



Universidade de Brasília
Departamento de Estatística

Aplicação de Modelos de Teoria de Resposta ao Item na Avaliação das
Questões da Prova de Língua Portuguesa do Vestibular da UnB

Gabriel Peixoto Veiga

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

Brasília
2022

Gabriel Peixoto Veiga

**Aplicação de Modelos de Teoria de Resposta ao Item na Avaliação das
Questões da Prova de Língua Portuguesa do Vestibular da UnB**

Orientador: Prof. Dr. Antonio Eduardo Gomes

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília
2022**

Resumo

O relatório apresentado tem por objetivo analisar, utilizando Teorias de Resposta ao Item, as questões da prova de português do vestibular da UnB (Universidade de Brasília) realizado no segundo semestre de 2014. Utilizando o software estatístico R, foi possível identificar o grau de dificuldade e discriminação dos itens, detectar itens possivelmente inadequados, ajustar o Modelo de Resposta Gradual e comparar as notas dos alunos obtidas pelo modelo convencional e pelo modelo de TRI (Teoria de Resposta ao Item).

Durantes as análises, foi utilizado o software estatístico R, tanto para a construção dos gráficos quanto para a construção dos modelos. Observou-se que a atual avaliação do CESPE com penalização gera consequências no momento de resposta. Observou-se que examinandos com maior nível de conhecimento apresentam maior tendência a responder as questões e eventualmente errar, ao passo que os examinandos que não respondem a uma questão apresentam um nível menor de conhecimento.

Palavras-chave: Vestibular, UnB, Teoria de Resposta ao Item, Modelo de Resposta Gradual.

Sumário

1 Introdução	5
2 Metodologia e Material	7
2.1 Material	7
2.2 Teoria de Resposta ao Item	7
2.2.1 Modelos Logísticos	8
2.2.2 Curva Característica do Item	9
2.2.3 Coeficiente de Correlação Ponto-Bisserial	9
2.2.4 Modelo de Resposta Gradual	10
2.2.5 Estimação de Parâmetros	10
3 Resultados	11
3.1 Análise Descritiva	11
3.2 Ajuste dos Modelos	19
3.2.1 Ajuste do Modelo Logístico de 3 Parâmetros	19
3.2.2 Ajuste do Modelo Logístico de 2 Parâmetros	26
3.3 Modelo de Resposta Gradual	28
3.3.1 Ajuste do Modelo de Resposta Gradual	41
4 Conclusões	49
Referências	50

1 Introdução

Ao longo da vida humana, do ensino fundamental às provas de Doutorado e certificações profissionais, os exames estiveram presentes para que se possa exercer alguma função em sociedade. Segundo a UnB, estima-se que cerca de 10.000 candidatos se inscrevem no vestibular todos os anos para disputar uma vaga na Universidade, as quais se apresentam em quantidade 5 vezes menor (2000 vagas por semestre). Em janeiro é utilizado o ENEM - Exame Nacional do Ensino Médio - como método de ingresso, e no meio do ano o Vestibular (CESPE/CEBRASPE), o qual será objeto de estudo deste trabalho.

A Teoria de Resposta ao Item é uma teoria aplicada para estimar a intensidade de traços latentes(características que não podem ser observadas diretamente), sendo o conhecimento o caso mais comum. Em outras palavras, a Teoria de Resposta ao Item (TRI) estuda o modelos matemáticos que relacionam variáveis observáveis (itens de teste) e traços não observáveis (aptidões).

Atualmente, cresce o número de avaliações educacionais que passam por aplicações de técnicas de TRI para estimação da intensidade de traços latentes. A TRI busca relacionar matematicamente a probabilidade de um indivíduo acertar um determinado item com suas características latentes, conhecimentos ou habilidades em determinada área.

Neste trabalho, foram aplicados modelos de TRI para analisar as questões do vestibular da UnB, eleita pelo Ranking Times Higher Education como a 7° melhor universidade do país. Sendo assim, é de responsabilidade do CEBRASPE - Centro Brasileiro de Pesquisa em Avaliação e Seleção e de Promoção de Eventos - uma prova que consiga relacionar acerto de itens e aptidões de conhecimento dos candidatos, o que será avaliado a seguir.

2 Metodologia e Material

2.1 Material

O material analisado neste trabalho será o banco de dados fornecidos pelo CE-BRASPE, referente ao vestibular da UnB 2014/2 cuja distribuição encontra-se na Figura 1. Foram estudadas as respostas de 7232 candidatos para as 48 questões tipo A (certo/errado) da prova de Língua Portuguesa.

Por meio do software R, serão utilizados gráficos e modelos baseados nas referências Anjos e Andrade (2012) e Baker e Kim (2017), para aplicar técnicas de Teoria de Resposta ao Item.

Data	Prova	Disciplinas-Foco	Nº de itens	Duração
1º DIA 7/6/2014	Conhecimentos – Parte I	Língua Espanhola, Língua Francesa ou Língua Inglesa	30	300min
	Conhecimentos – Parte II	Língua Portuguesa e Literaturas de Língua Portuguesa, Geografia e História, Artes (Artes Cênicas, Artes Visuais e Música), Filosofia e Sociologia	120	
	Redação em Língua Portuguesa	-	-	
2º DIA 8/6/2014	Conhecimentos – Parte III	Biologia, Física, Química e Matemática	150	300min

Figura 1: Distribuição das Questões do Vestibular da UnB 2014/2

2.2 Teoria de Resposta ao Item

Os Modelos de TRI se baseiam em dois postulados básicos. O primeiro estabelece que o desempenho de um indivíduo em um exame pode ser predito por um conjunto de fatores latentes, identificados pela letra grega θ (theta). A modelagem latente busca relacionar θ (efeito) e o desempenho final do indivíduo. O segundo postulado define que tal relação entre causa e resultado pode ser descrita por uma equação monotônica crescente Curva Característica do Item (CCI).

Segundo Andrade, Tavares e Valle (2000), essa relação é sempre expressa de tal forma que quanto maior a habilidade, maior a probabilidade de acerto no item. Além disso, os modelos dependem de 3 fatores: natureza do item, número de populações envolvidas e quantidade de traços latentes que estão sendo medidos.

Por meio do Modelo de Resposta Gradual (MRG), será possível avaliar a ordem

entre as saídas “não resposta”, “errado”, “certo”. Uma vez que, em itens do vestibular, uma resposta errada gera penalidade, ou seja, a não resposta pode significar tanto falta de conhecimento como medo da penalidade.

2.2.1 Modelos Logísticos

Os modelos logísticos a serem utilizados neste estudo são o de 2 parâmetros (ML2 - dificuldade e discriminação) e o de 3 parâmetros (ML3 - acerto ao acaso, dificuldade e discriminação). No ML2, a probabilidade do j -ésimo candidato acertar o i -ésimo item é dada por

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (2.2.1)$$

com $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, n$, onde

- $P(U_{ij} = 1|\theta_j)$ é a probabilidade do j -ésimo candidato, com proficiência θ_j , acertar o i -ésimo item.
- U_{ij} assume valores para o item i e candidato j , com 1 pra acerto e 0 pra erro.
- θ_j é a proficiência do j -ésimo indivíduo.
- a_i é o parâmetro de discriminação do i -ésimo item.
- b_i é o parâmetro de dificuldade do i -ésimo item.
- D é uma constante de escala que assume valor 1 ou 1,702, conforme se queira ajustar o modelo logístico ou uma aproximação para a ogiva da distribuição normal.

Para o modelo ML3, a probabilidade é dada por

$$P(U_{ij} = 1|\theta_j) = c_i + (1 - c_i) \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}, \quad (2.2.2)$$

em que c_i é o parâmetro que representa a probabilidade de acerto ao acaso do i -ésimo item.

2.2.2 Curva Característica do Item

Como citado anteriormente, os modelos de TRI buscam relacionar o desempenho do candidato com seus traços latentes. Na prática, tem-se a relação entre a probabilidade de se acertar o item e a proficiência do aluno. Um exemplo de Curva Característica do Item (CCI) pode ser visto na Figura 2.

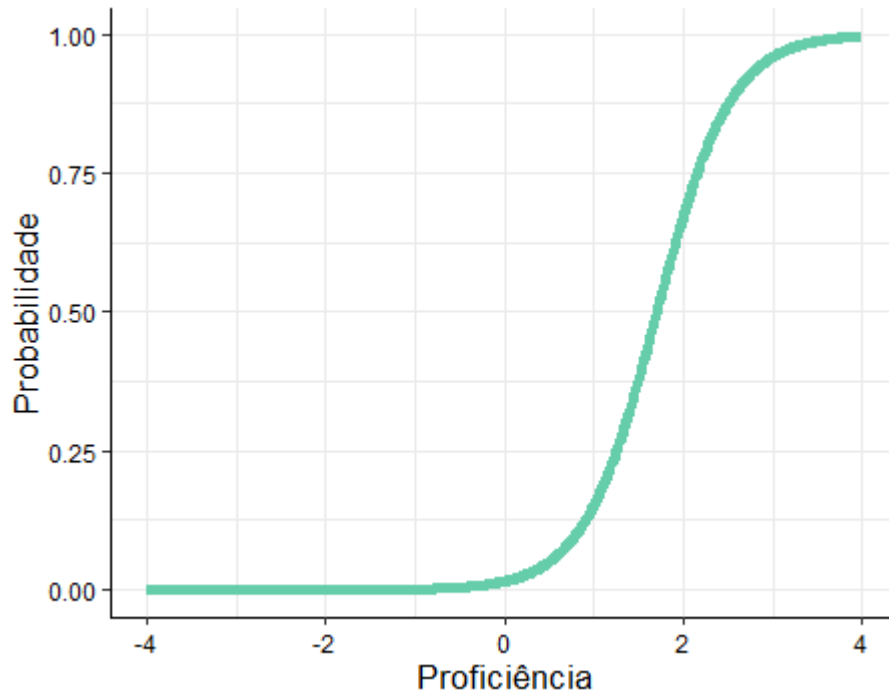


Figura 2: Curva Característica do Item

2.2.3 Coeficiente de Correlação Ponto-Bisserial

O coeficiente de correlação ponto-bisserial mede a correlação linear entre as respostas para um item e o escore dos respondentes. No caso do vestibular do CESPE, que categoriza as respostas em acerto, não resposta e erro, a dicotomização representa acerto e não acerto (erro ou sem resposta) de cada item com o escore do indivíduo na prova de português, gerando o coeficiente por item. Seu cálculo é dado por:

$$P_i^{PB} = \frac{\sum_{j=1}^n U_{ij}T_j - (\sum_{j=1}^n U_{ij})(\sum_{j=1}^n T_j)/n}{\sqrt{[\sum_{j=1}^n (U_{ij} - \bar{U}_i)^2] [\sum_{j=1}^n (T_j - \bar{T})^2]}} = \frac{n_i\bar{T}_A - n_i\bar{T}}{n\sqrt{p_i(1-p_i)}S_T} = \left(\frac{\bar{T}_A - \bar{T}}{S_T}\right)\sqrt{\frac{p_i}{1-p_i}} \quad (2.2.3)$$

onde

- p_i é proporção de candidatos com acerto no i -ésimo item;
- T_a é a média de escore dos candidatos com acerto no i -ésimo item;
- T é a média de escore dos candidatos;
- n_i é o número de respondentes com resposta positiva do i -ésimo item;
- S_T é o desvio padrão dos escores.

2.2.4 Modelo de Resposta Gradual

Por meio do MRG - Modelo de Resposta Gradual -, será possível avaliar a ordem entre as saídas “não resposta”, “errado”, “certo”, uma vez que, em itens do vestibular, uma resposta errada gera penalidade, ou seja, a não resposta pode significar tanto falta de conhecimento como medo da penalidade.

Desenvolvido por Samejima (1969), o MRG é utilizado para itens não dicotômicos, como é o vestibular da UnB (certo, errado e não resposta) e dado por

$$P_{i,k}^+(\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_{i,k})}}, \quad (2.2.4)$$

com $i = 1, 2, \dots, I$ e $j = 1, 2, \dots, n$, e $k = 1, 2, \dots, m_i$, onde

- $P_{i,k}^+(\theta_j)$ é a probabilidade de um indivíduo com proficiência θ_j escolher a categoria k ou superior no i -ésimo item.
- $b_{i,k}$ é o parâmetro de dificuldade da k -ésima categoria do i -ésimo item.

2.2.5 Estimação de Parâmetros

A estimação dos parâmetros nos Modelos de TRI é composta pelo cálculo de proficiência (habilidade) do candidato e pelo cálculo dos parâmetros do item. A estimação de proficiência dos candidatos ocorre utilizando métodos de Máxima Verossimilhança ou estimação Bayesiana. Por meio de cálculos iterativos é possível otimizar a estimação, pelos métodos Newton-Raphson ou Scoring de Fisher cuja exemplificação está representada na referência Anjos e Andrade (2012).

3 Resultados

3.1 Análise Descritiva

Nesta seção será apresentada a análise exploratória, com visualização geral dos dados e introduzindo as variáveis que serão estudadas ao longo do relatório.

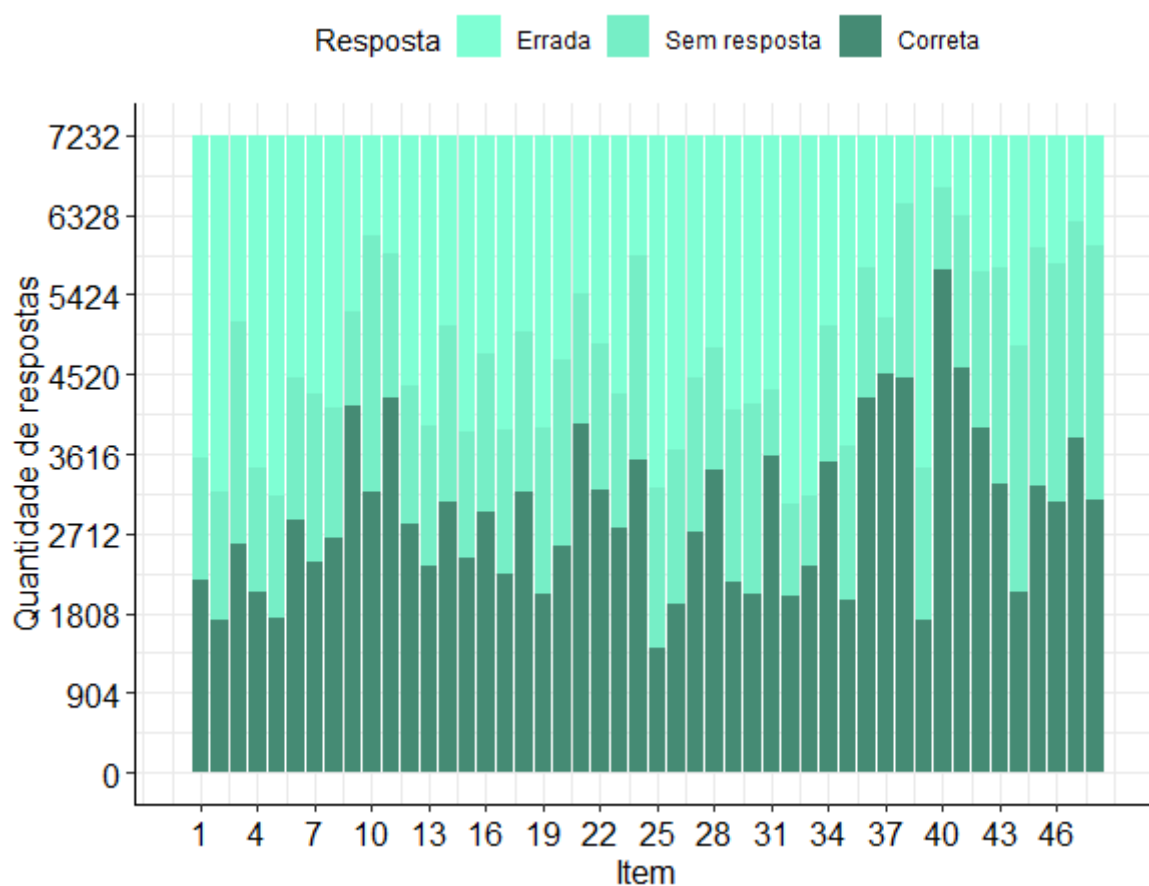


Figura 3: Distribuição das Respostas por Item

A Figura 3 apresenta a distribuição de respostas por item, com as categorias errada, sem resposta e correta. O item da prova com mais acertos foi o 40 (5718 acertos), o item com mais erros foi o item 32 (4187 erros) e o item que mais foi deixado em branco foi o item 10, o qual não foi respondido por 2893 dos 7232 candidatos.

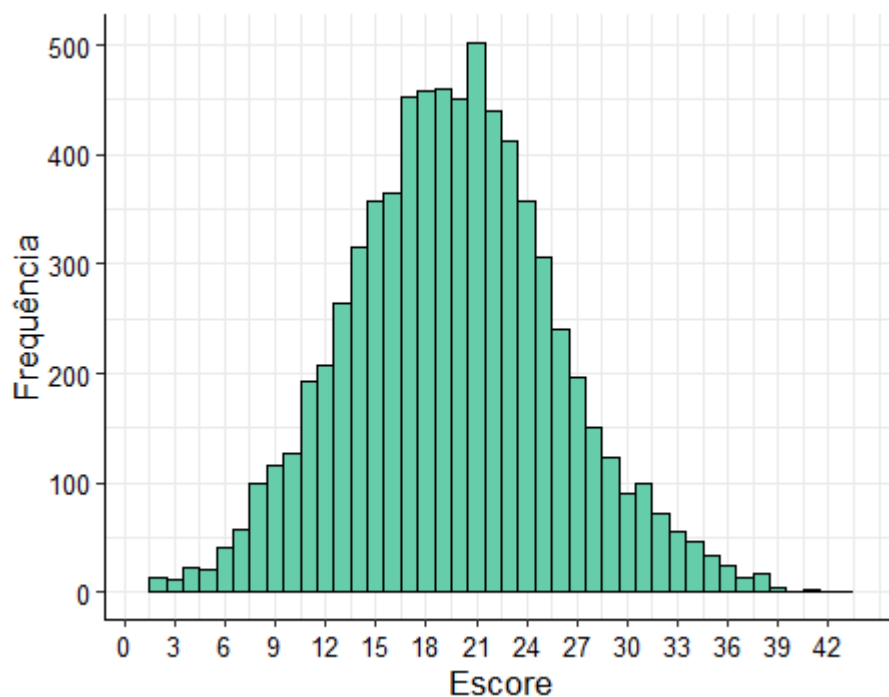


Figura 4: Distribuição do Escore

Na Figura 4 está apresentada a distribuição de frequências do escore dos candidatos na prova de língua portuguesa. Com 48 itens, o escore poderia variar entre 0 e 48, porém, o escore máximo obtido foi 44, com 1 observação. O escore mais frequente na prova foi 21, com frequência de 503 alunos.

