

CENTRO DE ESTATÍSTICA APLICADA - CEA
RELATÓRIO DE ANÁLISE ESTATÍSTICA - CÓDIGO 99P18

TÍTULO: Relatório de análise estatística sobre o projeto: “Avaliação do Aprendizado em Semiologia”.

PESQUISADORA: Célia Maria Kira

ORIENTADOR: Eduardo Marcondes

INSTITUIÇÃO: Faculdade de Medicina - USP

FINALIDADE DO PROJETO: Doutorado

RESPONSÁVEIS PELA ANÁLISE: Silvia Nagib Elian

Gustavo Henrique de Araujo Pereira

REFERÊNCIA DESTE TRABALHO: ELIAN, S. N. e PEREIRA, G. H. A. **Relatório de análise estatística sobre o projeto: “Avaliação do Aprendizado em Semiologia”**. São Paulo, IME-USP, 1999. (RAE – CEA - 99P18)

FICHA TÉCNICA

REFERÊNCIAS BIBLIOGRÁFICAS:

- BMDP – Statistical Software Manual, Release 7.0** (1990). Los Angeles: BMDP Statistical Software, Inc. 629p.
- BUSSAB, W. O. e MORETTIN, P. A. (1987). **Estatística Básica**. 4.ed. São Paulo: Atual. 320p.
- CONOVER, W. J. (1971). **Practical Nonparametric Statistics**. 1.ed. New York: John Wiley & Sons. 462p.
- LINDSEY, J. K. (1993). **Models for Repeated Measurement**. 1.ed. Oxford: Clarendon Press. 413p.
- NETER, J., KUTNER, M. H., NACHTSHEIM, C. J. and WASSERMAN, W. (1996). **Applied Linear Statistical Models**. 4. ed. Illinois: Richard D. Irwing. 1408p.
- SAS: Reference Manual, Versão 5.0** (1996). USA: SAS Institute. 940p.

PROGRAMAS COMPUTACIONAIS UTILIZADOS:

BMDP New System for Windows (versão 1.12);
 Excel for Windows (versão 97);
 Minitab for Windows (versão 11.12);
 SAS for Windows (versão 6.11);
 SPSS for Windows (versão 8.0);
 Word for Windows (versão 97).

TÉCNICAS ESTATÍSTICAS UTILIZADAS:

[Entre parênteses encontra-se a Classificação “Statistical Theory & Method Abstracts” (ISI)]

Análise Descritiva Unidimensional (03:010)

Estimação Paramétrica Unidimensional (04:010)

Testes de Hipóteses Paramétricas (05:010)

Testes de hipóteses Não Paramétricas (05:070)

Análise de Componentes de Variância (08:220)

ÁREA DE APLICAÇÃO:

Medicina - Epidemiologia (14:050)

ÍNDICE

RESUMO	6
1. Introdução.....	7
2. Descrição do estudo	7
3. Descrição das variáveis	9
4. Análise descritiva.....	10
5. Análise inferencial	15
5.1. Comparação do desempenho dos alunos dos grupos A e B.....	15
5.2. Comparação das notas de conceito oficial com as notas de conceito do questionário dentro de cada grupo	17
5.3. Comparações envolvendo as notas do questionário no grupo B	19
5.4. Comparações envolvendo as notas do questionário no grupo A	20
6. Conclusões.....	21
Apêndice A - Tabelas da análise descritiva	22
Apêndice B - Boxplots.....	34
Apêndice C - Gráficos de perfis.....	39
Apêndice D - Histogramas	64
Apêndice E - Modelo de efeitos aleatórios	67
Apêndice F - Tabelas da análise inferencial	74
Apêndice G - Questionário	95

RESUMO

Propedêutica clínica é uma disciplina do curso de graduação em medicina onde os alunos têm o primeiro contato com o paciente. Nessa disciplina, na Faculdade de Medicina da USP, notas de conceito são responsáveis por 60% da média final. Como essa avaliação possui um elevado grau de subjetividade, foi desenvolvido um questionário que procura dividir o conhecimento em Propedêutica Clínica em 21 questões.

Um objetivo do estudo era comparar o desempenho de um grupo de alunos que são avaliados mensalmente pelo questionário com um grupo de alunos que são avaliados apenas uma vez por semestre. Verificou-se que o desempenho médio é o mesmo nos dois grupos de alunos.

Desejava-se comparar também notas dadas pelo professor aos alunos com as notas dadas pelos alunos em uma auto avaliação. A conclusão nesse caso foi que a média das notas dadas pelo professor é maior que a média das notas dadas pelo aluno.

Outro objetivo do estudo era comparar notas dadas antes e depois do preenchimento do questionário. Verificou-se que não há diferença entre as médias dessas notas.

A última comparação de interesse era a das notas de conceito oficial, que participam da média final do aluno, com notas de conceito dadas antes e após o preenchimento do questionário. Concluiu-se que, em geral, a média das notas de conceito oficial são maiores.

A primeira comparação foi feita através de testes de igualdade de médias para amostras independentes, enquanto que para as demais foram utilizados modelos de efeitos aleatórios, preservando a estrutura de dependência dos dados.

1. Introdução

No currículo do curso de graduação em medicina da Faculdade de Medicina da Universidade de São Paulo há, no terceiro ano, uma disciplina denominada Propedêutica Clínica. Nesse curso, o aluno de medicina tem o primeiro contato com o paciente. Os objetivos da disciplina são capacitar o aluno a executar uma história clínica adequada do paciente, a realizar um exame físico completo e um correto diagnóstico de doenças simples, bem como prepará-lo para desenvolver uma boa relação médico-paciente.

A média final do curso de Propedêutica Clínica é composta de notas de prova teórica, de prova prática e de conceito. A última tem um peso de 60% na média final e, pelo fato de ser uma opinião quantitativa do professor sobre o aprendizado do aluno na disciplina, possui uma grande carga de subjetividade. Visando tornar essa nota mais objetiva, foi desenvolvido um questionário (Apêndice G), que divide o conhecimento em Propedêutica Clínica em 21 questões.

Os objetivos do estudo são:

- comparar as notas em anamnese (retirada da história clínica), exame físico, raciocínio clínico e conceito geral dado pelo aluno a si mesmo e pelo professor;
- comparar as notas em anamnese, exame físico, raciocínio clínico e conceito geral, dadas antes e após o preenchimento do questionário;
- comparar a nota de conceito oficial, com as notas de conceito dadas pelo professor antes e após o preenchimento do questionário;
- comparar a média final de um grupo de alunos avaliados mensalmente com o questionário, com a média de um grupo que preenche o questionário apenas uma vez por semestre.

2. Descrição do estudo

Nos cursos de Propedêutica Clínica da Faculdade de Medicina da Universidade de São Paulo, os alunos são divididos em 2 grupos, de acordo com a ordem alfabética da inicial de seus nomes. No primeiro semestre, um grupo tem aulas no Hospital das

Clínicas (HC) e o outro no Hospital Universitário (HU). No segundo semestre, esses grupos têm o seu local de aprendizado alternado. Em cada hospital, os alunos são divididos em 9 subgrupos e cada um deles tem aula com um professor.

Cada aluno, em cada semestre, recebe uma nota de conceito, uma nota de prova teórica e uma de prova prática. A nota de conceito depende da participação e interesse nas aulas, preparação e apresentação de seminários, participação nas atividades práticas e progresso do aluno observado durante o ano letivo. Os alunos com frequência entre 81% e 90% perdem 10% da nota de conceito, enquanto os alunos com frequência entre 70% e 80% perdem 20% dessa nota. Alunos com frequência inferior a 70% são reprovados. A média do curso de Propedêutica Clínica é composta pela média de conceito, que tem um peso de 60% na média final e pelas médias de provas teórica e prática, cada uma delas com 20% de peso.

No ano de 1998, o curso teve 176 alunos que cursaram a disciplina até o final. O grupo A, composto de 87 alunos, iniciou o curso no HC. Já o grupo B, com aulas iniciais no HU, tinha 89 alunos.

Por aleatorização, ficou decidido que os alunos do grupo A preencheriam mensalmente um questionário, que procurou dividir o conhecimento em Propedêutica Clínica em 21 questões. A cada uma delas, os alunos respondiam entre as opções “sempre”, “quase sempre”, “mais ou menos”, “quase nunca” e “nunca”. Antes de responder às questões, os alunos davam a si uma nota de anamnese, uma de exame físico, uma de raciocínio clínico e uma nota geral de conceito. Após o preenchimento do questionário, os alunos novamente se davam notas nas mesmas categorias. Esperava-se que alunos que tinham se dado notas boas, mas responderam “nunca” ou “quase nunca” a várias perguntas, diminuíssem sua nota. Efeito inverso era esperado em alunos com alta proporção de “sempre” e “quase sempre” e que tinham se dado notas baixas. As respostas do questionário serviam apenas para alunos e professores darem uma nota um pouco menos subjetiva após o preenchimento do questionário e portanto não foram analisadas.

Para os alunos do grupo B, o mesmo procedimento foi repetido, com a diferença de que eles preencheram o questionário apenas duas vezes, no final de cada semestre.

Simultaneamente à entrega aos alunos, era entregue aos professores o mesmo questionário, onde eles davam as notas e respondiam às questões referentes a cada um de seus alunos.

3. Descrição das variáveis

As variáveis resposta do estudo são notas que podem variar de 0 a 10. Essas notas foram dadas em cada uma das seguintes 4 categorias:

- anamnese: consiste na retirada da história clínica, que inclui, entre outras coisas, indagações ao paciente sobre quais as suas queixas, as durações dessas queixas e antecedentes pessoais e familiares de doenças;
- exame físico: exame feito pelo estudante que engloba, desde a retirada de sinais vitais, a exames relacionados com as queixas do paciente durante a anamnese;
- raciocínio clínico: é a obtenção de um diagnóstico a partir das informações coletadas na anamnese e o exame físico;
- conceito geral: é uma nota que leva em conta o desempenho dos estudantes nas 3 categorias acima (com peso de 40% para os dois primeiros e 20% para o último), além do desempenho em seminários e da participação em aulas e em atividades práticas.

Cada uma dessas notas pode depender ou dependia das seguintes variáveis explicativas:

- identificação do aluno;
- sexo;
- grupo: A para os alunos que estudaram o 1º semestre no HC e B para os alunos que estudaram inicialmente no HU;
- instante de aplicação do questionário: varia de 1 a 8 para os alunos do grupo A e 1 a 2 para os pertencentes ao grupo B;
- avaliador: professor ou aluno;
- condição: antes ou depois do preenchimento do questionário.

Além dessas variáveis relacionadas com o questionário, o estudo ainda possui as seguintes variáveis resposta:

- nota oficial de conceito: dada pelo professor a cada um de seus alunos no final de cada semestre. A média dessas notas oficiais de conceito representam 60% da média final;
- média final.

4. Análise descritiva

Estudou-se inicialmente a influência que a variável sexo poderia causar nas variáveis resposta para os dois grupos de alunos. Embora a variável sexo não seja de interesse do estudo, ela deve ser considerada, caso a distribuição por sexo seja muito diferente nos 2 grupos e haja diferença de desempenho entre os sexos. A Tabela A.1 mostra que a distribuição por sexo nos dois grupos é bastante semelhante. Além disso, os “boxplots” (Bussab e Morettin, 1987) para a média final (Gráfico B.1) e para o conceito oficial (Gráfico B.2) apresentam fortes indícios de que o desempenho médio não varia com o sexo. A variabilidade da média final e de conceito oficial parece ser um pouco maior para o sexo masculino. No entanto, essa pequena diferença de variabilidade não é um motivo suficiente para que se inclua na análise uma variável explicativa que não é de interesse do estudo.

O Gráfico B.3 e a Tabela A.2 mostram indícios de que a média da variável média final é a mesma para os dois grupos. Há apenas uma nota discrepante e, pelo fato dela não estar muito distante das demais observações, ela não deverá causar muita influência na análise. Como era esperado um desempenho superior do grupo A, buscou-se uma explicação para essa aparente igualdade de médias.

As Tabelas A.3 a A.6 mostram a distribuição observada do número de questionários preenchidos pelos alunos e pelos professores, em relação a cada um de seus alunos. Pode-se ver que apenas 3 alunos preencheram todos os questionários no grupo A e que 14 alunos deixaram de preencher 4 ou mais questionários. Já no grupo B, quase todos os alunos preencheram todos os questionários. Foram feitos “boxplots” da média final por grupo (Gráfico B.4), eliminando-se, no grupo A, os alunos que preencheram

menos de 5 questionários. O gráfico, assim como a Tabela A.7, continua mostrando indícios de que não há diferença na média da variável média final entre os dois grupos. As notas discrepantes observadas também não se mostram muito distantes das demais observações, não devendo causar modificações nas conclusões a serem obtidas. Por isso, a aparente igualdade de desempenho entre os dois grupos parece não ser devida a um desempenho inferior dos alunos que preencheram poucos questionários.

Investigou-se ainda a possibilidade do preenchimento mensal do questionário contribuir para um aumento na média do conceito, embora não tenha provocado um acréscimo na média final. O Gráfico B.5 e a Tabela A.8 sugerem que, no primeiro semestre, a média do conceito oficial é maior no grupo A, enquanto no segundo semestre o grupo B parece apresentar uma maior média de conceito oficial. O fato de cada grupo ter melhor desempenho em relação ao conceito oficial em um dos semestres é um indício de que os professores do hospital no qual o grupo está estudando influenciam na nota média de conceito. As notas discrepantes são poucas e suas distâncias às demais observações não são grandes. Portanto, o Grupo A também parece não ter melhor desempenho médio que o B em relação ao conceito oficial. O Gráfico B.5 sugere que a variabilidade do conceito oficial é maior no grupo B do que em A, nos dois semestres. No entanto, pelo fato da diferença de variabilidade observada não ter sido muito grande, apenas um teste estatístico poderá dizer se essa hipótese é verdadeira.

Para comparar as notas em anamnese, exame físico, raciocínio clínico e conceito dadas antes e após o preenchimento do questionário, foi construída a variável diferença. Assim, para cada aluno, foi calculada a diferença entre a nota dada depois e a correspondente nota dada antes do questionário. Essa variável foi construída tanto para as notas dadas pelo professor como para aquelas dadas pelos alunos. Observações com diferença positiva indicam que a nota depois do questionário é maior do que a nota antes do questionário. Diferenças negativas indicam que a nota antes do questionário é maior.

As figuras C.1 a C.4 mostram, para cada instante de aplicação do questionário, cada avaliador e cada grupo, a média da variável diferença, acrescida e subtraída de seu erro padrão. Para as notas dadas pelo aluno, parece não haver diferença entre as

médias das notas antes e depois do questionário, em todas as variáveis, nos dois grupos e para a maioria dos instantes de aplicação. Apenas em alguns poucos instantes, como o 2 da variável raciocínio clínico no grupo A (Figura C.3), parece haver uma diferença nas médias das notas antes e depois do questionário.

Para as notas dadas pelo professor, há um número maior de instantes de aplicação do questionário sugerindo uma possível diferença entre as médias das notas antes e depois. A Figura C.3, por exemplo, mostra que a variável raciocínio clínico para o grupo A, parece apresentar uma nota média depois do questionário superior à nota média antes, para os instantes de aplicação 4 e 6. No entanto, mesmo para o professor, na maioria dos instantes parece não haver diferença entre as médias das notas antes e depois do questionário. Além disso, mesmo nos instantes que apresentam maiores valores da diferença entre as médias das notas antes e depois, esta diferença nunca é maior que 0,3.

As Tabelas A.9 a A.12 apresentam as correlações entre as notas antes e depois da aplicação do questionário. Pode-se ver que para todas as variáveis, instantes de aplicação, avaliadores e grupos, a correlação é maior ou igual a 0,75. Essas altas correlações indicam que as notas antes e depois possuem uma forte associação linear positiva. Isso sugere que alunos com altas notas antes têm, em geral, notas altas depois. Por outro lado, alunos com baixas notas antes têm, em geral, notas baixas depois da aplicação do questionário.

Para comparar as médias das notas dadas pelo professor e pelo aluno, também foi construída a variável diferença, subtraindo-se a nota dada pelo aluno da correspondente nota dada pelo professor. A média da variável diferença, bem como esta média mais e menos seu erro padrão, são apresentados nas Figuras C.5 a C.8, para cada variável, instante de aplicação do questionário, grupo e condição (antes ou depois do preenchimento do questionário).

Os gráficos mostram indícios de que, para todas as variáveis, grupos, condições e maioria dos instantes, a nota média do professor é maior do que a nota média do aluno. Apenas alguns poucos instantes de aplicação do questionário, como o sexto da variável anamnese no grupo A antes do questionário (Figura C.5), parecem não apresentar diferença entre a média da nota dada pelo professor e a média da nota dada pelo

aluno. Além disso, os gráficos sugerem que, no grupo A, para todas as variáveis, durante o primeiro semestre, há uma tendência de crescimento da diferença entre a média das notas dadas pelo professor e pelo aluno. Imediatamente após a mudança de semestre, que faz com que os alunos mudem de hospital e professor, parece haver uma grande queda dessa diferença. Em seguida, no entanto, a diferença entre a média do professor e do aluno volta a apresentar tendência crescente, durante o segundo semestre.

Os gráficos também permitem compararmos o grupo A e B quanto à diferença entre as médias das notas dadas pelo professor e das notas dadas pelo aluno. O ideal é que se compare o instante de aplicação do questionário 4 e 8 do grupo A, respectivamente com os instantes 1 e 2 do grupo B, já que esses são os questionários que os dois grupos recebem no mesmo período. Para todas as variáveis, há indícios de que a média da diferença no grupo A é maior do que no grupo B, para os questionários do final do primeiro semestre. Já para os questionários do segundo semestre, essas médias parecem ser iguais.

As Tabelas A.13 a A.16 mostram que as correlações entre as notas dadas pelo professor e dadas pelo aluno são, em sua maioria, baixas. Mesmo nos instantes que apresentam uma correlação um pouco mais alta, essa estatística não supera o valor 0,5. Isso é um indício de que a associação linear entre as notas dos professores e as notas dos alunos, em geral, é fraca, ou até mesmo inexistente.

Diante dos resultados encontrados nas comparações entre as notas antes e depois do questionário e as notas do professor e do aluno, a pesquisadora levantou a hipótese de que os resultados dessas comparações seriam diferentes para alunos com bom e mau desempenho. Para verificar se essa hipótese era verdadeira, em cada grupo os alunos foram divididos em três categorias. Alunos com nota de conceito oficial do segundo semestre menor ou igual ao 1º quartil dessa variável formaram uma categoria denominada Q1. Alunos com nota de conceito oficial maior ou igual ao 3º quartil formaram a categoria Q3. Os demais foram agrupados em Q2. Devido à grande ocorrência de notas empatadas, nos dois grupos, as categorias Q1 e Q3 possuem mais de 25% das notas e, conseqüentemente, a categoria Q2 possui menos de 50% dos alunos.

A média e também o valor desta média mais e menos o erro padrão de cada uma das variáveis diferença (depois do questionário menos antes do questionário), em cada um dos níveis das variáveis explicativas, são apresentados nas Figuras C.9 a C.16. Em todas as variáveis, para todos os avaliadores, grupos e questionários, não há indícios de que haja diferença entre as categorias Q1, Q2 e Q3. Além disso, as Tabelas A.17 a A.20 sugerem que, de um modo geral, as correlações entre as notas antes e depois do questionário não diferem entre si para Q1, Q2 e Q3. Pelo fato dessas correlações serem altas, há indícios de que a associação linear entre as notas antes e depois é forte, independente do desempenho dos alunos. É importante ressaltar que, devido ao grande número de questionários não preenchidos, em alguns instantes de aplicação, a amostra em cada um das categorias é pequena.

Para comparar as notas dadas pelo professor e dadas pelo aluno por desempenho, foram construídas as figuras C.17 a C.24. Para algumas variáveis, nos dois grupos e condições (antes e depois do preenchimento do questionário), observa-se, nos últimos instantes, indícios de que a diferença entre a média do professor e a média do aluno varia com o desempenho deste. Isso fica bem caracterizado, por exemplo, na variável conceito geral, no grupo B, antes do questionário (Figura C.24). No segundo instante, os alunos com melhor desempenho (categoria Q3) parecem apresentar maior média da diferença das notas do aluno e do professor do que os de desempenho mediano (categoria Q2). Há indícios ainda de que a média dessa diferença é menor na categoria Q1 do que nas demais. Como apenas para a categoria Q1 essa média da diferença parece ser igual a 0, pode-se argumentar que os alunos de pior desempenho são os únicos que, em média, se avaliam da mesma forma que o professor os avalia. As correlações entre as notas dadas pelo professor e dadas pelo aluno são, predominantemente, baixas para os três grupos de desempenho (Tabelas A.21 a A.24). Isso é um indício de que a associação linear entre a nota do professor e do aluno é baixa, independente do desempenho do aluno.

Para comparar as notas de conceito oficial com as notas de conceito dadas antes e após o preenchimento do questionário no final de cada um dos semestres, foram construídas duas variáveis diferença. Elas foram obtidas subtraindo-se as notas de conceito antes do questionário, para a primeira variável e depois do questionário, para a

segunda, da nota de conceito oficial. Boxplots dessas variáveis (Gráficos B.6 e B.7) parecem indicar que a média do conceito oficial é maior do que as médias de conceito dadas antes e depois do preenchimento dos questionários. Pode-se ver que, nas duas variáveis, nos dois grupos e nos dois semestres, há observações cuja diferença é, em módulo maior ou igual a 1,5. Isso mostra que, para alguns alunos, os professores deram notas de conceito oficial bem diferente das notas de conceito do questionário. As tabelas de correlações entre a nota de conceito oficial e as notas de conceito do questionário (Tabelas A.25 e A.26) mostram indícios de forte associação linear entre essas variáveis.

5. Análise inferencial

5.1. Comparação do desempenho dos alunos dos grupos A e B

Com o objetivo de comparar as médias da variável média final nos grupos A e B, foi feito um teste t bicaudal de igualdade de médias de duas populações normais independentes (Bussab e Morettin, 1987). A versão mais simples desse teste supõe que a variância da média final é a mesma nos dois grupos. Para testar essa suposição, foi feito o teste de homogeneidade de variâncias de Levene (Neter et al., 1996). Não foi rejeitada a hipótese de que a variância da média final é a mesma nos dois grupos ($p=0,7240$). O teste t também supõe que as duas populações tenham distribuição normal. O teste de aderência de Anderson-Darling nos leva a concluir que a média final não tem distribuição normal nos dois grupos ($p= 0,0020$ para os dois grupos). Porém, observando-se os histogramas das médias finais nos dois grupos (Gráfico D.1), pode-se notar que o desvio na suposição de normalidade não é muito grande. Como o teste t é robusto com relação a pequenos desvios de normalidade e a amostra é grande, os resultados encontrados continuam válidos. A última suposição do teste é que as médias finais de dois alunos quaisquer são independentes. Pelo fato dessas médias serem associadas a indivíduos distintos, aceita-se a validade dessa suposição. Através do

teste t, conclui-se (Tabela F.1) que a média da variável média final é a mesma nos dois grupos ($p=0,8449$).

Conforme comentado na análise descritiva, 14 alunos do grupo A responderam menos de 5 questionários. Pelo fato desses estudantes terem preenchido menos da metade dos questionários, eles não podem ser considerados alunos típicos do grupo A. Para verificar se eram esses alunos atípicos que estavam levando à conclusão de que os dois grupos tinham, em média, o mesmo desempenho, foi refeito o teste t, excluindo-se, no grupo A, os estudantes que preencheram menos de 5 questionários. Foi aceita a homogeneidade de variâncias ($p=0,5480$) e a fuga de normalidade ($p= 0,0030$ para o grupo A e $0,0020$ para o B) não parece ser muito grande (Gráfico D.2). A conclusão do teste ($p=0,5615$) é que, mesmo excluindo-se os alunos atípicos de A, a média da variável média final é a mesma nos dois grupos.

