Gifted Status

- Students with Disabilities

- Limited English Proficiency

- Economically Disadvantaged

The table below includes a summary of all flagged items and categories (as above, General Item Response Issues, Differential Item Functioning, or both). The corresponding sections that follow present further detail regarding each item's response characteristics or degree of DIF.

| Item Number | Flag Type |
| --- | --- |
| 1 | General & DIF |
| 2 | General |
| 3 | General |
| 4 | General & DIF |
| 5 | General |
| 6 | General & DIF |
| 7 | General & DIF |
| 9 | General & DIF |
| 10 | General & DIF |
| 11 | General |
| 12 | General |
| 13 | General & DIF |
| 14 | General |
| 15 | General & DIF |
| 16 | General |
| 17 | General |
| 18 | General & DIF |
| 19 | General & DIF |
| 20 | General |
| 21 | General |
| 22 | General & DIF |

**Recommended Use of This Report**

As you review the analysis in sections 2 and 3, make note of test items that warrant concern and circle the basis for intervention in the table above. For example, if item 6 was flagged for both 'General', due to low difficulty, and 'DIF', but the DIF was minimal, you might circle 'General' but not 'DIF' in the table above so that when you review the test item content, you remember to examine qualities of the question that may have affected difficulty, but you avoid spending needless time exploring causes for DIF if the degree of DIF was in fact insubstantial.

# General Item Analysis

This section flags items for review based on discrimination (alpha value), difficulty (beta value), or aberrant response parameters (c value), as obtained from results using all student responses (i.e. prior to subgrouping for DIF analysis). Please review the listed items in each category.

## Discrimination

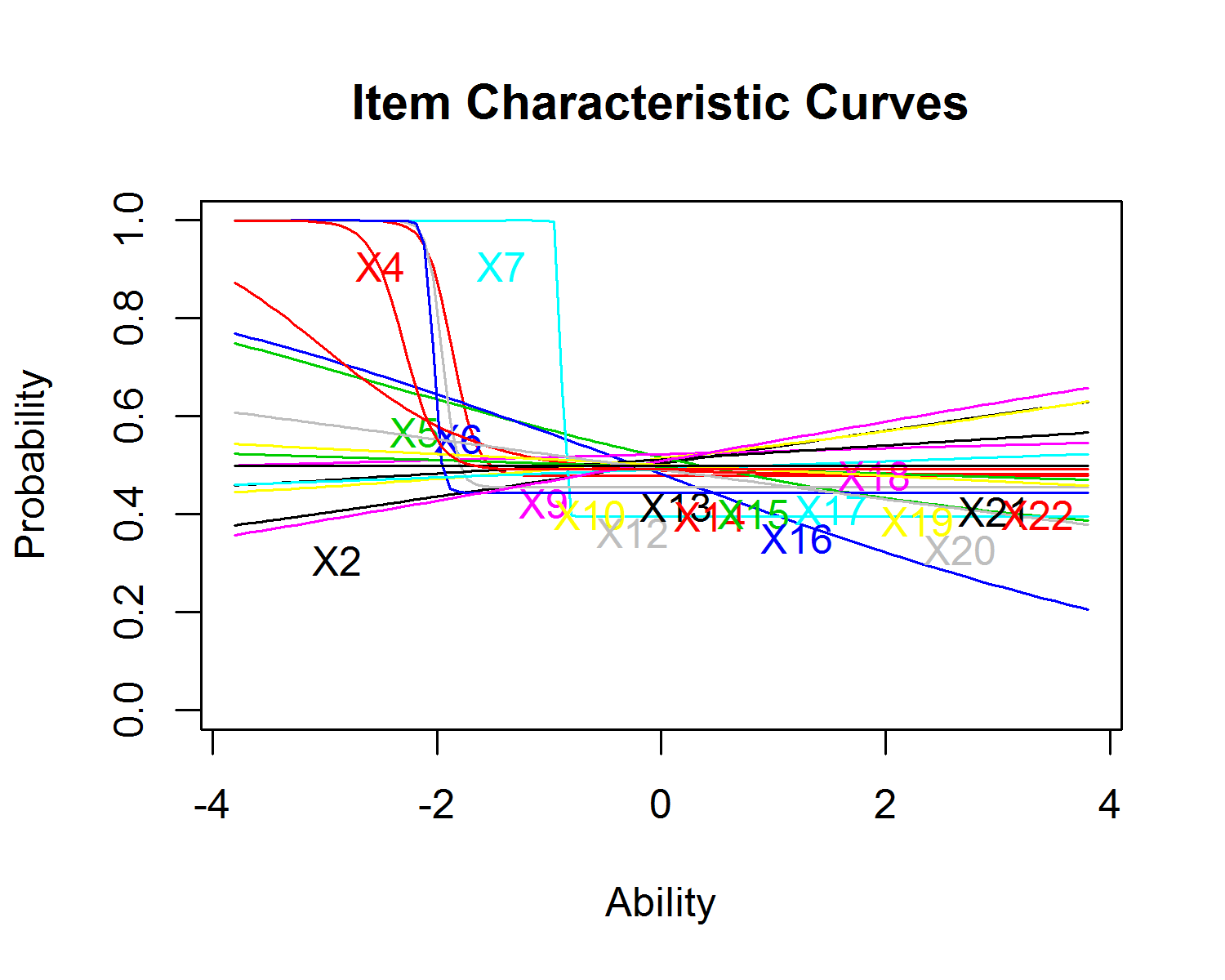
Items flagged here have relaxed 'discrimination' parameters, as estimated through a 3PL item analysis. Correct answers for these items are generally ineffective for determining a given student's ability level (i.e. correct vs. incorrect responses do not accurately indicate how much the student knows).

Item numbers: 2 4 5 6 7 9 10 12 13 14 15 16 17 18 19 20 21 22

Item Statistics:

| Item Number | Discrimination Score |
| --- | --- |
| 2 | 0.136253170 |
| 4 | -9.172751091 |
| 5 | -0.390337046 |
| 6 | -0.335587040 |
| 7 | -72.915985336 |
| 9 | 0.024395970 |
| 10 | 0.244812420 |
| 12 | -15.822057692 |
| 13 | 0.071262645 |
| 14 | -1.431657572 |
| 15 | -0.027809719 |
| 16 | -28.773323136 |
| 17 | 0.032935195 |
| 18 | 0.163312353 |
| 19 | -0.044680374 |
| 20 | -0.130517683 |
| 21 | -0.003275816 |
| 22 | -6.771489315 |

Item Characteristic Curves for Flagged Items:



## Difficulty

Items flagged here measure student ability more than 2 standard deviations away from the mean (i.e. the question is informative about a maximum of approximately 5% of students who complete this assessment). If an item's measured difficulty in the table below is above or below what was intended, please review the question content and answer options.