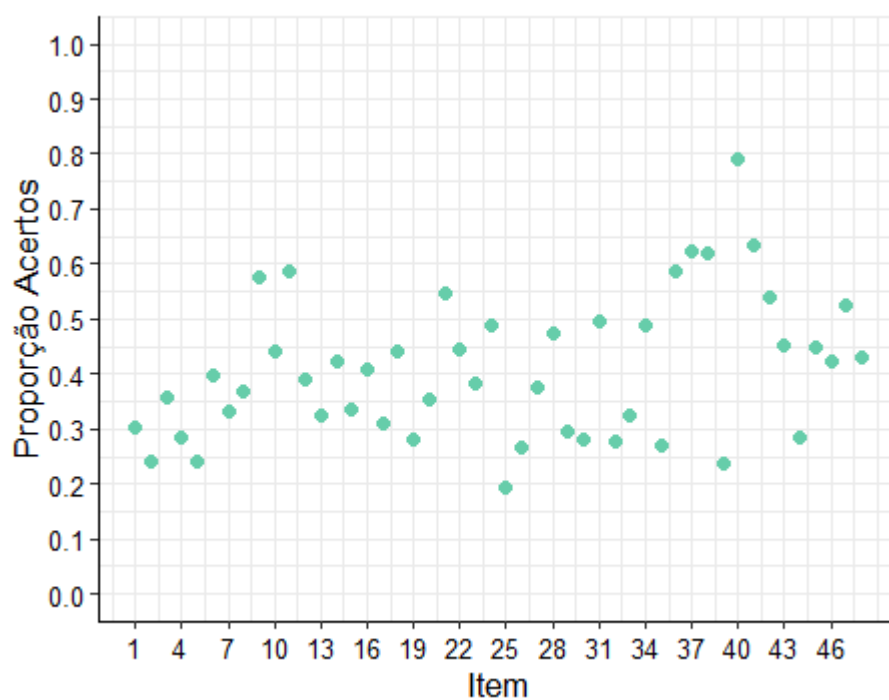
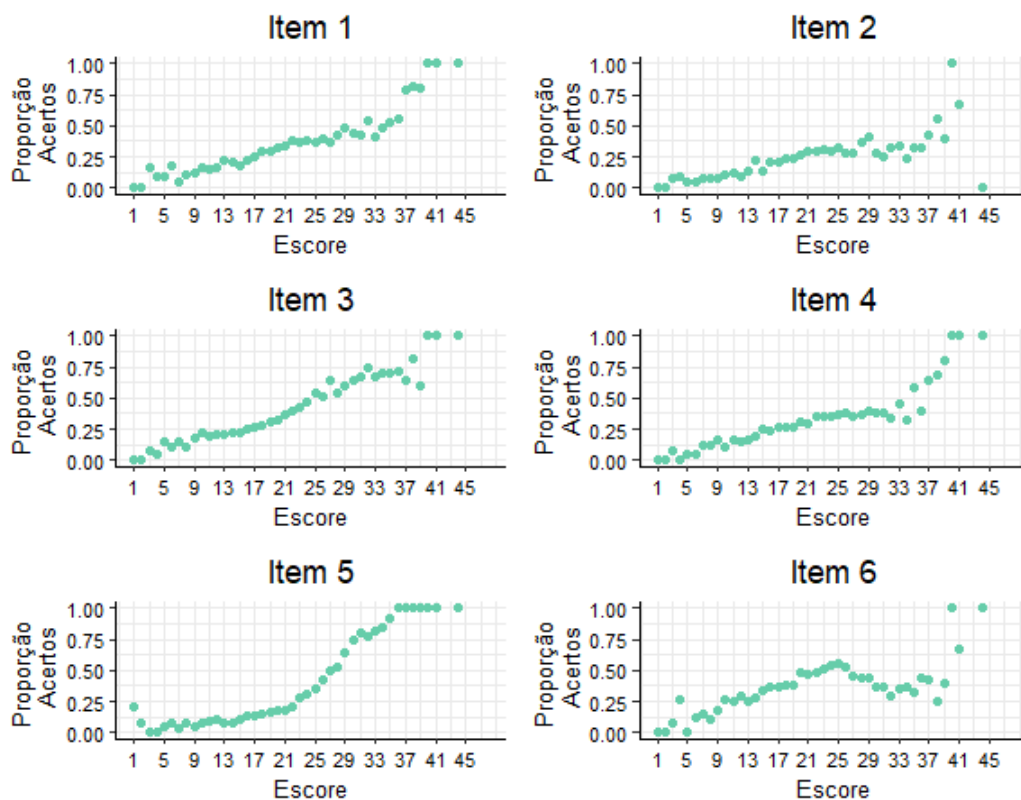
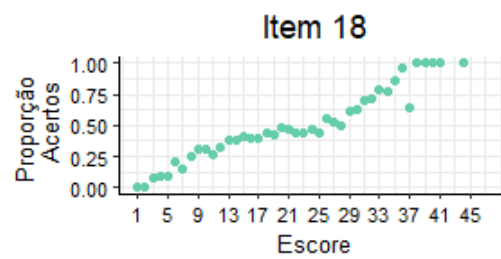
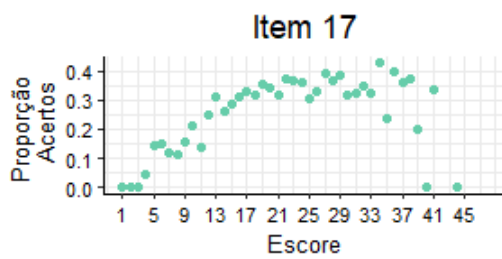
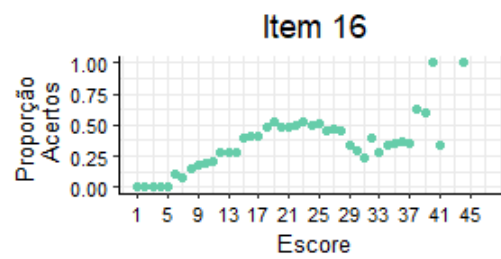
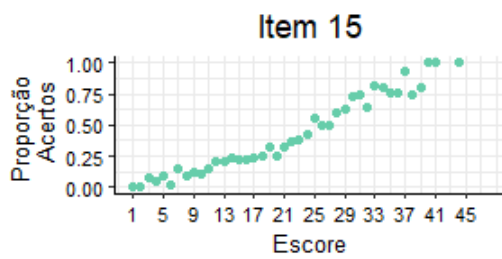
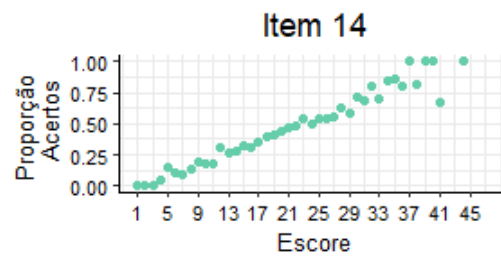
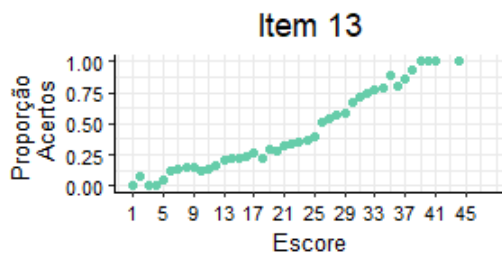
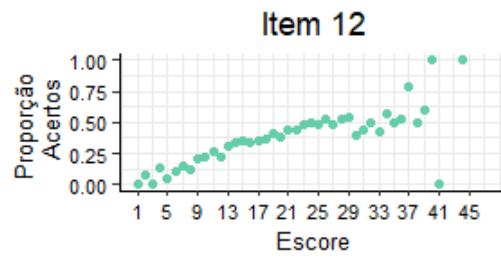
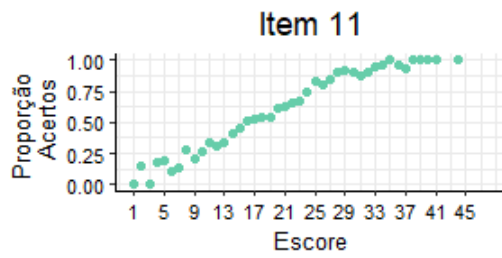
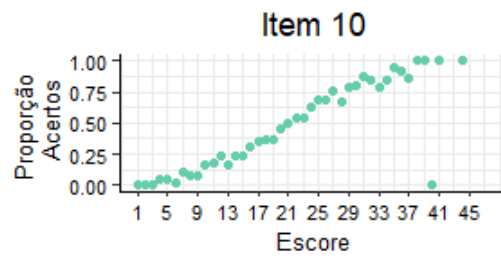
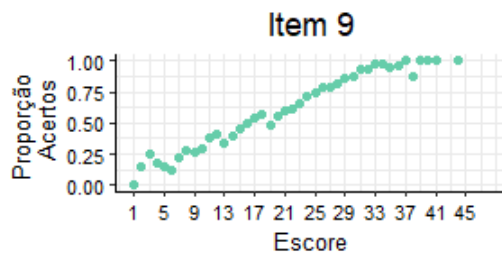
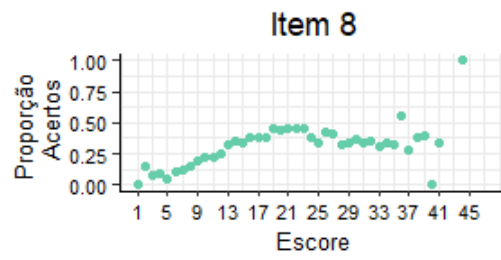
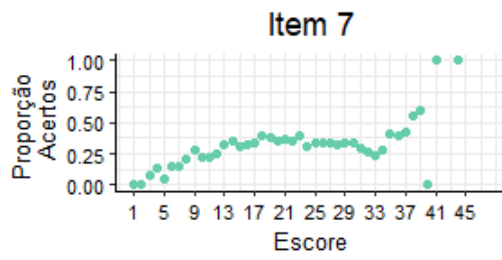
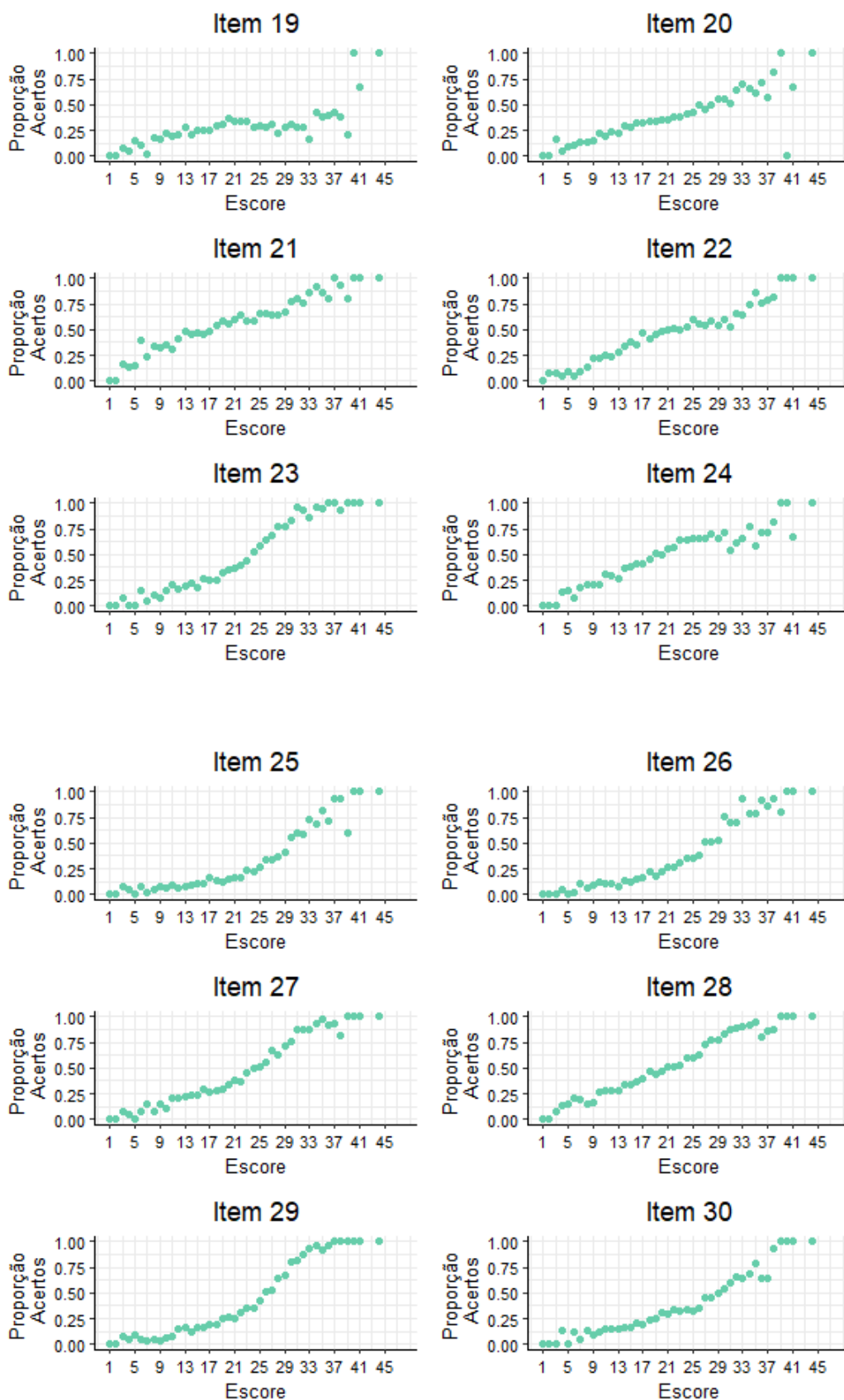


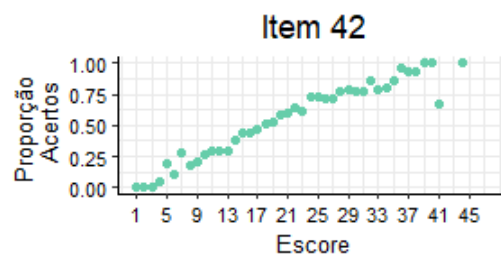
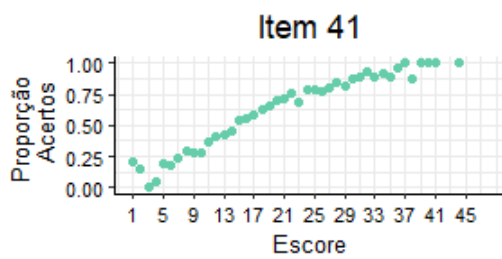
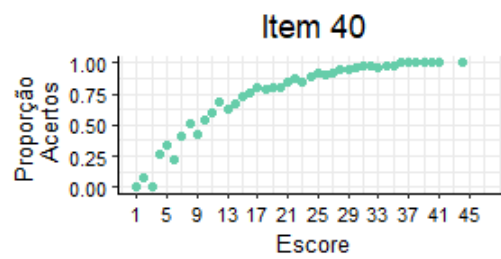
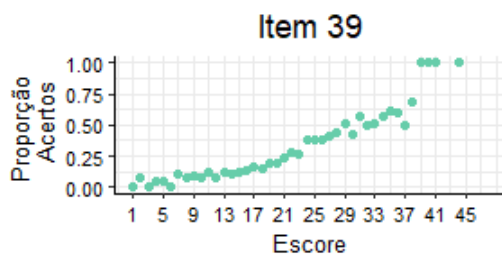
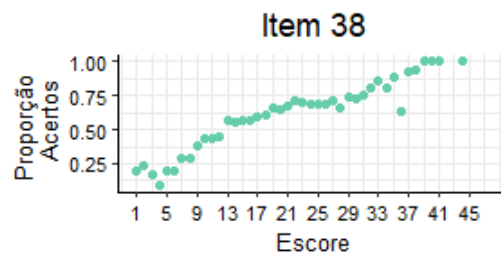
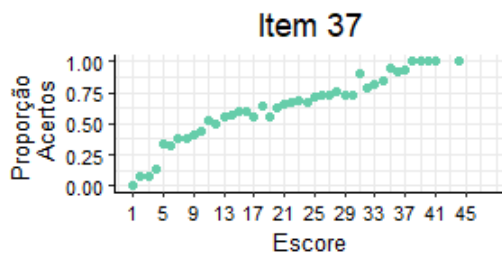
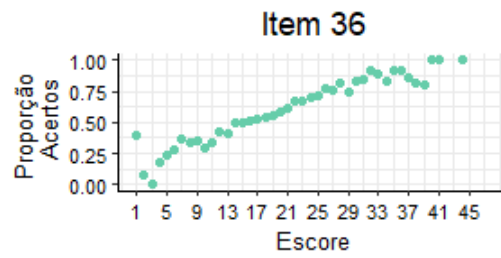
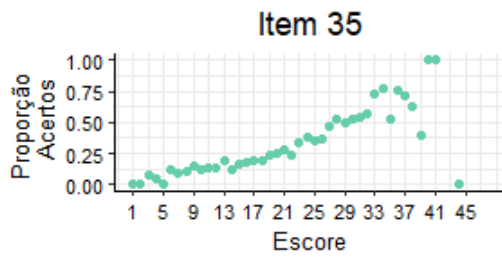
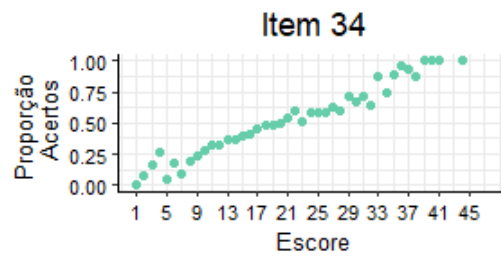
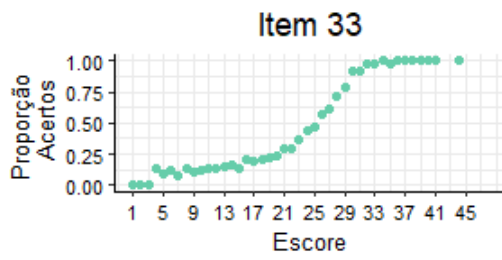
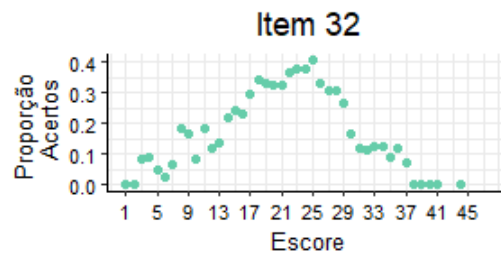
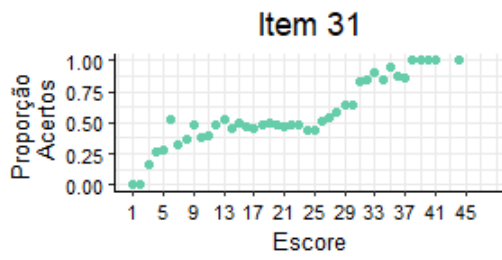
Figura 5: Proporção de Acertos Por Item

A Figura 5 representa a proporção de acertos de cada item da prova de português. Percebe-se que a maior parte dos itens obtiveram menos de 50% de acertos. Os itens que obtiveram mais de 50% de acertos foram: 47, 42, 21, 9, 36, 11, 39, 27, 41 e o item 40, que foi o mais acertado da prova com 79% de acerto.