Foi feito também um teste t para comparar as médias da variável média das notas de conceito oficial do primeiro e segundo semestre nos dois grupos. Não se comparou as médias da nota de conceito oficial para os dois grupos em cada semestre porque a análise descritiva sugeriu que os professores do Hospital das Clínicas dão, em média, notas mais altas que os do Hospital Universitário. Isso poderia levar à conclusão de diferença de desempenho entre os dois grupos devido à influência do hospital e não do questionário. A conclusão do teste é que não há evidências de que a média de conceito oficial seja diferente nos dois grupos ($p=0,5220$). Como aceitou-se que as variâncias são homogêneas nos dois grupos ($p=0,4310$) e o desvio da suposição de normalidade ($p= 0,0050$ para o grupo A e $0,0140$ para o grupo B) não é muito grande (Gráfico D.3), o resultado encontrado é confiável.

Nos três testes de comparação do desempenho médio dos dois grupos de alunos, foi observado um leve desvio da suposição de normalidade. Em virtude disso, para confirmar os resultados, repetiu-se a análise utilizando-se o teste não paramétrico de Mann-Whitney (Conover, 1971), que não exige que as populações tenham distribuição normal. Os níveis descritivos dos três testes são altos (Tabela F.2), confirmando a conclusão de que não há diferença no desempenho médio dos dois grupos.

5.2. Comparação das notas de conceito oficial com as notas de conceito do questionário dentro de cada grupo

A comparação das notas de conceito oficial com as notas de conceito do questionário dadas pelo professor foi feita através do ajuste de modelos de efeitos aleatórios (Lindsey, 1993). Essa técnica é recomendada para a comparação de tratamentos quando há dependência entre as observações, bem como a existência de dados incompletos. Nesse estudo, essa dependência existe pelo fato de um mesmo aluno receber tanto as notas de conceito oficial quanto às de conceito do questionário. Além disso, o fato do mesmo aluno receber essas notas ao longo do tempo (mensalmente no grupo A e semestralmente no grupo B) também produz uma estrutura de dependência entre as notas. Por outro lado, como não se dispõe de todas as notas para todos os instantes de aplicação do questionário, o problema envolve dados incompletos.

Inicialmente, comparou-se as notas no grupo A. Foram comparadas as notas de conceito dadas pelo professor nos questionários 4 e 8 com as notas de conceito oficial dadas respectivamente ao final do primeiro e do segundo semestre. Foram utilizados os questionários 4 e 8 pelo fato deles serem preenchidos no final do semestre e, portanto, logo após a avaliação de conceito oficial do professor.

O primeiro modelo obtido tem como variável resposta a nota e variáveis explicativas o tipo de nota (oficial ou dada pelo professor antes do questionário) e o semestre (primeiro ou segundo). As suposições básicas, bem como alguns detalhes da formulação do modelo, são apresentados no apêndice E. Conforme pode ser visto na Tabela F.3, rejeitamos a hipótese de que qualquer um dos parâmetros seja igual a 0. Isso significa que a diferença média entre a nota de conceito final e do questionário não é constante nos dois semestres. Portanto, deve-se comparar essas duas notas separadamente em cada um dos semestres.

Através do método de Bonferroni (Neter et al., 1996), controlou-se o nível de significância conjunto dos 4 testes de interesse em 0,0500. Desse modo, para rejeitarmos hipóteses de igualdade entre as médias, o nível decritivo do teste deve ser menor do que 0,0125. Concluiu-se (Tabela F.4) que a média da nota de conceito oficial é maior do que a média da nota de conceito do questionário para o segundo semestre

($p < 0,0001$). Porém, para o primeiro semestre, não rejeitamos a hipótese de que essas médias são iguais ($p = 0,0434$). Além disso, concluímos que a média no instante 1 é maior do que no 2, para os dois tipos de nota ($p < 0,0001$).

A Tabela F.5 fornece a estimativa pontual das diferenças entre essas médias, bem como o seu intervalo de confiança com coeficiente de confiança conjunto de 95% (método de Bonferroni). Utilizando-se esse método, estima-se, com 95% de confiança conjunta, que a diferença entre a média da nota de conceito oficial e a média da nota de conceito antes do questionário, no segundo semestre, está entre 0,25 e 0,64. Interpretações similares valem para os demais intervalos construídos.

Na comparação da média da nota de conceito oficial com a média da nota de conceito depois do questionário, no grupo A, foram obtidos resultados equivalentes. As estimativas dos parâmetros, os níveis descritivos associados aos testes de significância desses parâmetros, os resultados dos testes de interesse, bem como os intervalos de confiança para a diferença entre as médias são apresentados nas Tabelas F.6 a F.8.

Para o grupo B, ao compararmos as médias das notas de conceito antes do questionário e as médias das notas de conceito oficial (Tabela F.9), não rejeitamos a hipótese de que o parâmetro de interação é igual a 0 ($p = 0,6194$). Retirou-se então o parâmetro de interação e reestimou-se o modelo (Tabela F.10). As estimativas dos parâmetros finais mostram que há efeito tanto de tipo de nota quanto de semestre. Utilizando-se o método de Bonferroni com nível de significância conjunto de 5% (Tabela F.11), conclui-se que a média da nota de conceito oficial é maior do que a média da nota antes do questionário dentro de cada semestre. Já a média da nota no semestre 2 é maior do que no semestre 1, tanto para a nota de conceito oficial quanto para a nota de conceito do questionário. A diferença entre as médias da nota de conceito oficial e de conceito antes do questionário para cada semestre está entre 0,24 e 0,49 com coeficiente de confiança conjunto de 95% (Tabela F.12).

Nas tabelas F.13 a F.16 são apresentados os resultados para a comparação da média da nota de conceito oficial com a média de conceito depois do questionário, no grupo B. Os resultados obtidos pouco diferem da comparação anterior feita no grupo B.

5.3. Comparações envolvendo as notas do questionário no grupo B

Para cada uma das 4 notas do questionário no grupo B, foi ajustado um modelo de efeitos aleatórios. As variáveis explicativas envolvidas são o avaliador (professor ou aluno), o instante de aplicação do questionário e a condição (antes ou depois do preenchimento do questionário).

O modelo final para a nota de anamnese, obtido após a retirada uma por uma das variáveis não significantes, contém, além da constante, apenas os efeitos principais de avaliador e instante (Tabela F.17). Com um nível de significância conjunto de 5% (Tabela F.18), concluiu-se que a média da nota de anamnese no instante 2 é maior do que a média no instante 1 ($p=0,0004$). Outra conclusão fornecida pelo modelo é que a média de anamnese dada pelo professor é maior do que a média dessa nota dada pelo aluno ($p<0,0001$). Pelo fato de não haver no modelo final nenhuma variável que envolva a condição, conclui-se que não há diferença na média de anamnese entre as notas dadas antes e depois do preenchimento do questionário. Utilizando-se um nível de confiança conjunto de 95% (Tabela F.19), estima-se, por exemplo, que a diferença entre a média da nota de anamnese dada pelo professor e dada pelo aluno está entre 0,35 e 0,76.

Para a variável raciocínio clínico, foram encontrados resultados equivalentes. As Tabelas F.23, F.24 e F.25 mostram, respectivamente, o ajuste do modelo final, os níveis descritos associados às comparações de interesse e os intervalos de confiança para as diferenças significantes entre as médias, com coeficiente de confiança conjunto de 95%.

O modelo final para a variável exame físico contém, além dos efeitos principais de avaliador e instante, os efeitos de interação entre condição e avaliador e condição e instante (Tabela F.20). Construindo-se as comparações de interesse (Tabela F.21), concluiu-se que há diferença entre as médias das notas antes e depois do questionário para o professor no instante 2 e para o aluno no instante 1 ($p=0,0018$). Para o professor, a média da nota de exame físico depois do questionário é maior do que a média antes do questionário. Já para o aluno no instante 1, a média de exame físico é maior antes do questionário do que depois deste. A estimativa da diferença entre essas médias, porém é pequena (está entre 0,03 e 0,27) com 95% de confiança conjunta

(Tabela F.22). Para a nota de exame físico, concluiu-se também que a média da nota no instante 2 é maior do que no instante 1, qualquer que seja a condição ($p < 0,0040$). Além disso, foram detectadas diferenças entre as médias das notas dadas pelo professor e dadas pelo aluno para as duas condições ($p < 0,0001$). As estimativas das diferenças entre essas médias são grandes. Para a nota de exame físico depois do questionário, por exemplo, estima-se que a média da nota do professor é de 0,97 a 1,53 unidades a mais do que a média do aluno (Tabela F.22).

Para a nota de conceito geral (Tabelas F.26 a F.28), concluiu-se apenas que a média do professor é maior do que a média do aluno ($p < 0,0001$). Em todas as demais comparações de interesse, não rejeitamos as igualdades de médias.

5.4. Comparações envolvendo as notas do questionário no grupo A

Os modelos construídos para as notas do questionário no grupo A são similares aos modelos ajustados para o grupo B. A única diferença está no fato de que, por dispor de um número maior de instantes de aplicação do questionário, os modelos para o grupo B possuem um número maior de variáveis explicativas. Com isso, o número de comparações de interesse aumentou substancialmente. Para evitar que o nível de significância individual se tornasse muito pequeno, nas comparações entre os instantes, testou-se apenas a igualdade das médias no instante 1 com os demais instantes. Não se testou, por exemplo, se as médias de cada uma das 4 notas são iguais nos instantes 3 e 4.

Para a variável anamnese, o modelo final contém o efeito principal de avaliador, variáveis relacionadas com o efeito principal do instante e ainda variáveis para a interação entre avaliador e instante (Tabela F.29). Concluiu-se que (Tabela F.30), para todos os instantes, a média da nota dada pelo professor é maior do que a média da nota dada pelo aluno ($p < 0,0001$). Além disso, a média de anamnese no instante 1 é menor que essa média em qualquer um dos outros instantes, tanto para as notas dadas pelo professor como para aquelas dadas pelo aluno ($p < 0,0001$). Intervalos de confiança com coeficiente de confiança conjunto de 95% (Tabela F.31) revelam, por exemplo, que a diferença entre a média da nota de anamnese dada pelo professor e dada pelo

aluno, no instante 8, é de 0,62 a 1,19, com 95% de confiança conjunta. Por outro lado estima-se que a diferença das médias de anamnese nos instantes 1 e 8 para o aluno, está entre 1,13 e 1,80. Não foi rejeitada a hipótese de que a média de anamnese é a mesma antes e depois do questionário.

Para as demais variáveis, os resultados são bastante semelhantes. Em todas elas, foram observadas diferenças entre as médias das notas dadas pelo professor e dadas pelo aluno em todos os instantes. Aceitou-se ainda, para todas as variáveis, que a média da nota é a mesma antes e depois do preenchimento do questionário. Para quase todos os instantes, em todas as variáveis, verificou-se que média da nota no instante 1 é menor do que nos demais instantes, tanto para as notas dadas pelo professor como para as notas dadas pelo aluno. Os resultados obtidos para as variáveis exame físico, raciocínio clínico e conceito geral estão nas Tabelas F.32 a F.40.

6. Conclusões

A análise estatística mostrou que não há diferenças entre as médias de desempenho dos alunos avaliados mensalmente com o questionário e os avaliados semestralmente.

Para todos os níveis de instante e condição, em todas as variáveis, a média das notas dadas pelo professor é maior do que a média das notas dadas pelo aluno.

Com exceção da variável exame físico no grupo B, não há diferença entre as médias das notas antes e depois do questionário.

Quanto à variável tempo, observou-se que a média das notas não é constante em todos os instantes. Além disso, a nota média no último questionário é maior do que a nota média no primeiro questionário, para todas as variáveis, condições e avaliadores.

Com exceção do primeiro semestre no grupo A, a média das notas de conceito oficial é maior do que a média das notas de conceito do questionário dadas pelo professor.

Apêndice A

Tabelas da análise descritiva

Tabela A.1. Distribuição de sexo por grupo

	Grupo	
	A	B
homens	57%	61%
mulheres	43%	39%

Tabela A.2. Médias e desvios padrão amostrais para a variável média final

Grupo	Média	Desvio Padrão
A	8,00	0,71
B	7,98	0,73

Tabela A.3. Número de questionários preenchidos pelos alunos do grupo A

questionários preenchidos	freqüência
8	3
7	20
6	32
5	17
4	9
3	4
2	1
1	0
0	0

Tabela A.4. Número de questionários preenchidos pelos alunos do grupo B

questionários preenchidos	freqüência
2	88
1	2
0	0

Tabela A.5. Número de questionários preenchidos pelos professores em relação aos seus alunos do grupo A

questionários preenchidos	freqüência
8	32
7	2
6	9
5	9
4	35
3	0
2	0
1	0
0	0

Tabela A.6. Número de questionários preenchidos pelos professores em relação aos seus alunos do grupo B

questionários preenchidos	freqüência
2	43
1	46
0	1

Tabela A.7. Médias e desvios padrão amostrais para a variável média final (sem os alunos do grupo A que preencheram menos de 5 questionários)

Grupo	Média	Desvio Padrão
A	8,04	0,71
B	7,98	0,73

Tabela A.8. Médias e desvios padrão amostrais para a variável conceito oficial

Grupo	Média	Desvio Padrão
A no 1º semestre	8,95	0,76
A no 2º semestre	8,58	0,85
B no 1º semestre	8,62	1,05
B no 2º semestre	9,05	0,86

Tabela A.9. Correlações entre as notas de anamnese antes e depois do questionário

instante	grupo	professor	aluno
1	A	0,93	0,90
2	A	0,96	0,80
3	A	0,97	0,78
4	A	0,95	0,92
5	A	0,93	0,88
6	A	0,88	0,90
7	A	0,95	0,86
8	A	0,92	0,93
1	B	0,92	0,85
2	B	0,83	0,83

Tabela A.10. Correlações entre as notas de exame físico antes e depois do questionário

instante	grupo	professor	aluno
1	A	0,96	0,90
2	A	0,94	0,92
3	A	0,91	0,89
4	A	0,85	0,93
5	A	0,82	0,88
6	A	0,89	0,91
7	A	0,94	0,92
8	A	0,83	0,90
1	B	0,95	0,75
2	B	0,86	0,89

Tabela A.11. Correlações entre as notas de raciocínio clínico antes e depois do questionário

instante	grupo	professor	aluno
1	A	0,94	0,92
2	A	0,90	0,94
3	A	0,86	0,93
4	A	0,92	0,97
5	A	0,82	0,93
6	A	0,80	0,92
7	A	0,93	0,94
8	A	0,93	0,92
1	B	0,93	0,89
2	B	0,80	0,91

Tabela A.12. Correlações entre as notas de conceito geral antes e depois do questionário

instante	grupo	professor	aluno
1	A	0,96	0,88
2	A	0,95	0,92
3	A	0,97	0,91
4	A	0,97	0,96
5	A	0,90	0,87
6	A	0,91	0,92
7	A	0,91	0,95
8	A	0,93	0,94
1	B	0,96	0,85
2	B	0,93	0,87

Tabela A.13. Correlações entre as notas de anamnese dadas pelo professor e pelo aluno

instante	grupo	antes do preenchimento	depois do preenchimento
1	A	0,35	0,37
2	A	-0,04	0,01
3	A	0,27	0,18
4	A	0,02	0,02
5	A	0,37	0,28
6	A	0,26	0,19
7	A	0,01	0,08
8	A	0,07	0,05
1	B	0,35	0,29
2	B	-0,29	-0,26

Tabela A.14. Correlações entre as notas de exame físico dadas pelo professor e pelo aluno

instante	grupo	antes do preenchimento	depois do preenchimento
1	A	0,41	0,42
2	A	-0,04	-0,07
3	A	0,15	0,22
4	A	0,08	-0,01
5	A	0,34	0,27
6	A	0,02	0,02
7	A	-0,03	0,05
8	A	0,20	< 0,01
1	B	0,20	0,15
2	B	-0,16	0,05

Tabela A.15. Correlações entre as notas de raciocínio clínico dadas pelo professor e pelo aluno

instante	grupo	antes do preenchimento	depois do preenchimento
1	A	0,44	0,42
2	A	0,09	0,15
3	A	0,22	0,07
4	A	-0,38	-0,17
5	A	0,25	0,19
6	A	< 0,01	0,13
7	A	-0,04	0,02
8	A	0,02	-0,01
1	B	0,11	0,12
2	B	-0,08	-0,05

Tabela A.16. Correlações entre as notas de conceito geral dadas pelo professor e pelo aluno

instante	grupo	antes do preenchimento	depois do preenchimento
1	A	0,40	0,37
2	A	-0,26	-0,23
3	A	0,16	0,17
4	A	0,18	-0,01
5	A	0,49	0,29
6	A	0,13	0,10
7	A	-0,05	0,02
8	A	0,04	-0,07
1	B	0,18	0,15
2	B	-0,23	-0,12

Tabela A.17. Correlações entre as notas de anamnese antes e depois do questionário por desempenho dos alunos

instante	grupo	professor			aluno		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,94	0,81	0,94	0,94	0,80	0,93
2	A	0,94	0,96	0,97	0,85	0,92	0,72
3	A	0,99	0,93	0,96	0,91	0,83	0,63
4	A	1,00	0,98	0,96	0,91	0,98	0,93
5	A	0,89	0,96	0,95	0,94	0,62	0,91
6	A	0,95	0,86	0,83	0,97	0,60	0,92
7	A	0,96	0,88	0,93	0,93	0,88	0,82
8	A	0,97	0,82	0,88	0,94	0,81	0,95
1	B	0,80	0,99	0,97	0,80	0,88	0,87
2	B	0,49	0,82	0,89	0,88	0,80	0,84

Tabela A.18. Correlações entre as notas de exame físico antes e depois do questionário por desempenho dos alunos

instante	grupo	professor			aluno		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,96	0,90	0,97	0,95	0,79	0,88
2	A	0,95	1,00	0,95	0,95	1,00	0,91
3	A	0,93	0,64	0,94	0,91	0,90	0,87
4	A	0,97	0,98	0,81	0,98	0,79	0,96
5	A	0,83	0,79	0,80	0,94	0,73	0,88
6	A	0,91	0,74	0,93	0,95	0,95	0,85
7	A	0,95	0,87	0,90	0,97	0,89	0,87
8	A	0,86	0,44	0,85	0,89	0,91	0,90
1	B	0,91	0,96	0,95	0,59	0,92	0,58
2	B	0,81	0,65	0,96	0,89	0,86	0,90

Tabela A.19. Correlações entre as notas de raciocínio clínico antes e depois do questionário por desempenho dos alunos

instante	grupo	professor			aluno		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,91	0,98	0,95	0,98	0,79	0,94
2	A	0,95	1,00	0,72	0,94	0,98	0,96
3	A	0,82	0,90	0,93	0,96	0,95	0,89
4	A	0,99	0,92	0,80	0,98	1,00	0,94
5	A	0,84	0,55	0,91	0,95	0,77	0,94
6	A	0,90	0,62	0,87	0,94	0,81	0,93
7	A	0,98	0,96	0,83	0,94	0,96	0,94
8	A	0,99	0,96	0,88	0,95	0,83	0,92
1	B	0,90	0,97	0,93	0,89	0,90	0,88
2	B	0,54	0,57	0,35	0,85	0,93	0,94

Tabela A.20. Correlações entre as notas de conceito geral antes e depois do questionário por desempenho dos alunos

instante	grupo	professor			aluno		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,95	0,99	0,97	0,96	0,91	0,78
2	A	0,94	1,00	0,96	0,95	0,96	0,87
3	A	0,98	0,90	0,98	0,94	0,94	0,89
4	A	0,97	1,00	0,96	0,99	0,98	0,90
5	A	0,90	0,80	0,91	0,92	0,55	0,87
6	A	0,96	0,82	0,86	0,98	0,74	0,93
7	A	0,95	0,70	0,83	0,99	0,91	0,93
8	A	0,94	0,75	0,94	0,97	0,90	0,92
1	B	0,96	0,97	0,95	0,91	0,85	0,75
2	B	0,76	0,86	1,00	0,90	0,81	0,89

Tabela A.21. Correlações entre as notas de anamnese dadas pelo professor e pelo aluno por desempenho dos alunos

instante	grupo	antes do questionário			depois do questionário		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,35	0,12	0,31	0,39	0,24	0,28
2	A	-0,03	0,88	-0,24	-0,30	0,69	0,01
3	A	0,38	0,43	0,14	0,42	-0,15	-0,03
4	A	0,18	ND *	-0,30	0,30	ND *	-0,21
5	A	0,40	0,65	0,30	0,22	0,35	0,29
6	A	0,26	0,63	0,21	0,11	0,32	0,18
7	A	0,13	0,19	-0,18	0,10	0,51	-0,17
8	A	0,24	-0,04	-0,12	0,32	-0,09	-0,24
1	B	0,37	0,39	0,21	0,19	0,44	0,14
2	B	-0,11	-0,22	0,08	0,13	-0,28	0,11

* Não foi apresentado o valor da correlação por haver apenas 2 notas nesse instante para essa categoria

Tabela A.22. Correlações entre as notas de exame físico dadas pelo professor e pelo aluno por desempenho dos alunos

instante	grupo	antes do questionário			depois do questionário		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,44	0,54	0,41	0,44	0,13	0,39
2	A	-0,26	0,29	0,15	-0,33	0,29	0,19
3	A	0,37	0,08	-0,09	0,51	-0,32	-0,12
4	A	-0,05	ND *	0,11	-0,05	ND *	0,17
5	A	0,27	0,38	0,40	0,09	0,44	0,37
6	A	0,02	-0,23	0,17	0,14	-0,05	-0,06
7	A	0,07	0,21	-0,15	0,07	0,49	-0,08
8	A	0,05	0,20	0,33	-0,03	0,04	0,06
1	B	0,12	0,28	0,29	-0,13	0,21	0,01
2	B	-0,28	-0,17	0,24	0,17	0,22	0,18

* Não foi apresentado o valor da correlação por haver apenas 2 notas nesse instante para essa categoria

Tabela A.23. Correlações entre as notas de raciocínio clínico dadas pelo professor e pelo aluno por desempenho dos alunos

instante	grupo	antes do questionário			depois do questionário		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,46	0,43	0,38	0,54	0,04	0,41
2	A	0,03	-0,01	0,16	0,03	0,09	0,35
3	A	0,61	-0,25	0,01	0,42	< 0,01	-0,35
4	A	-0,43	ND *	-0,32	-0,15	ND *	-0,02
5	A	0,19	0,27	0,29	0,10	0,40	0,12
6	A	0,03	-0,10	0,07	0,19	0,17	0,02
7	A	0,02	0,15	-0,17	0,02	0,18	-0,12
8	A	-0,10	-0,14	0,16	-0,07	0,37	0,01
1	B	0,05	0,15	-0,01	0,14	-0,02	-0,05
2	B	0,13	-0,09	-0,19	0,07	0,01	0,24

* Não foi apresentado o valor da correlação por haver apenas 2 notas nesse instante para essa categoria

Tabela A.24. Correlações entre as notas de conceito geral dadas pelo professor e pelo aluno por desempenho dos alunos

instante	grupo	antes do questionário			depois do questionário		
		Q1	Q2	Q3	Q1	Q2	Q3
1	A	0,47	0,37	0,49	0,56	0,37	0,38
2	A	-0,58	-0,15	-0,09	-0,56	-0,21	0,11
3	A	0,46	-0,30	0,02	0,57	-0,25	-0,14
4	A	0,24	ND *	0,24	0,09	ND *	0,25
5	A	0,34	0,53	0,61	0,15	0,29	0,37
6	A	0,06	0,32	0,20	0,10	0,16	0,11
7	A	0,02	0,41	-0,24	0,06	0,63	-0,16
8	A	0,14	0,08	0,01	0,09	-0,14	-0,15
1	B	0,18	0,31	-0,16	< 0,01	0,27	-0,16
2	B	-0,13	-0,05	-0,16	0,25	-0,03	-0,13

* Não foi apresentado o valor da correlação por haver apenas 2 notas nesse instante para essa categoria