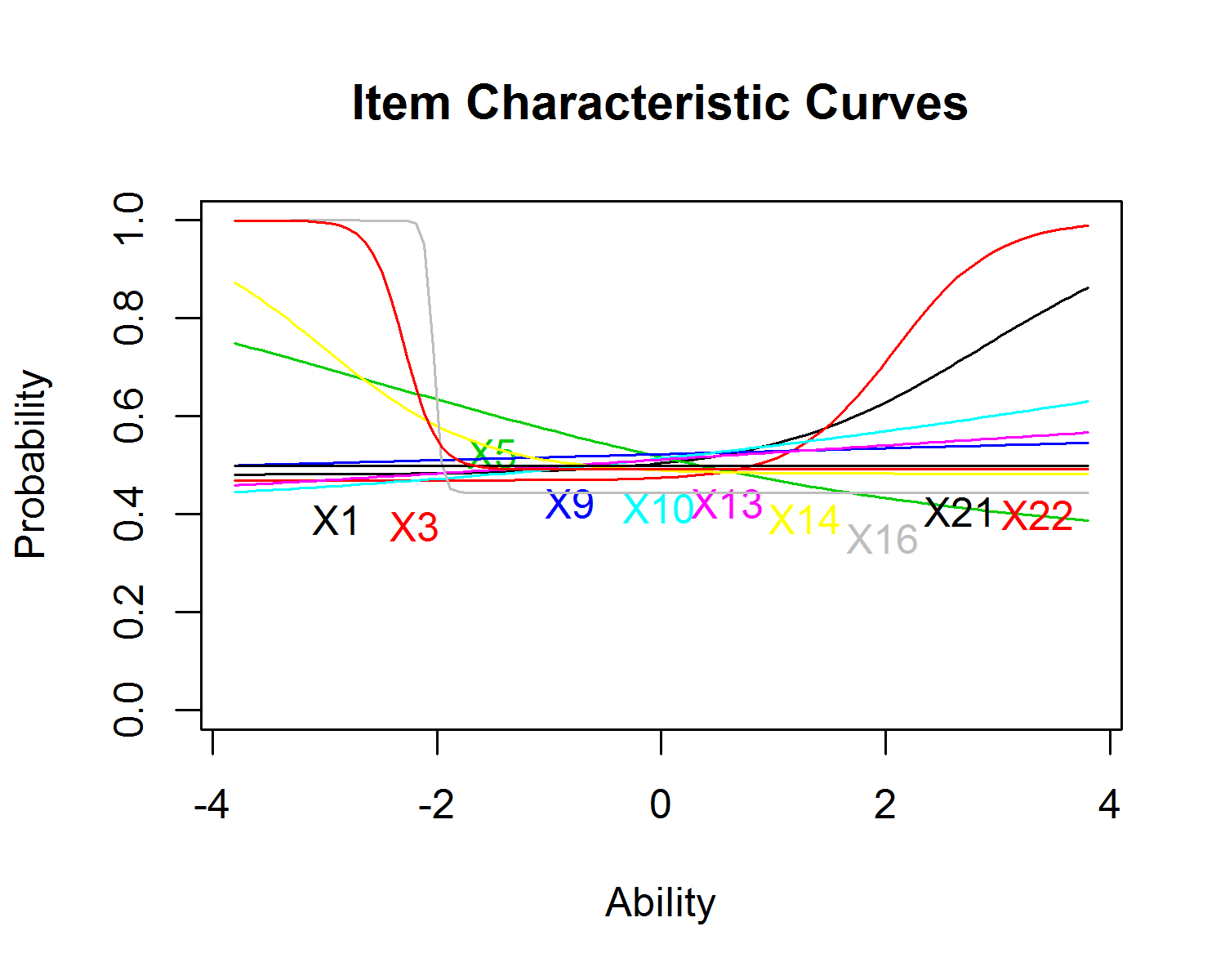
\*\*\*(Note: difficulty metrics displayed here will vary somewhat from the difficulty values presented in Unify (Performance Matters) due to a difference in parameters/degrees of freedom permitted in one model vs. the other (for those interested, this analysis uses a 3-parameter (3PL) model to allow for bottom-asymptote aberrant responses (often called 'guessing') while Unify uses a 2PL model, which forces the bottom-asymptote value to 0 for all items). In nearly all cases, items flagged as >|2| standard deviations here are also identified as >|2| standard deviations in Unify, even if the specific numbers differ.)\*\*\*

Item numbers: 1 3 5 9 10 13 14 16 21 22

Item Statistics:

| Item Number | Difficulty (beta) Score |
| --- | --- |
| 1 | 2.855854 |
| 3 | 2.076710 |
| 5 | -2.519792 |
| 9 | -3.767323 |
| 10 | 5.546024 |
| 13 | 5.032759 |
| 14 | -3.016720 |
| 16 | -2.029579 |
| 21 | -1364.674470 |
| 22 | -2.293854 |

Item Characteristic Curves for Flagged Items:



## Aberrant Response

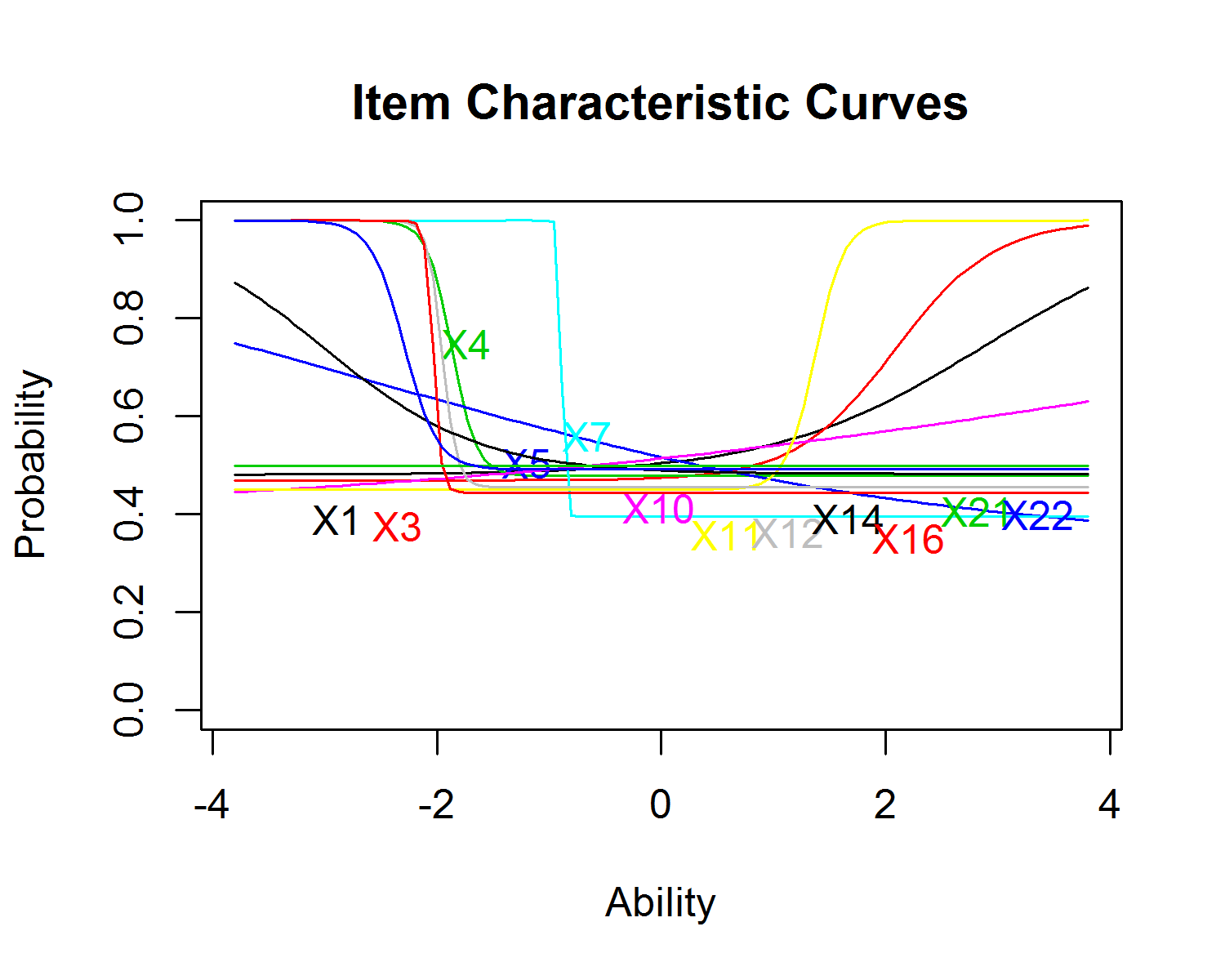
Items flagged here have an abberant response parameter greater than .25. This indicates that students with the lowest possible ability level (i.e. no content knowledge whatsoever) answered the question correctly >25% of the time, based on responses to all items on this assessment. Please review the flagged items below.

Item numbers: 1 3 4 5 7 10 11 12 14 16 21 22

Item Statistics:

| Item Number | Aberrant Response Score |
| --- | --- |
| 1 | 0.4811340 |
| 3 | 0.4697250 |
| 4 | 0.4783672 |
| 5 | 0.3362171 |
| 7 | 0.3957867 |
| 10 | 0.3893663 |
| 11 | 0.4511061 |
| 12 | 0.4556295 |
| 14 | 0.4829857 |
| 16 | 0.4445823 |
| 21 | 0.4937864 |
| 22 | 0.4919249 |

Item Characteristic Curves for Flagged Items:



# Differential Item Functioning (DIF) Analysis

## Summary

Items described in this section have been flagged for exhibiting DIF. Reviewers should note that the analysis behind this program raises a DIF flag if the difference between subgroup responses for two given students of the same ability level is statistically significant. That is, items are first identified if they exhibit low probability of the observed DIF happening by chance (i.e. low statistical p-value (not to be confused with the 'percent correct' p-value listed in Unify)), as opposed to the actual direction or degree of DIF. For any flagged item, a visualization of the DIF is provided below; **reviewers should examine the visualizations for flagged items to determine whether the degree of DIF is problematic. If so, item content should be subject to review.**

The table below summarizes items flagged for DIF, as well as the subgroup associated with the given flag.

| Item Number | Flag |
| --- | --- |
| 1 | Gifted |
| 4 | Ethnicity |
| 6 | Gifted, LEP |
| 7 | SWD, ED |
| 9 | Ethnicity |
| 10 | Gifted, SWD |
| 13 | LEP |
| 15 | ED |
| 18 | Ethnicity |
| 19 | Ethnicity |
| 22 | Gender |

Note: Sample size among the following subgroups was insufficient to identify DIF (or lack of DIF): .