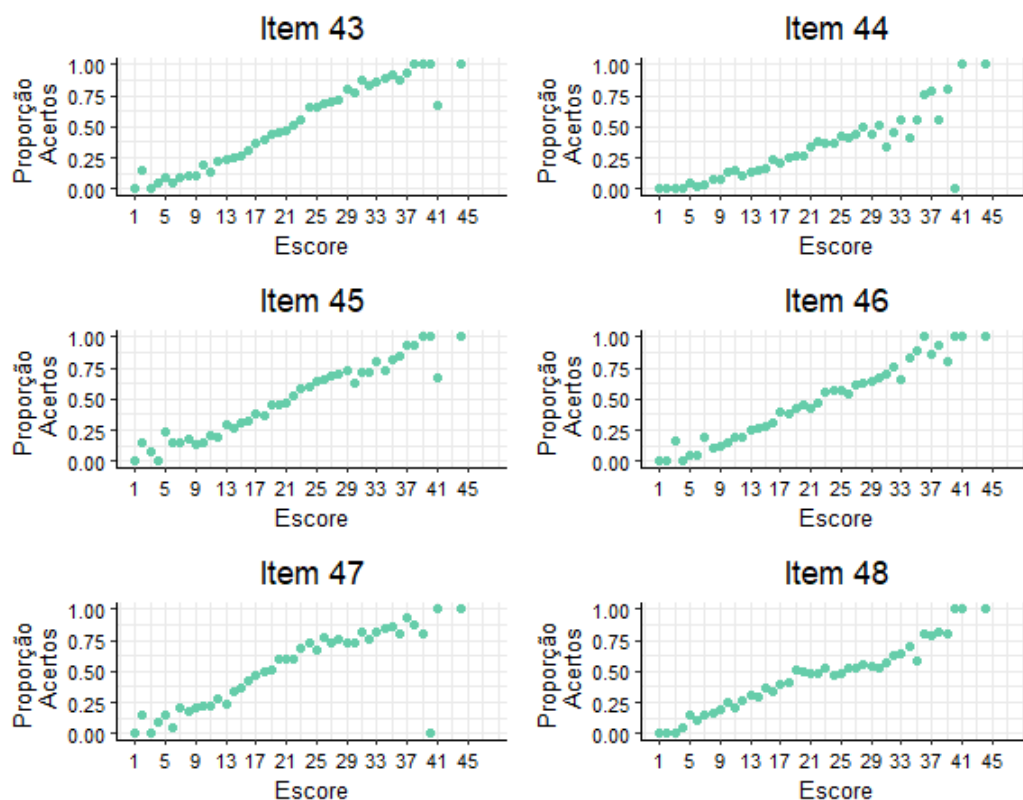


Figura 6: Curva Característica não Paramétrica

Na Figura 6, temos uma representação da CCI (Curva Característica do Item), de forma não paramétrica. Os gráficos estão separados por item, de forma que no eixo das abscissas os candidatos foram separados por escore, e no eixo das ordenadas está representada a proporção de acerto que o grupo de alunos teve no respectivo item.

Por meio dele é possível perceber se alunos com escore maior na prova de língua portuguesa tiveram uma maior taxa de acerto, item por item. Nos itens 7, 16, 17, 31 e 32 tal fato não pode ser notado com clareza e foram observados separadamente durante a construção dos modelos. Como representado na Figura 4 poucos alunos obtiveram escore igual ou superior a 33, portanto é notado um comportamento irregular da proporção de acerto do itens para tais escores.

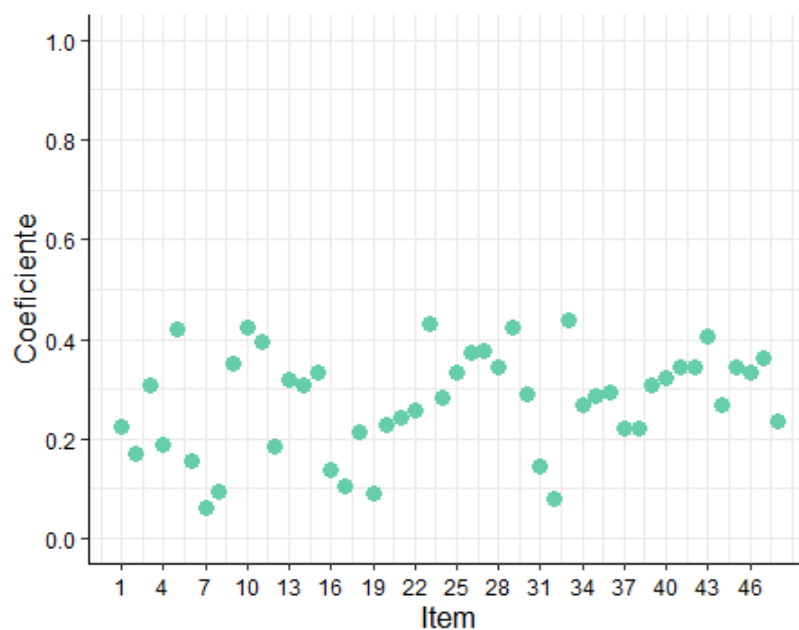


Figura 7: Coeficiente de Correlação Ponto-Biserial por Item

Na Figura 7 representa-se o coeficiente de correlação ponto-biserial por item. No caso foi correlacionado acerto e não acerto (erro ou sem resposta) de cada item com o escore do indivíduo na prova de português, gerando o coeficiente por item.

Percebe-se que nenhum item obteve coeficiente menor que 0, e os itens com menor coeficiente são os itens 7 e 32, que também apresentaram comportamento irregular na CCI não paramétrica (Figura 6).

3.2 Ajuste dos Modelos

Nesta seção será apresentada a análise dos modelos de TRI.

3.2.1 Ajuste do Modelo Logístico de 3 Parâmetros

Item	Acerto ao Acaso	Dificuldade	Discriminação
1	0,28	2,05	4,81
2	0,23	2,46	3,13
3	0,26	-1,45	-2,77
4	0,00	-1,68	-0,63
5	0,09	-1,20	-4,72
6	0,00	-1,30	-0,35
7	0,20	1,40	1,68
8	0,26	1,35	2,78
9	0,00	0,18	-1,81
10	0,00	-0,35	-1,17
11	0,07	0,12	-1,58
12	0,02	-2,49	-0,21
13	0,24	-1,46	-3,28
14	0,37	-1,60	-3,87
15	0,15	-1,36	-1,35
16	0,00	1,05	0,34
17	0,23	1,62	2,70
18	0,36	1,60	1,95
19	0,17	1,38	3,32
20	0,00	-7,26	-0,08
21	0,46	1,23	9,54
22	0,00	-0,57	-0,48
23	0,20	-0,99	-4,79
24	0,01	-0,23	-0,38
25	0,09	-1,50	-2,88
26	0,00	-1,17	-1,23
27	0,10	-0,95	-1,34
28	0,38	-1,29	-15,94
29	0,18	-1,31	-17,58
30	0,21	-1,54	-3,77
31	0,34	0,74	16,59
32	0,00	-3,03	-0,34
33	0,09	-0,87	-5,57
34	0,00	50,00	0,00
35	0,18	-1,66	-2,15
36	0,00	0,58	-0,60
37	0,30	-22,29	0,00
38	0,57	1,48	3,66
39	0,00	-1,66	-0,86
40	0,00	2,62	-0,53
41	0,56	-1,44	-1,64
42	0,04	0,05	-0,85
43	0,00	-0,35	-0,83
44	0,05	-3,72	-0,31
45	0,39	-1,57	-3,01
46	0,00	-0,70	-0,52
47	0,22	-0,68	-0,87
48	0,37	1,70	2,43

Tabela 1: Parâmetros Modelo ML3

A Tabela 1 mostra os resultados do modelo ajustado de 3 parâmetros (Acerto ao Acaso, Dificuldade e Discriminação). Os valores estimados para os parâmetros dos itens indicam valores aberrantes (módulo superior a 7) para estimativas do parâmetro de dificuldade para os itens 20, 34, 37, e para o parâmetro de discriminação para os itens de 1 a 6, 9 a 15, 20 a 30, 32 a 47, e de acerto ao acaso extremamente alto para os itens de 1 a 3, 7, 8, 13 a 15, 17 a 19, 21, 23, 27 a 31, 35, 37, 38, 41, 45, 47 e 48, indicando ajuste inadequado do modelo ML3.

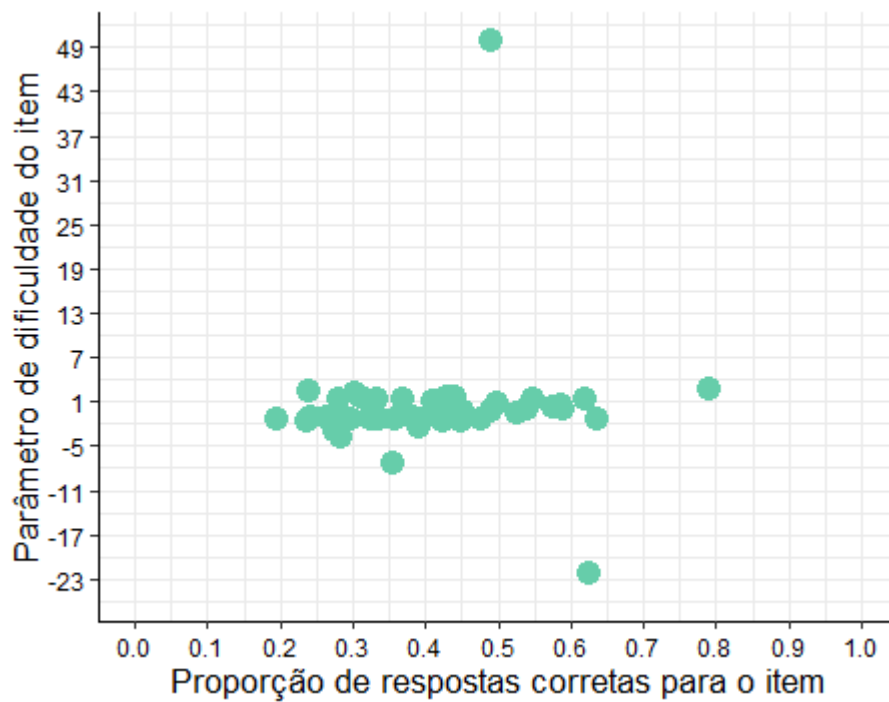
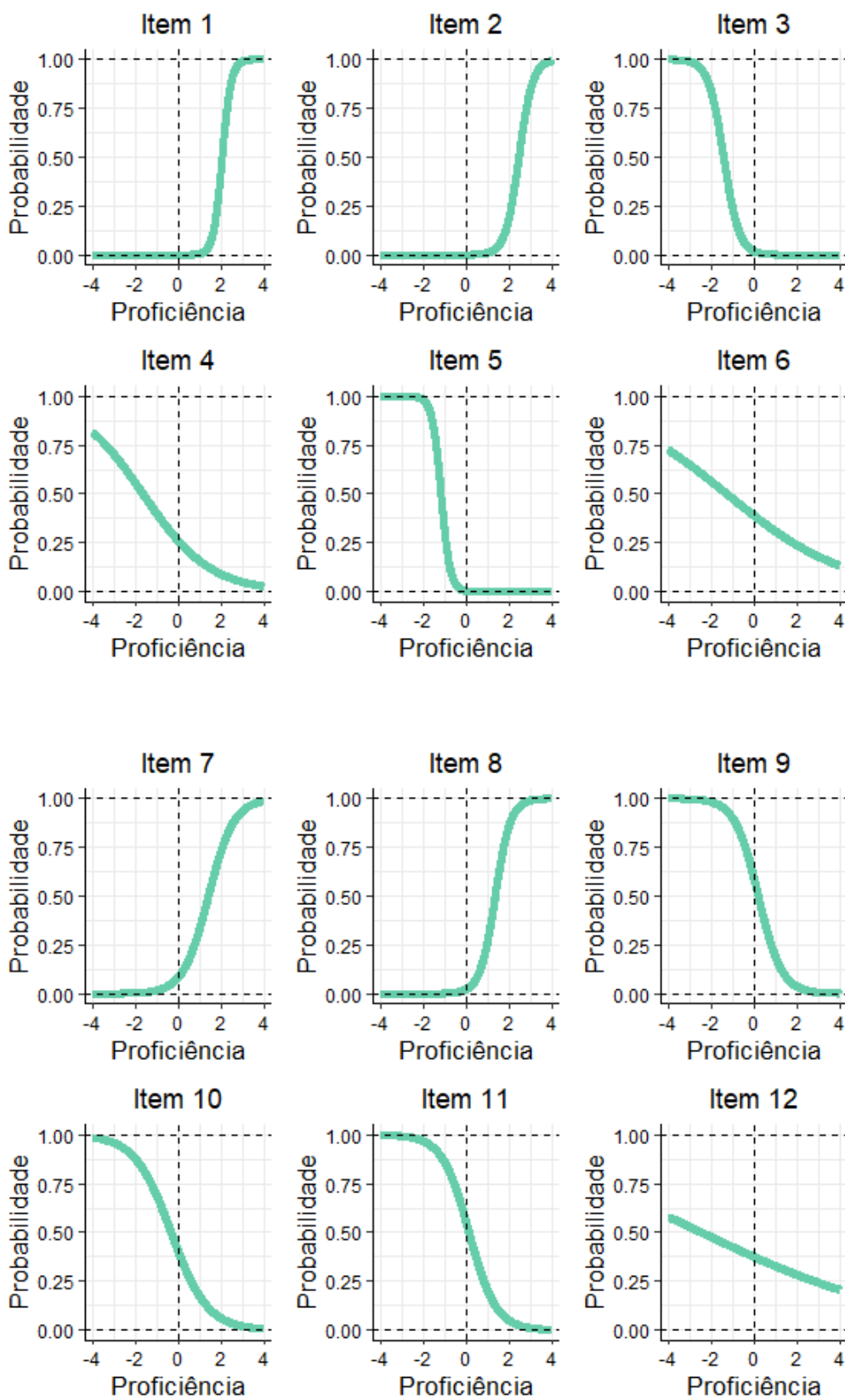
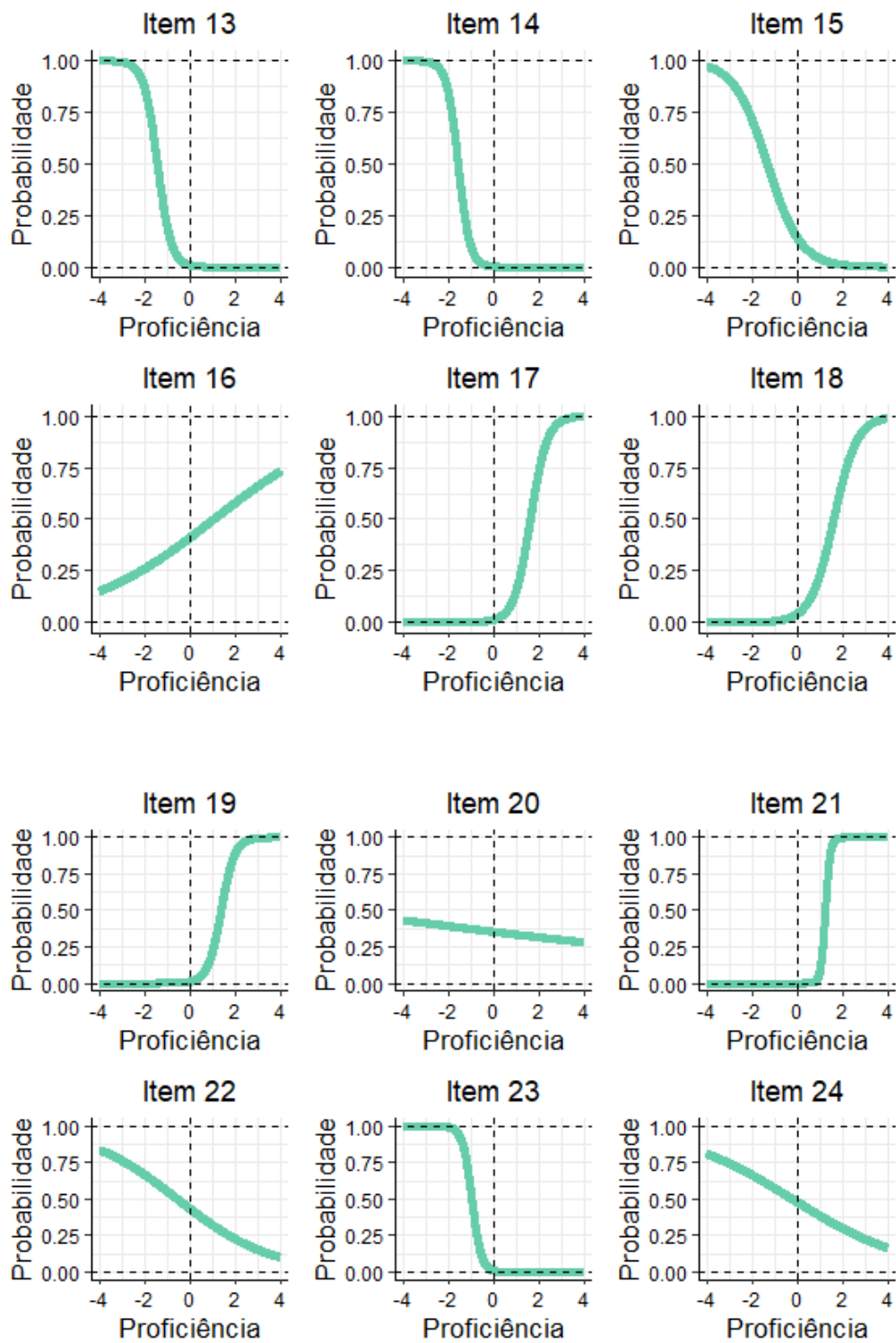
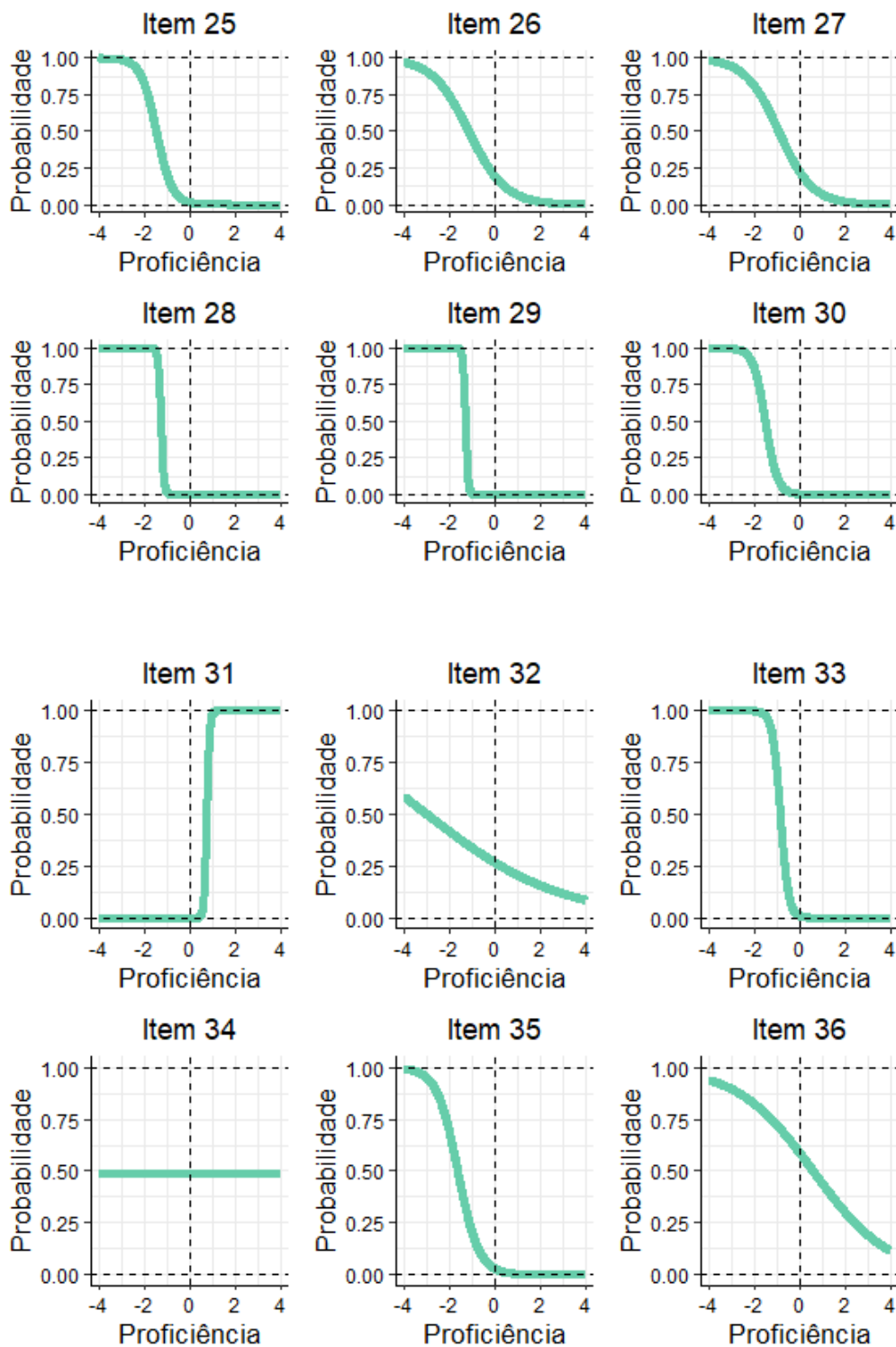