Tabela A.25. Correlações entre o conceito oficial e o conceito antes do questionário

semestre	grupo A	grupo B
1	0,77	0,76
2	0,61	0,82

Tabela A.26. Correlações entre o conceito oficial e o conceito depois do questionário

semestre	grupo A	grupo B
1	0,78	0,75
2	0,63	0,83

Apêndice B

Boxplots

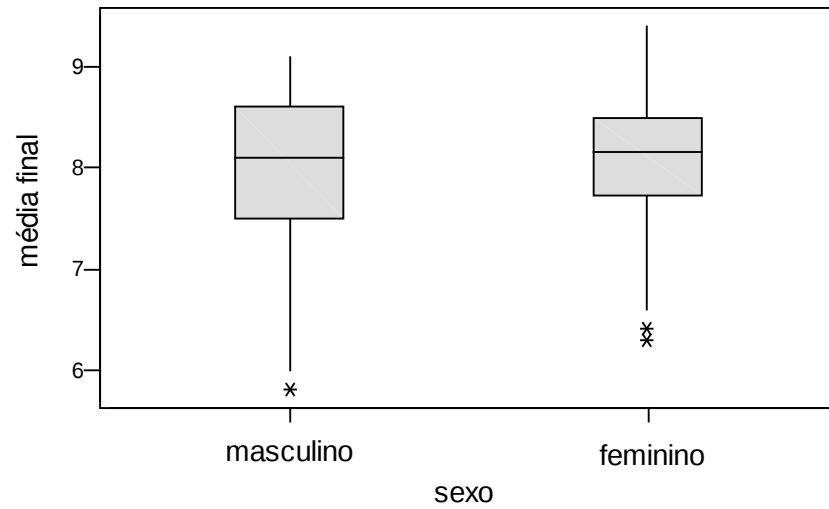
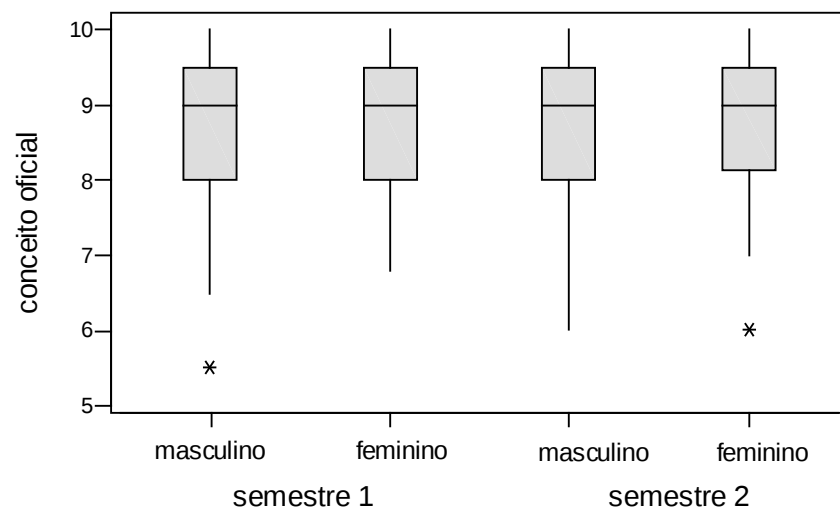
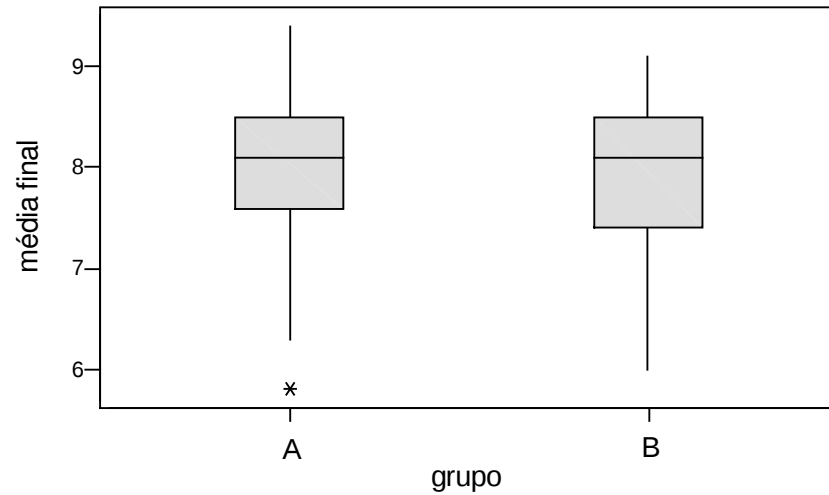
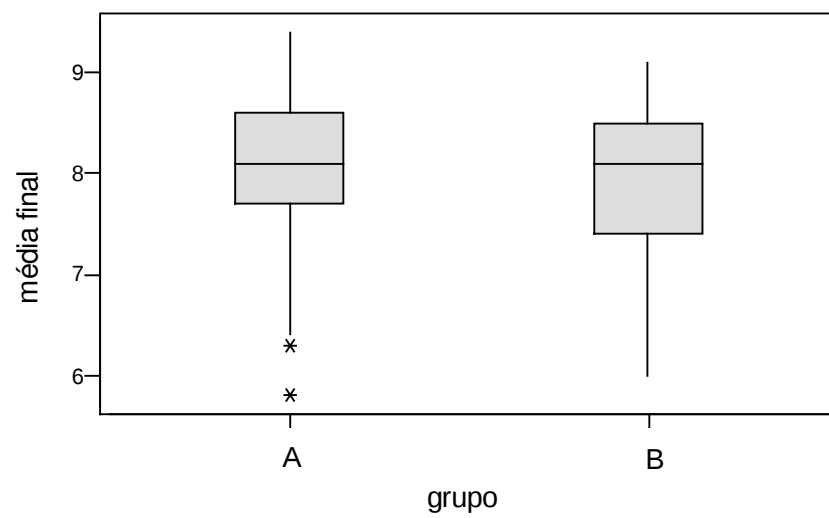
Gráfico B.1. Boxplots da média final por sexo**Gráfico B.2.** Boxplots do conceito oficial por sexo

Gráfico B.3. Boxplots da média final por grupo**Gráfico B.4.** Boxplots da média final por grupo (sem os alunos atípicos do grupo A)*

* Alunos que preencheram menos de 5 questionários

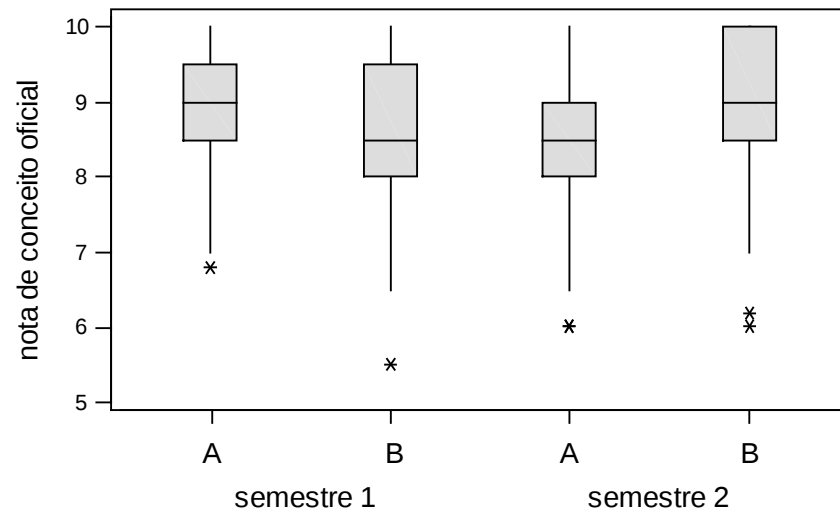
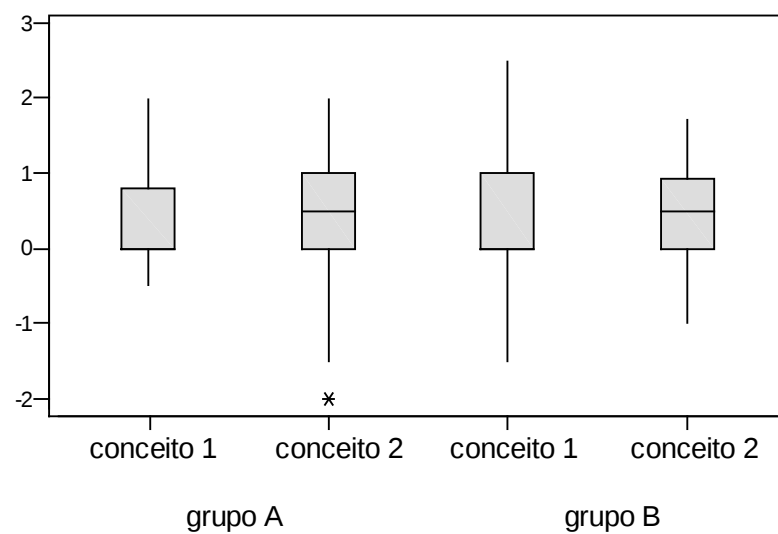
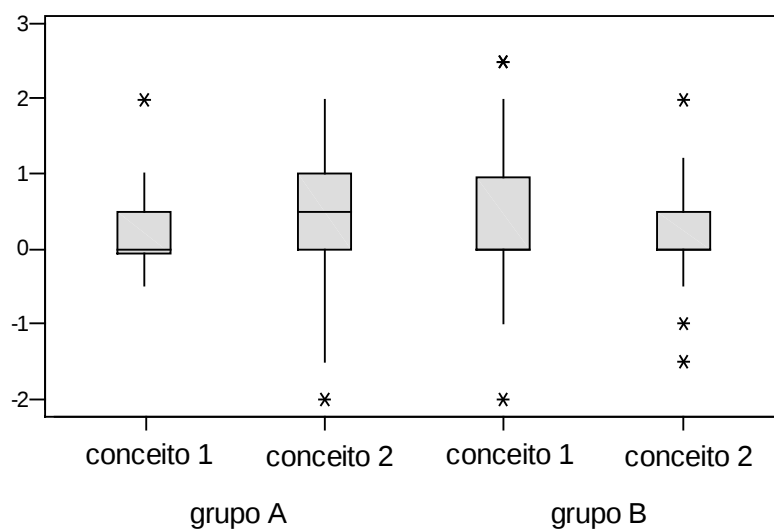
Gráfico B.5. Boxplots do conceito oficial por grupo**Gráfico B.6.** Boxplots da variável diferença (conceito oficial – conceito antes do questionário)

Gráfico B.7. Boxplots da variável diferença (conceito oficial – conceito depois do questionário)



Apêndice C

Gráficos de perfis

Figura C.1. Média para a variável diferença (depois do questionário – antes do questionário) em anamnese por grupo

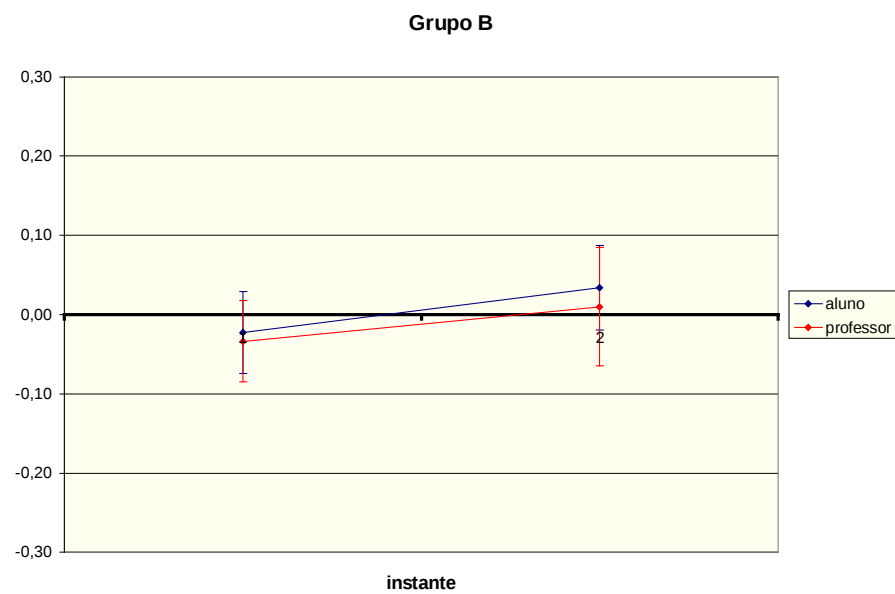
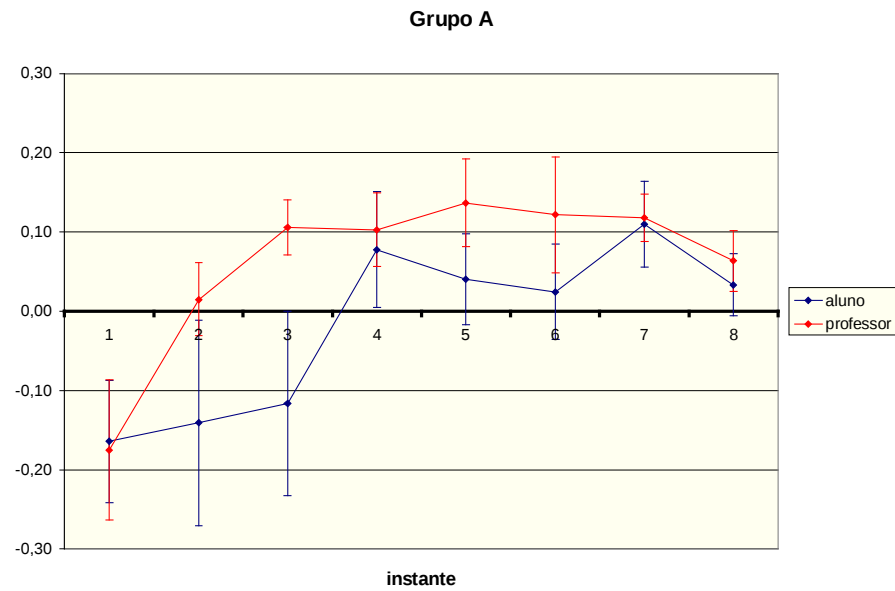


Figura C.2. Média para a variável diferença (depois do questionário – antes do questionário) em exame físico por grupo

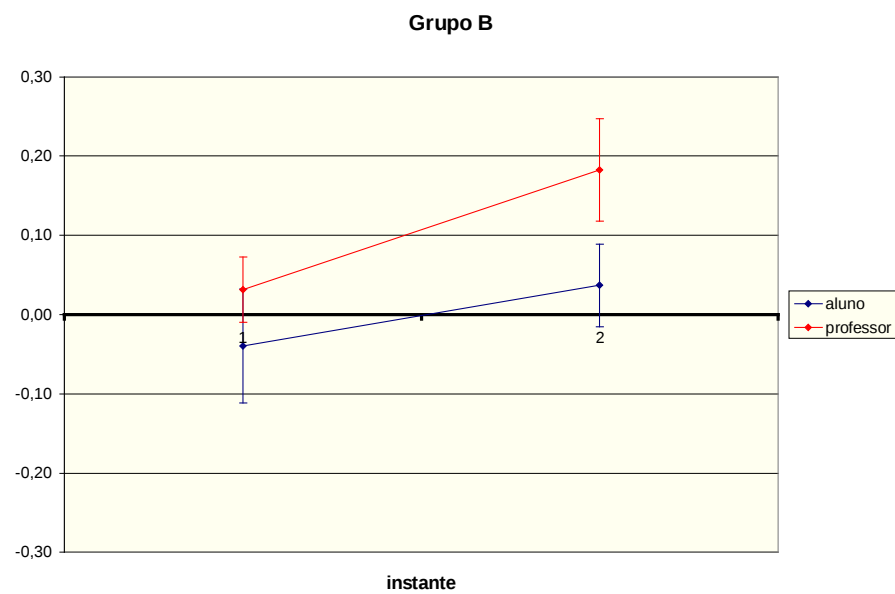
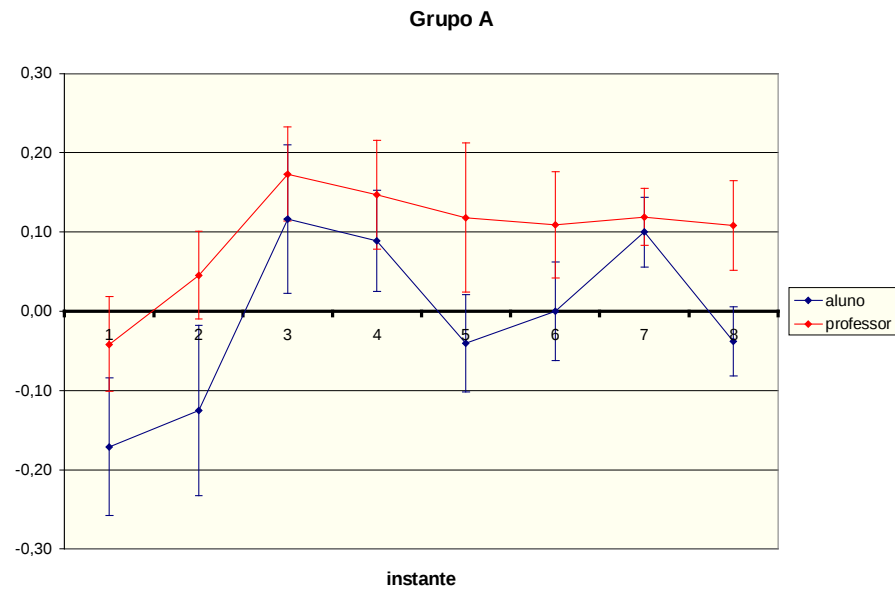


Figura C.3. Média para a variável diferença (depois do questionário – antes do questionário) em raciocínio clínico por grupo

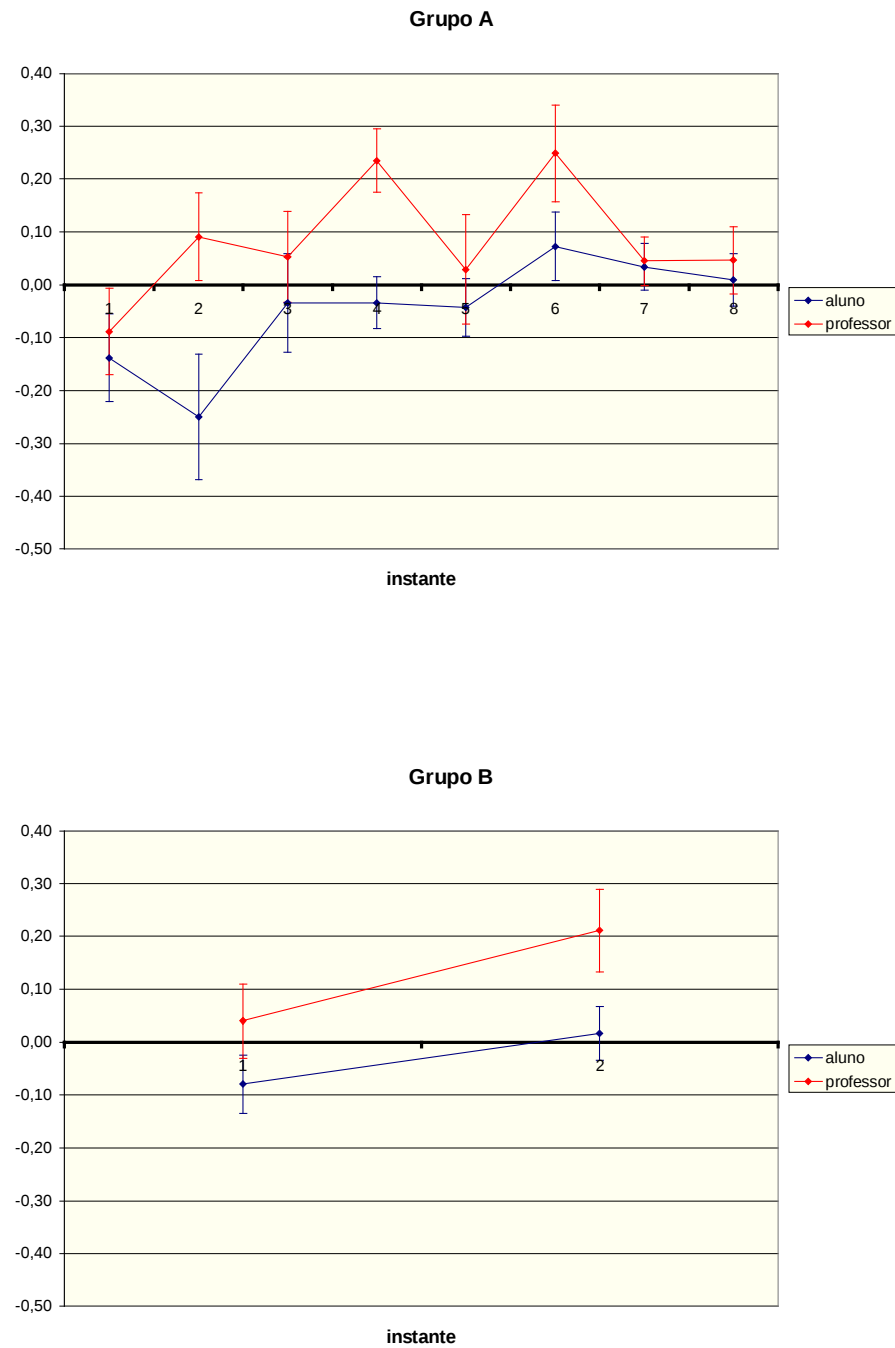


Figura C.4. Média para a variável diferença (depois do questionário – antes do questionário) em conceito geral por grupo

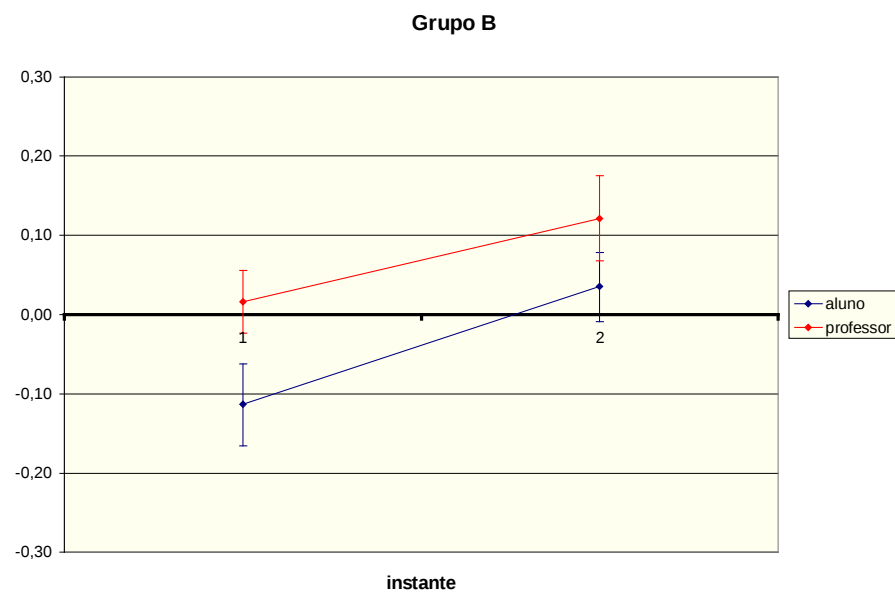
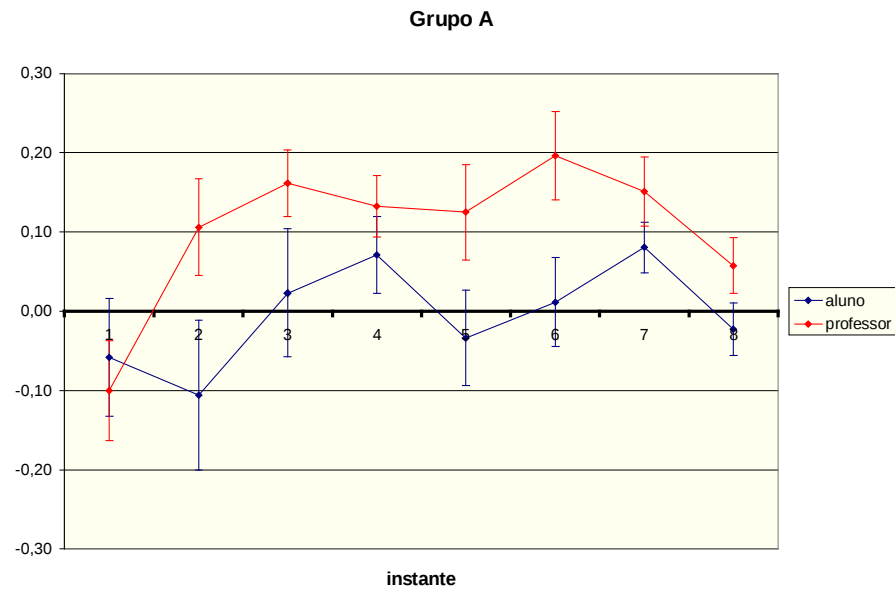