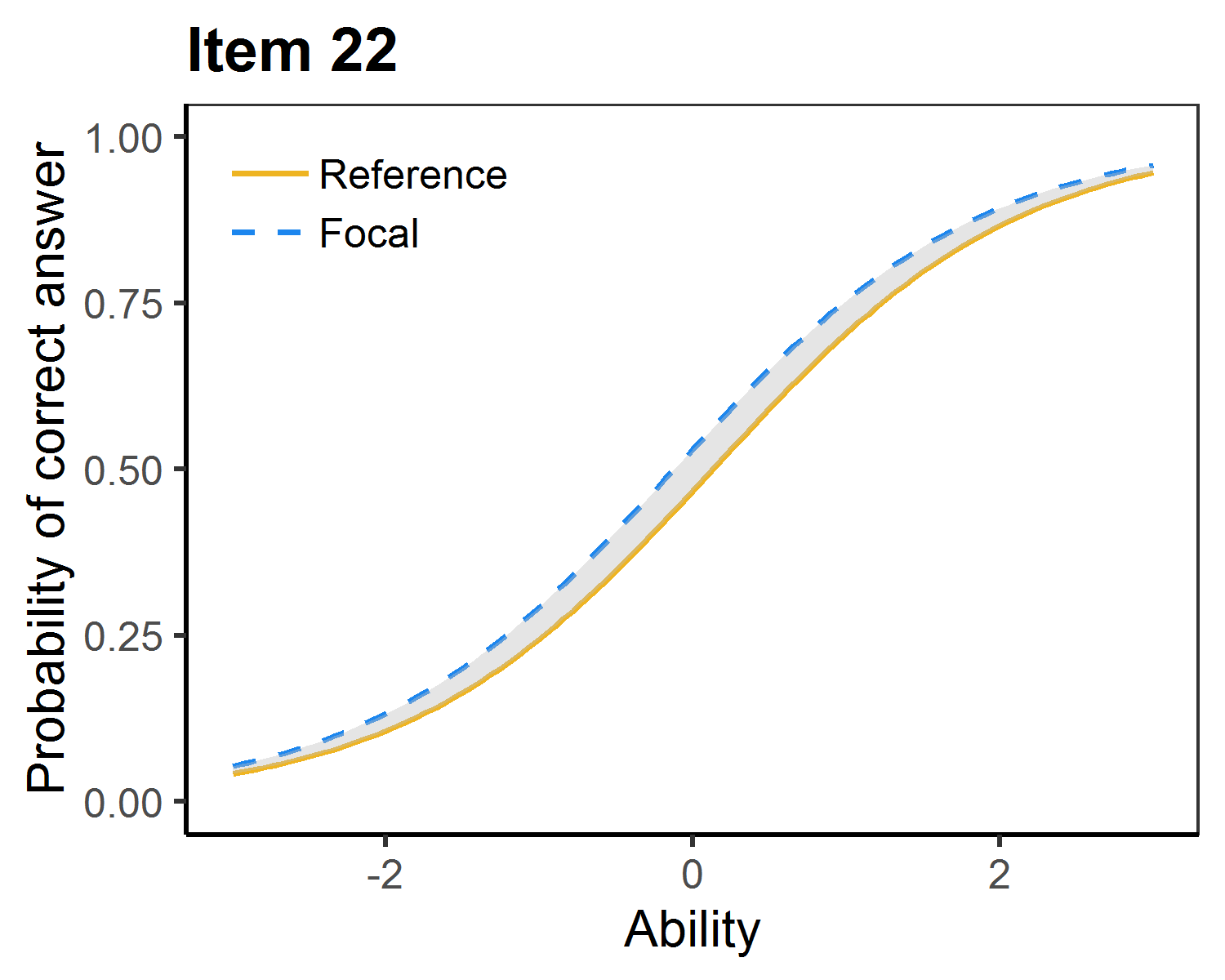
## Gender

Items flagged in this section have exhibited differential item functioning (DIF) by gender. That is, two students of the same ability level but different genders perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

The following items have been flagged for DIF based on gender. In the charts that follow, 'Reference Group' refers to male students; 'Focal Group' refers to female students.

Item numbers: 22

DIF Plots:



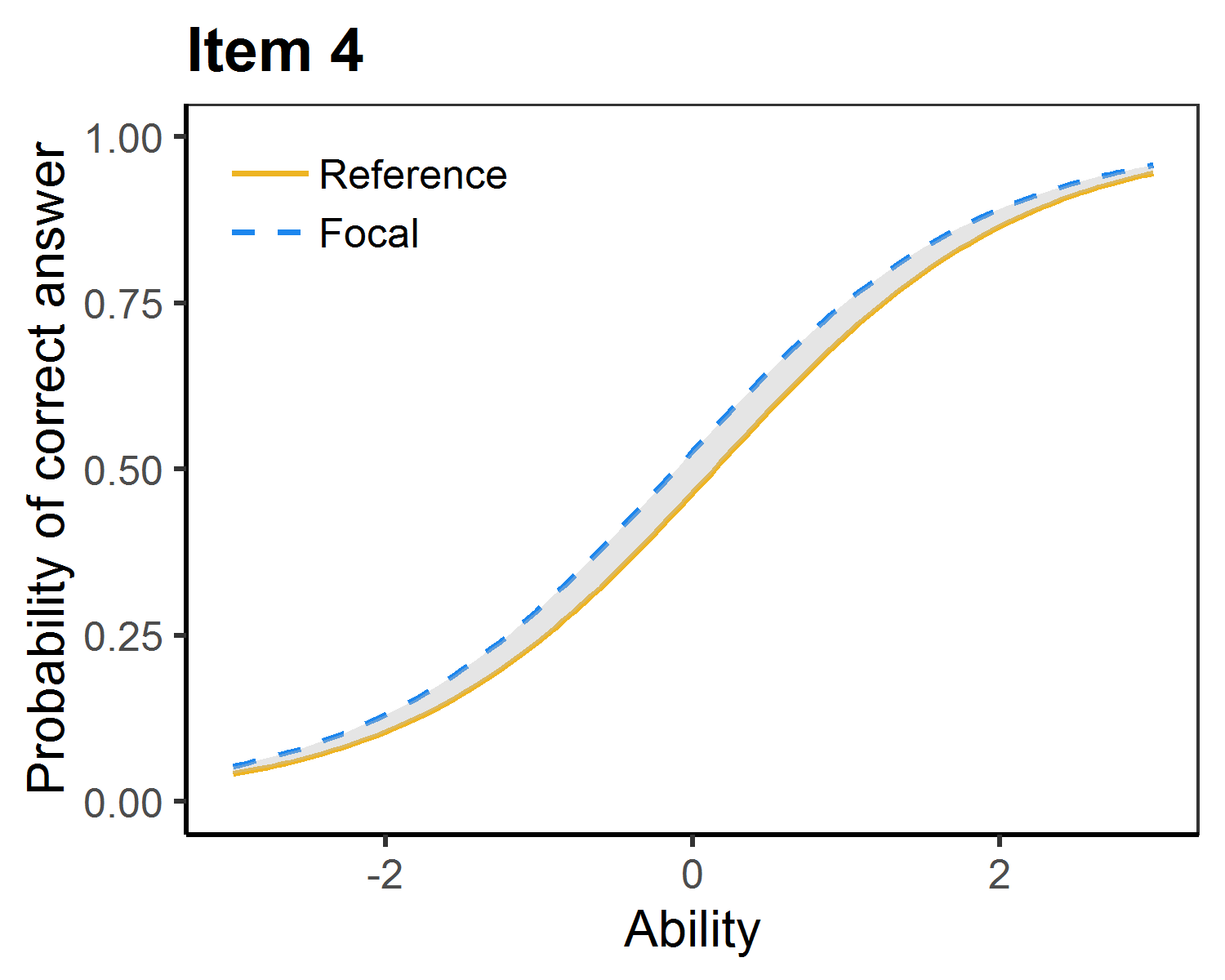
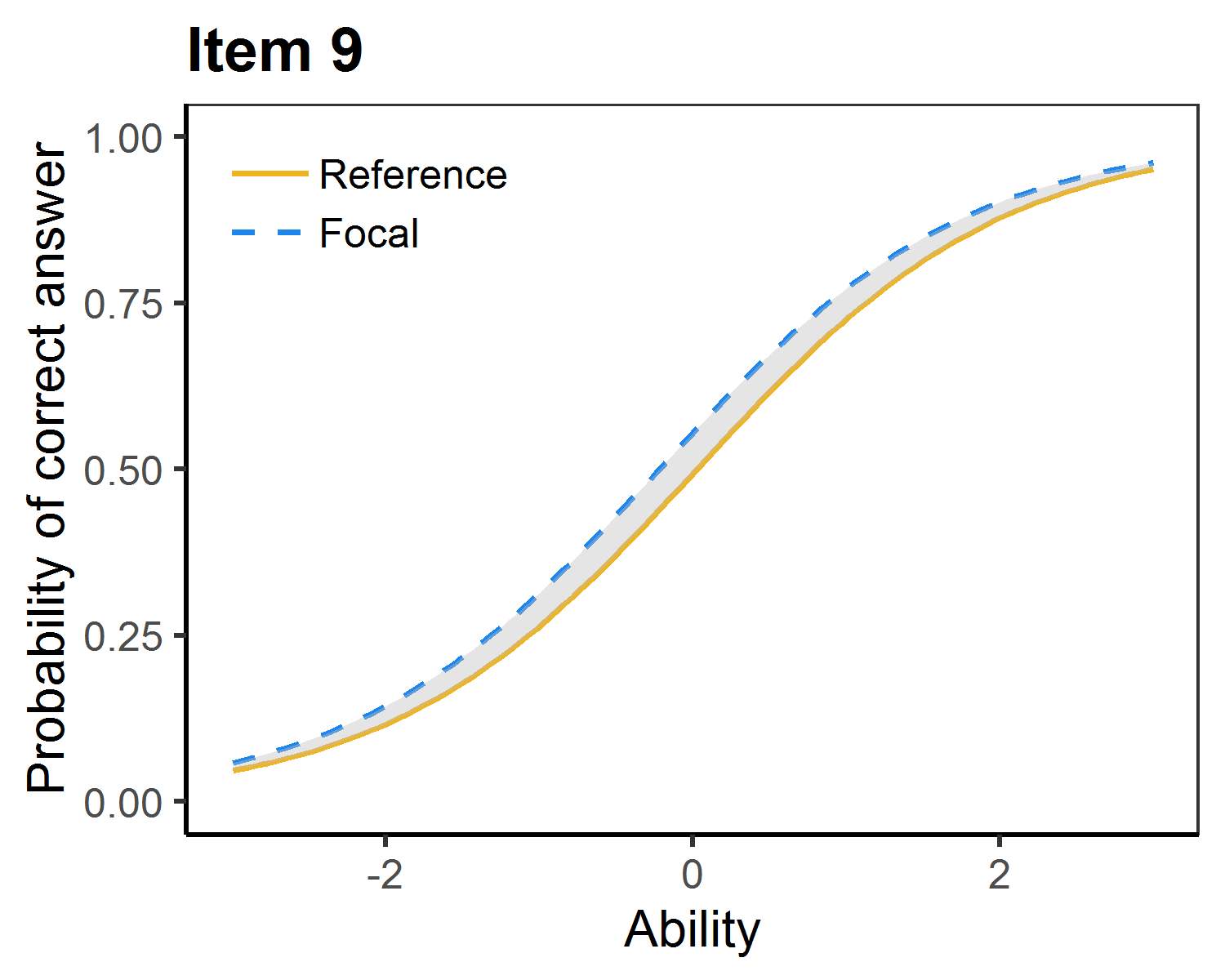