Figura 8: Proporção de acertos por Parâmetro de Dificuldade

Na Figura 8 está representada a proporção de acertos pelo parâmetro de dificuldade do item. Os itens 20, 34 e 37 destacam-se pela dificuldade e o item 40 destaca-se pela grande quantidade de acertos. A não relação entre as duas variáveis é mais um indicativo de que o modelo ML3 não se adequou completamente aos dados.







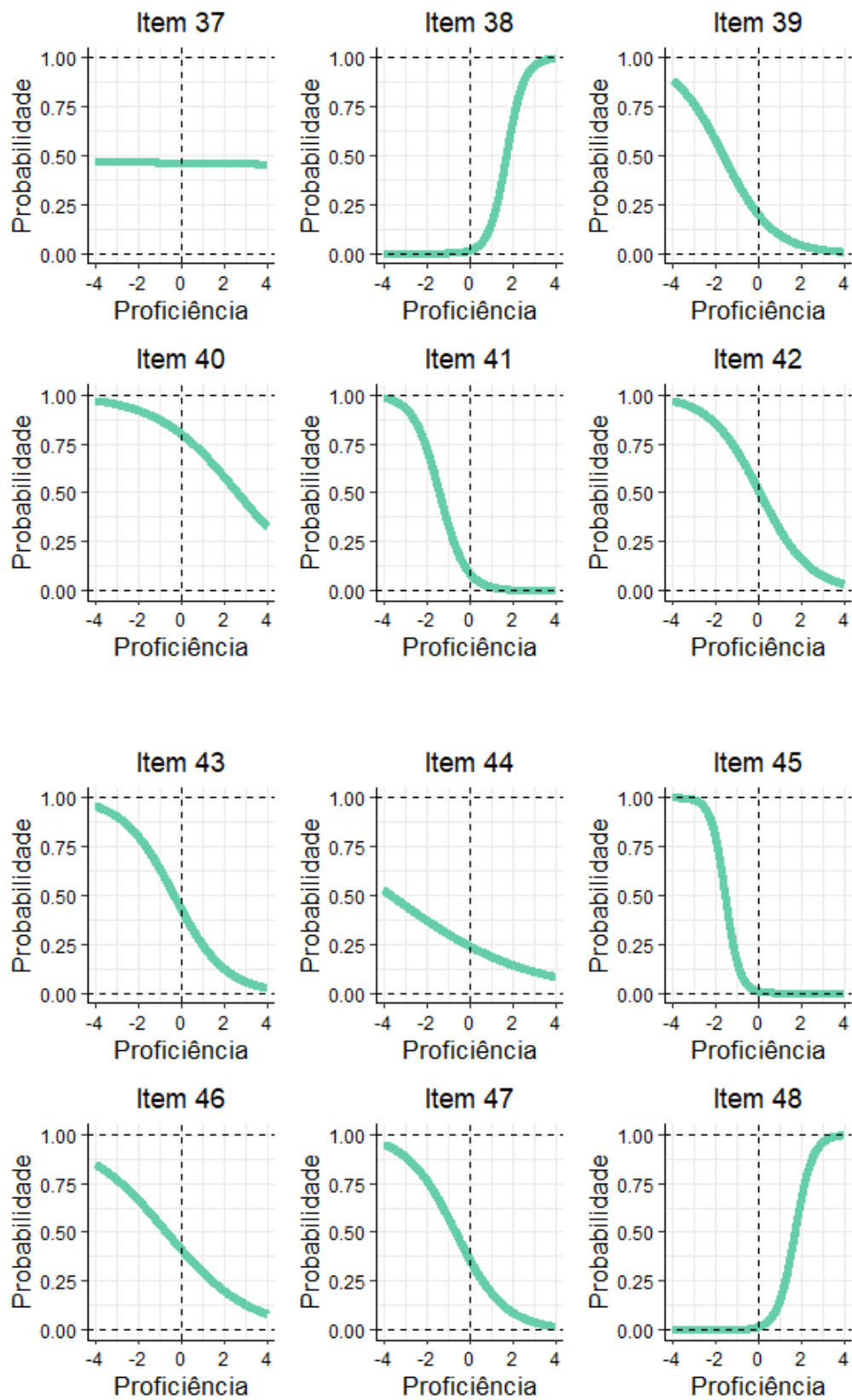


Figura 9: Curvas Características dos Itens para o modelo ML3

Os gráficos da Figura 9 apresentam a CCI de cada item para o modelo logístico de 3 parâmetros. Nela percebe-se o que havia sido destacado com a tabela - valores incongruentes de discriminação (negativos) para os quais a probabilidade do candidato acertar o item diminui à medida que sua proficiência aumenta, o que mostra que o modelo ML3 é inadequado.

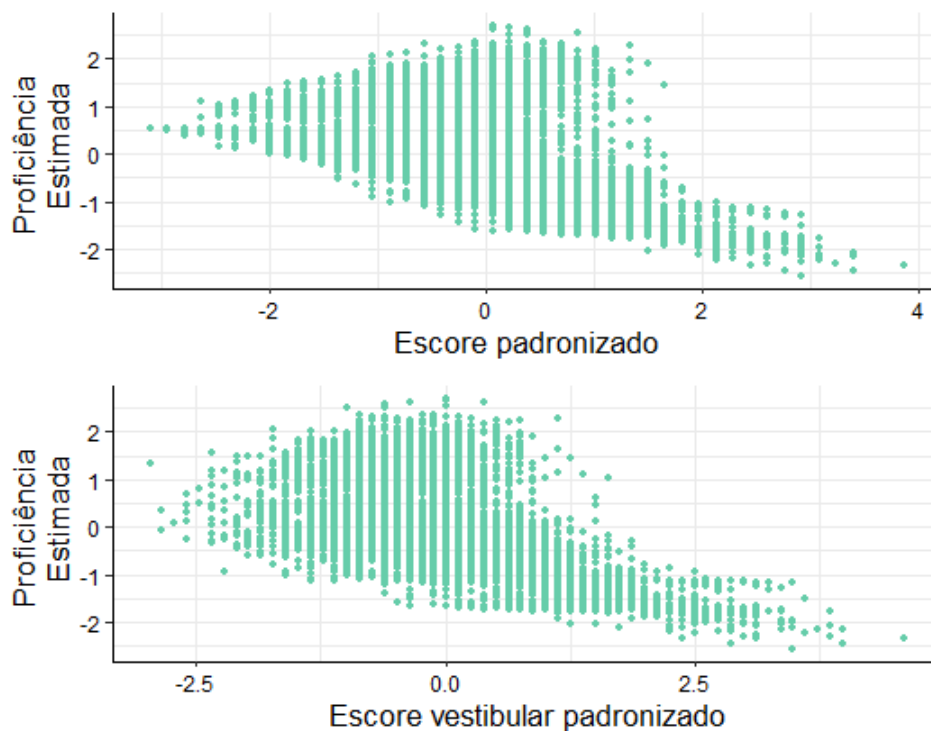


Figura 10: Proficiência Estimada por Escore e Escore do Vestibular Padronizados do Modelo ML3

Na Figura 10 está representada a Proficiência Estimada (do modelo ML3) por Escore Padronizado dos alunos, no gráfico superior o escore na prova de língua portuguesa e abaixo o escore padronizado do vestibular (pontuação do vestibular do candidato subtraído da média e dividido pelo desvio padrão). Percebe-se que o modelo não se ajustou adequadamente aos dados com relação negativa entre as variáveis proficiência e escore, quando o esperado era que proficiências maiores estejam acompanhadas de escores maiores.

3.2.2 Ajuste do Modelo Logístico de 2 Parâmetros

Nesta seção, será analisado o modelo retirando-se o Acerto ao Acaso e considerando os parâmetros Dificuldade e Discriminação.

Item	Dificuldade	Discriminação
1	22,94	0,04
2	35,56	0,03
3	1,09	0,59
4	1,69	0,59
5	1,05	1,55
6	1,20	0,36
7	-1,29	-0,58
8	-1,16	-0,48
9	-0,24	1,95
10	0,25	1,36
11	-0,31	1,64
12	1,79	0,25
13	1,39	0,56
14	1,66	0,19
15	0,96	0,81
16	-1,65	-0,22
17	-2,35	-0,35
18	-1,61	-0,15
19	-1,75	-0,58
20	3,26	0,18
21	1,06	-0,18
22	0,46	0,53
23	0,52	1,21
24	0,08	0,50
25	1,63	1,05
26	1,02	1,33
27	0,59	1,07
28	0,23	0,45
29	0,98	1,08
30	2,70	0,36
31	-0,03	-0,38
32	3,51	0,28
33	0,60	2,19
34	0,38	0,12
35	1,85	0,57
36	-0,55	0,71
37	-5,89	0,09
38	4,00	-0,12
39	1,54	0,88
40	-2,10	0,70
41	-1,43	0,40
42	-0,20	0,93
43	0,25	0,97
44	2,63	0,36
45	0,59	0,36
46	0,54	0,63
47	-0,15	0,74
48	-2,74	-0,10

Tabela 2: Parâmetros Modelo ML2

Na Tabela 2, estão representados os resultados do modelo de 2 parâmetros (ML2). Percebe-se que houve diferença em relação à Tabela 1, uma vez que o parâmetro Acerto ao Acaso apresentou valores consideráveis (módulo superior a 10) e incongruentes.

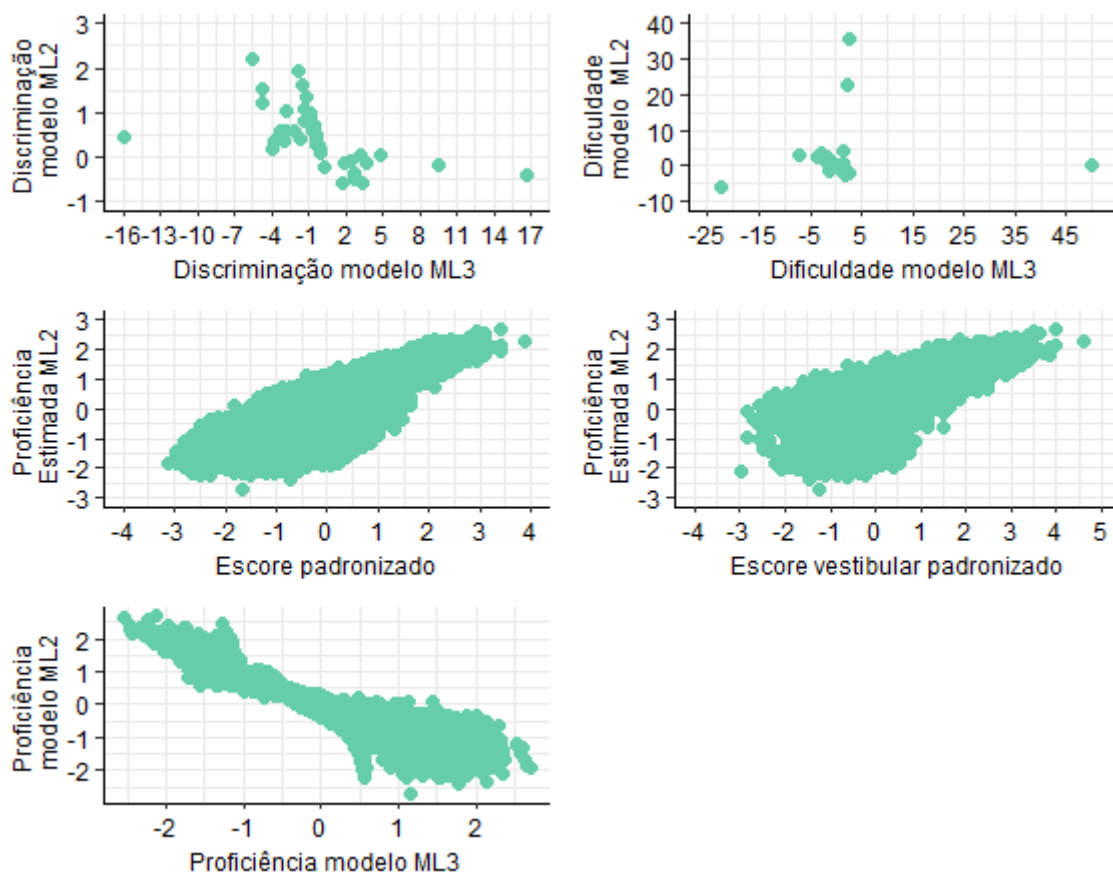


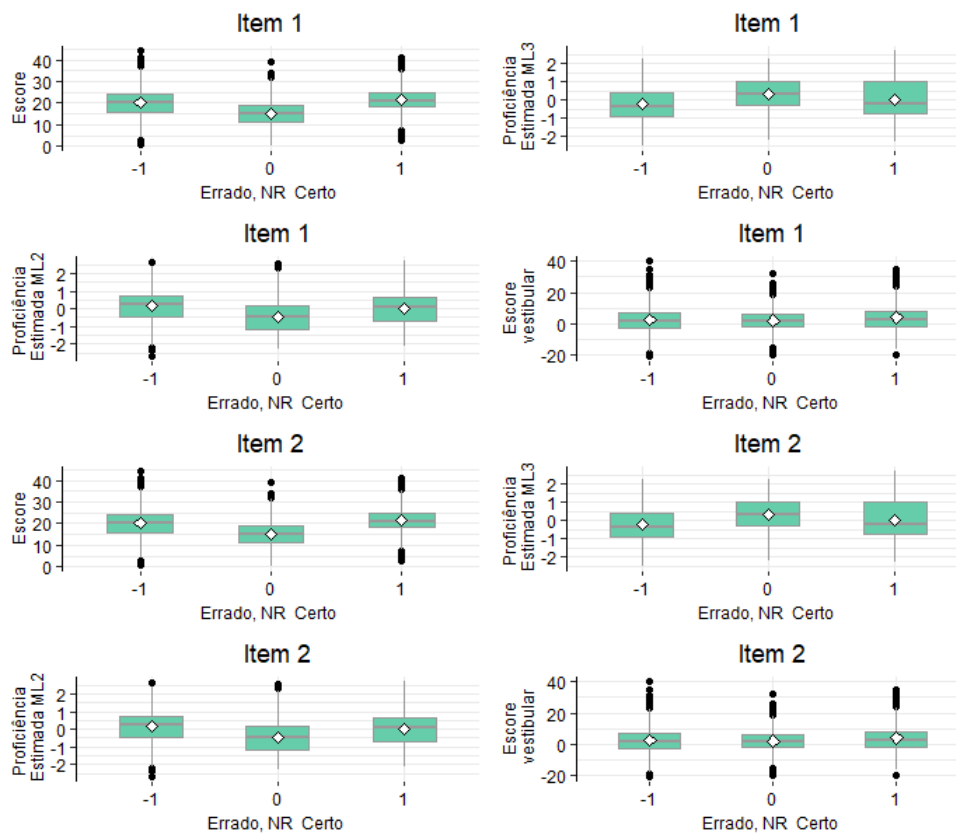
Figura 11: Comparativo dos Parâmetros para os Modelos com 2 e 3 Parâmetros