Figura C.5. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em anamnese por grupo

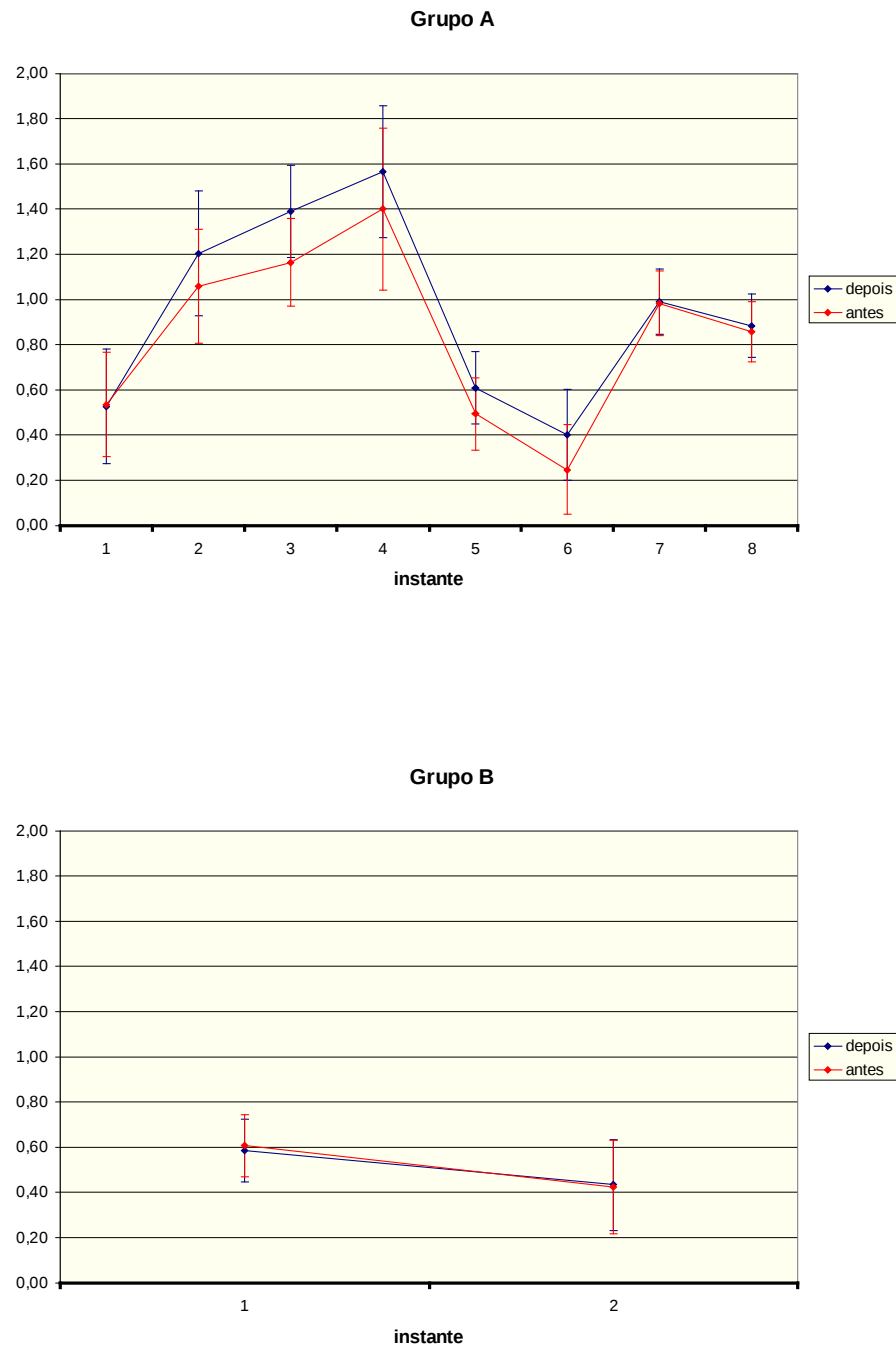


Figura C.6. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em exame físico por grupo

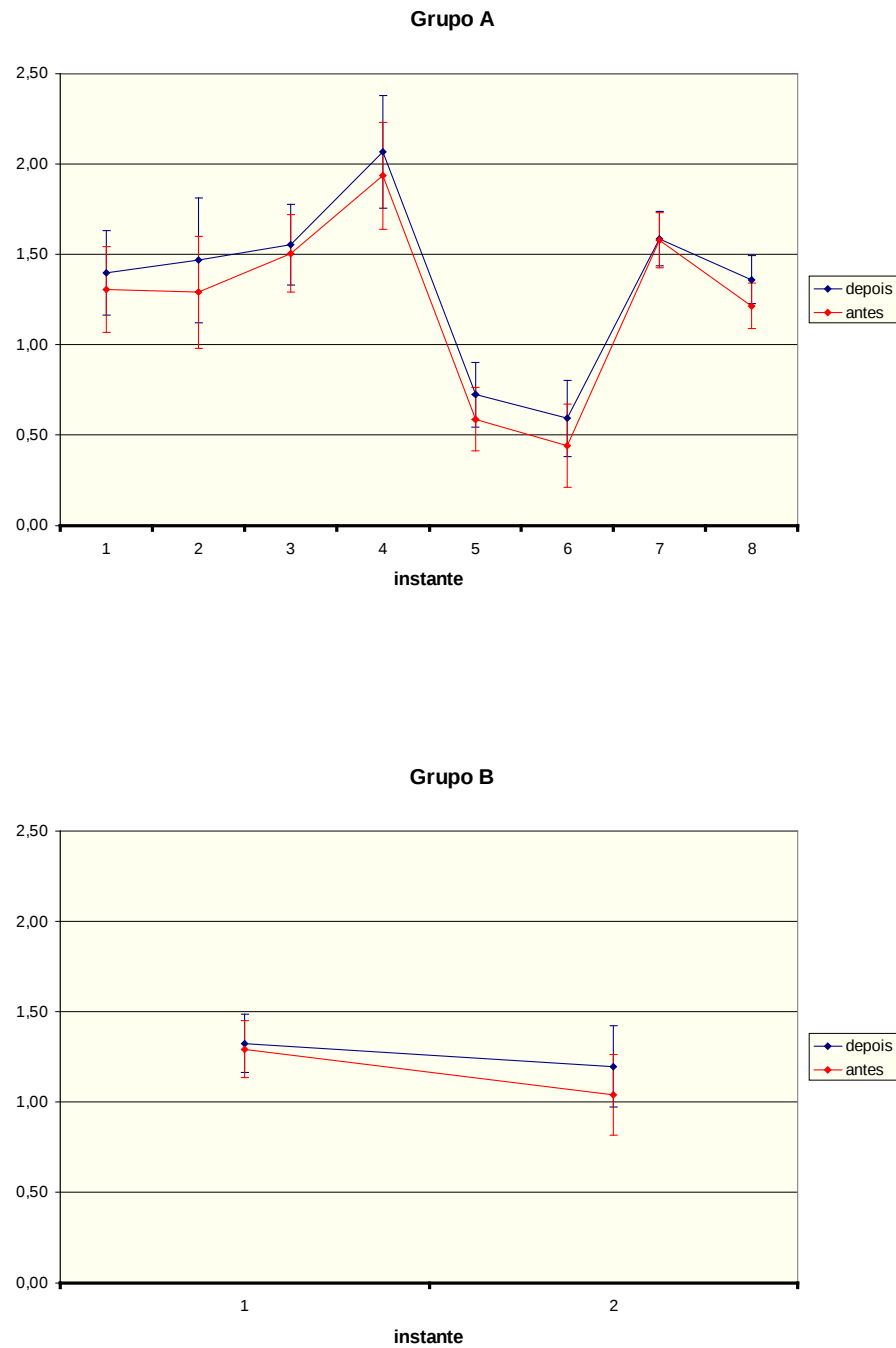


Figura C.7. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em raciocínio clínico por grupo

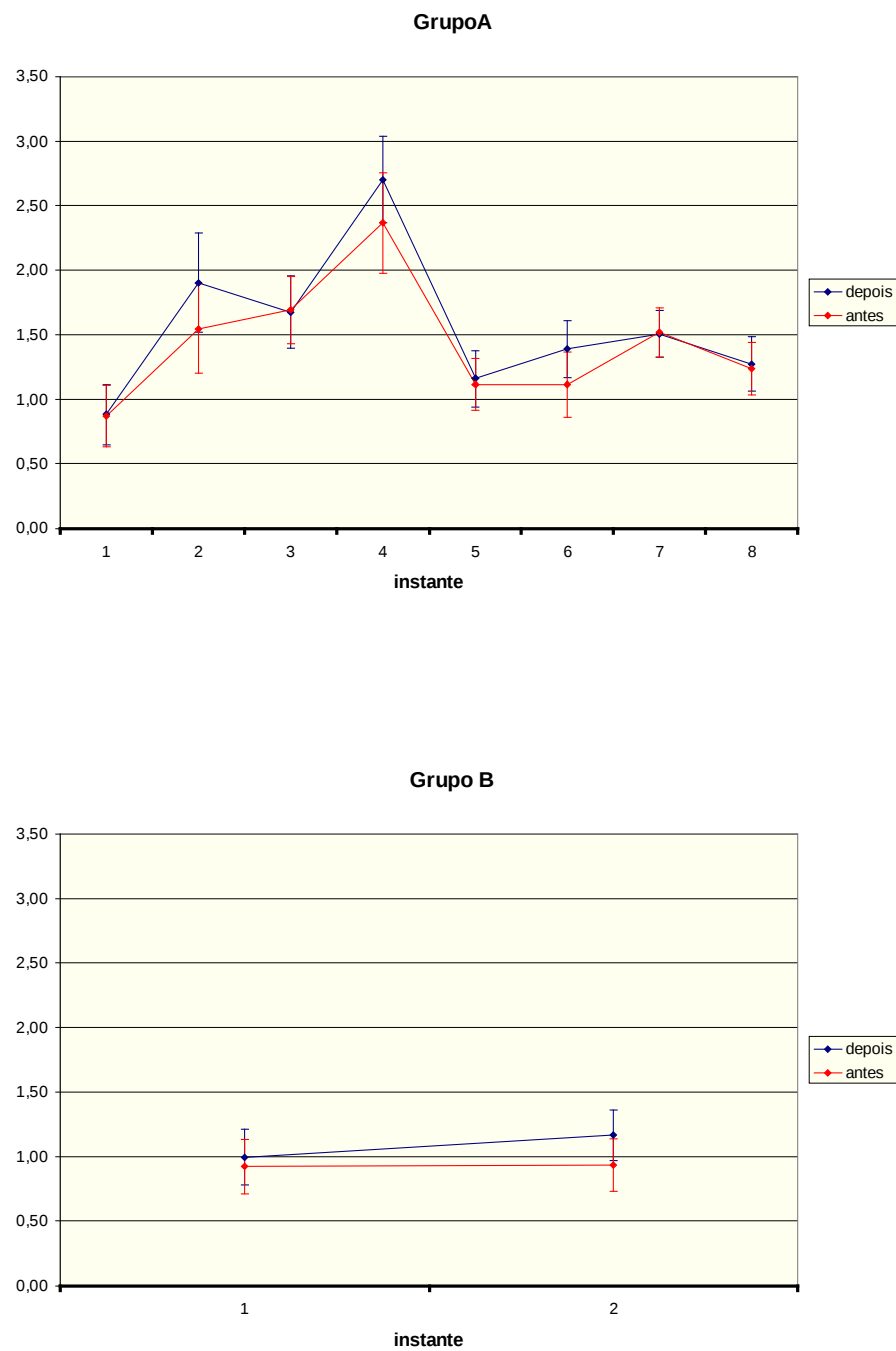


Figura C.8. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em conceito geral por grupo

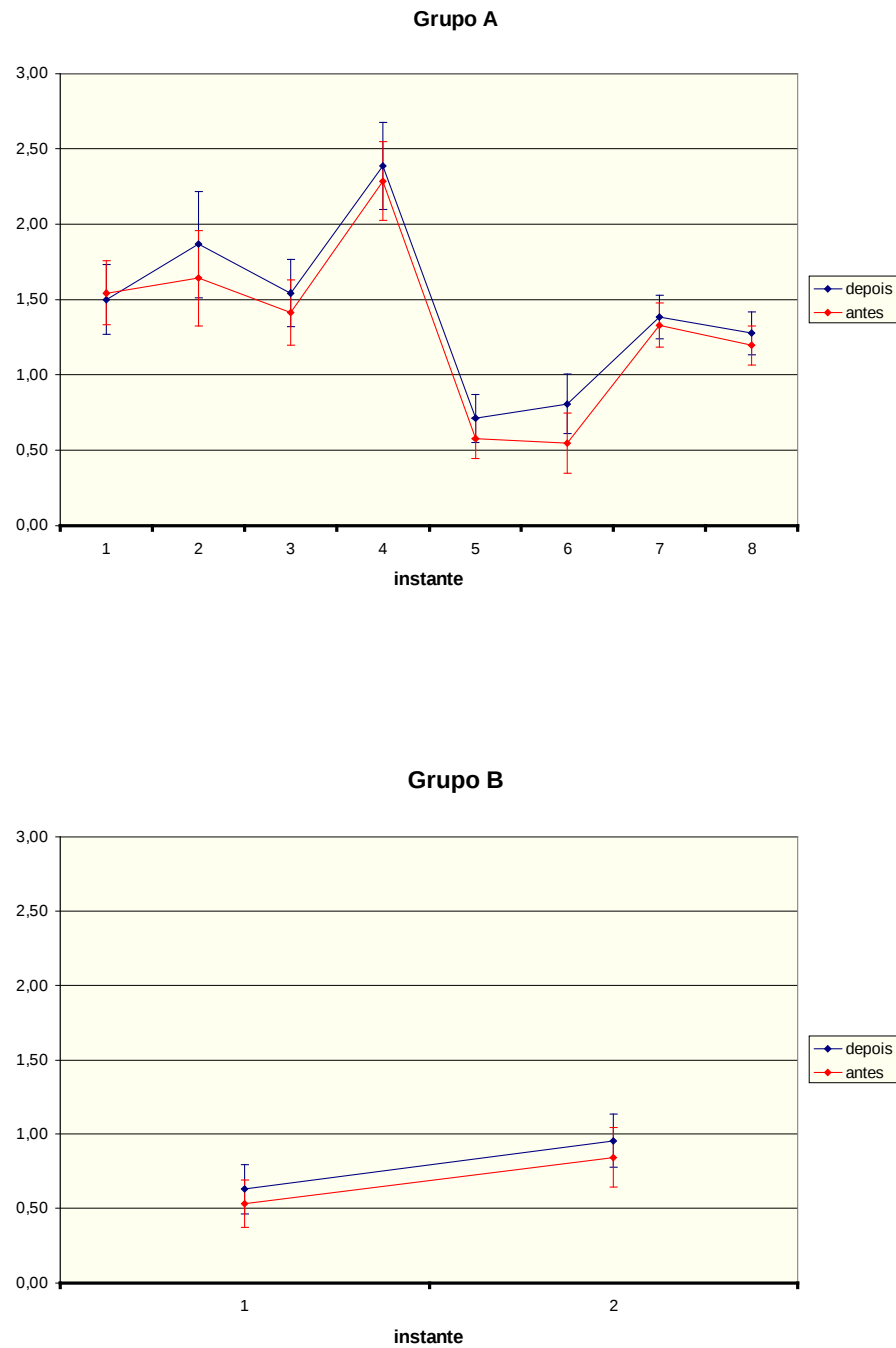


Figura C.9. Média para a variável diferença (depois do questionário – antes do questionário) em anamnese por avaliador no grupo A

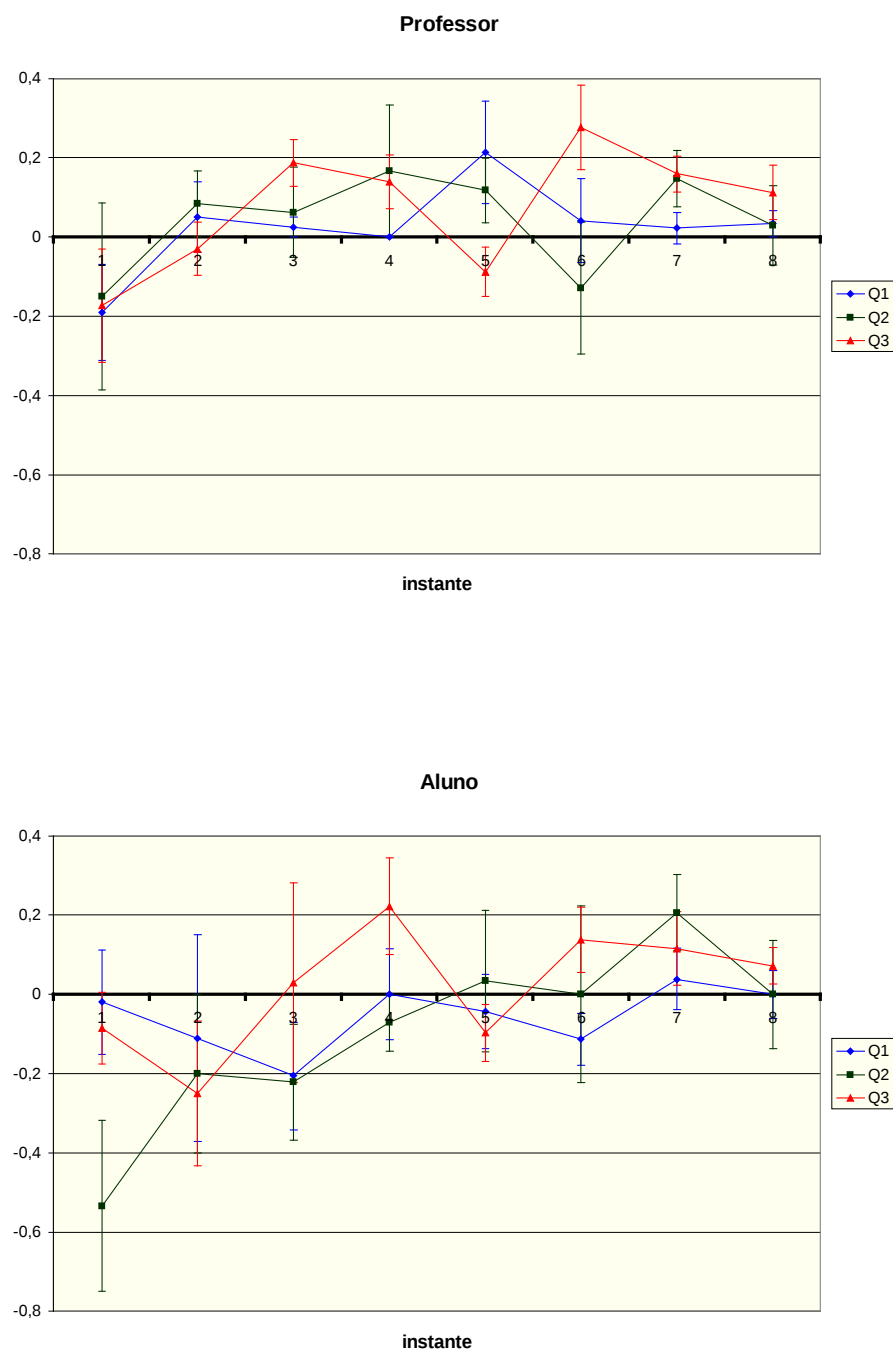


Figura C.10. Média para a variável diferença (depois do questionário – antes do questionário) em anamnese por avaliador no grupo B

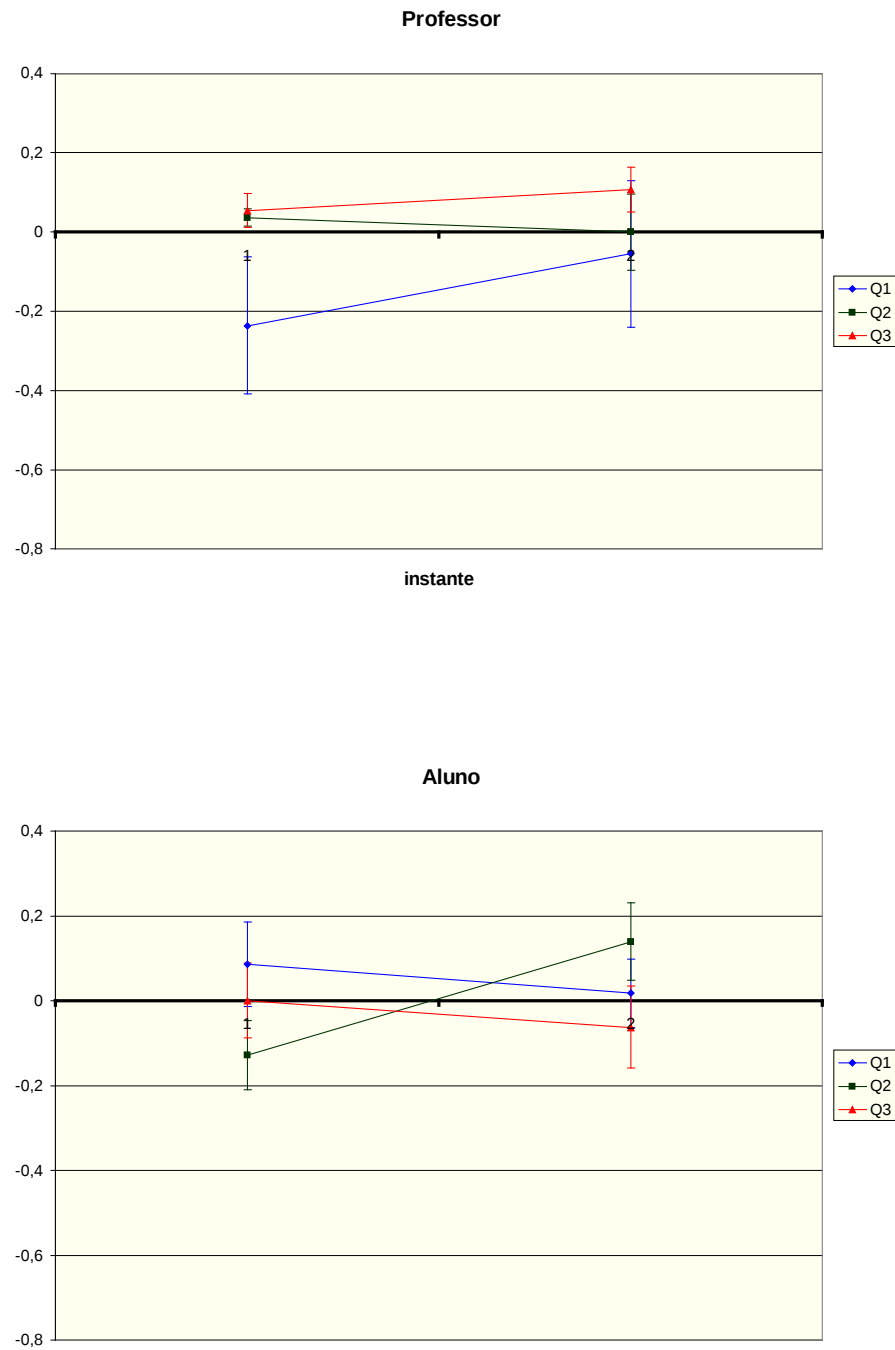


Figura C.11. Média para a variável diferença (depois do questionário – antes do questionário) em exame físico por avaliador no grupo A