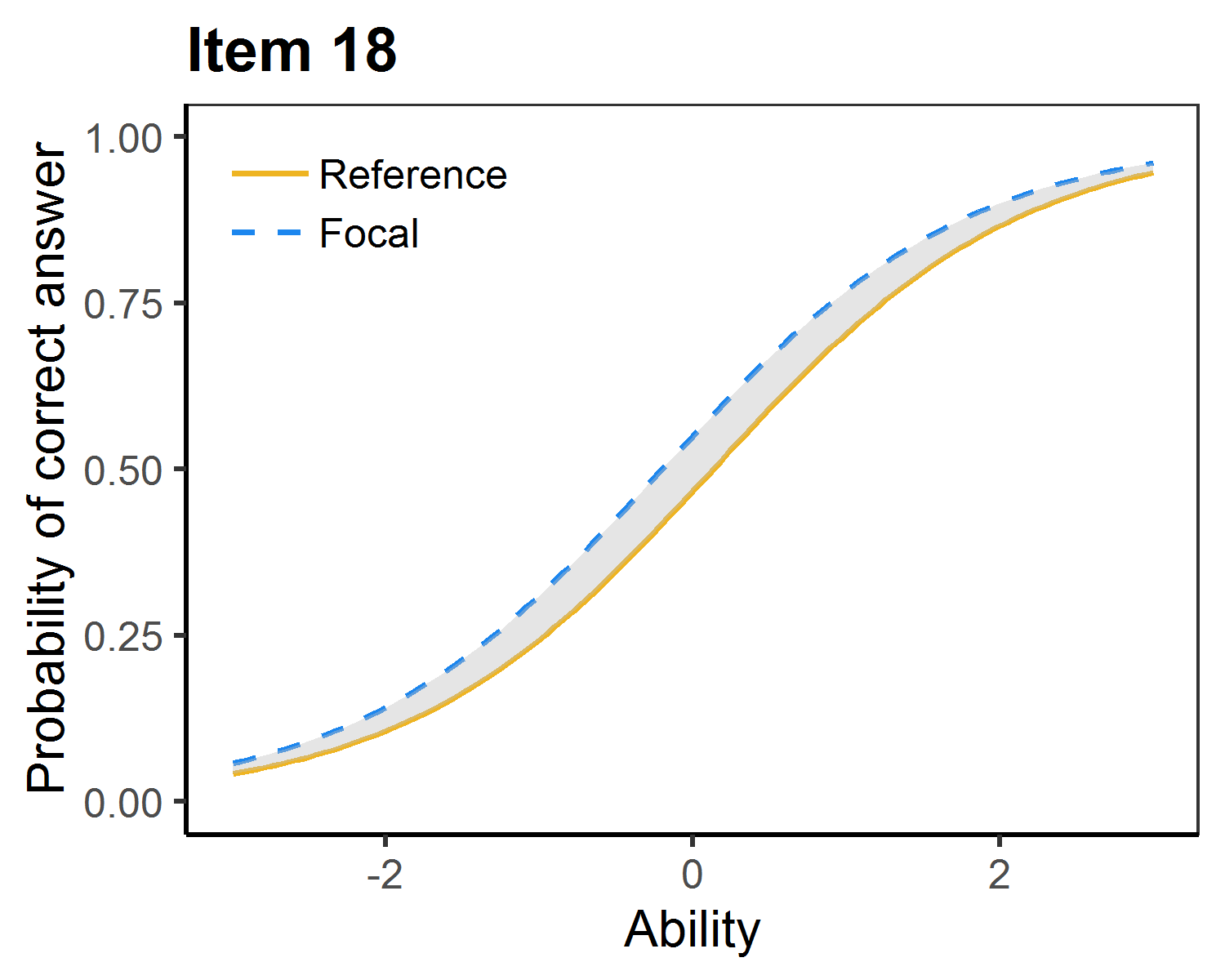
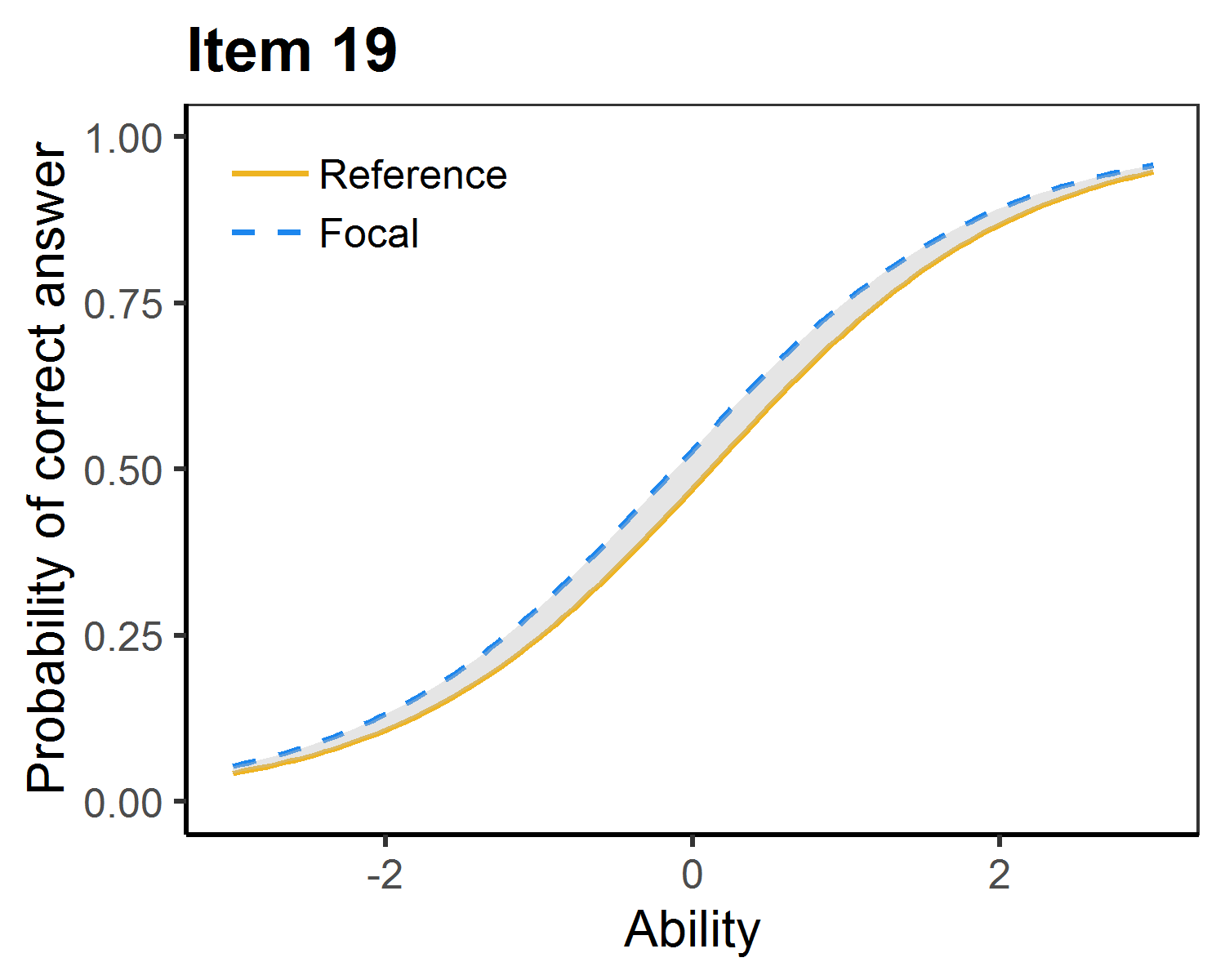
## Ethnicity