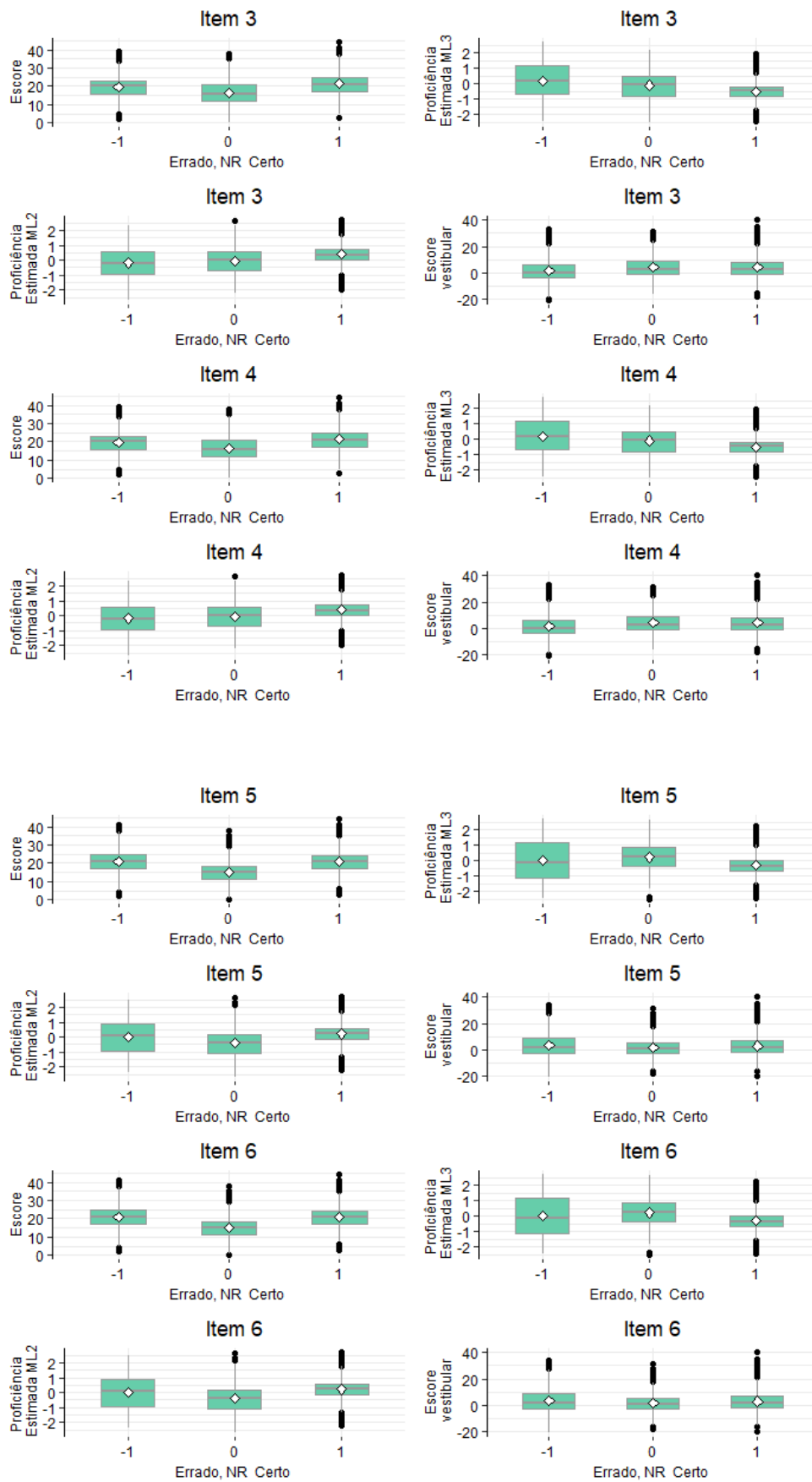
Como citado, pela Figura 11 percebe-se diferença significativa entre os parâmetros nos modelos ML2 e ML3. Nos dois primeiros gráficos da figura estão reveladas as diferenças entre as variáveis Discriminação e Dificuldade, respectivamente, causadas pelo mal ajuste dos modelos.

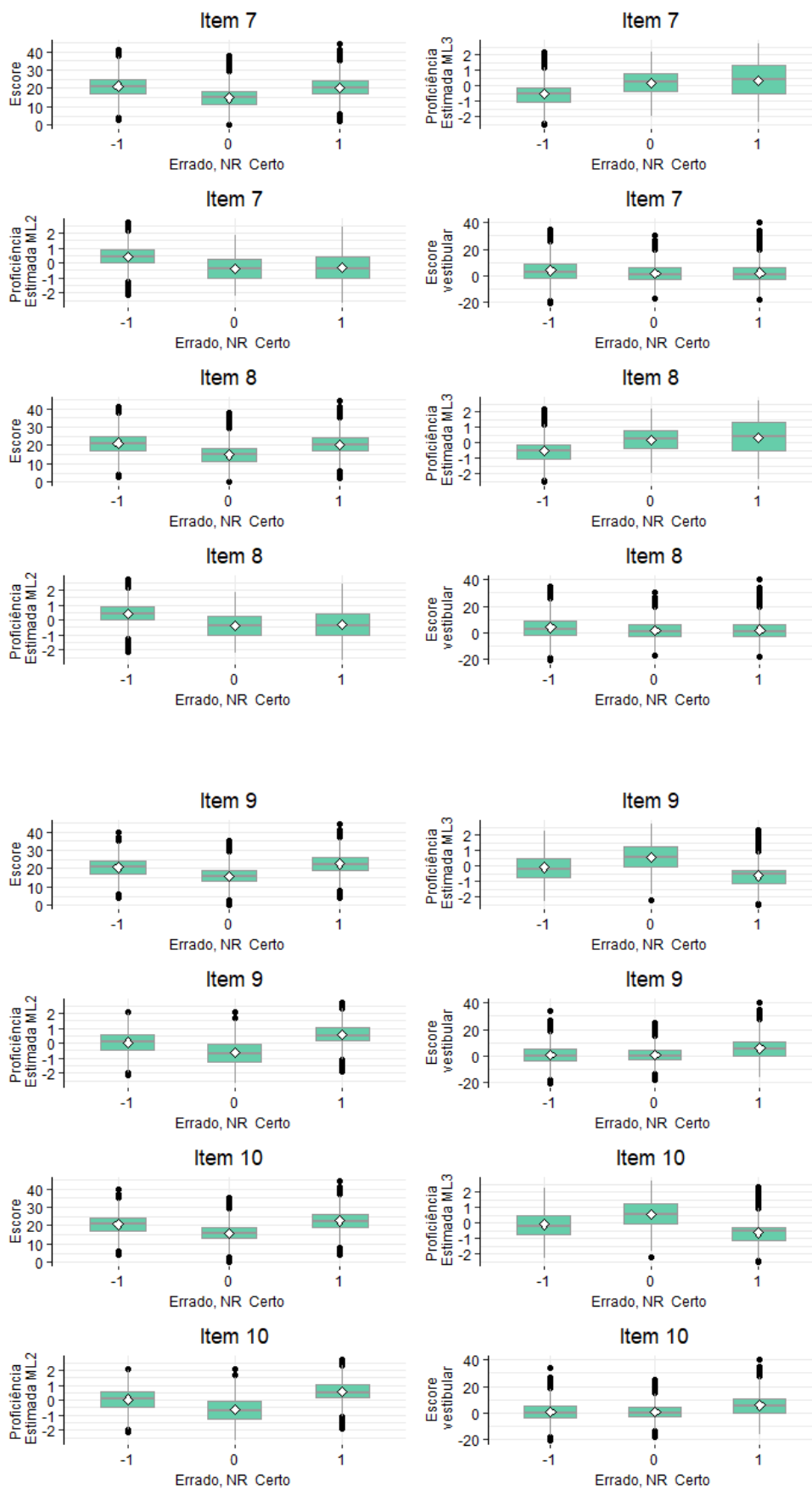
Nos gráficos 3 e 4 estão representadas as proficiências estimadas do Modelo ML2 pelo escore padronizado e escore do vestibular padronizado, respectivamente. Embora grande dispersão dos dados, percebe-se uma melhora em relação ao Modelo ML3, com relação positiva entre proficiência e escore.

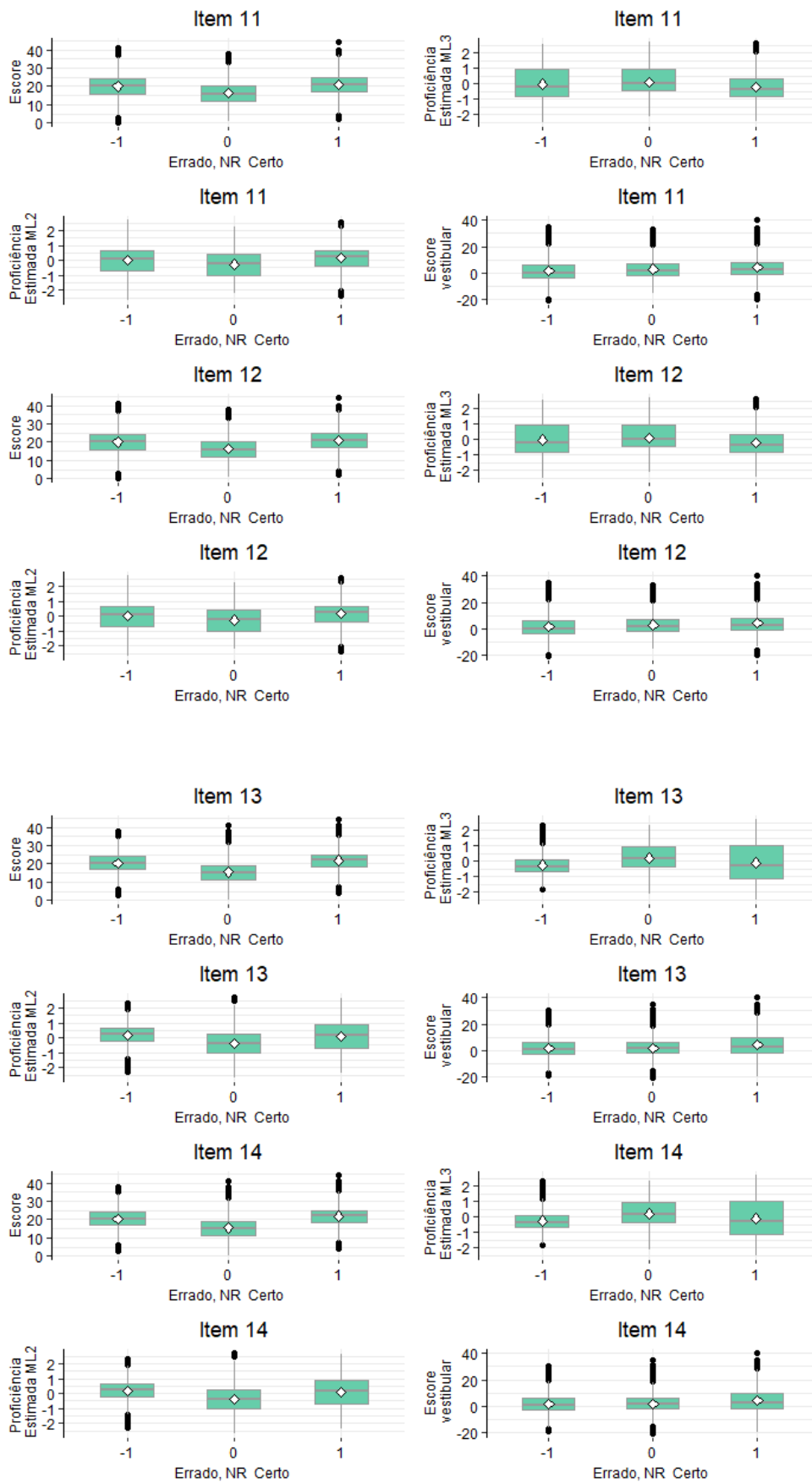
3.3 Modelo de Resposta Gradual

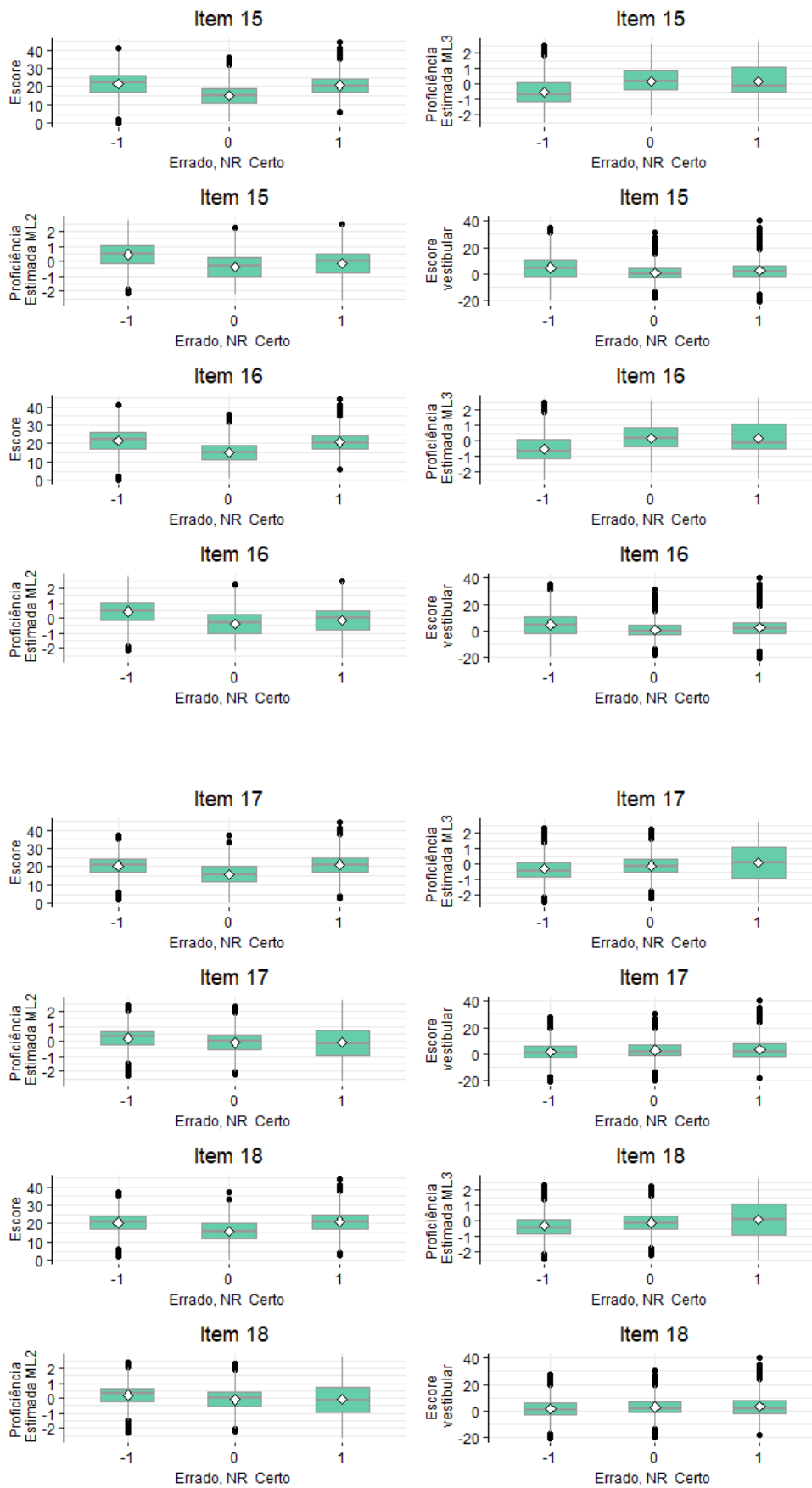
Nessa seção será apresentado o resultado do Modelo de Resposta Gradual. Inicialmente a introdução ao modelo e posteriormente o ajuste.

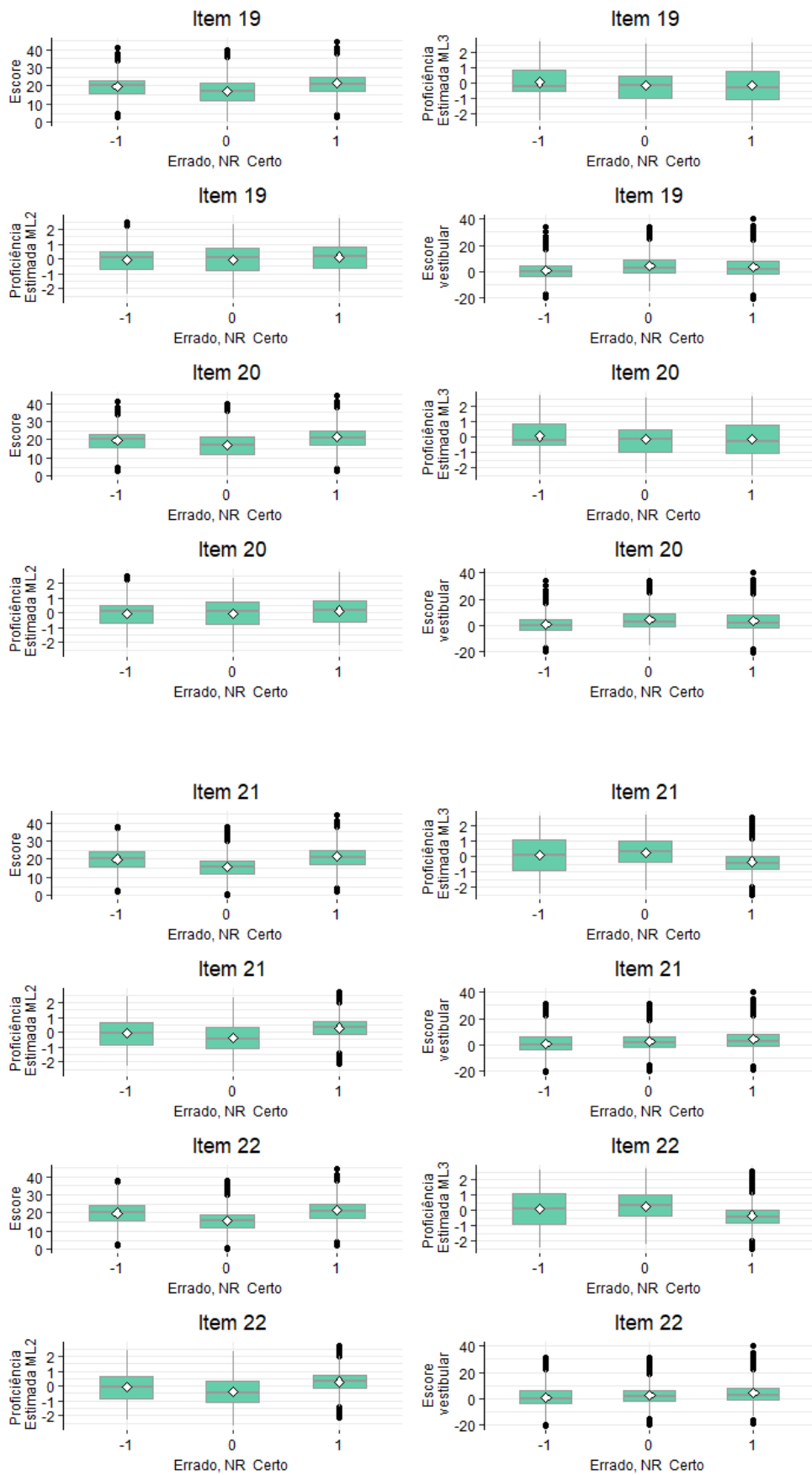


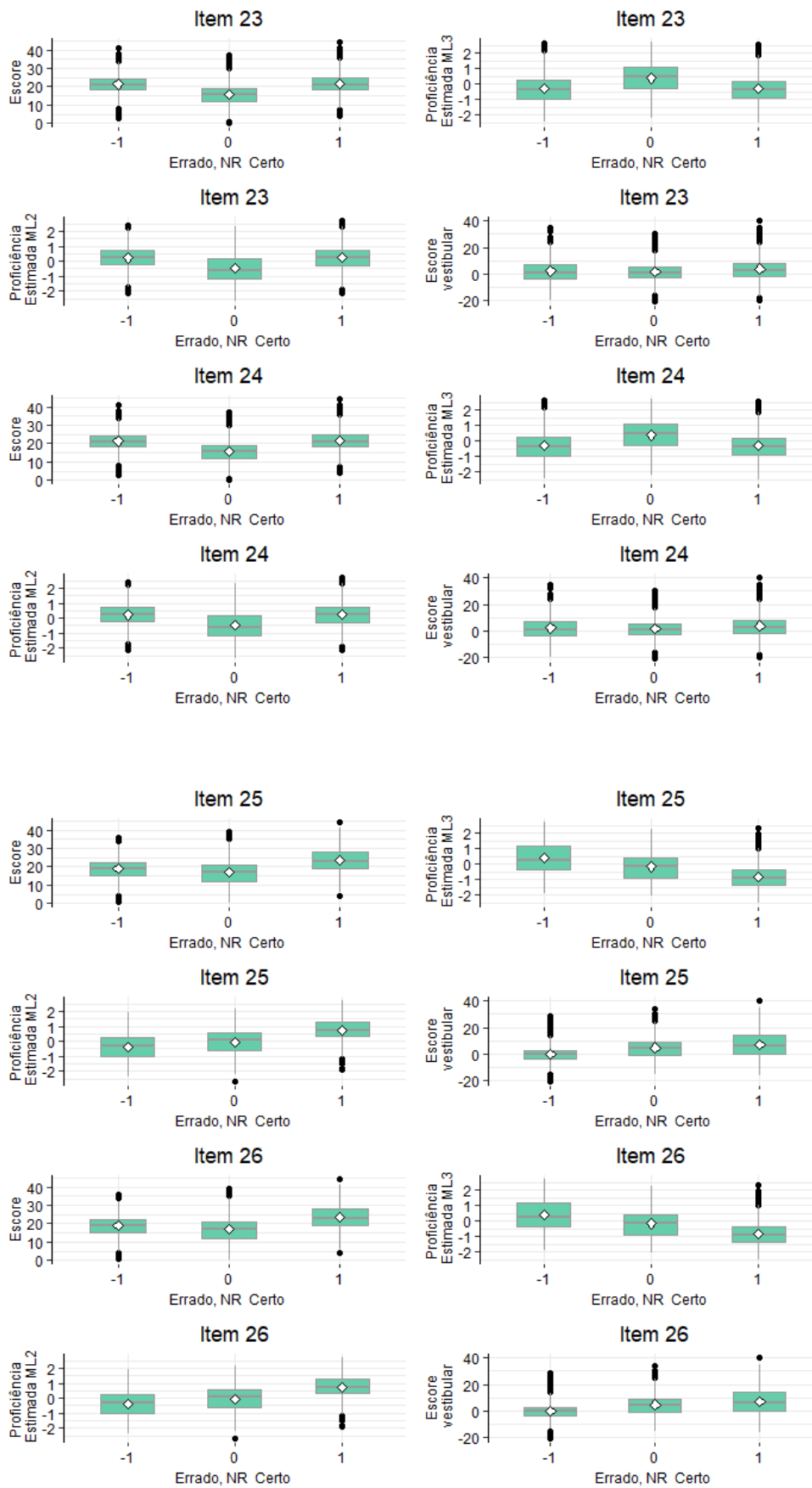


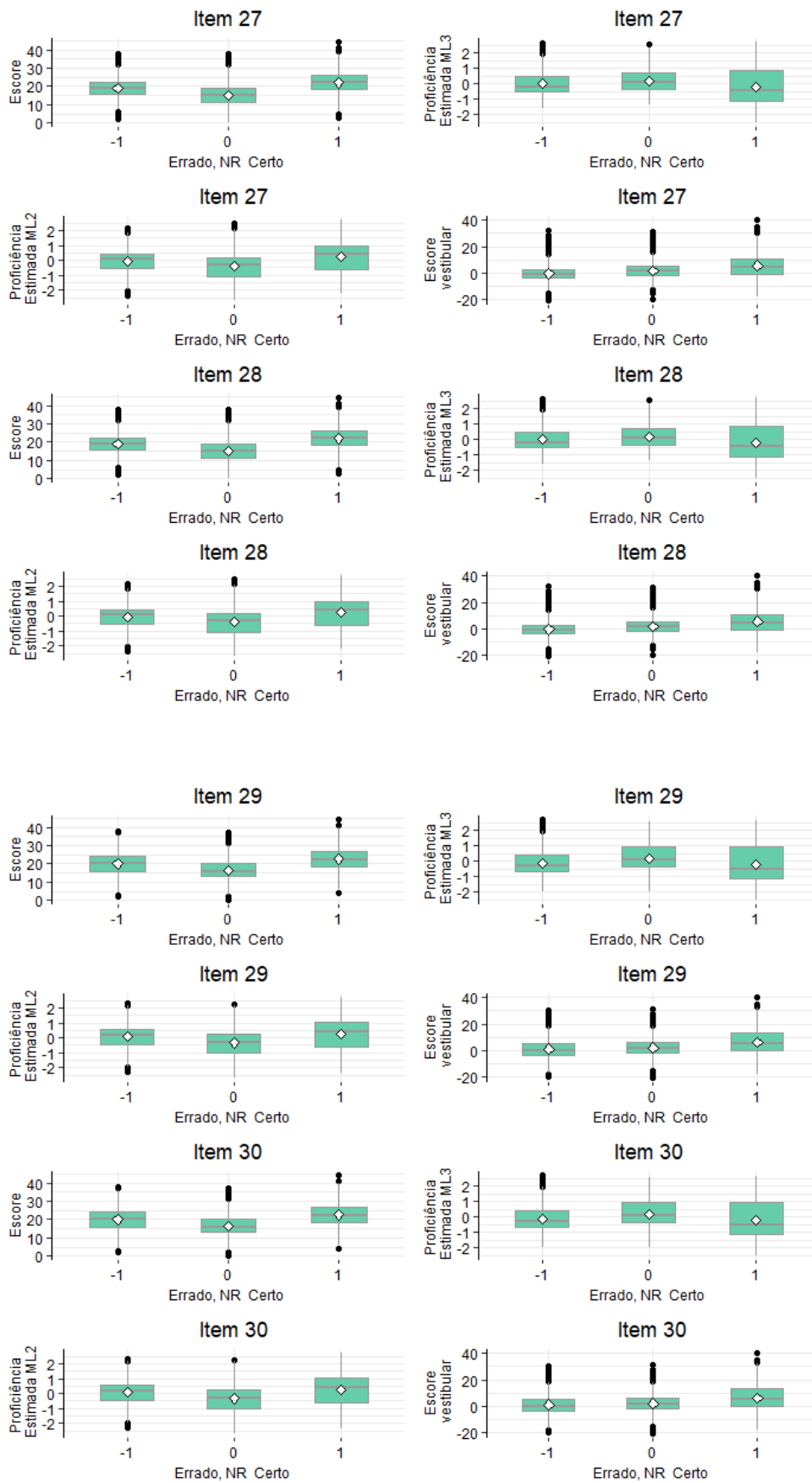


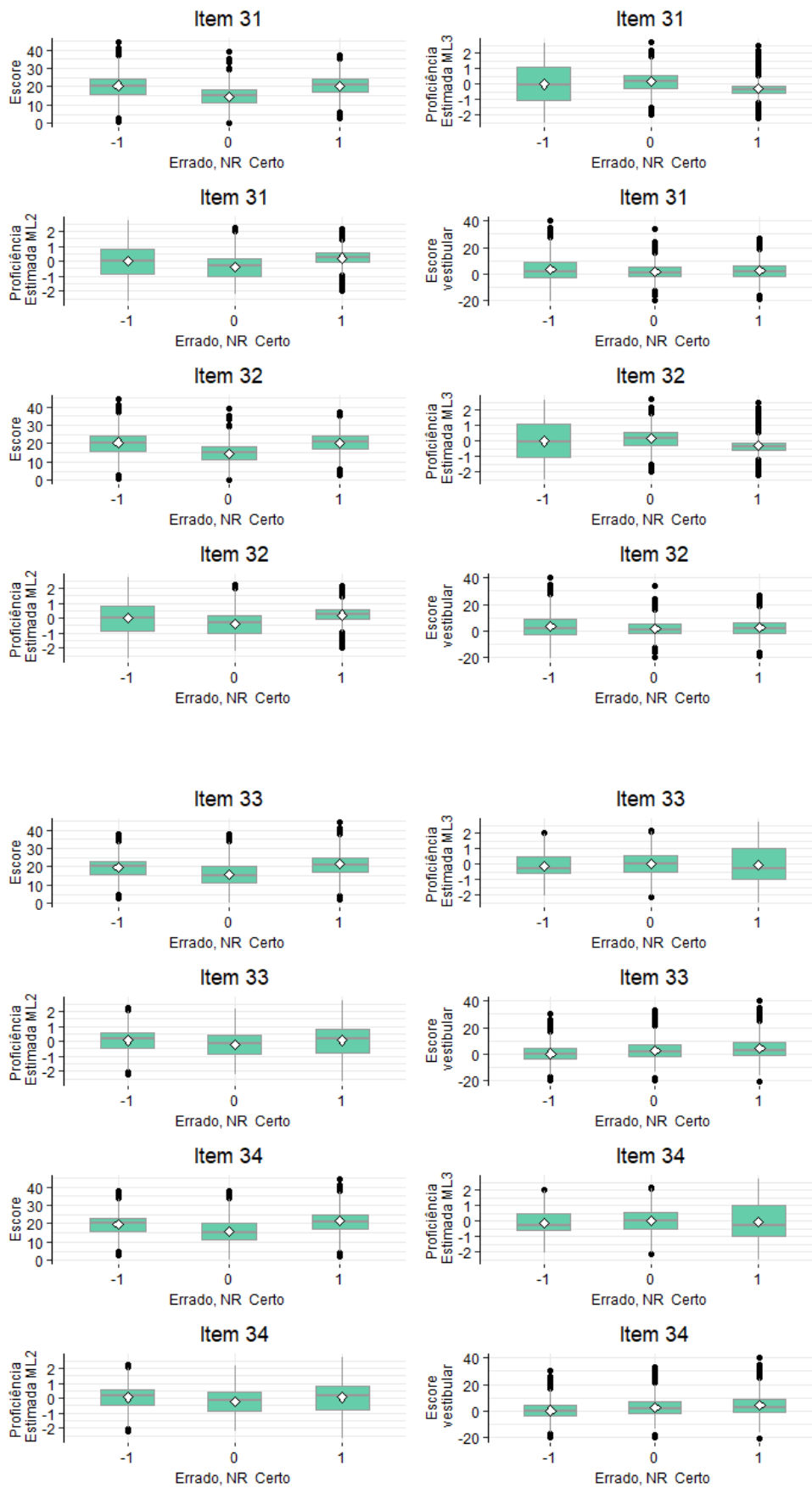


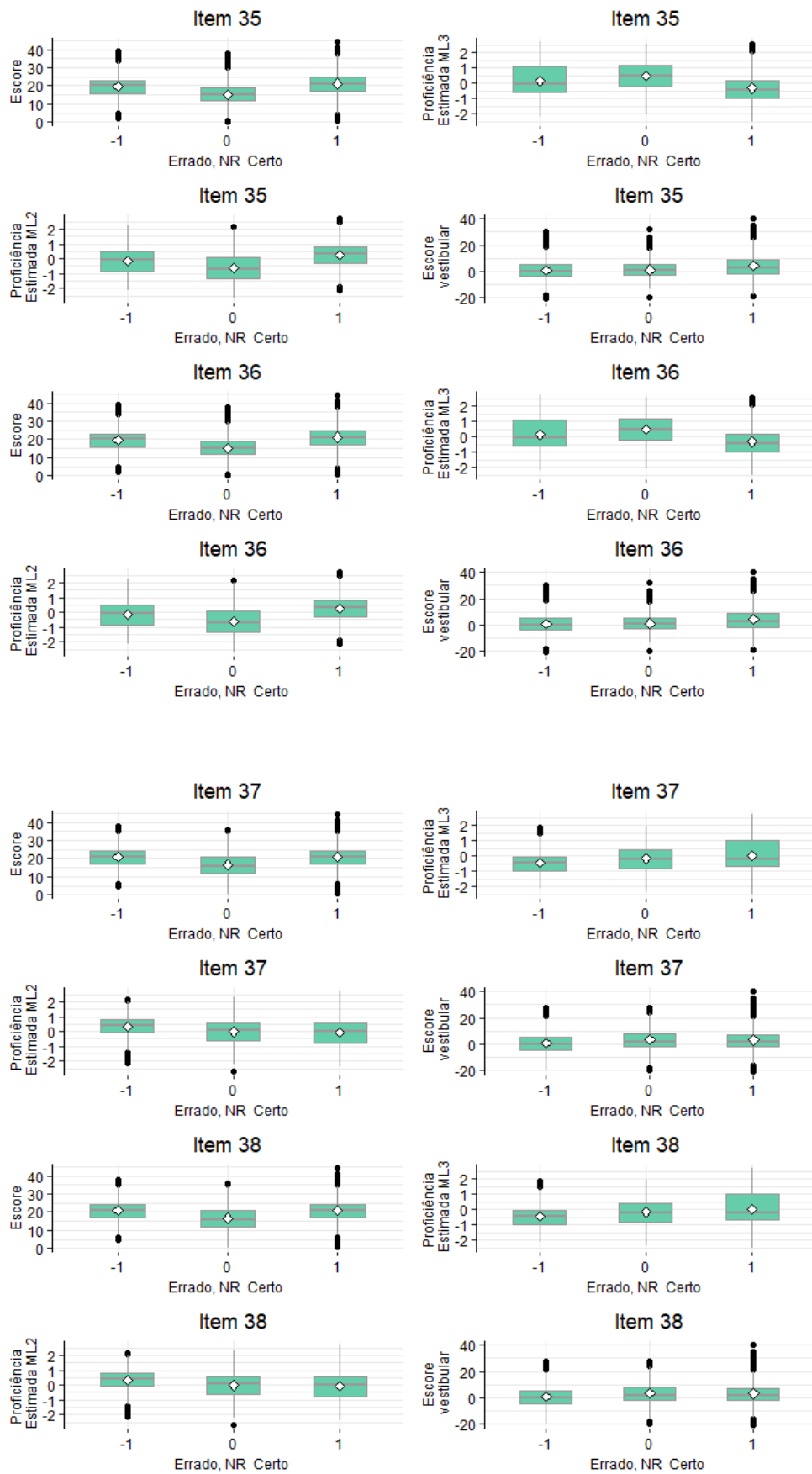


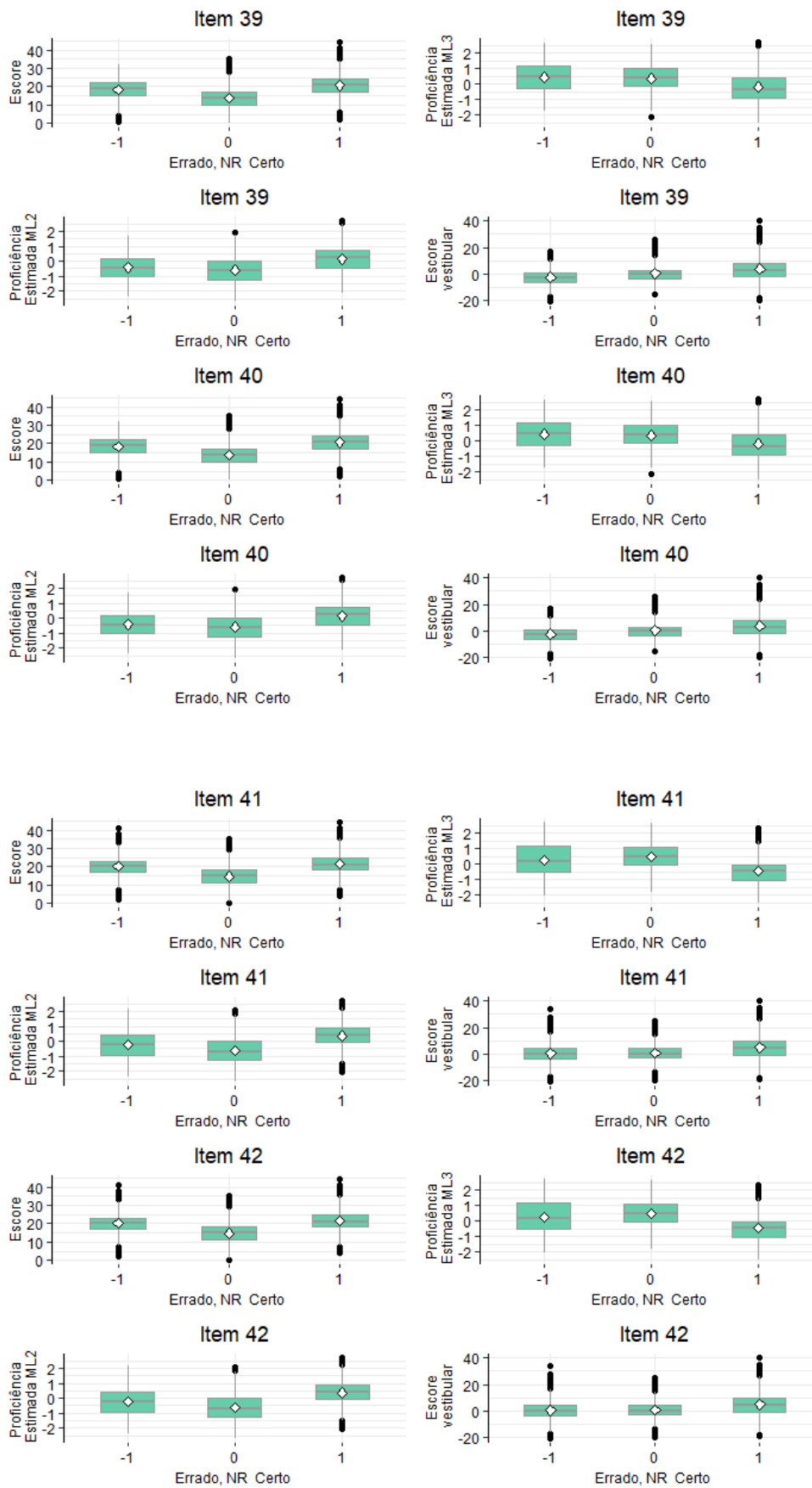


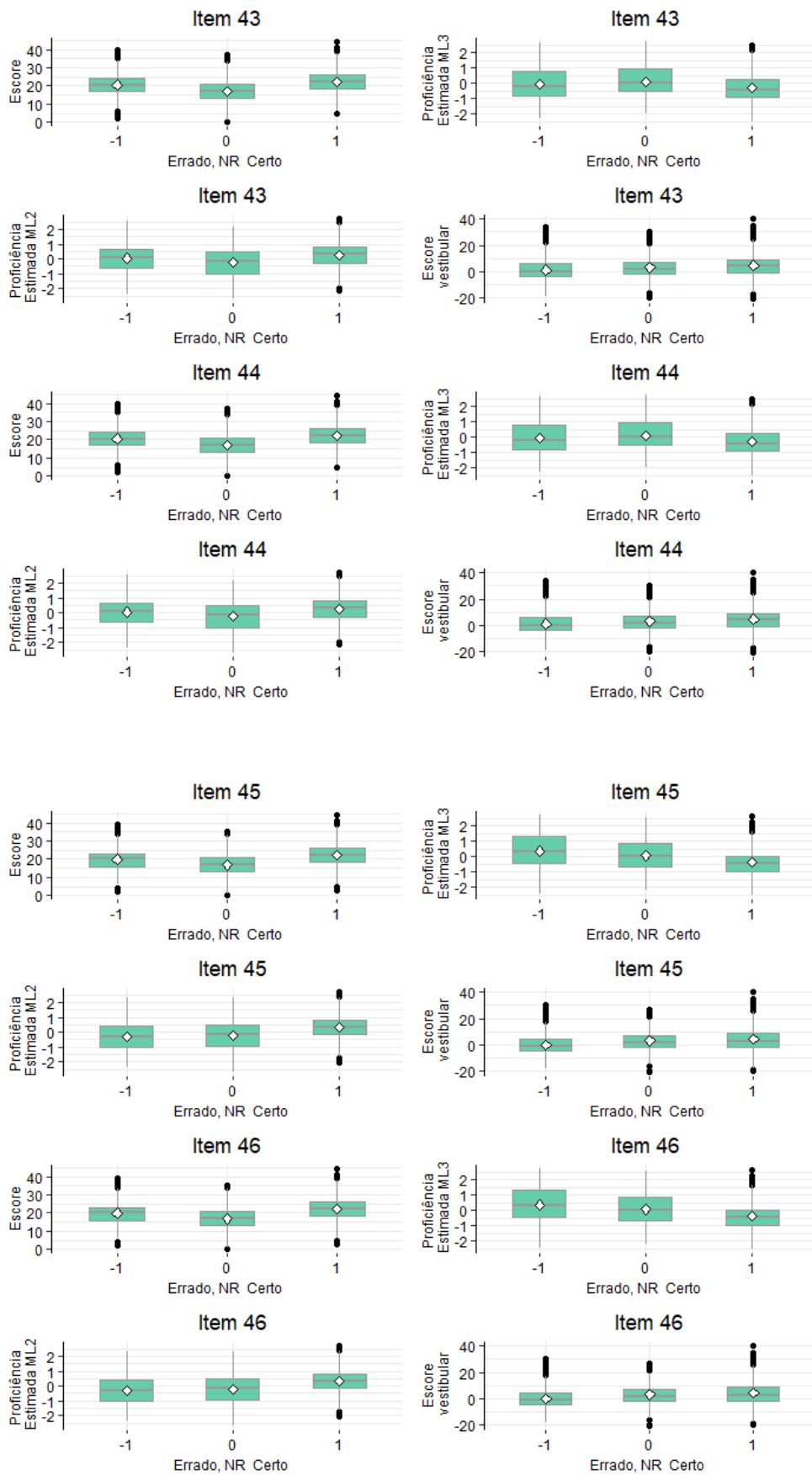












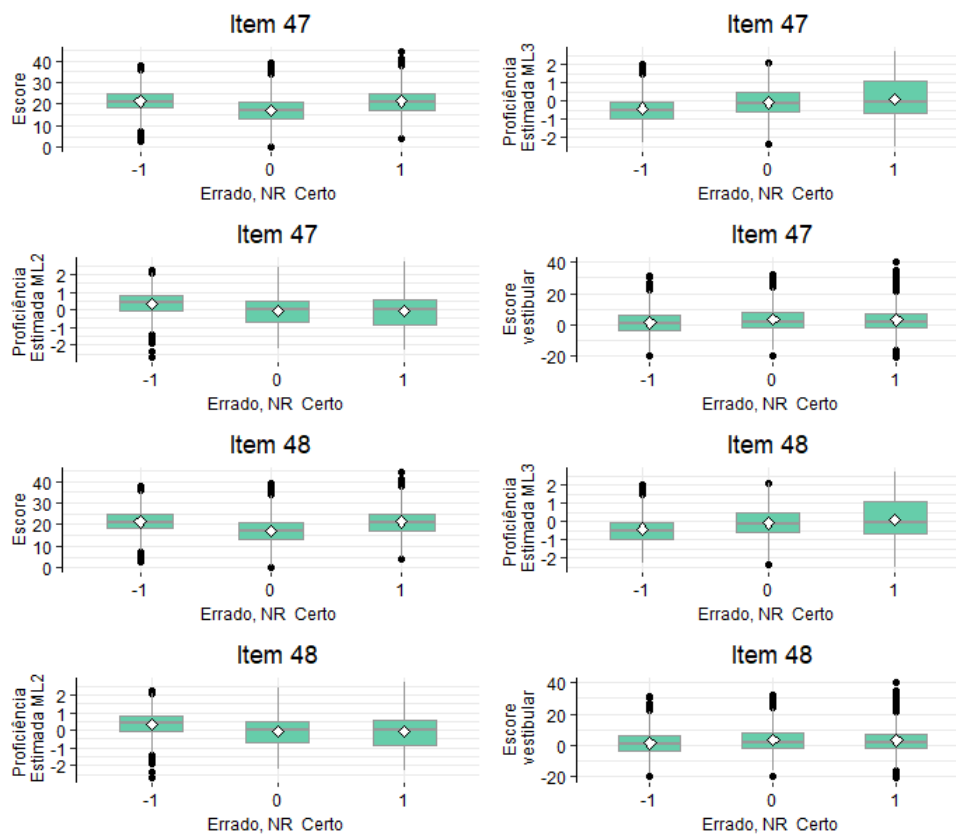


Figura 12: Boxplots do Escore, Proficiência Estimada do ML2, Proficiência Estimada do ML3 e Escore do Vestibular por Item

Na Figura 12, estão representados os Boxplots do Escore (prova de língua portuguesa), a Proficiência do Modelo ML2, Proficiência do Modelo ML3 e o Escore do Vestibular, respectivamente, por Item. Analisando o primeiro Boxplot de cada item, percebe-se que o grupo que acertou o item tem o maior escore na prova de português, e o grupo que errou tem o escore maior que os alunos que deixaram o item sem resposta.

Percebe-se que a Proficiência do ML2 possui o mesmo comportamento na maioria dos itens. No modelo ML3, percebe-se a Proficiência dos alunos que erraram é maior que a dos alunos que acertaram, em grande parte dos itens, indicando que o modelo ML3 não se ajustou corretamente aos dados. Em relação ao Escore do Vestibular, percebe-se uma alternância entre os alunos que erraram com os que não responderam, no segundo lugar, sendo o primeiro ocupado pelos candidatos que acertaram os itens.

3.3.1 Ajuste do Modelo de Resposta Gradual

Para o primeiro ajuste do Modelo, utilizaremos a lógica utilizada pelo CESPE, sendo o primeiro nível o erro (-1 ponto), o segundo a não resposta (0 pontos) e o terceiro o acerto (1 ponto).

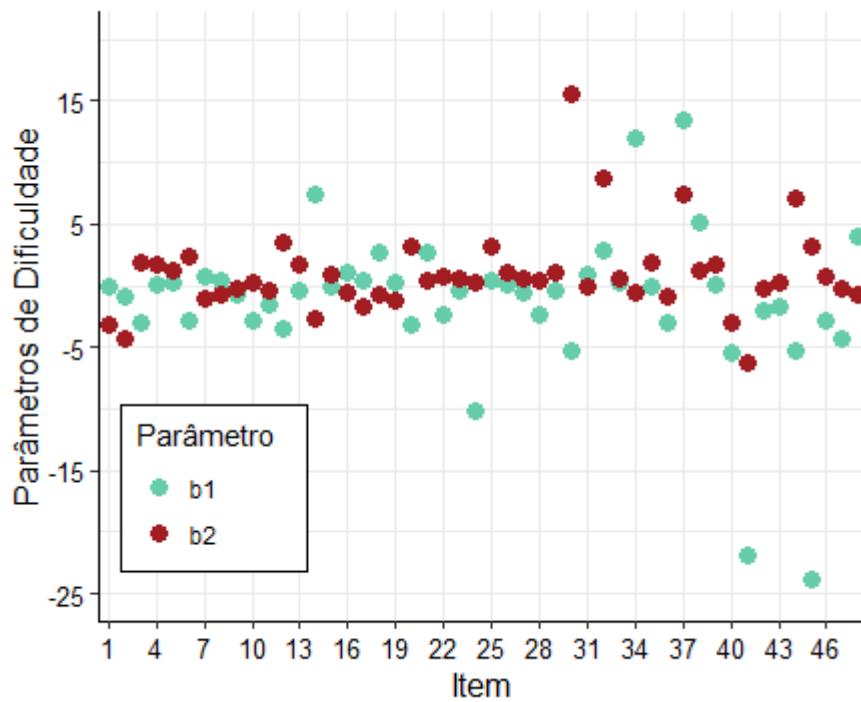


Figura 13: Parâmetro de Dificuldade do MRG na ordem erro, sem resposta e acerto

Por meio da Figura 13, encontramos valores destoantes de dificuldade (itens 30, 41 e 45), além de valores do parâmetro b_1 superiores ao parâmetro b_2 , demonstrando que o modelo não se ajustou bem aos dados.

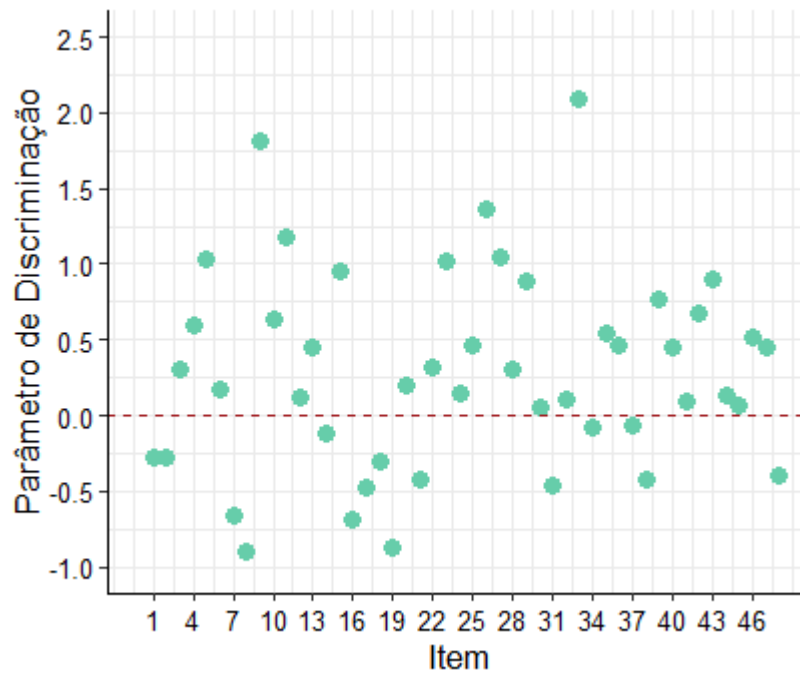


Figura 14: Parâmetro de Discriminação do MRG na ordem erro, sem resposta e acerto

Na Figura 14, temos a confirmação da Figura 13, com valores negativos de discriminação, o que nos leva a testar a ordenação sugerida nos Boxplots: Não Resposta, Resposta Errada e Resposta Certa.

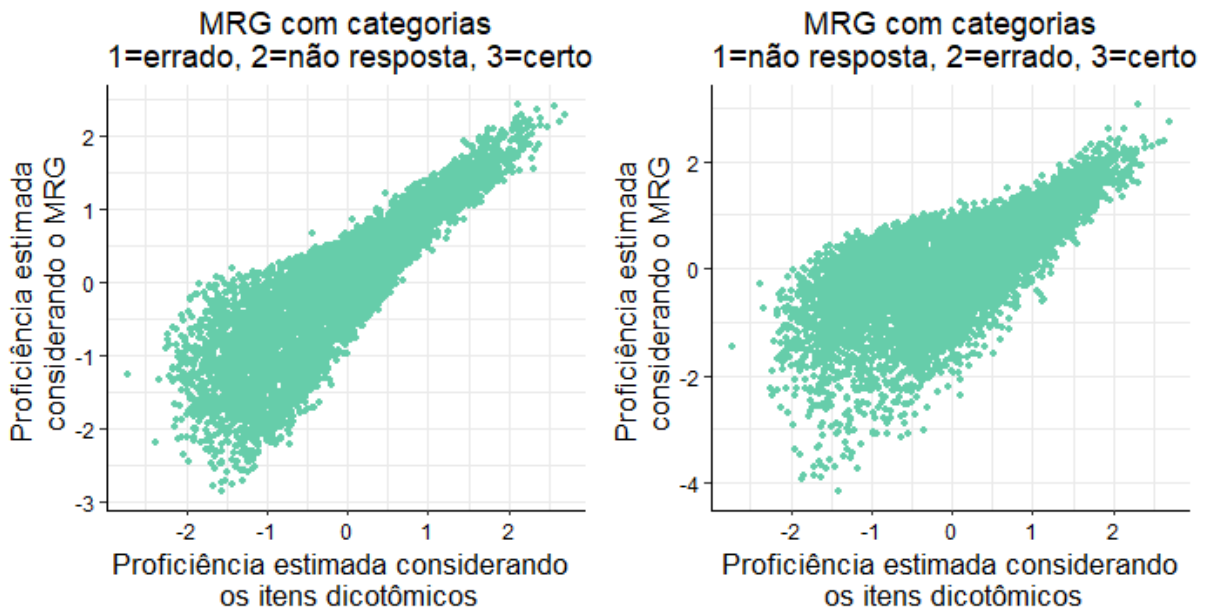


Figura 15: Proficiência estimada considerando itens dicotômicos pela Proficiência estimada considerando os modelos MRG

Na Figura 15, representa-se na primeira imagem a proficiência estimada considerando itens dicotômicos (acerto e não acerto) vs proficiência estimada considerando o

MRG com a ordem errado, não resposta e certo, enquanto no lado direito temos a ordem não resposta, errado e certo. Os gráficos apresentaram poucas diferenças portanto será aprofundada a análise da segunda ordenação.

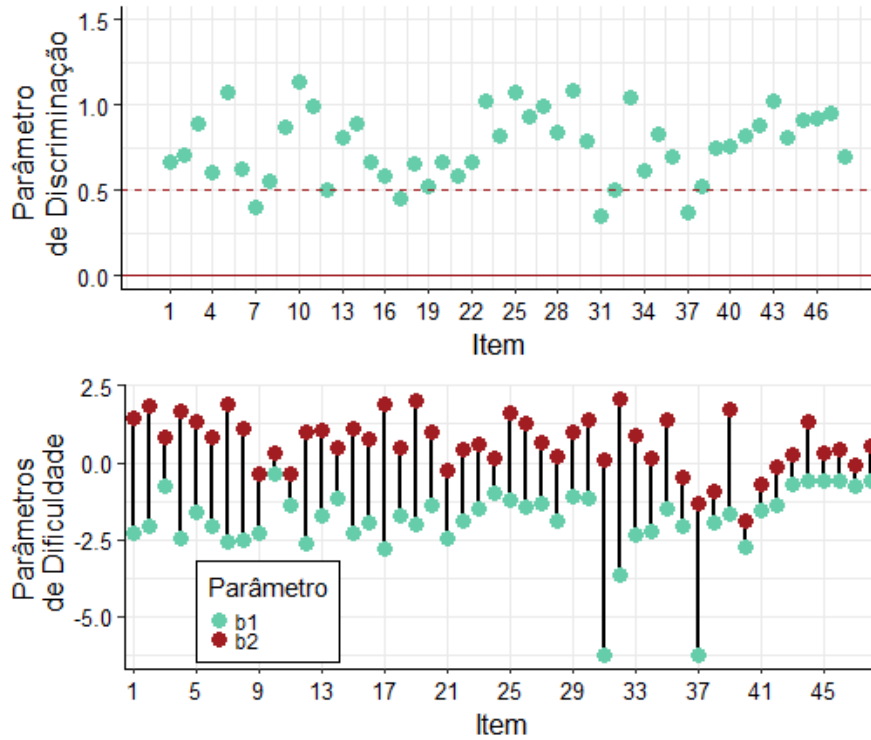
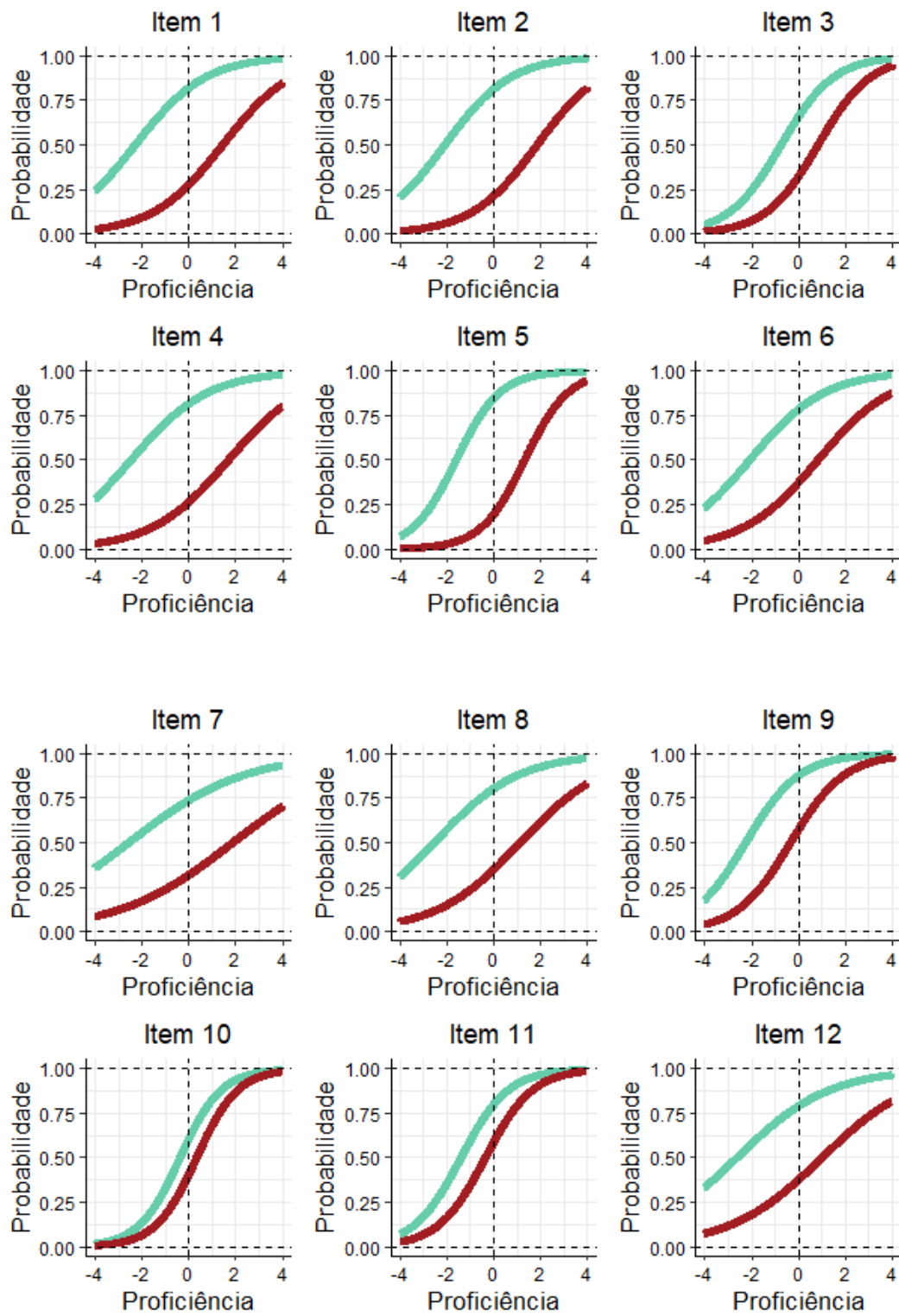
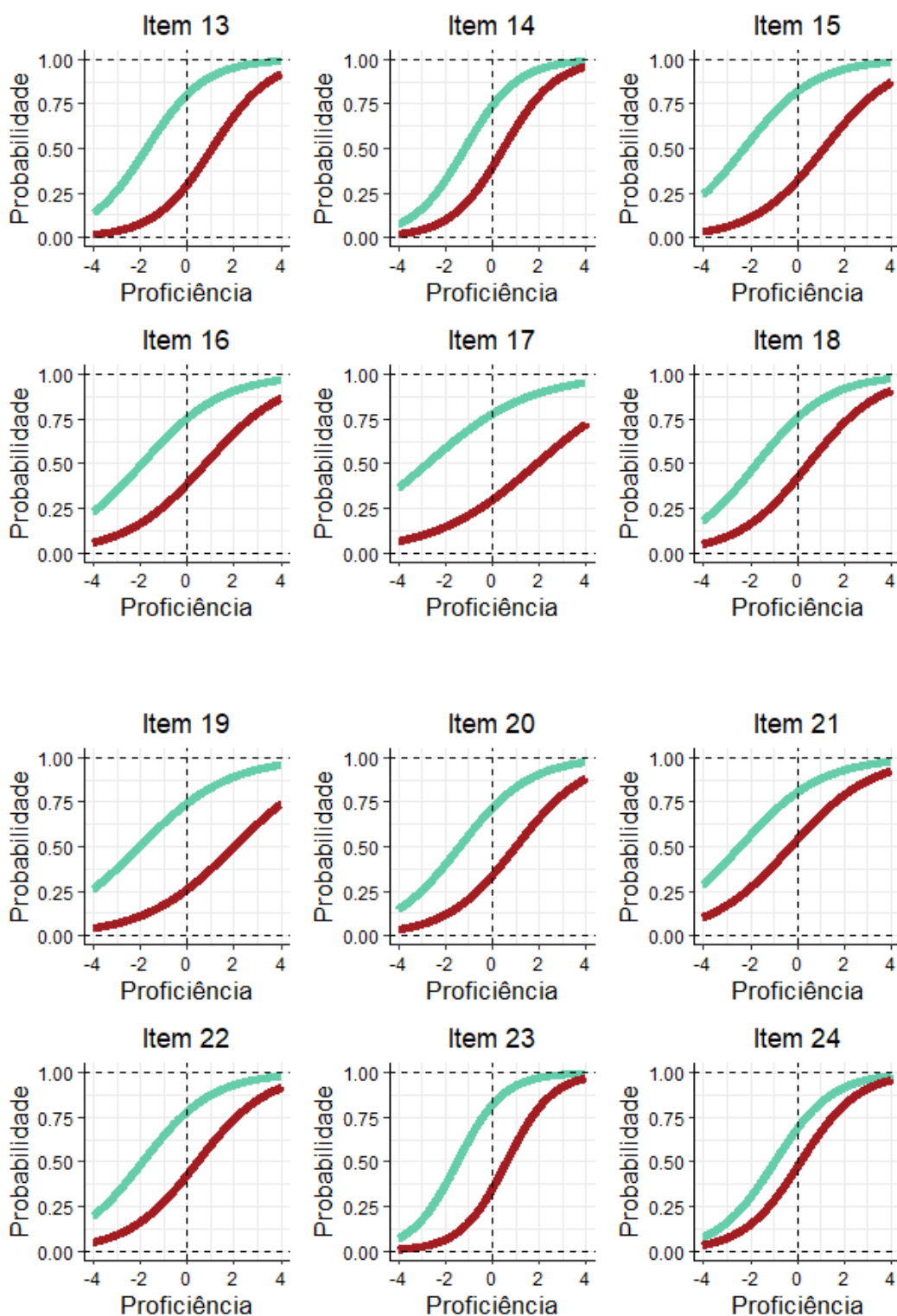


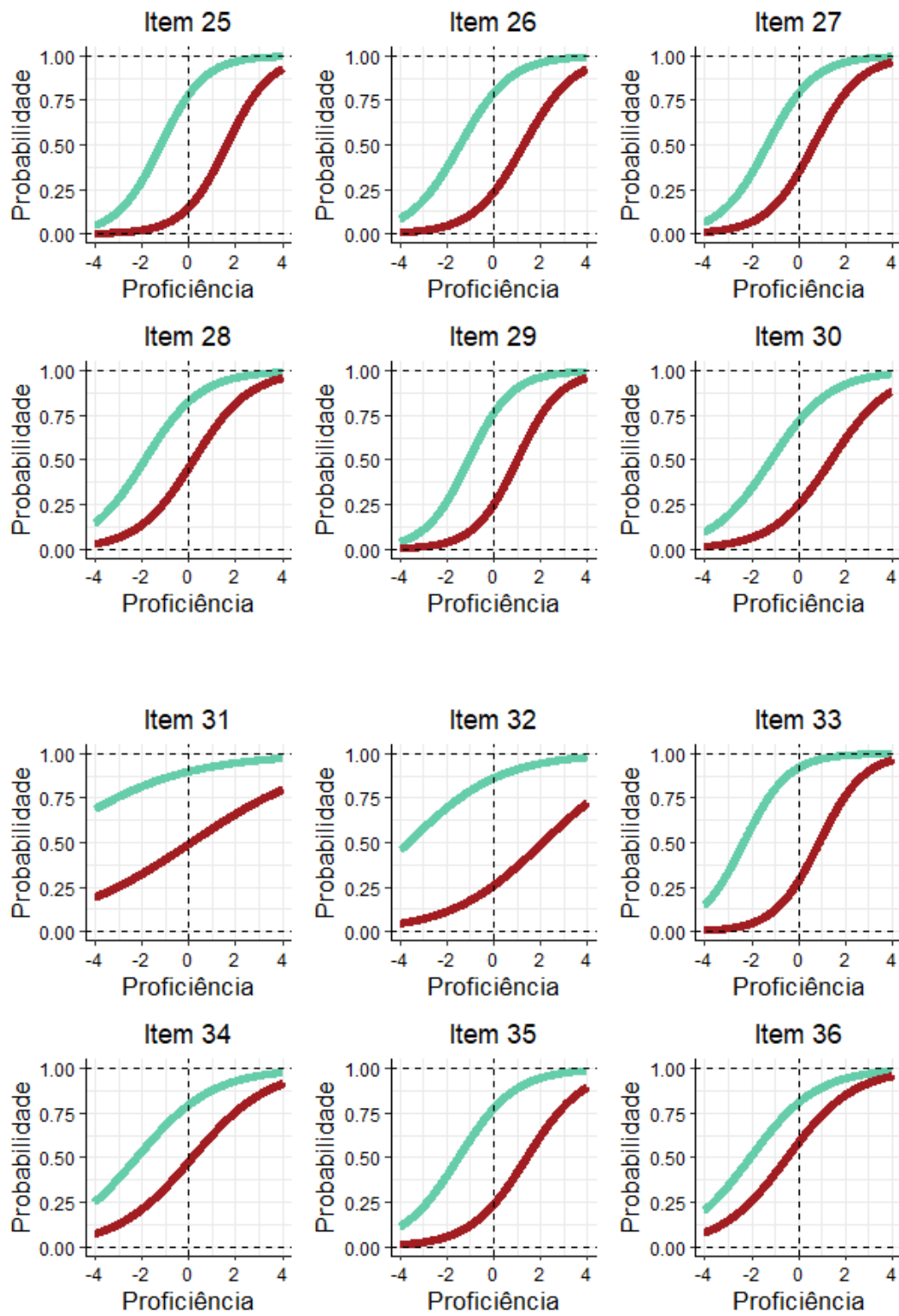
Figura 16: Distribuição dos Parâmetros de Discriminação e Dificuldade por Item para o MRG com a ordem não resposta, erro e acerto

Na Figura 16, percebe-se um melhor ajuste do MRG, com a nova ordenação. O Parâmetro de discriminação apresenta valores positivos e próximos de 0.5 (os itens 31, 37, 7 e 17 foram os únicos a apresentarem discriminação menor que 0.5), sendo o item 10 o item com maior valor.

Para o parâmetro de Dificuldade, o ordenação não resposta, erro e acerto também mostra melhor ajuste aos dados, com valores regulares e parâmetro b_2 superior ao b_1 em todos os itens. Os itens 31 e 37 apresentaram novamente valores baixos, agora para o parâmetro b_1 de dificuldade.







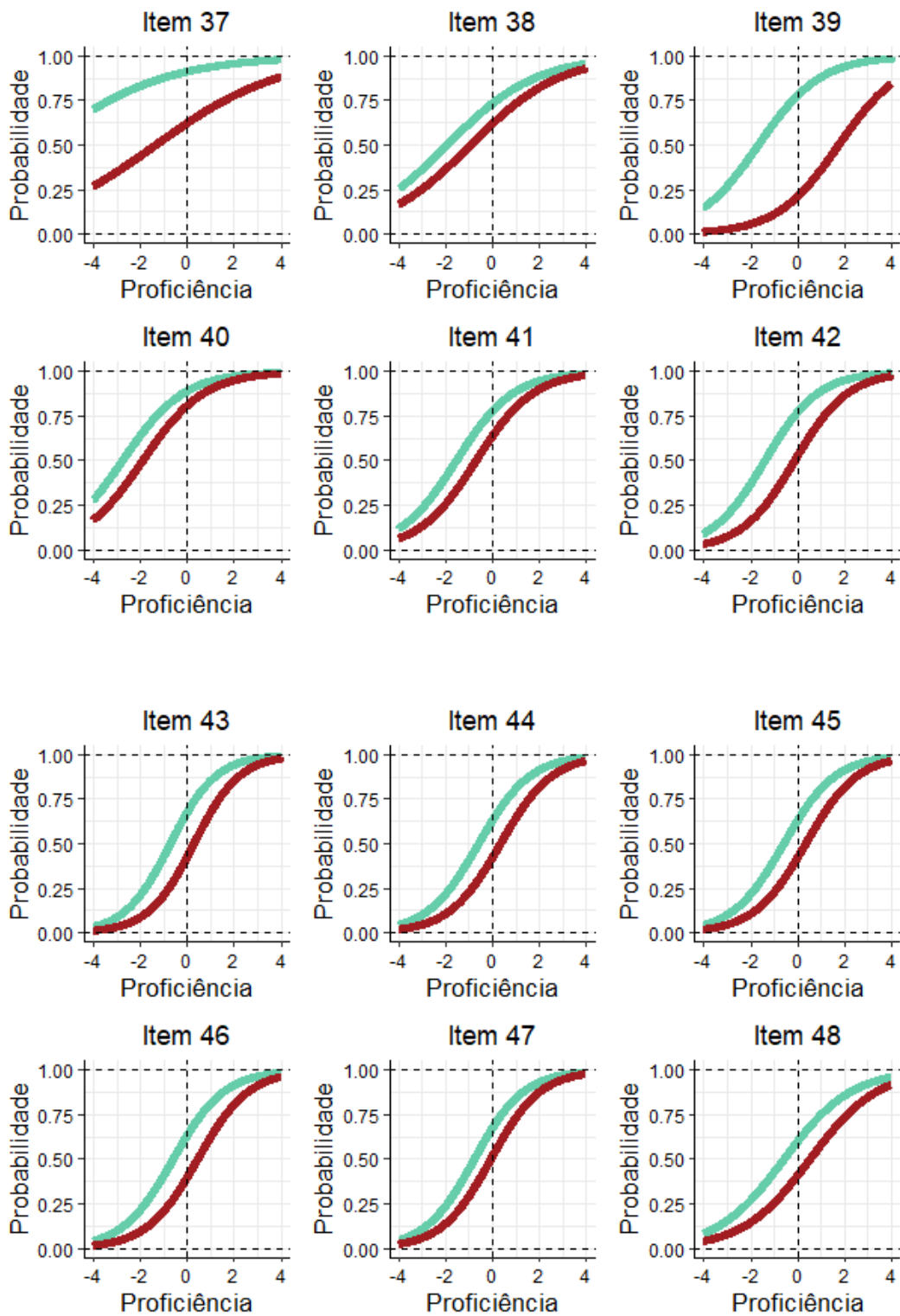


Figura 17: Curva Característica do Item para o Modelo de Resposta Gradual com a ordenação não resposta, erro e acerto

Na Figura 17 está representada a CCI para o Modelo de Resposta Gradual considerando a ordem não resposta, resposta errada e acerto, ou seja, considerando os alunos que não responderam como nível inferior aos que erraram.

Percebe-se que os itens 7, 17, 31 e 37 apresentaram tanto na Figura 16 quanto na Figura 17, um baixo poder de discriminação dos candidatos, indicando a reavaliação de tais itens pela banca examinadora.

4 Conclusões

As análises apresentadas representam o estudo da Prova de Português (48 itens tipo A) do Vestibular da UnB, realizado no segundo semestre de 2014. A banca examinadora (CESPE/CEBRASPE) utiliza a metodologia de punição dos itens errados no cálculo da nota final dos candidatos.

Ao utilizar a mesma ordenação (erro, não resposta e acerto), o estudo identificou ajuste inadequado do modelo aos dados, com valores de parâmetros incongruentes. Sendo assim, percebeu-se que os alunos com menor proficiência tendem a deixar o item em branco, enquanto os alunos que possuem algum tipo de conhecimento do item, tendem a responder o item.

Sendo assim, a ordenação que trouxe melhor ajuste aos dados foi não resposta, resposta incorreta e resposta correta, com valores efetivos dos parâmetros. Além disso, os itens 7, 17, 31 e 37 apresentaram baixa discriminação, indicando baixo poder de separar candidatos com diferentes níveis de proficiência.

Referências

ANDRADE, D. F. de; TAVARES, H. R.; VALLE, R. da C. Teoria da resposta ao item: conceitos e aplicações. *ABE, Sao Paulo*, 2000.

ANJOS, A. dos; ANDRADE, D. F. de. Teoria da resposta ao item com uso do r. 2012.

BAKER, F. B.; KIM, S.-H. *The basics of item response theory using R*. [S.l.]: Springer, 2017.