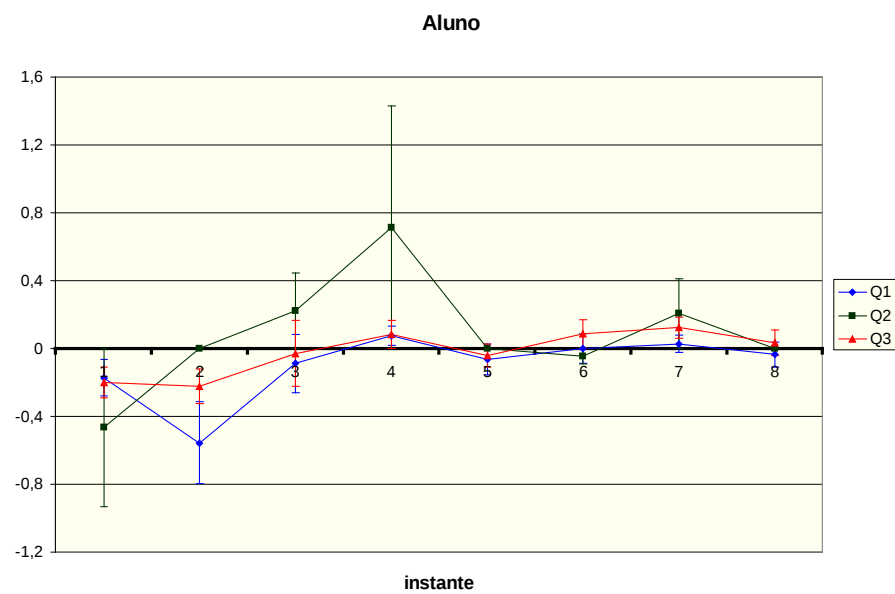
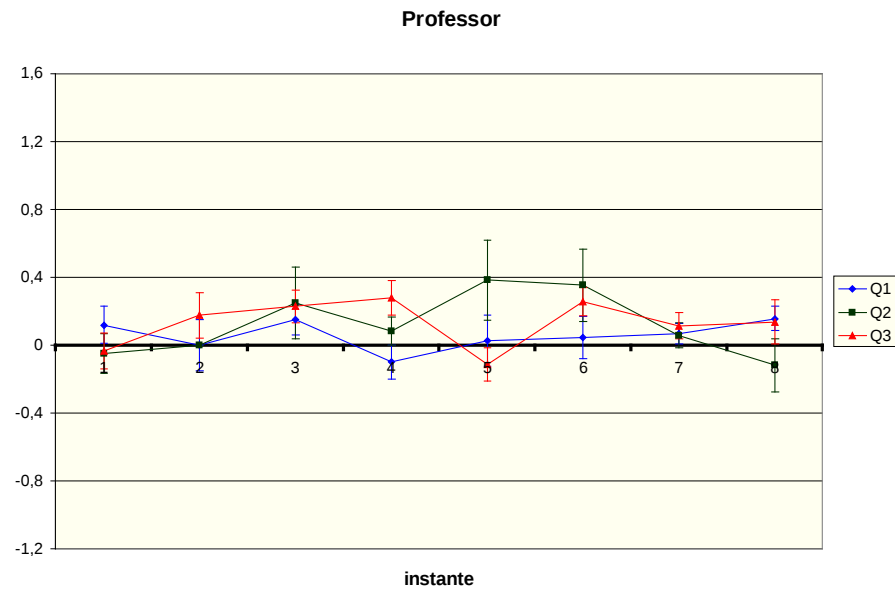


Figura C.12. Média para a variável diferença (depois do questionário – antes do questionário) em exame físico por avaliador no grupo B

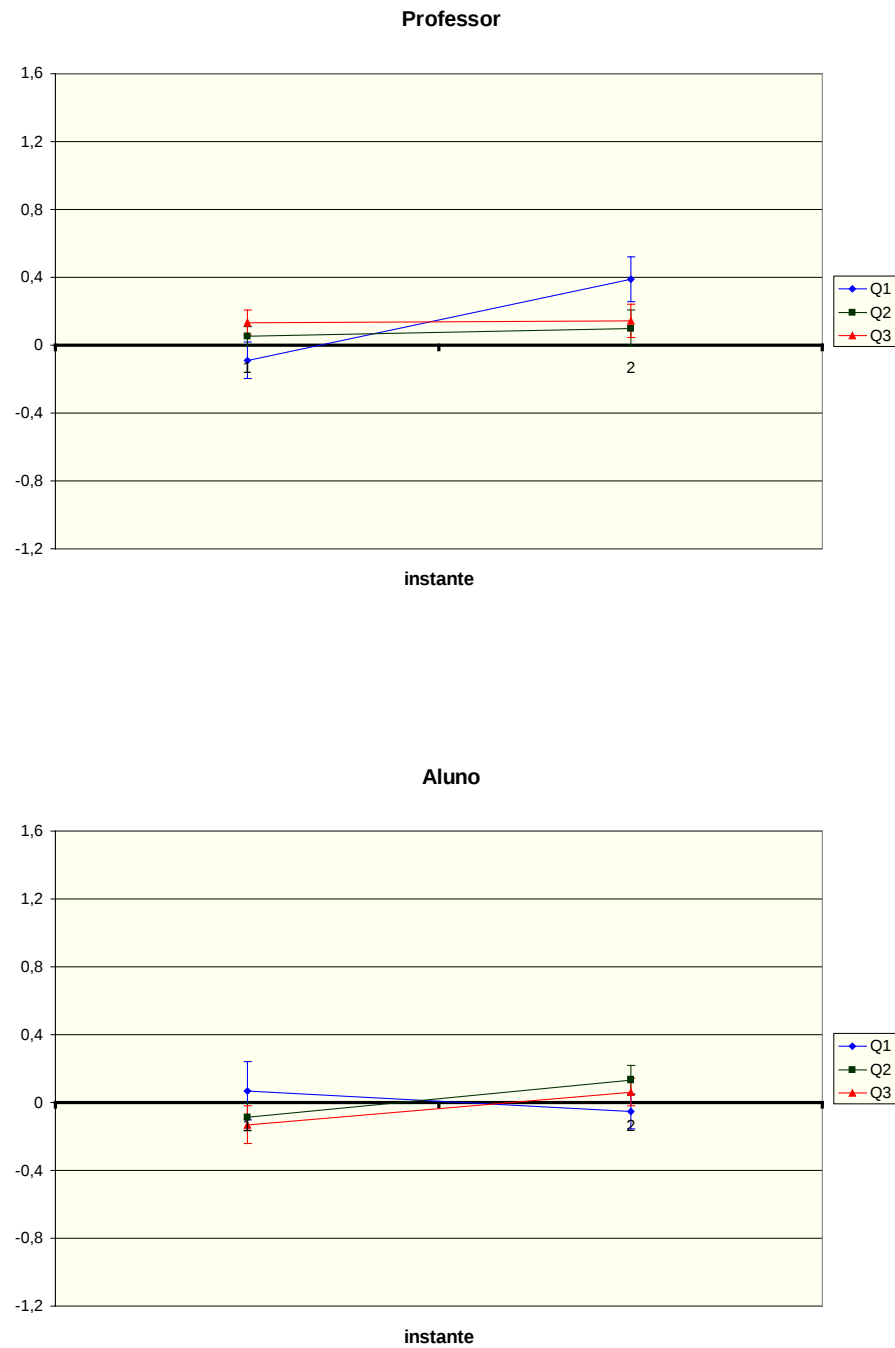


Figura C.13. Média para a variável diferença (depois do questionário – antes do questionário) em raciocínio clínico por avaliador no grupo A

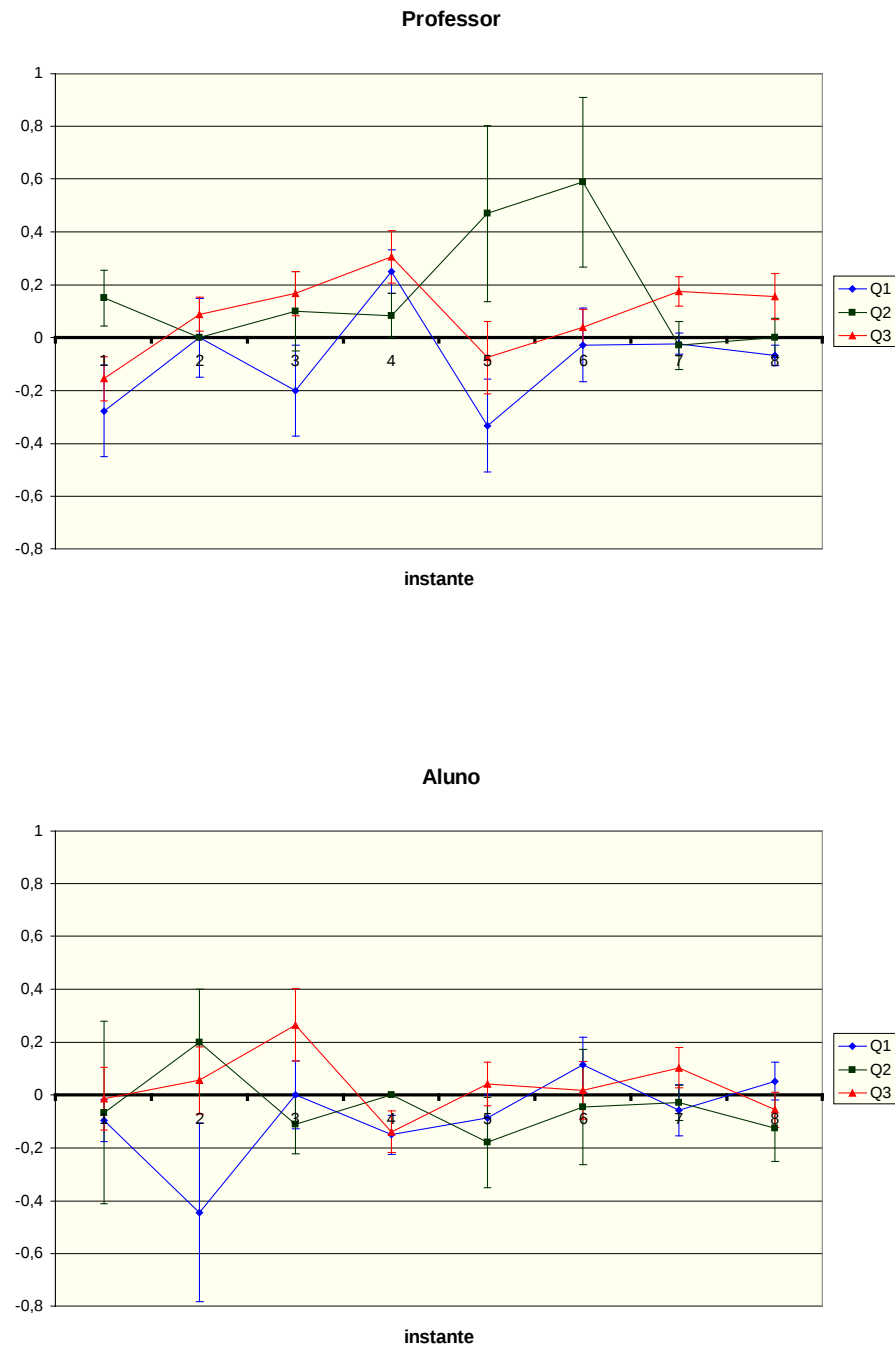


Figura C.14. Média para a variável diferença (depois do questionário – antes do questionário) em raciocínio clínico por avaliador no grupo B

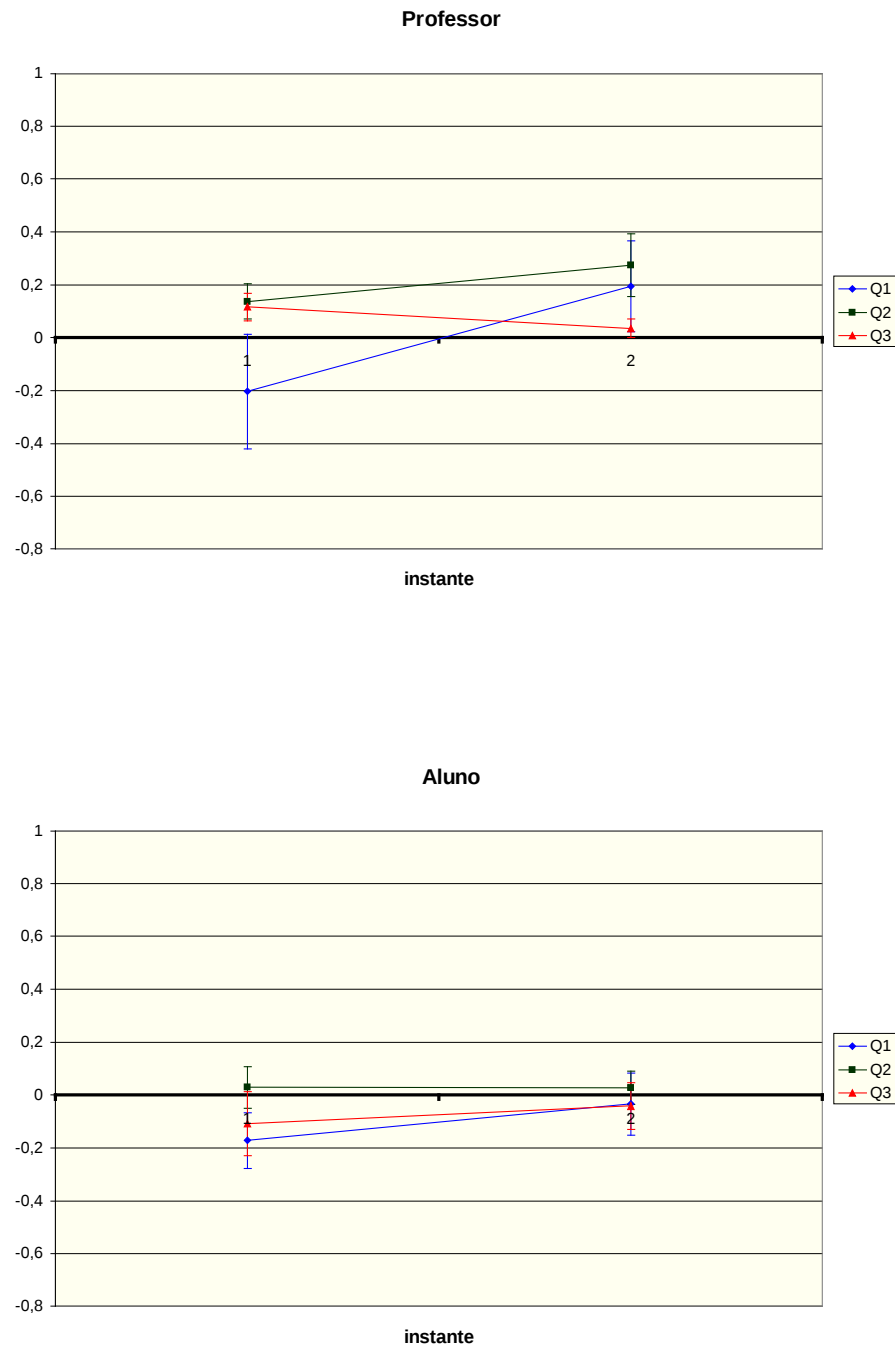


Figura C.15. Média para a variável diferença (depois do questionário – antes do questionário) em conceito geral por avaliador no grupo A

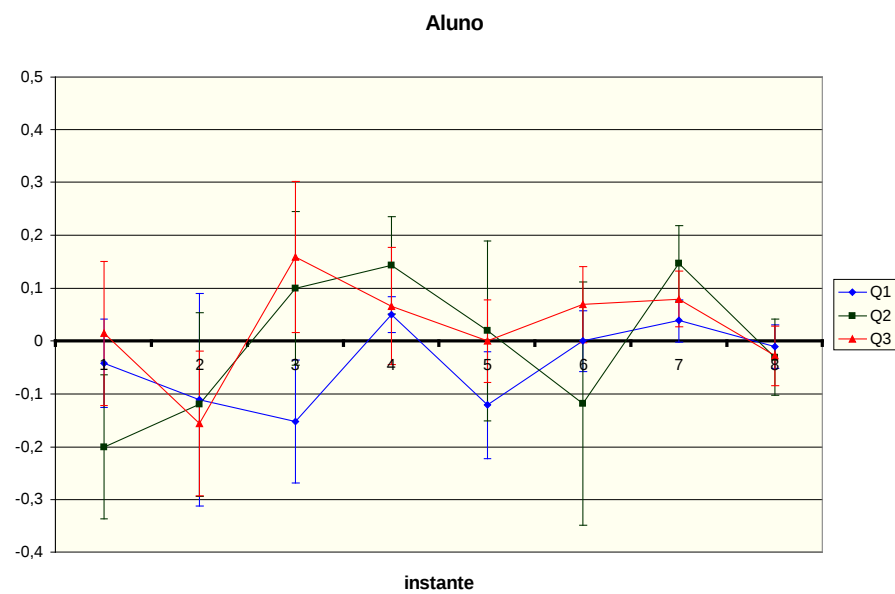
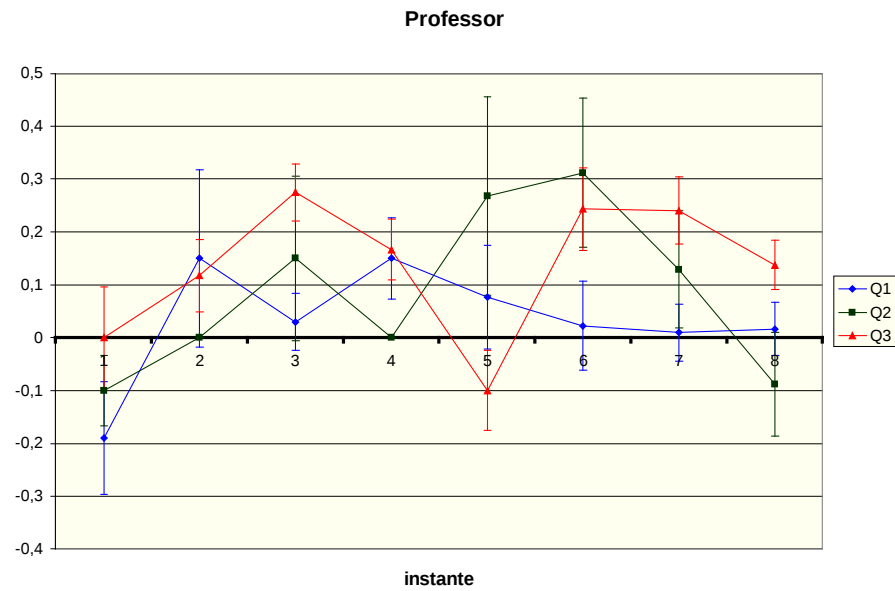


Figura C.16. Média para a variável diferença (depois do questionário – antes do questionário) em conceito geral por avaliador no grupo B

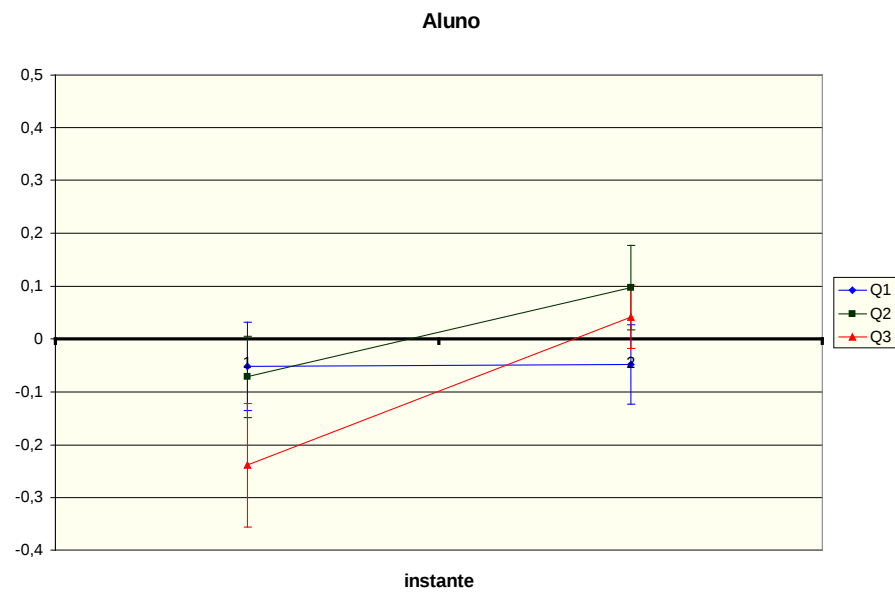
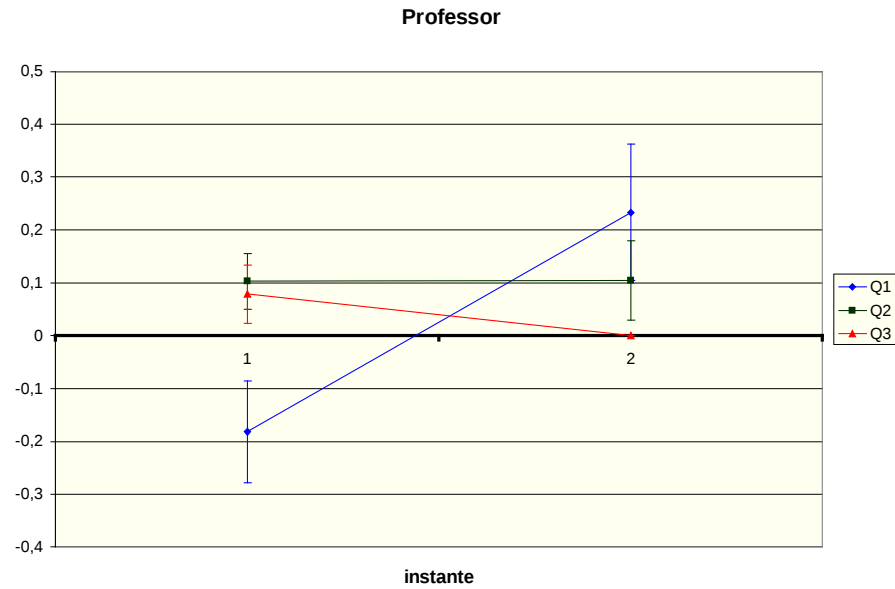


Figura C.17. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em anamnese por condição no grupo A

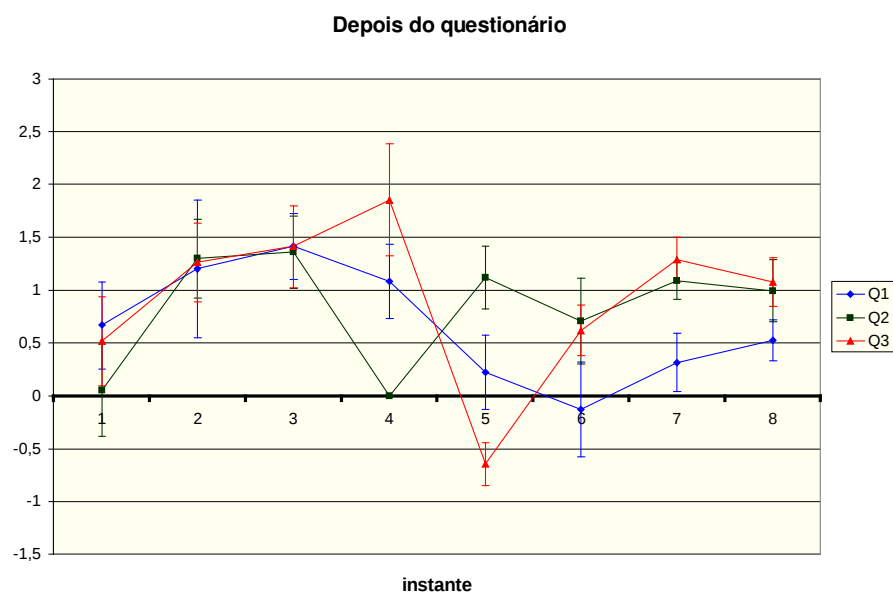
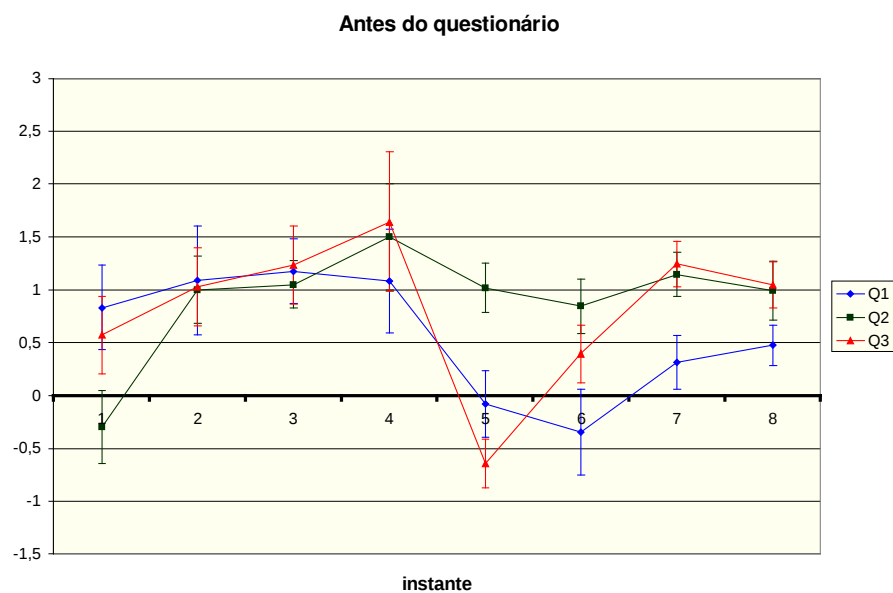


Figura C.18. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em anamnese por condição no grupo B

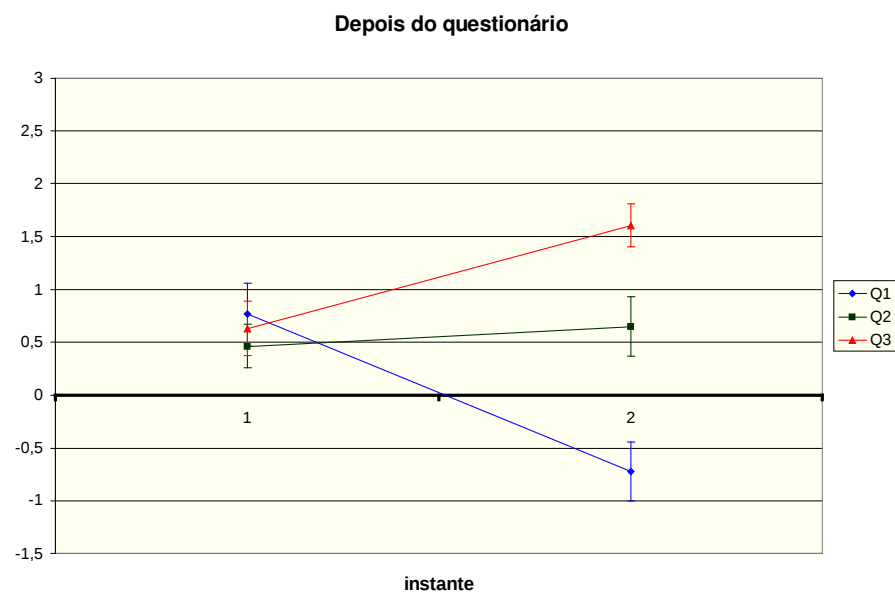
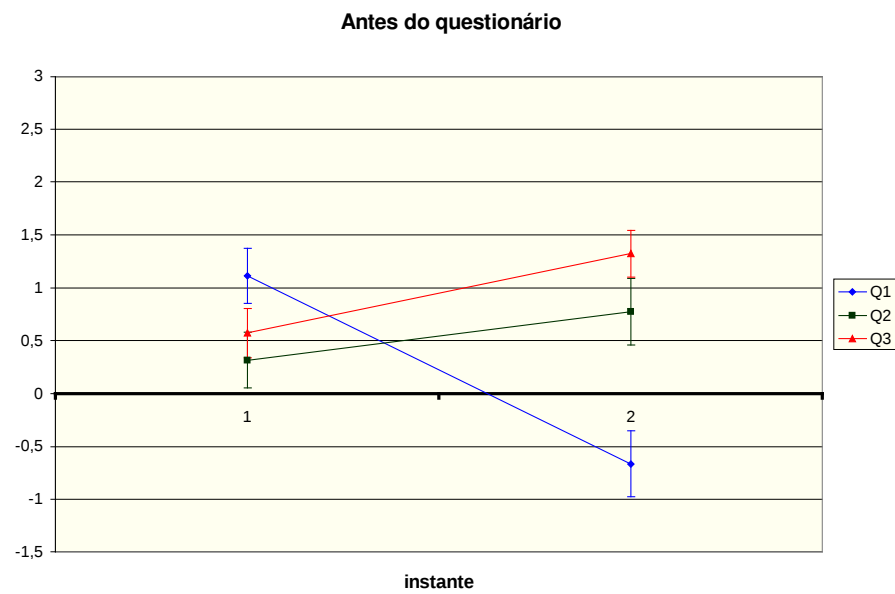


Figura C.19. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em exame físico por condição no grupo A

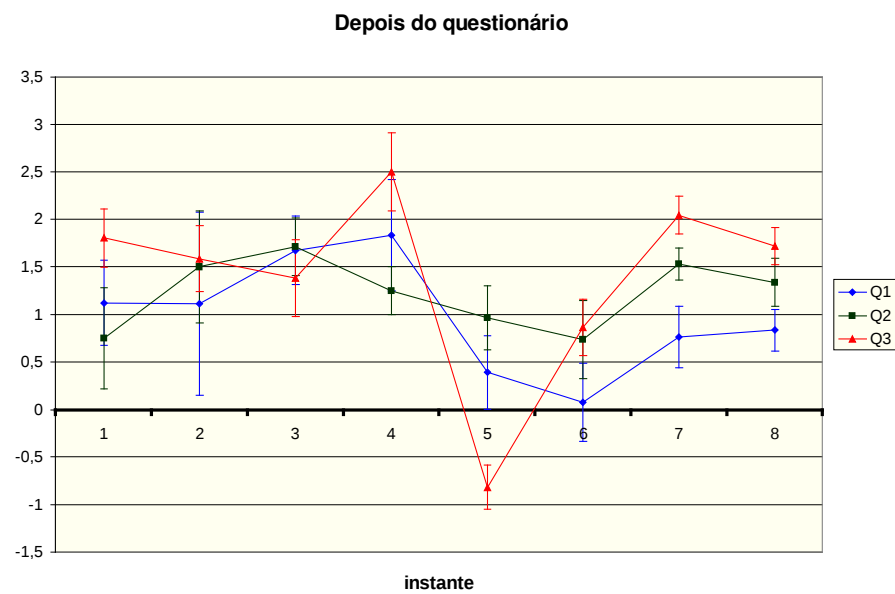
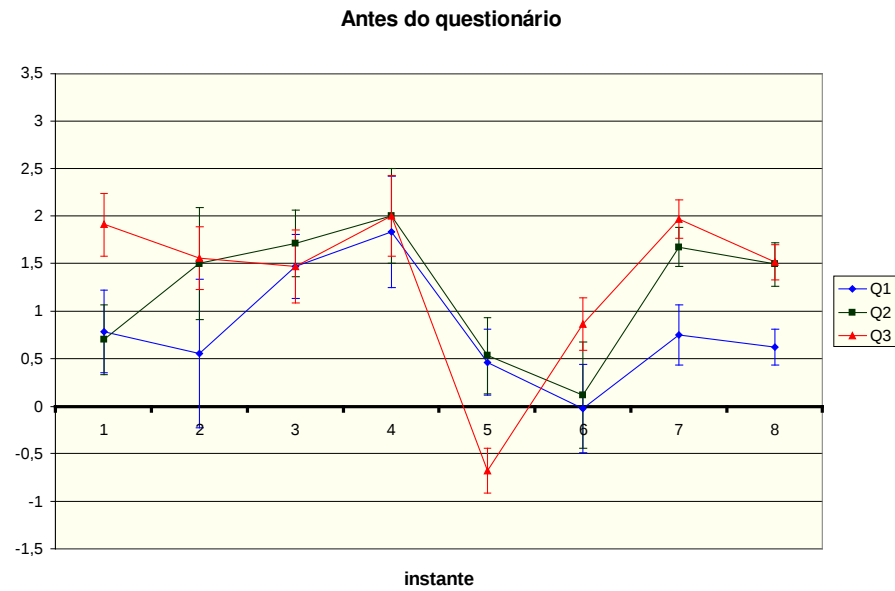


Figura C.20. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em exame físico por condição no grupo B

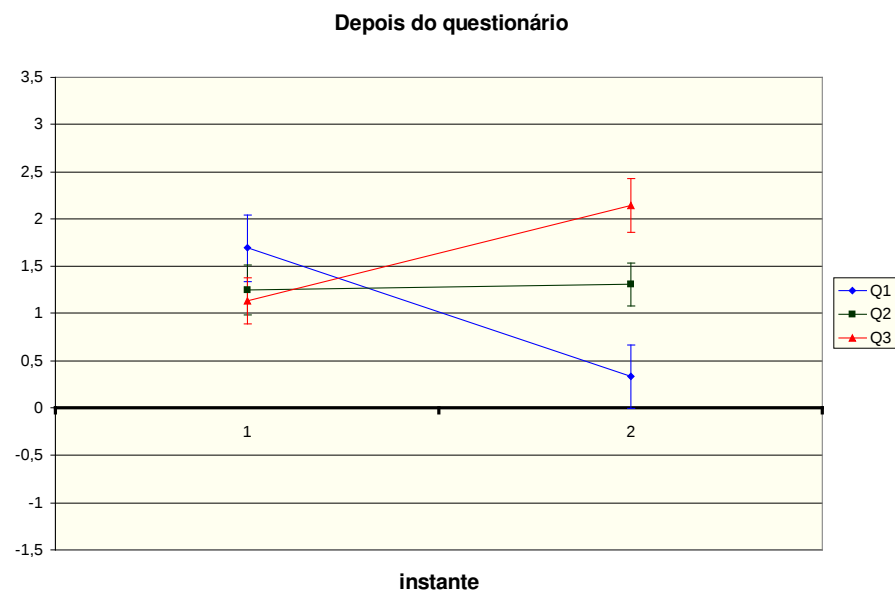
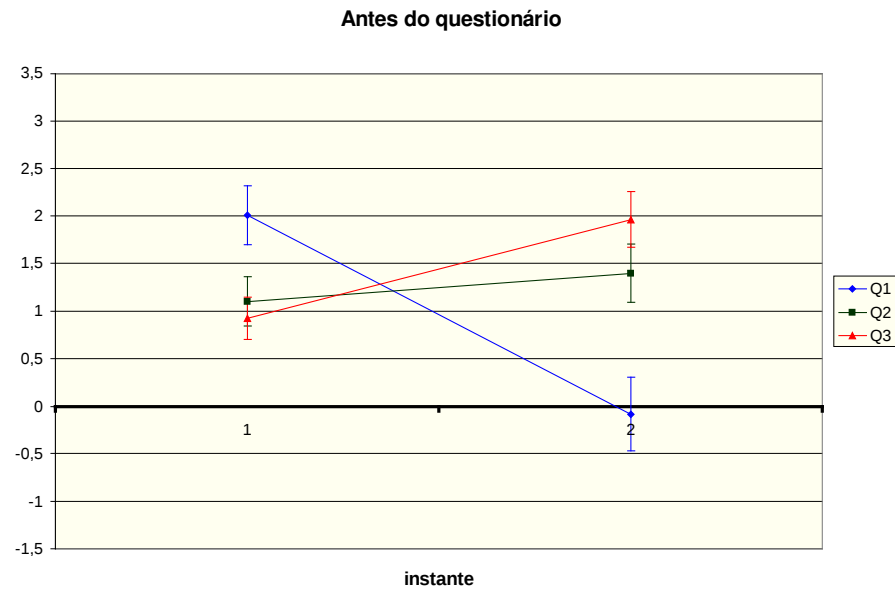


Figura C.21. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em raciocínio clínico por condição no grupo A

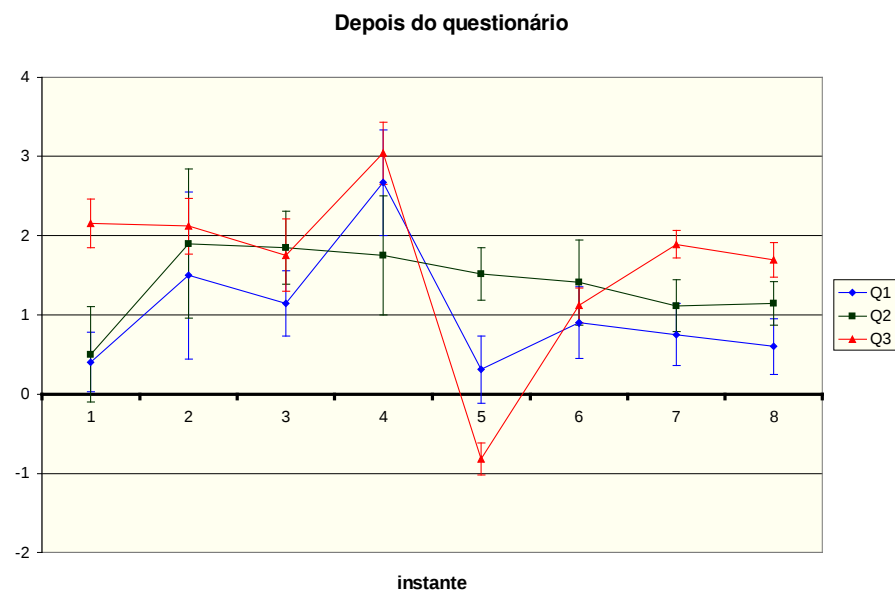
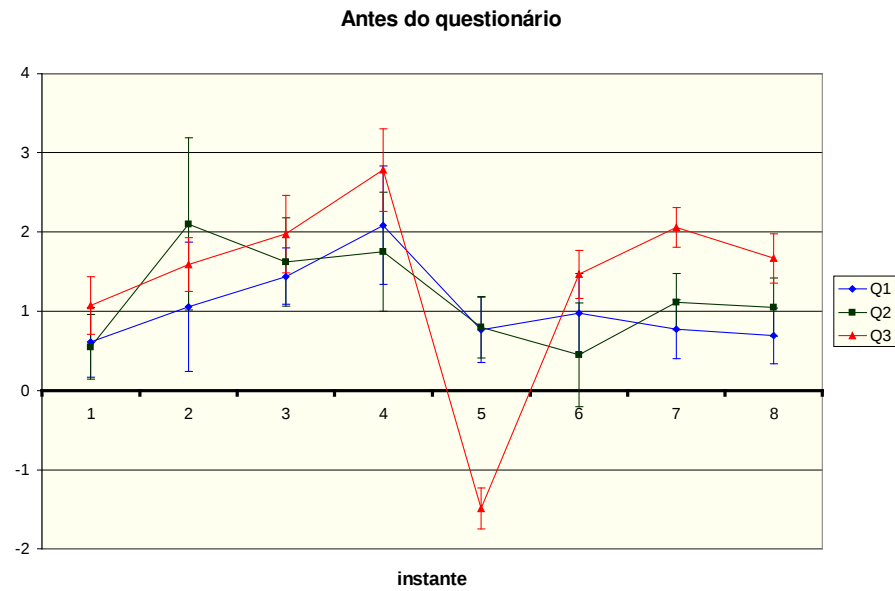


Figura C.22. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em raciocínio clínico por condição no grupo B

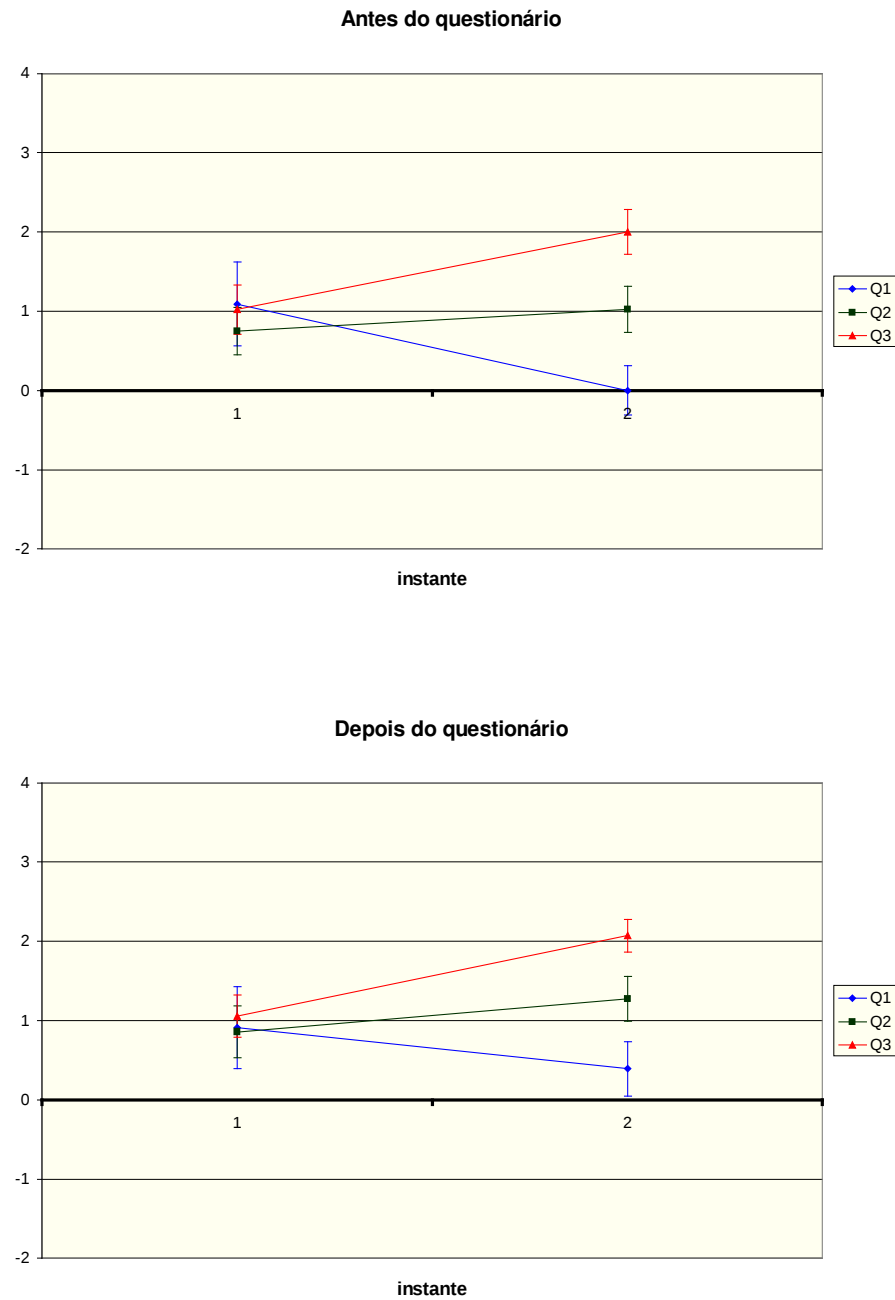


Figura C.23. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em conceito geral por condição no grupo A

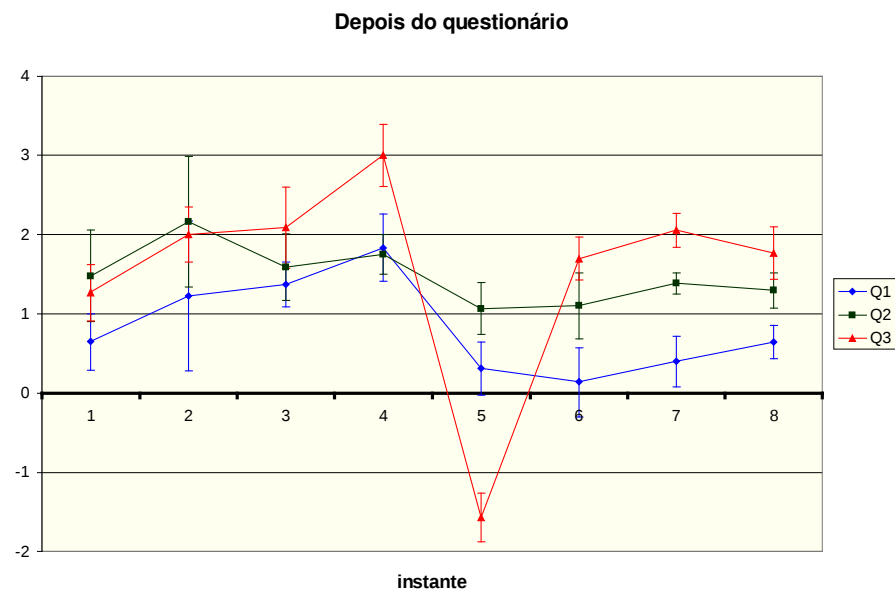
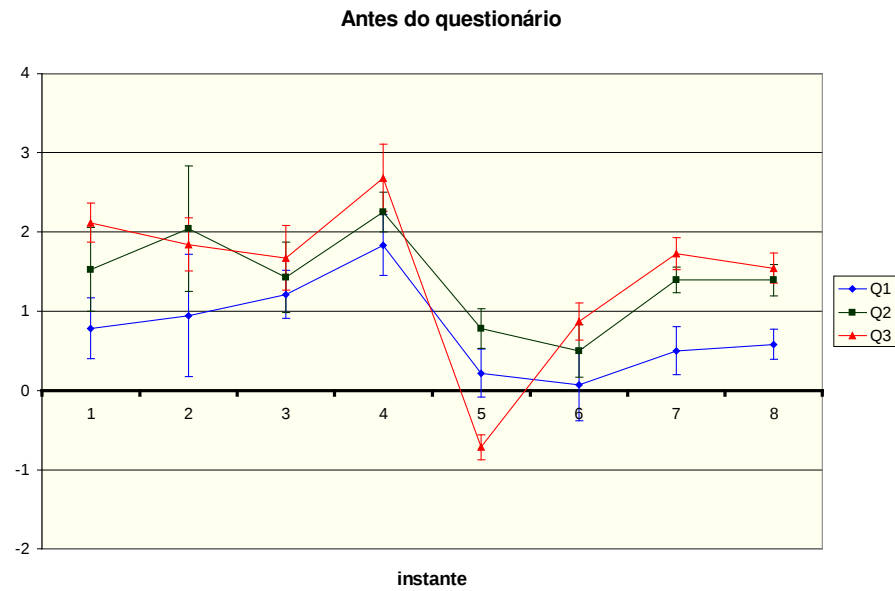
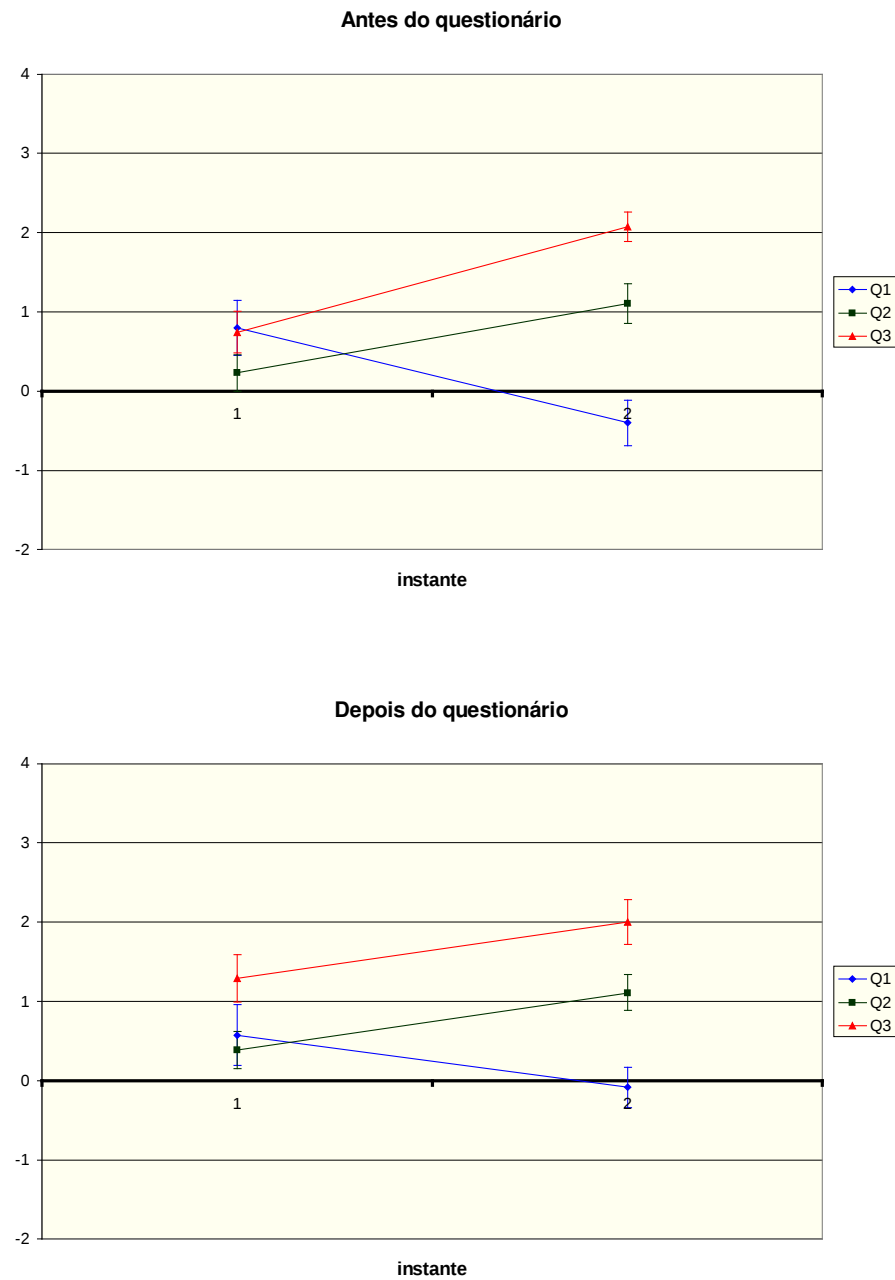


Figura C.24. Média para a variável diferença (nota dada pelo professor – nota dada pelo aluno) em conceito geral por condição no grupo B



Apêndice D

Histograma

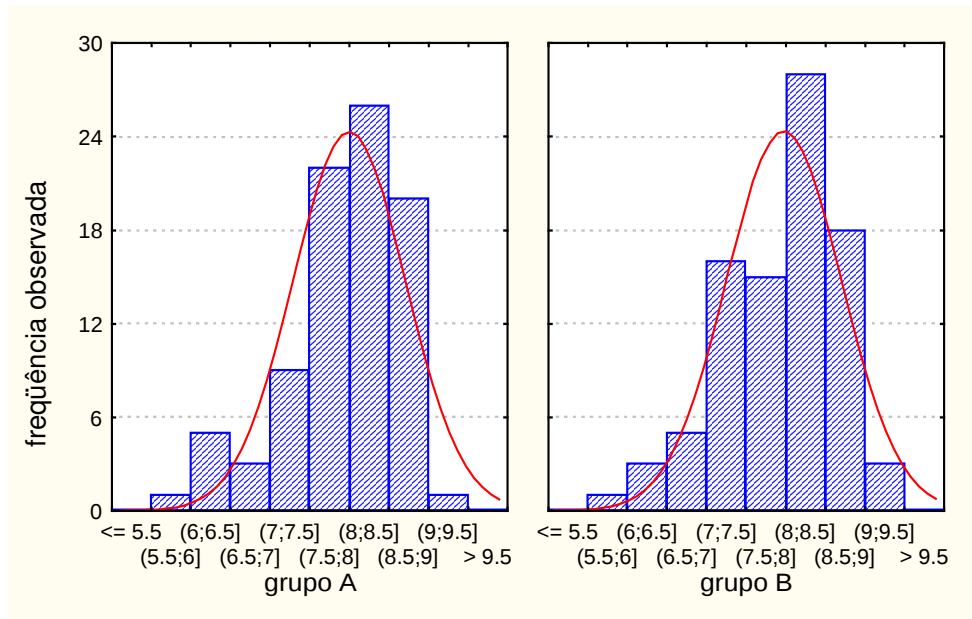
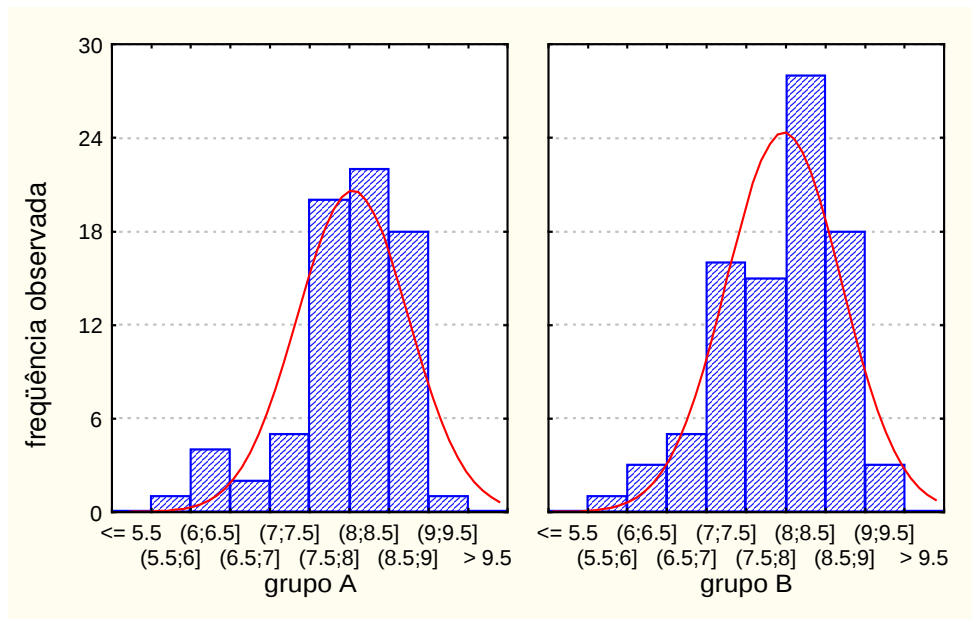
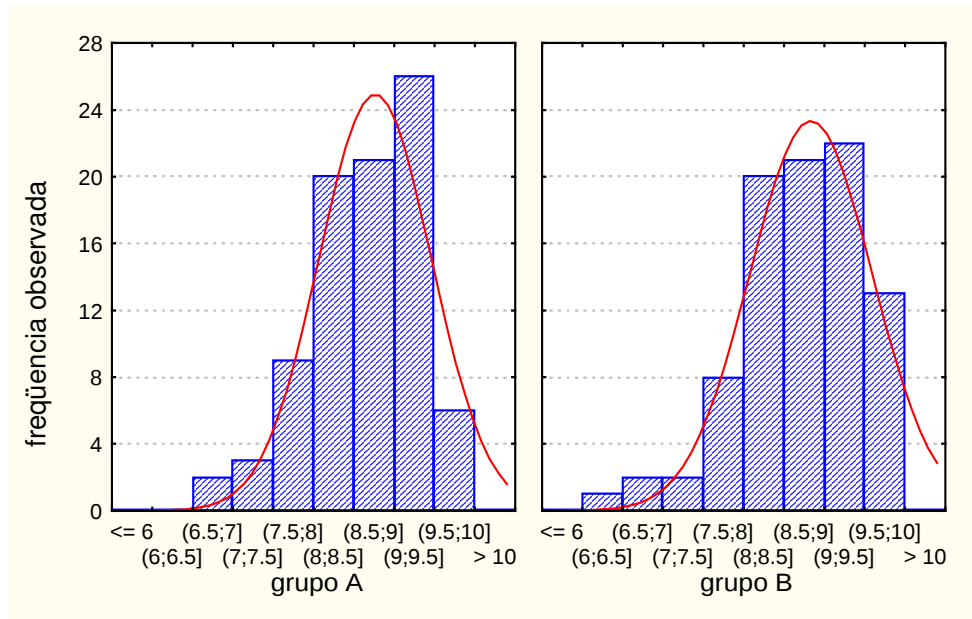
Gráfico D.1. Histograma da variável média final**Gráfico D.2.** Histograma da variável média final (sem os indivíduos do grupo A que responderam a menos de 5 questionários)

Gráfico D.3. Histograma da variável média de conceito oficial



Apêndice E

Modelo de efeitos aleatórios

Modelo de efeitos aleatórios

Há situações em que uma variável resposta Y está relacionada com p variáveis explicativas $X_1, X_2, \dots, X_p$ e o mesmo elemento amostral é observado em mais de um instante ou situação. Nesse caso, ficam então definidos, para o j -ésimo elemento amostral, os seguintes vetores de observações:

$(X_{j11}, X_{j21}, \dots, X_{jp1}, Y_{j1})$ instante ou situação 1

$(X_{j12}, X_{j22}, \dots, X_{jp2}, Y_{j2})$ instante ou situação 2

:

:

$(X_{j1k_j}, X_{j2k_j}, \dots, X_{jp k_j}, Y_{jk_j})$ instante ou situação K_j

As variáveis descritas acima podem ser analisadas através do particular modelo de efeitos aleatórios:

$$Y_j = X_j \beta + X_j b_j + \varepsilon_j$$

onde:

$$Y_j = \begin{bmatrix} Y_{j1} \\ Y_{j2} \\ \dots \\ Y_{jk_j} \end{bmatrix} \text{ é o vetor de respostas do indivíduo } j;$$

$$X_j = \begin{bmatrix} X_{j11} & X_{j21} & \dots & X_{jp1} \\ X_{j12} & X_{j22} & \dots & X_{jp2} \\ \dots & \dots & \dots & \dots \\ X_{j1k_j} & X_{j2k_j} & \dots & X_{jp k_j} \end{bmatrix} \text{ é o vetor de variáveis explicativas do indivíduo } j;$$

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_p \end{bmatrix} \text{ é o vetor de parâmetros;}$$

$b_j = \begin{bmatrix} b_{j1} \\ b_{j2} \\ \dots \\ b_{jp} \end{bmatrix}$ é o vetor de efeitos aleatórios associados ao j-ésimo elemento amostral;

$\varepsilon_j = \begin{bmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \dots \\ \varepsilon_{jkj} \end{bmatrix}$ é o vetor de erros casuais associados ao j-ésimo elemento amostral.

As suposições básicas deste modelo são:

- i) $b_j \sim N_p(0, \Sigma_b)$ $j = 1, 2, \dots, n$;
- ii) $\varepsilon_j \sim N_{kj}(0, \sigma^2 I_{kj})$ $j = 1, 2, \dots, n$;
- iii) $b_1, b_2, \dots, b_n, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ são vetores aleatórios independentes, e como consequência tem-se :
 - a) $E(Y_j) = X_j \beta$
 - b) $\text{Var}(Y_j) = X_j \Sigma_b X_j' + \sigma^2 I_{kj} = \Sigma_j$
 - c) $Y_j \sim N_{kj}(X_j \beta, \Sigma_j)$

O módulo 5V do software estatístico BMDP, assim como o programa SAS, estimam simultaneamente β e Σ_j no modelo de efeitos aleatórios. Neste trabalho, estimou-se esses parâmetros por máxima verossimilhança através do procedimento iterativo de Newton-Raphson.

Uma grande vantagem dessa técnica estatística é permitir a existência de dados incompletos. Nesses casos, pode-se estimar seus valores e imputá-los aos dados (solução do BMDP), ou não utilizar imputação, trabalhando com matrizes de variáveis explicativas de dimensões diferentes para cada indivíduo (solução do SAS).

Na análise inferencial, foram ajustados diversos modelos de efeitos aleatórios. Naqueles utilizados para comparar as notas de conceito do questionário com as notas de conceito oficial, a variável resposta é a nota e as variáveis explicativas consideradas foram:

- $X_0 = 1$ para a constante;

- semestre = $\begin{cases} 1 & \text{para o semestre 1} \\ -1 & \text{para o semestre 2;} \end{cases}$
- tipo de nota = $\begin{cases} 1 & \text{para o conceito oficial} \\ -1 & \text{para o conceito do questionário;} \end{cases}$
- semestre * tipo de nota, correspondendo ao efeito de interação entre semestre e tipo de nota.

Os modelos construídos para comparar as notas dadas pelo professor e dadas pelo aluno e as notas antes e depois do preenchimento do questionário têm como variável resposta uma das quatro notas do questionário (anamnese, exame físico, raciocínio clínico ou conceito geral). Para o grupo de alunos B, foram utilizadas as seguintes variáveis explicativas:

- $X_0 = 1$ para a constante;
- avaliador = $\begin{cases} 1 & \text{para o professor} \\ -1 & \text{para o aluno;} \end{cases}$
- instante = $\begin{cases} 1 & \text{para o primeiro semestre} \\ -1 & \text{para o segundo semestre;} \end{cases}$
- condição = $\begin{cases} 1 & \text{para antes do questionário} \\ -1 & \text{para depois do questionário;} \end{cases}$
- avaliador * instante;
- avaliador * condição;
- instante * condição;
- avaliador * instante * condição.

Para o grupo A, os alunos são avaliados em 8 instantes. Como não se está assumindo nenhuma função para o crescimento da nota em função do instante, é necessária a construção de 7 variáveis do tipo “dummy” para representar esses instantes. As variáveis construídas foram as seguintes:

$$\bullet \text{ instante}_i = \begin{cases} 1 & \text{para o instante } i \\ -1 & \text{para o instante 8} \\ 0 & \text{para os demais instantes;} \end{cases} \quad i=1, 2, \dots, 7.$$

O número de variáveis de interação também aumenta, já que deve-se combinar cada uma das variáveis do instante com as variáveis avaliador e condição.

Com exceção dos modelos para comparação das notas do questionário no grupo A, não se restringiu a estrutura da matriz Σ_j cujos parâmetros foram estimados pelos softwares utilizados na análise. Para o grupo A, devido ao grande número de parâmetros envolvidos e a restrições computacionais, foi necessária uma suposição para a estrutura de Σ_j . Foi suposto que Σ_j pode ser modelada por apenas dois parâmetros, um para as variâncias e outro para as covariâncias. Essa suposição talvez não seja muito razoável, pois as correlações das notas antes e depois do questionário são altas, enquanto as correlações entre as notas dadas pelo professor e pelo aluno são baixas. Já o padrão definido para Σ_j supõe que essas correlações são iguais. Porém, mesmo adotando-se esse padrão para Σ_j , as conclusões estão de acordo com a sugerida pela análise descritiva.

Os softwares utilizados para o ajuste de modelos de efeitos aleatórios foram o BMDP e o SAS. Os modelos para comparar as notas de conceito oficial e do questionário foram ajustados no BMDP. Os demais foram construídos no SAS.

Em todos os casos, inicialmente eram ajustados modelos com todas as variáveis explicativas. Retirava-se então do modelo a variável não significativa (adotando-se um nível de significância de 5%) de maior nível descritivo e reestimava-se os parâmetros. A variável menos significativa era então retirada e o processo se repetia até que o modelo contivesse apenas variáveis significantes.

Para que os testes de interesse possam ser feitos, é necessário obter a esperança da nota dada cada uma das condições em teste. Como exemplo, será apresentada a comparação entre as médias das notas dadas pelo professor e dadas pelo aluno, no grupo B, para a variável anamnese.

O modelo para nota de anamnese é dado por

anamnese = $\beta_0 X_0 + \beta_1 \text{avaliador} + \beta_2 \text{instante} + \beta_3 \text{condição} + \beta_4 \text{avaliador} * \text{instante} + \beta_5 \text{avaliador} * \text{condição} + \beta_6 \text{instante} * \text{condição} + \beta_7 \text{avaliador} * \text{instante} * \text{condição} + b_0 X_0 + b_1 \text{avaliador} + b_2 \text{instante} + b_3 \text{condição} + b_4 \text{avaliador} * \text{instante} + b_5 \text{avaliador} * \text{condição} + b_6 \text{instante} * \text{condição} + b_7 \text{avaliador} * \text{instante} * \text{condição} + \varepsilon$.

Como, $E(b_1) = E(b_2) = E(b_3) = E(b_4) = E(b_5) = E(b_6) = E(b_7) = E(\varepsilon) = 0$, tem-se que, por exemplo,

Nota média do professor no instante 1 antes do questionário = $\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7$ e

Nota média do aluno no instante 1 antes do questionário = $\beta_0 - \beta_1 + \beta_2 + \beta_3 - \beta_4 - \beta_5 + \beta_6 - \beta_7$.

Para verificar se as médias das notas dadas pelo professor e dadas pelo aluno são iguais, no instante 1 antes do questionário, deve-se testar

$\beta_0 + \beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_5 + \beta_6 + \beta_7 = \beta_0 - \beta_1 + \beta_2 + \beta_3 - \beta_4 - \beta_5 + \beta_6 - \beta_7$.

Portanto, esse teste é equivalente a testar se $2\beta_1 + 2\beta_4 + 2\beta_5 + 2\beta_7 = 0$.

Utilizando-se o mesmo raciocínio para as outras combinações de níveis de instante e condição, chega-se à Tabela F.1.

Tabela E.1. Comparação entre as notas dadas pelo professor e dadas pelo aluno

Comparação	Nível	Teste
1	Instante 1, antes do questionário	$2\beta_1 + 2\beta_4 + 2\beta_5 + 2\beta_7 = 0$
2	Instante 1, depois do questionário	$2\beta_1 + 2\beta_4 - 2\beta_5 - 2\beta_7 = 0$
3	Instante 2, antes do questionário	$2\beta_1 - 2\beta_4 + 2\beta_5 - 2\beta_7 = 0$
4	Instante 2, depois do questionário	$2\beta_1 - 2\beta_4 - 2\beta_5 + 2\beta_7 = 0$

Pode-se notar que, se $\beta_5 = \beta_7 = 0$, as comparações 1 e 2 e 3 e 4 são equivalentes. Da mesma forma, se $\beta_4 = \beta_7 = 0$, as comparações 1 e 3 e 2 e 4 são as mesmas. Caso $\beta_4 = \beta_5 = \beta_7 = 0$, pode-se fazer uma única comparação ($2\beta_1 = 0$) onde está se testando se as médias das notas dadas pelo professor e dadas pelo aluno são iguais, independente do nível do instante e condição.

Para controlar o nível de significância conjunto, quando foram feitas várias comparações simultaneamente, utilizou-se o critério de Bonferroni. Nas comparações das notas de conceito oficial e do questionário e em todas as comparações envolvendo as notas do questionário no grupo A, foram utilizados níveis de significância conjunto de 5% para cada modelo construído. Nas comparações das notas do questionário no grupo B, utilizou-se um nível de significância conjunto de 5% por comparação de interesse (professor e aluno, antes e depois e comparações entre os instantes). Isso foi feito para se evitar que o nível de significância individual, devido ao elevado número de instantes, se torne muito pequeno. Para se obter o nível de significância individual, dividiu-se 5% pelo número de comparações não equivalentes. Portanto, se para algum modelo, deve-se comparar se $2\beta_2=0$ e se $-2\beta_2=0$, esses dois testes são considerados como se fossem apenas um. Nos testes com níveis descritivos menores que o nível de significância individual, rejeita-se a hipótese de que as médias são iguais.

Para as comparações em que as diferenças entre as médias são estatisticamente significantes, construiu-se, utilizando-se o método de Bonferroni, intervalos de confiança com coeficiente de confiança conjunto de 95%. Isso significa que se fossem realizados um grande número de estudos equivalentes, em 95% deles todos os verdadeiros valores das diferenças entre as médias pertenceriam aos intervalos construídos. Esses intervalos permitiram quantificar as diferenças entre as médias de interesse.

Conforme mostrado no início desse apêndice, os modelos de efeitos aleatórios supõem que os erros ε_j tenham distribuição normal. Como a amostra é grande e os testes de hipóteses utilizados são robustos com relação à pequenos desvios de normalidade, julgou-se desnecessário a verificação dessa hipótese do modelo.

Apêndice F

Tabelas da análise inferencial

Tabela F.1. Níveis descritivos dos testes t para duas amostras independentes

teste	nível descritivo
Igualdade das médias da média final nos dois grupos	0,8500
Igualdade das médias da média final nos dois grupos (excluindo-se os indivíduos atípicos do grupo A) *	0,5600
Igualdade das médias da média de conceito oficial nos dois grupos	0,5200

* Alunos que preencheram menos de 5 questionários

Tabela F.2. Níveis descritivos dos testes de Mann-Whitney

teste	nível descritivo
Igualdade das médias da média final nos dois grupos	0,8900
Igualdade das médias da média final nos dois grupos (excluindo-se os indivíduos atípicos do grupo A) *	0,6200
Igualdade das médias da média de conceito oficial nos dois grupos	0,4500

Alunos que preencheram menos de 5 questionários

Tabela F.3. Modelo ajustado para a nota antes do questionário no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	8,61	0.07	< 0,0001
semestre	0,25	0.04	< 0,0001
tipo de nota	0,16	0.03	< 0,0001
semestre * tipo de nota	-0.07	0.02	0,0033

Tabela F.4. Níveis descritivos dos testes de comparação de médias antes do questionário no grupo A

Nível descritivo de rejeição: $p < 0,0125$	
resultado	nível descritivo
μ instante 1, oficial = μ instante 1, questionário	0,0434
μ instante 2, oficial > μ instante 2, questionário	< 0,0001
μ instante 1, oficial > μ instante 2, oficial	< 0,0001
μ instante 1, questionário > μ instante 2, questionário	< 0,0001

Tabela F.5. Estimativa e intervalo de confiança para as diferenças entre as médias antes do questionário no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança
μ instante 2, oficial - μ instante 2, questionário	0,45	[0,25;0,64]
μ instante 1, oficial - μ instante 2, oficial	0,37	[0,16;0,57]
μ instante 1, questionário - μ instante 2, questionário	0,64	[0,40;0,88]

Tabela F.6. Modelo ajustado para a nota depois do questionário no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	8.65	0.08	< 0,0001
semestre	0.27	0.04	< 0,0001
tipo de nota	0.11	0.03	0.0008
semestre * tipo de nota	-0.08	0.02	0.0004

Tabela F.7. Níveis descritivos dos testes de comparação de médias depois do questionário no grupo A

Nível descritivo de rejeição: $p < 0,0125$	
resultado	nível descritivo
μ instante 1, oficial = μ instante 1, questionário	0,5145
μ instante 2, oficial > μ instante 2, questionário	< 0,0001
μ instante 1, oficial > μ instante 2, oficial	< 0,0001
μ instante 1, questionário > μ instante 2, questionário	< 0,0001

Tabela F.8. Estimativa e intervalo de confiança para as diferenças entre as médias depois do questionário no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança
μ instante 2, oficial - μ instante 2, questionário	0,39	[0,19;0,59]
μ instante 1, oficial - μ instante 2, oficial	0,37	[0,16;0,57]
μ instante 1, questionário - μ instante 2, questionário	0,71	[0,46;0,95]

Tabela F.9. Modelo inicial ajustado para a nota antes do questionário no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.65	0.08	< 0,0001
semestre	-0.23	0.07	0.0008
tipo de nota	0.18	0.02	< 0,0001
semestre * tipo de nota	0.02	0.03	0.6194

Tabela F.10. Modelo final ajustado para a nota antes do questionário no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.65	0.08	< 0,0001
semestre	-0.22	0.06	0.0004
tipo de nota	0.18	0.03	< 0,0001

Tabela F.11. Níveis descritivos dos testes de comparação de médias antes do questionário no grupo B

Nível descritivo de rejeição: $p < 0,025$	
resultado	nível descritivo
$\mu_{\text{oficial}} > \mu_{\text{questionário}}$	$< 0,0001$
$\mu_{\text{instante 2}} > \mu_{\text{instante 1}}$	$0,0004$

Tabela F.12. Estimativa e intervalo de confiança para as diferenças entre as médias antes do questionário no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança
$\mu_{\text{oficial}} - \mu_{\text{questionário}}$	0,36	[0,24;0,49]
$\mu_{\text{instante 2}} - \mu_{\text{instante 1}}$	0,44	[0,16;0,71]

Tabela F.13. Modelo inicial ajustado para a nota depois do questionário no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.68	0.08	$< 0,0001$
semestre	-0.25	0.07	$< 0,0001$
tipo de nota	0.15	0.03	$< 0,0001$
semestre * tipo de nota	0.04	0.03	0.2313

Tabela F.14. Modelo final ajustado para a nota depois do questionário no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.7	0.08	$< 0,0001$
semestre	-0.21	0.06	0.0004
tipo de nota	0.15	0.03	$< 0,0001$

Tabela F.15. Níveis descritivos dos testes de comparação de médias depois do questionário no grupo B

Nível descritivo de rejeição: $p < 0,025$	
resultado	nível descritivo
$\mu_{\text{oficial}} > \mu_{\text{questionário}}$	0,0004
$\mu_{\text{instante 2}} > \mu_{\text{instante 1}}$	$< 0,0001$

Tabela F.16. Estimativa e intervalo de confiança para as diferenças entre as médias depois do questionário no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança
$\mu_{\text{oficial}} - \mu_{\text{questionário}}$	0,29	[0,17;0,41]
$\mu_{\text{instante 2}} - \mu_{\text{instante 1}}$	0,44	[0,16;0,71]

Tabela F.17. Modelo final ajustado para a nota de anamnese no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.26	0.05	$< 0,0001$
avaliador	0.28	0.05	$< 0,0001$
instante	-0.16	0.04	0.0004

Tabela F.18. . Níveis descritivos dos testes de comparação de médias para a nota de anamnese no grupo B

Nível descritivo de rejeição: $p < 0,0250$	
resultado	nível descritivo
$\mu_{\text{professor}} > \mu_{\text{aluno}}$	$< 0,0001$
$\mu_{\text{instante 2}} > \mu_{\text{instante 1}}$	0,0004

Tabela F.19. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de anamnese no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
$\mu_{\text{professor}} - \mu_{\text{aluno}}$	0,55	0,35	0,76
$\mu_{\text{instante 2}} - \mu_{\text{instante 1}}$	0,31	0,12	0,50

Tabela F.20. Modelo final ajustado para a nota de exame físico no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	7,86	0,06	< 0,0001
avaliador	0,59	0,05	< 0,0001
instante	-0,18	0,05	0,0003
avaliador * condição	-0,04	0,02	0,0111
instante * condição	0,03	0,01	0,0119

Tabela F.21. Níveis descritivos dos testes de comparação de médias para a nota de exame físico no grupo B

Nível descritivo de rejeição: $p < 0,0083$

resultado	nível descritivo
$\mu_{\text{professor, instante 1, antes}} = \mu_{\text{professor, instante 1, depois}}$	0,7781
$\mu_{\text{professor, instante 2, antes}} < \mu_{\text{professor, instante 2, depois}}$	0,0018
$\mu_{\text{aluno, instante 1, antes}} > \mu_{\text{aluno, instante 1, depois}}$	0,0018
$\mu_{\text{aluno, instante 2, antes}} = \mu_{\text{aluno, instante 2, depois}}$	0,7781
$\mu_{\text{professor, antes}} > \mu_{\text{aluno, antes}}$	< 0,0001
$\mu_{\text{professor, depois}} > \mu_{\text{aluno, depois}}$	< 0,0001
$\mu_{\text{instante 2, antes}} > \mu_{\text{instante 1, antes}}$	0,0033
$\mu_{\text{instante 2, depois}} > \mu_{\text{instante 1, depois}}$	< 0,0001

Tabela F.22. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de exame físico no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		inferior	superior
μ professor, instante 2, antes - μ professor, instante 2, depois	0,15	0,03	0,27
μ aluno, instante 1, depois - μ aluno, instante 1, antes	0,15	0,03	0,27
μ professor, antes - μ aluno, antes	1,10	0,79	1,40
μ professor, depois - μ aluno, depois	1,25	0,97	1,53
μ instante 2, antes - μ instante 1, antes	0,30	0,04	0,56
μ instante 2, depois - μ instante 1, depois	0,43	0,17	0,70

Tabela F.23. Modelo final ajustado para a nota de raciocínio clínico no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	7.68	0.06	< 0,0001
avaliador	0.54	0.07	< 0,0001
instante	-0.13	0.05	0.0113

Tabela F.24. Níveis descritivos dos testes de comparação de médias para a nota de raciocínio clínico no grupo B

Nível descritivo de rejeição: $p < 0,0250$

resultado	nível descritivo
μ professor > μ aluno	< 0,0001
μ instante 2 > μ instante 1	0,0113

Tabela F.25. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de raciocínio clínico no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
$\mu_{\text{professor}} - \mu_{\text{aluno}}$	1,07	0,77	1,38
$\mu_{\text{instante 2}} - \mu_{\text{instante 1}}$	0,26	0,03	0,48

Tabela F.26. Modelo final ajustado para a nota de conceito geral no grupo B

efeito	estimativa	erro padrão	nível descritivo
constante	8.13	0.06	< 0,0001
avaliador	0.39	0.05	< 0,0001
instante * condição	0.02	0.01	0.0252

Tabela F.27. Níveis descritivos dos testes de comparação de médias para a nota de conceito geral no grupo B

Nível descritivo de rejeição: $p < 0,0250$

resultado	nível descritivo
$\mu_{\text{professor}} > \mu_{\text{aluno}}$	< 0,0001
$\mu_{\text{instante 1, antes}} = \mu_{\text{instante 2, antes}}$	0,0252
$\mu_{\text{instante 1, depois}} = \mu_{\text{instante 2, depois}}$	0,0252
$\mu_{\text{instante 1, antes}} = \mu_{\text{instante 1, depois}}$	0,0252
$\mu_{\text{instante 2, antes}} = \mu_{\text{instante 2, depois}}$	0,0252

Tabela F.28. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de conceito geral no grupo B (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
$\mu_{\text{professor}} - \mu_{\text{aluno}}$	0,78	0,54	1,01

Tabela F.29. Modelo final ajustado para a nota de anamnese no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	7.30	0.06	< 0,0001
avaliador	0.44	0.02	< 0,0001
instante1	-0.30	0.08	0.0003
instante2	0.55	0.07	< 0,0001
instante4	-0.16	0.05	0.0027
instante5	0.24	0.06	< 0,0001
instante6	0.32	0.05	< 0,0001
instante7	-0.20	0.06	0.0005
avaliador * instante1	0.23	0.07	0.0005
avaliador * instante3	0.37	0.07	< 0,0001
avaliador * instante4	-0.20	0.05	0.0002
avaliador * instante5	-0.21	0.06	0.0003

Tabela F.30. Níveis descritivos dos testes de comparação de médias para a nota de anamnese no grupo A

Nível descritivo de rejeição: $p < 0,0083$ para a comparação entre avaliadores e $0,0036$ para a comparação entre instantes

resultado	nível descritivo
μ professor, instante 1 $>$ μ aluno, instante 1	$< 0,0001$
μ professor, instante 2 $>$ μ aluno, instante 2	$< 0,0001$
μ professor, instante 3 $>$ μ aluno, instante 3	$< 0,0001$
μ professor, instante 4 $>$ μ aluno, instante 4	$< 0,0001$
μ professor, instante 5 $>$ μ aluno, instante 5	$< 0,0001$
μ professor, instante 6 $>$ μ aluno, instante 6	$< 0,0001$
μ professor, instante 7 $>$ μ aluno, instante 7	$< 0,0001$
μ professor, instante 8 $>$ μ aluno, instante 8	$< 0,0001$
μ instante 2, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 3, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 4, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 5, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 6, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 7, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 8, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 2, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 3, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 4, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 5, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 6, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 7, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 8, aluno $>$ μ instante 1, aluno	$< 0,0001$

Tabela F.31. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de anamnese no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
μ professor, instante 1 - μ aluno, instante 1	0,48	0,16	0,81
μ professor, instante 2 - μ aluno, instante 2	0,89	0,76	1,01
μ professor, instante 3 - μ aluno, instante 3	1,34	0,97	1,72
μ professor, instante 4 - μ aluno, instante 4	1,62	1,20	2,05
μ professor, instante 5 - μ aluno, instante 5	0,49	0,20	0,79
μ professor, instante 6 - μ aluno, instante 6	0,48	0,16	0,80
μ professor, instante 7 - μ aluno, instante 7	0,89	0,76	1,01
μ professor, instante 8 - μ aluno, instante 8	0,91	0,62	1,19
μ instante 2, professor - μ instante 1, professor	1,07	0,70	1,44
μ instante 3, professor - μ instante 1, professor	1,59	1,26	1,93
μ instante 4, professor - μ instante 1, professor	2,28	1,83	2,74
μ instante 5, professor - μ instante 1, professor	1,00	0,65	1,36
μ instante 6, professor - μ instante 1, professor	1,40	1,04	1,76
μ instante 7, professor - μ instante 1, professor	1,69	1,38	2,00
μ instante 8, professor - μ instante 1, professor	1,89	1,54	2,24
μ instante 2, aluno - μ instante 1, aluno	0,66	0,31	1,02
μ instante 3, aluno - μ instante 1, aluno	0,73	0,42	1,05
μ instante 4, aluno - μ instante 1, aluno	1,15	0,75	1,54
μ instante 5, aluno - μ instante 1, aluno	0,99	0,65	1,34
μ instante 6, aluno - μ instante 1, aluno	1,41	1,05	1,77
μ instante 7, aluno - μ instante 1, aluno	1,29	1,00	1,57
μ instante 8, aluno - μ instante 1, aluno	1,46	1,13	1,80

Tabela F.32. Modelo final ajustado para a nota de exame físico no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	6.90	0.06	< 0,0001
avaliador	0.62	0.02	< 0,0001
instante1	-1.33	0.06	< 0,0001
instante2	-0.26	0.08	0.0017
instante4	0.59	0.08	< 0,0001
instante5	-0.25	0.06	< 0,0001
instante6	0.17	0.06	0.0044
instante7	0.43	0.06	< 0,0001
avaliador * instante4	0.39	0.07	< 0,0001
avaliador * instante5	-0.33	0.05	< 0,0001
avaliador * instante6	-0.28	0.06	< 0,0001
avaliador * instante7	0.18	0.05	0.0011

Tabela F.33. Níveis descritivos dos testes de comparação de médias para a nota de exame físico no grupo A

Nível descritivo de rejeição: $p < 0,0100$ para a comparação entre avaliadores e $0,0036$ para a comparação entre instantes

resultado	nível descritivo
μ professor, instante 1 $>$ μ aluno, instante 1	$< 0,0001$
μ professor, instante 2 $>$ μ aluno, instante 2	$< 0,0001$
μ professor, instante 3 $>$ μ aluno, instante 3	$< 0,0001$
μ professor, instante 4 $>$ μ aluno, instante 4	$< 0,0001$
μ professor, instante 5 $>$ μ aluno, instante 5	$< 0,0001$
μ professor, instante 6 $>$ μ aluno, instante 6	$< 0,0001$
μ professor, instante 7 $>$ μ aluno, instante 7	$< 0,0001$
μ professor, instante 8 $>$ μ aluno, instante 8	$< 0,0001$
μ instante 2, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 3, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 4, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 5, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 6, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 7, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 8, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 2, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 3, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 4, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 5, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 6, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 7, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 8, aluno $>$ μ instante 1, aluno	$< 0,0001$

Tabela F.34. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de exame físico no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de		confiança
		limite inferior	limite superior	
μ professor, instante 1 - μ aluno, instante 1	1,24	1,12	1,37	
μ professor, instante 2 - μ aluno, instante 2	1,24	1,12	1,37	
μ professor, instante 3 - μ aluno, instante 3	1,24	1,12	1,37	
μ professor, instante 4 - μ aluno, instante 4	2,02	1,60	2,43	
μ professor, instante 5 - μ aluno, instante 5	0,59	0,29	0,88	
μ professor, instante 6 - μ aluno, instante 6	0,69	0,37	1,00	
μ professor, instante 7 - μ aluno, instante 7	1,60	1,30	1,89	
μ professor, instante 8 - μ aluno, instante 8	1,32	1,04	1,61	
μ instante 2, professor - μ instante 1, professor	1,07	0,74	1,39	
μ instante 3, professor - μ instante 1, professor	1,33	1,16	1,50	
μ instante 4, professor - μ instante 1, professor	2,31	1,92	2,70	
μ instante 5, professor - μ instante 1, professor	0,76	0,47	1,05	
μ instante 6, professor - μ instante 1, professor	1,22	0,92	1,52	
μ instante 7, professor - μ instante 1, professor	1,94	1,64	2,24	
μ instante 8, professor - μ instante 1, professor	2,01	1,73	2,30	
μ instante 2, aluno - μ instante 1, aluno	1,07	0,74	1,39	
μ instante 3, aluno - μ instante 1, aluno	1,33	1,16	1,50	
μ instante 4, aluno - μ instante 1, aluno	1,53	1,18	1,89	
μ instante 5, aluno - μ instante 1, aluno	1,41	1,11	1,71	
μ instante 6, aluno - μ instante 1, aluno	1,78	1,46	2,10	
μ instante 7, aluno - μ instante 1, aluno	1,58	1,29	1,88	
μ instante 8, aluno - μ instante 1, aluno	1,93	1,64	2,22	

Tabela F.35. Modelo final ajustado para a nota de raciocínio clínico no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	6.74	0.07	< 0,0001
avaliador	0.71	0.03	< 0,0001
instante1	-1.04	0.07	< 0,0001
instante2	-0.43	0.10	< 0,0001
instante4	0.57	0.09	< 0,0001
instante5	0.30	0.07	< 0,0001
instante7	0.30	0.07	< 0,0001
avaliador * instante1	-0.33	0.06	< 0,0001
avaliador * instante4	0.57	0.08	< 0,0001
avaliador * instante5	-0.18	0.06	0.004

Tabela F.36. Níveis descritivos dos testes de comparação de médias para a nota de raciocínio clínico no grupo A

Nível descritivo de rejeição: $p < 0,0125$ para a comparação entre avaliadores e $0,0036$ para a comparação entre instantes

resultado	nível descritivo
μ professor, instante 1 $>$ μ aluno, instante 1	$< 0,0001$
μ professor, instante 2 $>$ μ aluno, instante 2	$< 0,0001$
μ professor, instante 3 $>$ μ aluno, instante 3	$< 0,0001$
μ professor, instante 4 $>$ μ aluno, instante 4	$< 0,0001$
μ professor, instante 5 $>$ μ aluno, instante 5	$< 0,0001$
μ professor, instante 6 $>$ μ aluno, instante 6	$< 0,0001$
μ professor, instante 7 $>$ μ aluno, instante 7	$< 0,0001$
μ professor, instante 8 $>$ μ aluno, instante 8	$< 0,0001$
μ instante 2, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 3, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 4, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 5, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 6, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 7, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 8, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 2, aluno $=$ μ instante 1, aluno	$0,0578$
μ instante 3, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 4, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 5, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 6, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 7, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 8, aluno $>$ μ instante 1, aluno	$< 0,0001$

Tabela F.37. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de raciocínio clínico no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
μ professor, instante 1 - μ aluno, instante 1	0,77	0,42	1,11
μ professor, instante 2 - μ aluno, instante 2	1,43	1,29	1,57
μ professor, instante 3 - μ aluno, instante 3	1,43	1,29	1,57
μ professor, instante 4 - μ aluno, instante 4	2,56	2,11	3,02
μ professor, instante 5 - μ aluno, instante 5	1,07	0,74	1,40
μ professor, instante 6 - μ aluno, instante 6	1,43	1,29	1,57
μ professor, instante 7 - μ aluno, instante 7	1,43	1,29	1,57
μ professor, instante 8 - μ aluno, instante 8	1,31	0,99	1,63
μ instante 2, professor - μ instante 1, professor	0,94	0,50	1,37
μ instante 3, professor - μ instante 1, professor	1,37	1,08	1,65
μ instante 4, professor - μ instante 1, professor	2,50	1,97	3,04
μ instante 5, professor - μ instante 1, professor	1,19	0,83	1,55
μ instante 6, professor - μ instante 1, professor	1,66	1,30	2,03
μ instante 7, professor - μ instante 1, professor	1,67	1,31	2,03
μ instante 8, professor - μ instante 1, professor	1,61	1,19	2,02
μ instante 3, aluno - μ instante 1, aluno	0,71	0,44	0,98
μ instante 4, aluno - μ instante 1, aluno	0,71	0,23	1,19
μ instante 5, aluno - μ instante 1, aluno	0,89	0,54	1,24
μ instante 6, aluno - μ instante 1, aluno	1,00	0,65	1,36
μ instante 7, aluno - μ instante 1, aluno	1,01	0,67	1,35
μ instante 8, aluno - μ instante 1, aluno	1,06	0,66	1,47

Tabela F.38. Modelo final ajustado para a nota de conceito geral no grupo A

efeito	estimativa	erro padrão	nível descritivo
constante	7.14	0.07	< 0,0001
avaliador	0.66	0.02	< 0,0001
instante1	-0.93	0.05	< 0,0001
instante2	-0.27	0.08	0.0006
instante4	0.63	0.07	< 0,0001
instante5	-0.30	0.05	< 0,0001
instante6	0.19	0.06	0.0006
instante7	0.27	0.05	< 0,0001
avaliador * instante2	0.20	0.07	0.0057
avaliador * instante4	0.49	0.07	< 0,0001
avaliador * instante5	-0.37	0.05	< 0,0001
avaliador * instante6	-0.28	0.05	< 0,0001

Tabela F.39. Níveis descritivos dos testes de comparação de médias para a nota de conceito geral no grupo A

Nível descritivo de rejeição: $p < 0,0100$ para a comparação entre avaliadores e $0,0036$ para a comparação entre instantes

resultado	nível descritivo
μ professor, instante 1 $>$ μ aluno, instante 1	$< 0,0001$
μ professor, instante 2 $>$ μ aluno, instante 2	$< 0,0001$
μ professor, instante 3 $>$ μ aluno, instante 3	$< 0,0001$
μ professor, instante 4 $>$ μ aluno, instante 4	$< 0,0001$
μ professor, instante 5 $>$ μ aluno, instante 5	$< 0,0001$
μ professor, instante 6 $>$ μ aluno, instante 6	$< 0,0001$
μ professor, instante 7 $>$ μ aluno, instante 7	$< 0,0001$
μ professor, instante 8 $>$ μ aluno, instante 8	$< 0,0001$
μ instante 2, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 3, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 4, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 5, professor $=$ μ instante 1, professor	$0,0050$
μ instante 6, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 7, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 8, professor $>$ μ instante 1, professor	$< 0,0001$
μ instante 2, aluno $>$ μ instante 1, aluno	$0,0002$
μ instante 3, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 4, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 5, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 6, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 7, aluno $>$ μ instante 1, aluno	$< 0,0001$
μ instante 8, aluno $>$ μ instante 1, aluno	$< 0,0001$

Tabela F.40. Estimativa e intervalo de confiança para as diferenças entre as médias das notas de conceito geral no grupo A (coeficiente de confiança conjunto de 95%)

diferença	estimativa	intervalo de confiança	
		limite inferior	limite superior
μ professor, instante 1 - μ aluno, instante 1	1,31	1,20	1,43
μ professor, instante 2 - μ aluno, instante 2	1,71	1,30	2,11
μ professor, instante 3 - μ aluno, instante 3	1,31	1,20	1,43
μ professor, instante 4 - μ aluno, instante 4	2,30	1,91	2,68
μ professor, instante 5 - μ aluno, instante 5	0,57	0,30	0,84
μ professor, instante 6 - μ aluno, instante 6	0,75	0,46	1,04
μ professor, instante 7 - μ aluno, instante 7	1,31	1,20	1,43
μ professor, instante 8 - μ aluno, instante 8	1,25	0,98	1,51
μ instante 2, professor - μ instante 1, professor	0,86	0,50	1,23
μ instante 3, professor - μ instante 1, professor	0,93	0,78	1,09
μ instante 4, professor - μ instante 1, professor	2,06	1,70	2,42
μ instante 6, professor - μ instante 1, professor	0,84	0,57	1,12
μ instante 7, professor - μ instante 1, professor	1,21	0,98	1,44
μ instante 8, professor - μ instante 1, professor	1,31	1,04	1,58
μ instante 2, aluno - μ instante 1, aluno	0,47	0,10	0,84
μ instante 3, aluno - μ instante 1, aluno	0,93	0,78	1,09
μ instante 4, aluno - μ instante 1, aluno	1,07	0,75	1,40
μ instante 5, aluno - μ instante 1, aluno	1,00	0,72	1,28
μ instante 6, aluno - μ instante 1, aluno	1,41	1,11	1,70
μ instante 7, aluno - μ instante 1, aluno	1,21	0,98	1,44
μ instante 8, aluno - μ instante 1, aluno	1,38	1,11	1,64

Apêndice G

Questionário

Nº DE IDENTIFICAÇÃO:

DATA:

AVALIAÇÃO DO PROFESSOR

NOTA DE 0 A 10	anamnese adequada	exame físico adequado	raciocínio clínico	nota de CONCEITO
nota que você dá ao aluno no final deste mês				

CHECAGEM DOS OBJETIVOS	SEMPRE	QUASE SEMPRE	MAIS OU MENOS	QUASE NUNCA	NUNCA
POSTURA E ATENÇÃO:					
apresentação ao paciente/postura e educação					
explica ao paciente o que irá executar					
perguntas abertas na anamnese					
exploração adequada da QD					
ordenação da história e capacidade de síntese					
exploração do ISDA, Apess e Hábitos, Afamiliars concernentes ao caso					
linguagem correta/linguagem médica					
HABILIDADES:					
medidas qualitativas (Egeral, Enutr, grau de hidratação, mucosas, etc)					
medidas quantitativas (medida de PA, FC, P e demais parâmetros vitais					
exame neuropsíquico					
exame especial (ggllos, ORL, OFT, tireóide, pulsos, articular), se concernente ao caso					
exame pulmonar - todas as etapas					
exame cardíaco - todas as etapas					
exame abdominal - todas as etapas					
uso adequado do tempo					
CONHECIMENTO:					
preparo dos seminários e leitura recomendada					
participação nos seminários/discussão de casos					
integração de dados da Anamnese e Ex FÍSICO					
forneendo raciocínio clínico e DF sindrômica					
capacidade de fazer diagnósticos diferenciais					
autonomia e capacidade de adaptação					
Tem em mente os objetivos a serem cumpridos quando executa suas tarefas					

NOTA DE 0 A 10	anamnese adequada	exame físico adequado	raciocínio clínico	nota de CONCEITO
nota que você dá ao aluno no final deste mês				