Items flagged in this section have exhibited differential item functioning (DIF) by ethnicity. That is, two students of the same ability level but different ethnicities perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

The following items have been flagged for DIF based on ethnicity. In the charts that follow, 'Reference Group' refers to African American students; 'Focal Group' refers to non-African American students. (Due to limited sample size among student responses, DIF analysis cannot be accurately conducted based on other ethnicities at this time.)

Item numbers: 4 9 18 19

DIF Plots:

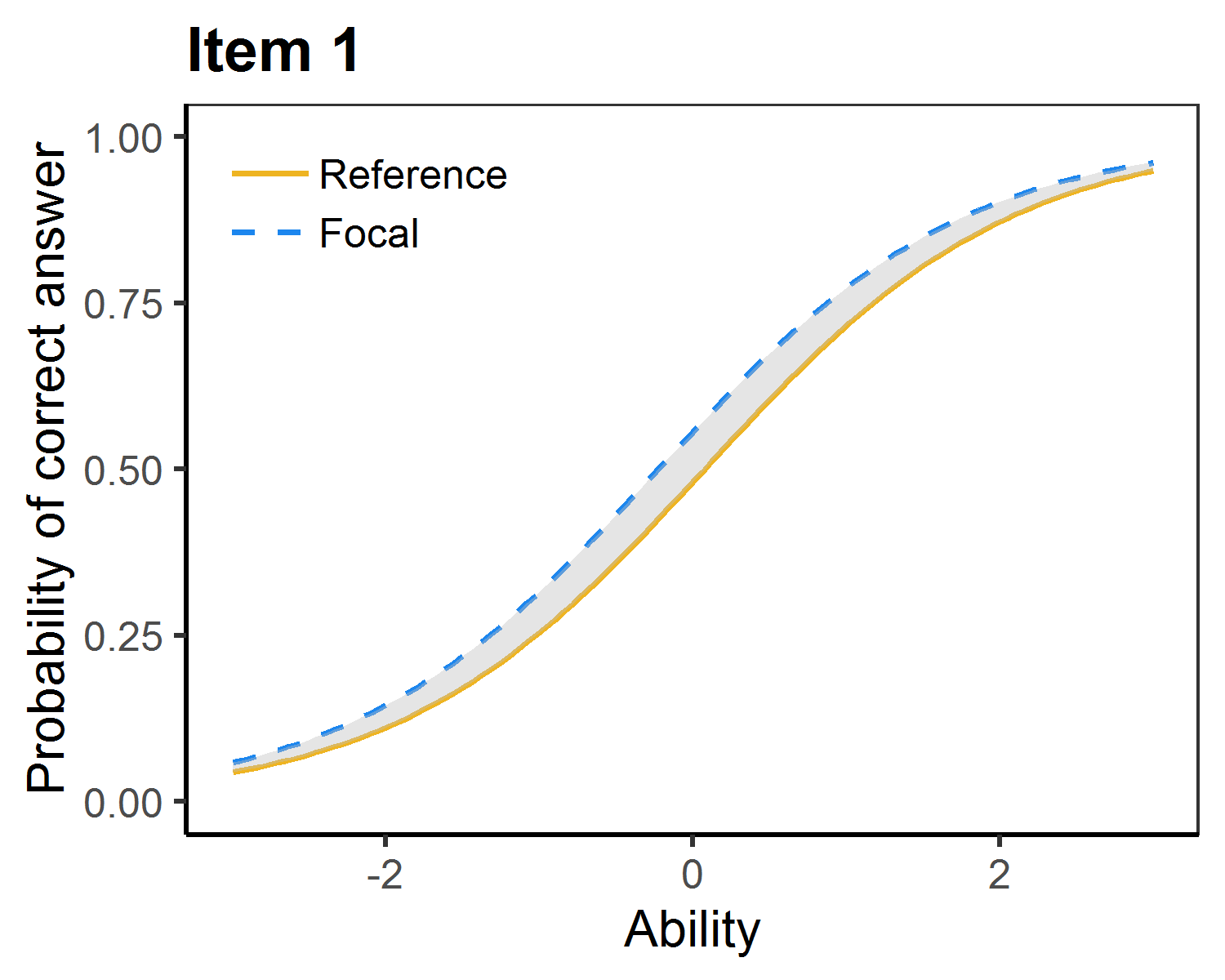
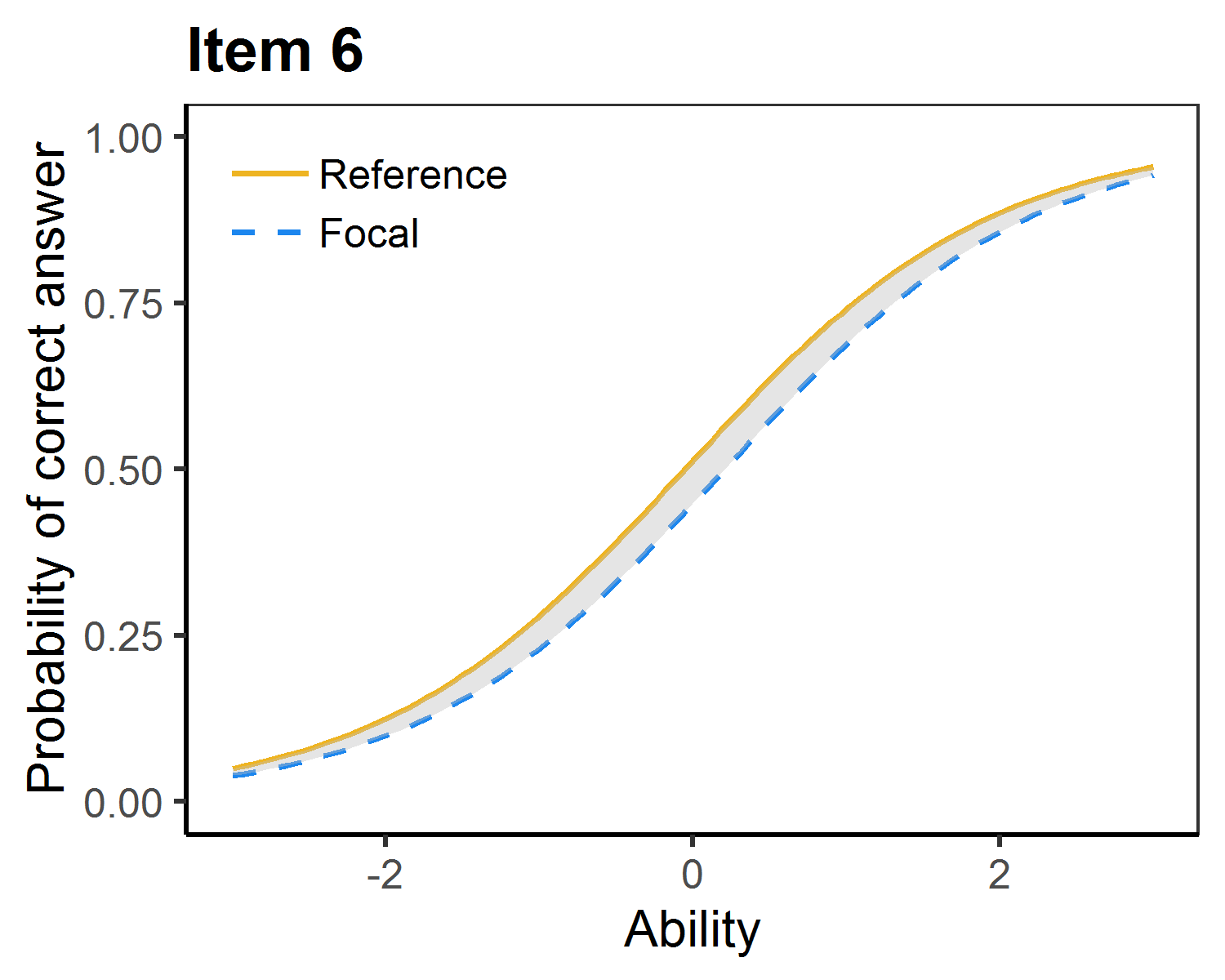
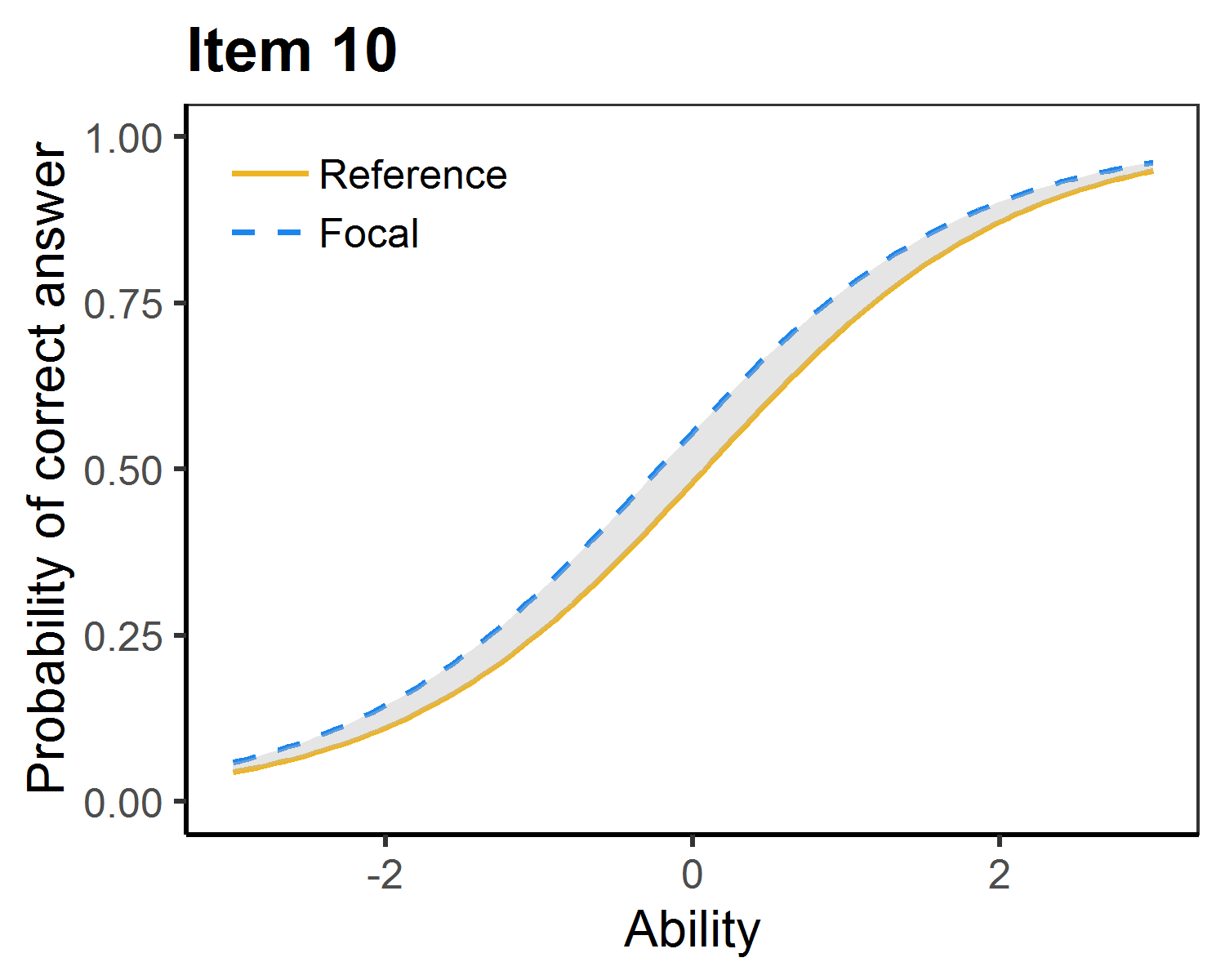
## Gifted

Items flagged in this section have exhibited differential item functioning (DIF) by gifted status. That is, for two students of the same ability level wherein one is labeled 'gifted' and the other is not, the students perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

The following items have been flagged for DIF based on gifted status. In the charts that follow, 'Reference Group' refers to students NOT labeled as gifted; 'Focal Group' refers to students labeled as gifted.

Item numbers: 1 6 10

DIF Plots:

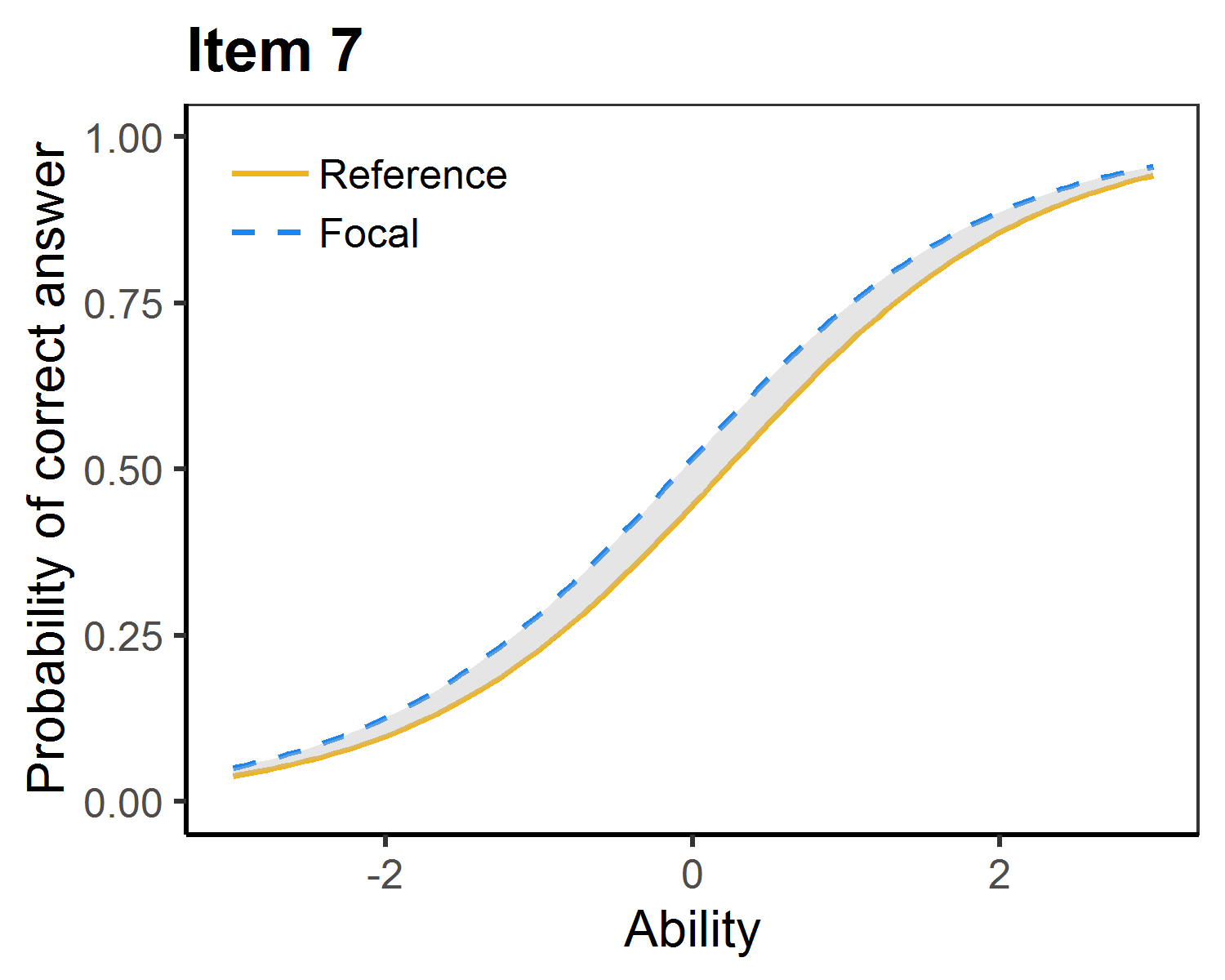
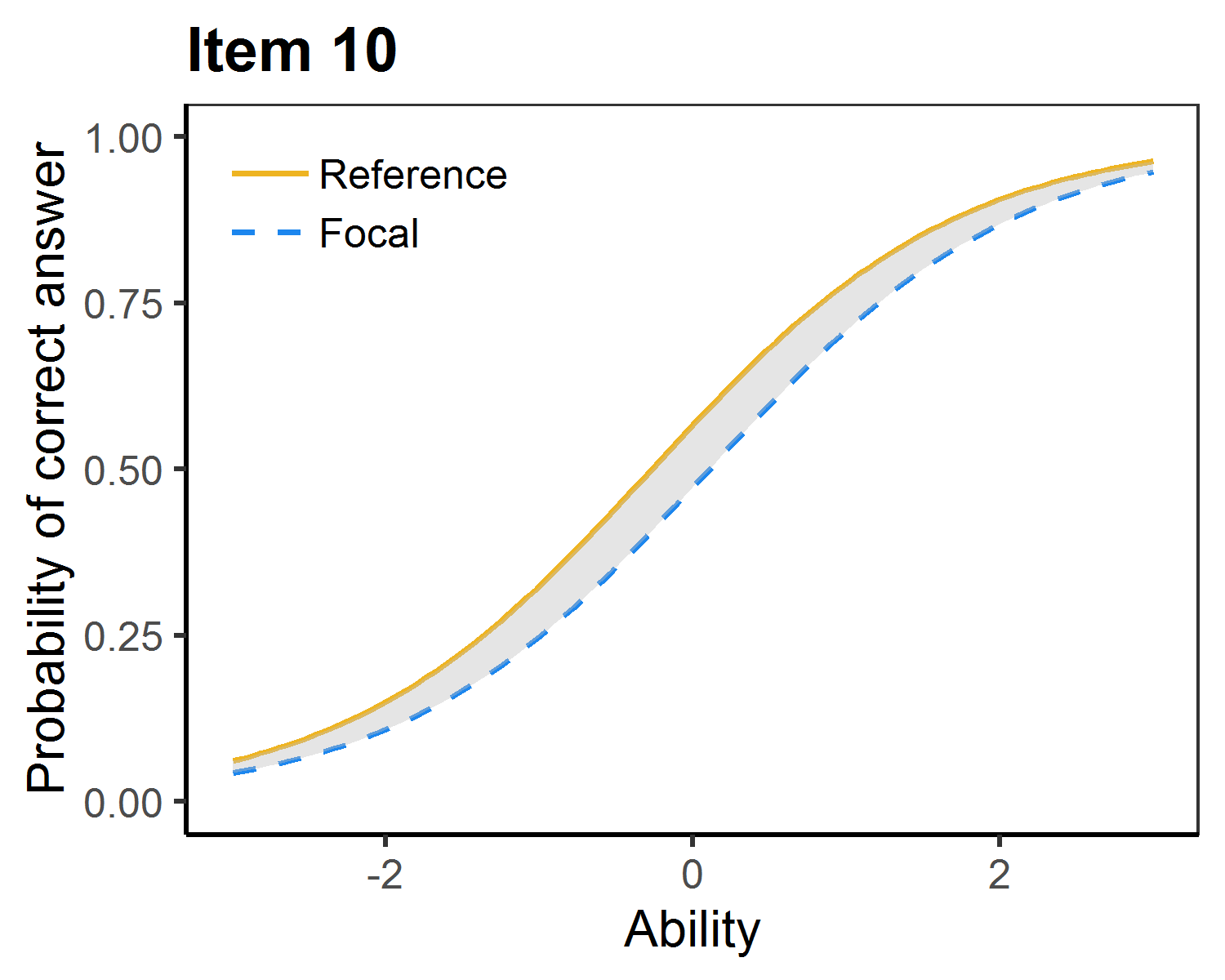
## Students with Disabilities

Items flagged in this section have exhibited differential item functioning (DIF) by disability status. That is, for two students of the same ability level wherein one is labeled as a 'student with disability' (SWD) and the other is not, the students perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

The following items have been flagged for DIF based on SWD status. In the charts that follow, 'Reference Group' refers to students NOT labeled as SWD; 'Focal Group' refers to students labeled as SWD.

Item numbers: 7 10

DIF Plots:

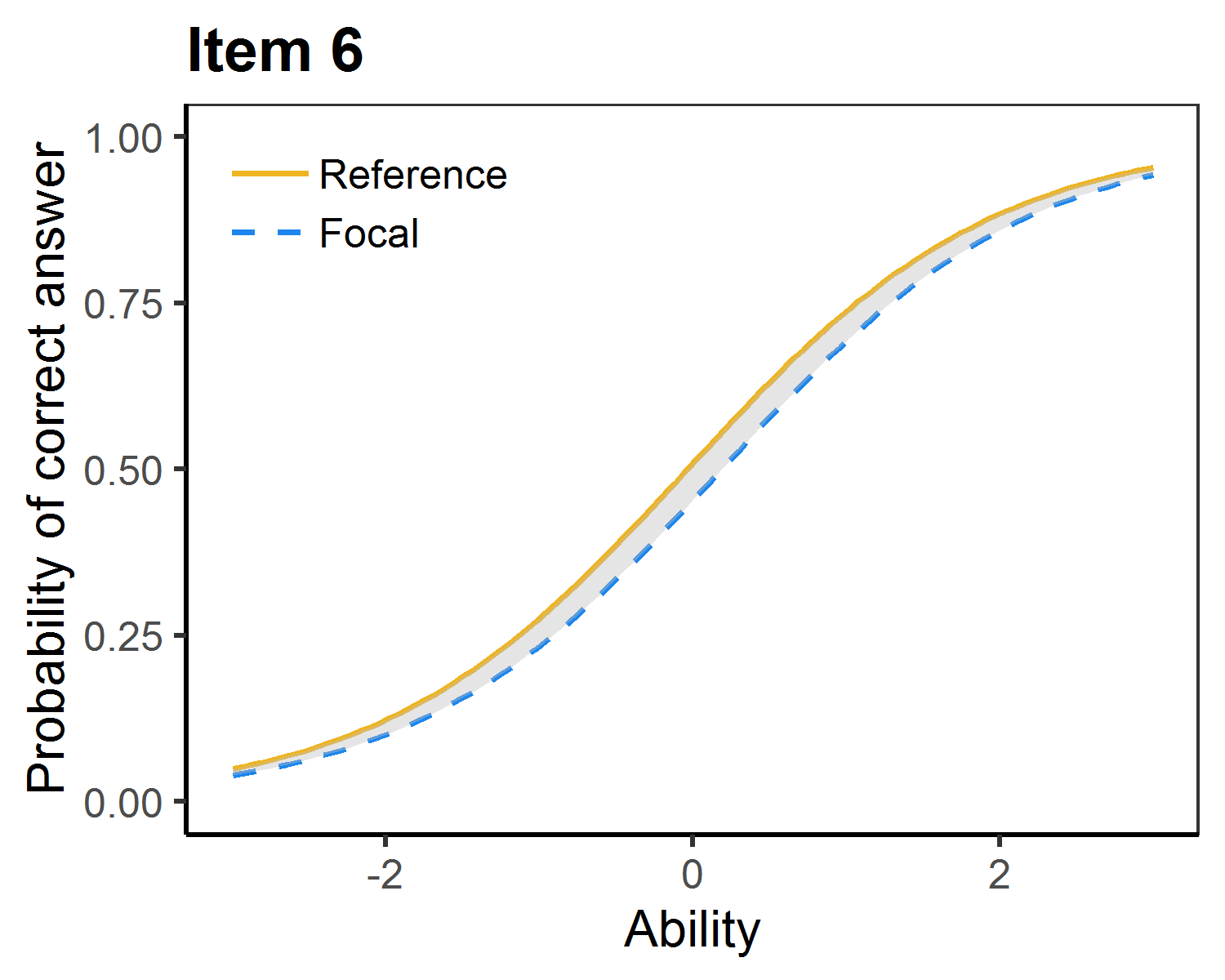
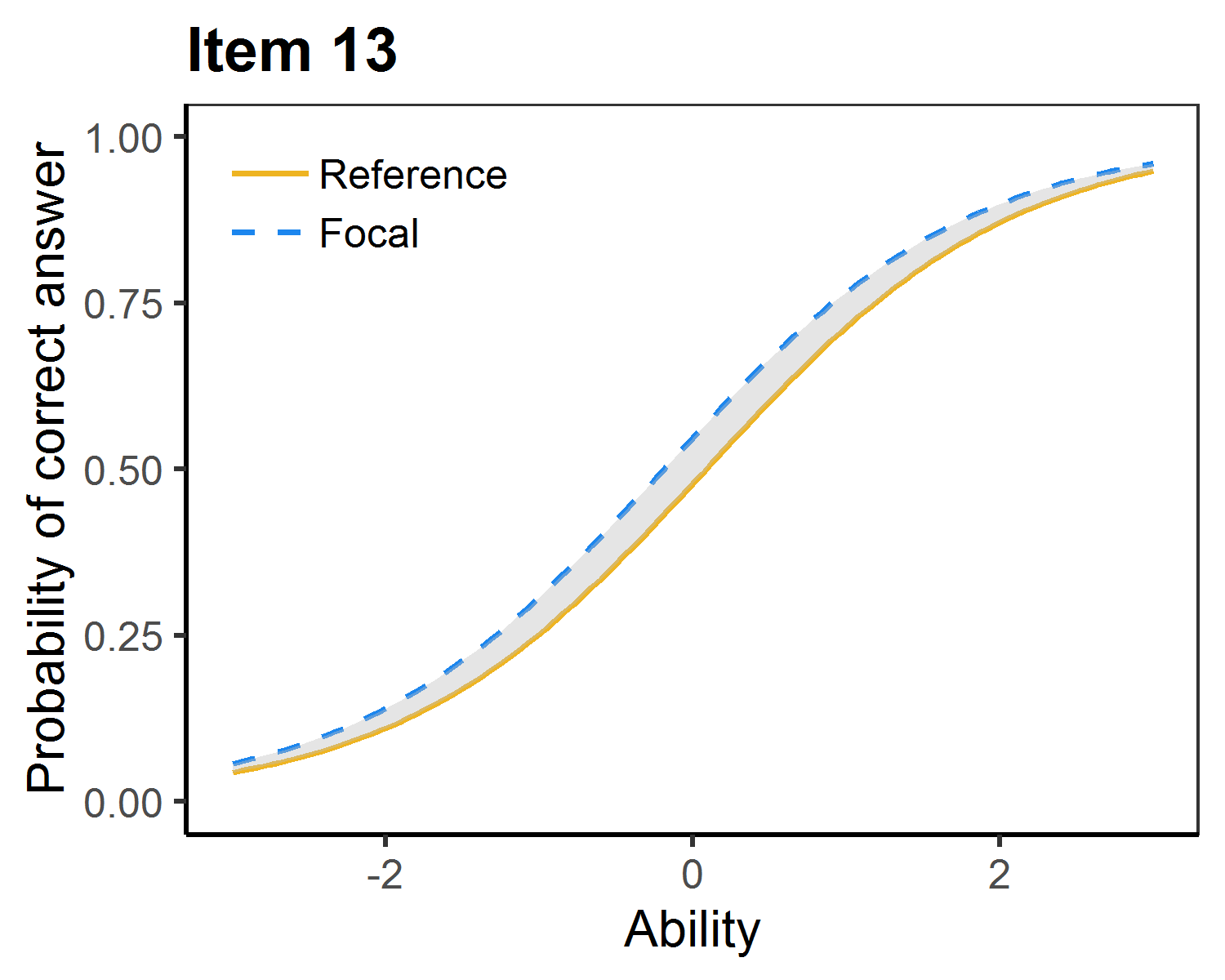
## Students with Limited English Proficiency

Items flagged in this section have exhibited differential item functioning (DIF) by 'Limited English Proficiency' (LEP) status. That is, for two students of the same ability level wherein one is labeled as a 'student with limited English proficiency' and the other is not, the students perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

The following items have been flagged for DIF based on LEP status. In the charts that follow, 'Reference Group' refers to students NOT labeled as LEP; 'Focal Group' refers to students labeled as LEP.

Item numbers: 6 13

DIF Plots:

## 'Economically Disadvantaged' Students

Items flagged in this section have exhibited differential item functioning (DIF) by 'Economically Disadvantaged' (ED) status. That is, for two students of the same ability level wherein one is living below the Federal Poverty Rate (FPR) and the other is not, the students perform differently on the item in question. As mentioned above, items are initially flagged based upon the statistical significance--not the degree or direction--of the DIF. Therefore, it is important that reviewers examine the charts below to determine whether the degree of DIF is substantial enough to warrant further evaluation of the assessment item.

(Note: at this time, the FPR is the most accurate indicator of socioeconomic status available at the district level. This analysis acknowledges, however, that the low threshold for FPR qualification makes it a crude indicator of economic distress. I.e., families earning, e.g., $15,000/year above the FPR may still fall well within a better qualitative definition of economic distress.)

The following items have been flagged for DIF based on ED status. In the charts that follow, 'Reference Group' refers to students NOT labeled as ED; 'Focal Group' refers to students labeled as ED.

Item numbers: 7 15

DIF Plots:

