

Análise de Rasch de uma Medida de Problemas Decorrentes do Beber em Estudantes Universitários

Nelson Hauck Filho
Marco Antonio Pereira Teixeira

*Universidade Federal do Rio Grande do Sul
Porto Alegre, RS, Brasil*

RESUMO

O objetivo do presente estudo foi investigar as propriedades psicométricas de um instrumento de avaliação de problemas decorrentes do beber típicos de estudantes universitários. Os participantes foram 401 estudantes universitários usuários de álcool (média de idades = 22,44; DP = 3,43), que responderam a um conjunto de 24 itens preliminares. O modelo de Rasch foi empregado para avaliar os itens quanto à dificuldade (δ_i), à independência local e ao Funcionamento Diferencial do Item (DIF) para o sexo dos participantes. Após a exclusão de dois itens com DIF, obteve-se uma medida unidimensional com 20 itens, com fidedignidade para os itens de 0,99 e fidedignidade para as pessoas de 0,73. As evidências sugerem que o novo instrumento avalia adequadamente um amplo contínuo de problemas relacionados ao uso de bebidas alcoólicas presentes no contexto universitário. Orientações de uso, uma escala *logit* e uma escala *logit* padronizada são apresentadas para a interpretação dos escores.

Palavras-chave: Uso de álcool; teoria de resposta ao item; funcionamento diferencial do item.

ABSTRACT

Rasch Analysis of a Measure of Alcohol-Related Problems in Undergraduate Students

This study aimed to evaluate the psychometric properties of an instrument measuring alcohol-related problems typical of undergraduate students. The participants were 401 undergraduate students that reported using alcohol at least occasionally (mean age = 22.44; SD = 3.43), which also answered a preliminary 24-item measure. Rasch modeling was used to estimate items regarding in terms of difficulties (δ_i), local independence and and Differential Item Functioning (DIF) among items for each gender. After excluding two items with DIF, results showed adequate psychometric properties for a relatively non-biased, unidimensional 20-item measure. Item reliability was .99 and person reliability was .73. Evidence suggests that the new instrument adequately addresses a large continuum of drinking problems that present themselves in the university context. Recommendations for use, along with a logit and standardized logit scales are presented for the interpretation of the scores.

Keywords: Alcohol use; item response theory; differential item functioning.

RESUMEN

Rasch Análisis de una Medida de Problemas con la Bebida en Estudiantes Universitarios

El objetivo de este estudio fue evaluar las propiedades psicométricas de un instrumento para la evaluación de los problemas derivados de la bebida en los estudiantes universitarios. Los participantes fueron 401 estudiantes universitarios que utilizan el alcohol (edad media = 22,44, DT = 3,43), y que respondieron a una escala preliminar de 24 ítems. El modelo de análisis de Rasch fue utilizado para estimar las dificultades de los ítems (δ_i), la independencia local y el Funcionamiento Diferencial del Ítem para el sexo (DIF). Después de excluir dos ítems con DIF, los resultados mostraron propiedades psicométricas adecuadas para una medida unidimensional con 20 ítems. La fiabilidad para los ítems fue 0.99 y 0.73 para las personas. Las evidencias sugieren que el nuevo instrumento aborda adecuadamente un amplio contínuo de problemas relacionados con el uso de alcohol presente en el contexto universitario. Al final, se presentan recomendaciones de uso, una escala Logit y una escala Logit estandarizada para la interpretación de los resultados.

Palabras clave: Uso de alcohol; teoría de respuesta al ítem; funcionamiento diferencial del ítem.

Estudos apontam para uma alta prevalência do uso e do abuso de bebidas alcoólicas entre estudantes universitários em diversos contextos culturais ao redor do mundo (Clements, 1999; Peuker, Fogaça, e Bizarro, 2006; Silveira et al., 2008). Um levantamento feito em nível nacional com 12711 estudantes universitários revelou que cerca de 30% dos participantes foram classificados nas faixas de moderado a grave para o risco do desenvolvimento de dependência de álcool (Andrade, Duarte e Oliveira, 2010). Além disso, verificou-se que, entre os participantes com idade até 30 anos, a prevalência de episódios de uso pesado no último mês foi de aproximadamente 35%. O uso pesado episódico, frequentemente definido como toda situação de consumo de seis ou mais doses de álcool, além de se estar associado ao desenvolvimento crônico da dependência, também pode levar a uma série de consequências negativas em curto prazo (Silveira et al., 2008). Essas consequências podem variar de eventos relativamente amenos, como ter uma ressaca, a outros com maior repercussão para a integridade do indivíduo, como envolver-se em brigas ou ter problemas com a lei (Perkins, 2002). Em virtude disso, tem havido um crescente interesse, na literatura científica, pelo desenvolvimento de instrumentos para avaliar problemas decorrentes do uso de bebidas alcoólicas em indivíduos jovens (Engs e Hanson, 1994; Hurlbut e Sher, 1992; Maddock, Laforge, Rossi e O'Hare, 2001; Read, Kahler, Strong e Colder, 2006; White e Labouvie, 1989).

Instrumentos que avaliam problemas decorrentes do beber abordam uma variedade de consequências negativas tanto do uso agudo quanto do uso crônico do álcool, como aspectos interpessoais, legais, emocionais, e relacionados à saúde física dos indivíduos. Por exemplo, o instrumento *College Alcohol Problems Scale* avalia problemas de ordem social, como comportamentos sexuais de risco e dirigir embriagado, e problemas de ordem emocional, como humor deprimido e irritável (CAPS; Maddock et al., 2001). No entanto, um aspecto de utilidade para a pesquisa e a prática clínica é que muitos desses recursos metodológicos foram desenvolvidos com o foco em estudantes universitários (e também escolares, em alguns casos), de modo que há itens cujo conteúdo é específico para problemas próprios do contexto acadêmico ou escolar. Assim, o *Young Adult Alcohol Problems Screening Test*, além de avaliar consequências genéricas do uso e do abuso de bebidas, como não se lembrar de eventos relacionados ao uso de bebidas e dirigir embriagado, investiga repercussões do beber no contexto acadêmico, como perder aulas e ter um baixo desempenho acadêmico (YAAPST; Hurlbut

e Sher, 1992). Da mesma forma, outros instrumentos, como o *Rutgers Alcohol Problem Index* (RAPI; White e Labouvie, 1989), o *Young Adult Alcohol Consequences Questionnaire* (YAACQ; Read, Kahler, Strong e Colder, 2006), e o *Student Alcohol Questionnaire* (SAQ; Engs e Hanson, 1994) também avaliam prejuízos de natureza acadêmica relacionados ao uso de álcool. O instrumento YAACQ, especificamente, apresenta uma subescala apenas sobre problemas acadêmicos e ocupacionais, como absentismo e baixo desempenho. Portanto, esse tipo de instrumento, geralmente, é mais adequado para o uso no contexto universitário do que os tradicionais questionários para rastreio de transtornos (como o *Alcohol Use Disorders Identification Test*; Babor e Grant, 1989), que podem não conter (ou conter poucos) itens direcionados à avaliação das consequências situacionais de curto prazo do uso da substância (Larimer e Crouse, 2002).

No entanto, embora os instrumentos acima mencionados cumpram um papel importante na estimação do nível do envolvimento com o álcool de universitários, sendo empregados em diversos estudos internacionais, todos foram desenvolvidos tendo em vista os problemas decorrentes do beber típicos de estudantes da realidade norte-americana. Nesse sentido, comportamentos de uso de bebidas alcoólicas são altamente influenciados por uma série de variáveis contextuais (Cox e Klinger, 1988). Há, dessa forma, a possibilidade da existência de características presentes no Brasil que não são encontradas em outros lugares e que podem não ser devidamente avaliadas, no caso da utilização de um instrumento que não representa os comportamentos típicos da realidade universitária brasileira. Assim, no contexto nacional, os autores têm sido levados a empregar questionários com itens que refletem aspectos específicos do beber dos estudantes brasileiros (por exemplo, Caetano, Mills, Pinsky, Zaleski e Laranjeira, 2011; Souza; Landim, Perdigão; Morais e Carneiro Filho, 1999). Todavia, não tem havido um esforço sistemático no sentido da construção de instrumentos com adequadas qualidades psicométricas e que estimem, com a menor quantia de erro possível, o verdadeiro nível de problemas decorrentes do beber enfrentados pelos universitários. De fato, um instrumento de avaliação de problemas decorrentes do beber para a utilização no Brasil deve não apenas ser representativo dos comportamentos típicos do contexto nacional, mas também proporcionar uma avaliação que: (1) seja o menos enviesada possível para grupos populacionais específicos, como homens e mulheres, e (2) considere o nível de severidade específico de cada problema particular relatado pelos indivíduos bebedores.

O tópico do viés nos itens dos instrumentos psicométricos diz respeito a um fenômeno conhecido como Funcionamento Diferencial do Item (*Differential Item Functioning* ou apenas DIF). O DIF ocorre quando existe uma diferença de probabilidade de resposta a um item para dois indivíduos pertencentes a grupos populacionais distintos (por exemplo, homens e mulheres), mas que possuem mesmo nível no contínuo da medida (Embretson e Reise, 2000). A presença de DIF é indicativa de que a variação nos escores nos itens não é influenciada apenas pelo aspecto investigado pelo instrumento, mas também por uma segunda característica, que interage com a propriedade que é o alvo da avaliação (Fidalgo e Scalon, 2012). O DIF é um aspecto indesejável em um instrumento psicométrico porque pode produzir diferenças de significado entre os escores de grupos populacionais distintos, mesmo quando não existem diferenças reais, subestimando o aspecto avaliado em alguns indivíduos e, em outros, superestimando-o. Uma vez que a avaliação de problemas decorrentes do beber pode apresentar implicações na tomada de decisão na prática clínica, é recomendável que esse aspecto seja devidamente investigado em instrumentos desenvolvidos na área. Em um estudo ilustrativo, Earleywine, LaBrie e Pedersen (2008) investigaram a presença do DIF para o sexo dos respondentes no RAPI (White e Labouvie, 1989). Os resultados mostraram que, dos 23 itens, seis apresentaram viés para o sexo, sendo recomendada a sua exclusão. Em outro estudo, os autores avaliaram o DIF para o sexo em um conjunto de indicadores de abuso e dependência do álcool. Diversos aspectos apresentaram DIF, incluindo beber rápido demais, que se mostrou um aspecto mais indicativo de patologia para adolescentes do sexo feminino do que para aqueles do sexo masculino (Hoepfner, Kahler e Jackson, 2011).

O segundo aspecto a ser abordado na construção de instrumentos para a avaliação de problemas decorrentes do beber remete à escala de medida proporcionada pelo instrumento. O objetivo de uma medida psicométrica deve ser proporcionar uma escala linear, sendo possível o estabelecimento de distâncias regulares entre os indivíduos quanto a uma característica mensurável (Bond e Fox, 2007). Em contraste, usualmente, os instrumentos de problemas decorrentes do beber têm sido desenvolvidos e utilizados de modo a apenas permitir o cômputo de escores brutos a partir dos itens endossados (por exemplo, alguém que marcou “sim” para dois itens recebe “2”, e assim por diante). Esse procedimento permite obter um somatório da quantidade total de problemas que um indivíduo experienciou em um determinado período e também calcular médias e desvios-padrão para grupos de

indivíduos. No entanto, nesse somatório, um item que descreve um evento como “ter tido dor de cabeça após beber” recebe o mesmo valor que um item como “ter tido a necessidade de atendimento médico”, embora ambos descrevam situações com um nível de gravidade bastante distinto. O mero somatório dos itens, nesse caso, não permite a compreensão exata do grau de prejuízos experienciados por um indivíduo e em que posição provável do contínuo ele se encontra, a partir dos itens que endossou. Para esse propósito, é necessário que o ajuste dos itens a um *continuum* de problemas decorrentes do beber seja devidamente investigado com técnicas estatísticas adequadas. Por exemplo, Kahler, Strong e Read (2005) realizaram uma análise de Rasch (Rasch, 1960) dos itens do YAACQ. Os resultados mostraram que uma escala intervalar de problemas somente poderia ser obtida após a exclusão de 24 dos 48 itens do instrumento. Resultados semelhantes também foram relatados para o YAAPST por Kahler, Strong, Read, Palfai e Wood (2004).

Portanto, identifica-se, a necessidade do desenvolvimento de um instrumento de avaliação de problemas decorrentes do beber que seja representativo do fenômeno no contexto universitário brasileiro e que supere algumas dificuldades metodológicas enfrentadas na área. Dessa maneira, o objetivo do presente estudo foi empregar a modelagem de Rasch (1960) para avaliar as propriedades psicométricas de uma nova medida de problemas decorrentes do beber em estudantes universitários brasileiros. Procedimentos de análise de DIF foram usados para manter um conjunto de itens menos sujeito a vieses em função do sexo do indivíduo. Além disso, foram investigadas estimativas de dificuldade dos itens, sendo buscada a construção de um sistema de pontuação que respeitasse o significado (nível de severidade) específico de cada item de um ponto de vista clínico.

MÉTODOS

Participantes e procedimentos de coleta

A amostra para o presente estudo foi composta por 401 estudantes universitários usuários de bebidas alcoólicas de uma universidade federal no Rio Grande do Sul. A idade dos participantes variou de 18 a 33 anos (média=22,44; DP=3,43), sendo 54,5% do sexo feminino. Os participantes eram provenientes de 24 cursos de graduação, principalmente medicina veterinária (27,6%), administração (12,5%), enfermagem (7,9%), agronomia (6,9%) e ciências biológicas (6,6%). Entre os participantes, 49,7% relataram consumir bebidas alcoólicas de duas a quatro vezes por mês e

41,1% afirmaram consumir pelo menos cinco doses em uma ocasião típica de uso.

A definição da amostra se deu a partir da conveniência, sendo feitas aplicações dos instrumentos de forma coletiva, padronizada e em sala de aula de turmas de graduação para as quais houve o aceite das coordenações de curso e dos professores. Foram combinados horários com os professores para que a equipe de pesquisa pudesse comparecer durante o período de aula para fazer um convite aos alunos para participar da pesquisa. Todos os participantes assinaram termos de consentimento livre e esclarecido antes de responder aos instrumentos de pesquisa. As aplicações levaram, em média, 30 minutos. O projeto de pesquisa foi previamente aprovado pelo Comitê de Ética em Pesquisa do Instituto de Psicologia da Universidade Federal do Rio Grande do Sul (registro nº 25000.089325/2006-58). As coletas fizeram parte da dissertação de mestrado do primeiro autor, que teve como foco a temática da motivação para uso de bebidas alcoólicas em estudantes universitários. Outros instrumentos não pertinentes a esse estudo também foram empregados durante as coletas dos dados.

Instrumentos

Os participantes responderam a um questionário sociodemográfico e ao instrumento relevante para o presente estudo, composto por 24 problemas decorrentes do beber característicos do contexto universitário nacional. Os itens do instrumento foram construídos a partir de uma revisão da literatura sobre o assunto (por exemplo, Hurlbut e Sher, 1992; Perkins, 2002; Read et al., 2006; White e Labouvie, 1989) e de discussões com um grupo de estudantes universitários. Esse último procedimento buscou investigar que problemas decorrentes do beber poderiam ser específicos dos estudantes brasileiros, não sendo contemplados pelos itens de instrumentos construídos em outros contextos. Os indicadores selecionados para compor a lista variam de problemas relativamente leves e frequentes (por exemplo, ter dores de cabeça após beber) até outros relacionados a situações de alto risco para a saúde e a vida do indivíduo (por exemplo, ter problemas com a lei). O formato de avaliação empregado foi o de resposta dicotômica, sendo os participantes solicitados a assinalar “Sim” ou “Não” para cada um dos 24 itens, tendo em vista que os itens se referiam exclusivamente a episódios de bebida e considerando os últimos 12 meses.

Análise dos dados

Foi utilizado o software Winsteps 3.72.0 para avaliar o ajuste dos itens ao modelo de Rasch (Rasch,

1960), estimar os parâmetros da dificuldade (δ_i) dos itens e avaliar a possível presença de DIF para o sexo. O modelo de Rasch (Rasch, 1960) propõe que uma resposta positiva a um item qualquer é função da habilidade θ_n do indivíduo e da dificuldade δ_i do item respondido. “Habilidade”, nesse contexto, designa o nível de traço latente do indivíduo ou da característica a ser avaliada, sendo um termo empregado para se referir a qualquer construto psicológico avaliado em pesquisas científicas. O modelo, de natureza probabilística, prediz a ocorrência de respostas positivas por parte de indivíduos com um nível de patologia acima daquele expresso pelo item (dicotômico) e de respostas negativas por parte de indivíduos com um nível de patologia menor do que aquele representado pelo item.

A escala para as medidas de habilidade dos indivíduos e dificuldade dos itens é uma escala de *logits*, que varia teoricamente de infinito negativo a infinito positivo, mas que, na prática, abrange valores entre -3,00 e +3,00 (Bond e Fox, 2007). Usualmente, no modelo de Rasch, as estimativas de dificuldade (δ_i) dos itens têm sua média fixada em 0,00, sendo que valores abaixo indicam que o item é mais fácil (representa um nível menos intenso do contínuo do construto, nesse caso) do que a média do conjunto de itens e, inversamente, um item com *logit* positivo é mais difícil do que a média do conjunto (Bond e Fox, 2007). A dificuldade dos itens, representada usualmente pela letra grega delta (δ) por autores relacionados à modelagem de Rasch (Andrich, 1988; Bond e Fox, 2007), corresponde ao chamado parâmetro “b”, tal como discutido em textos mais amplos sobre Teoria de Resposta ao Item (Embretson e Reise, 2000). Quanto à habilidade das pessoas (isto é, nível de problemas, nesse caso), sua média não é fixada, apresentando uma distribuição empírica cuja interpretação pode ser feita mediante um contraste com a dificuldade estimada para os itens.

O ajuste individual dos itens ao modelo de Rasch foi avaliado pelos índices *infit* e *outfit*, que quantificam resíduos (quadrados médios) para os itens com relação ao modelo (Bond e Fox, 2007; Linacre, 1991). O valor ideal de ajuste é 1,00, sendo considerados aceitáveis valores entre 0,50 a 1,50 (Linacre, 1991). A unidimensionalidade do instrumento, por sua vez, foi avaliada pela análise de componentes principais dos resíduos do modelo, que é considerado o procedimento mais adequado para investigar possíveis dimensões salientes presentes nos dados (Linacre, 1991). Adotou-se o critério de considerar como violação substancial da unidimensionalidade todo autovalor > 2,0 para um contraste principal. Também foram avaliadas correlações entre os resíduos dos itens acima de 0,30, que indicam dependência local e possíveis violações da unidimensionalidade.

Para a análise do DIF, foi utilizado o software GMHDIF (Fidalgo, 2011), que implementa o método *Generalized Mantel-Haenszel*, aplicável tanto a itens de natureza nominal quanto ordinal. Basicamente, o procedimento divide a amostra em estratos, a partir de um critério – que, nesse caso, foi o escore bruto total. A seguir, para cada estrato, comparam-se as diferenças nas distribuições das respostas a cada item entre os grupos. A estatística qui-quadrado de Mantel-Haenszel é utilizada para testar a significância estatística dessa diferença. Mais detalhes sobre o procedimento podem ser encontrados em Fidalgo e Scalón (2012). Optou-se por um nível de significância conservador, de $p < 0,01$, em função da grande quantidade de testes estatísticos sem hipóteses prévias, o que aumenta a taxa de erro de Tipo I (correção de Bonferroni). Por fim, a fidedignidade dos itens e das pessoas (isto é, a replicabilidade das estimativas de dificuldade e de

habilidade obtidas, respectivamente) foi estimada. Esses coeficientes variam de 0,00 a 1,00 sendo o valor de referência $> 0,70$ (Bond e Fox, 2007). Diferente do tradicional alpha de Cronbach, a fidedignidade dos itens não é influenciada pelo número de itens. De maneira análoga, a fidedignidade das pessoas não é influenciada pela quantidade de casos.

RESULTADOS

A análise da unidimensionalidade mediante componentes principais dos resíduos mostrou que apenas 1,5 autovalor da variância total dos 24 itens foi explicada pelo primeiro contraste. Também não foram observadas correlações residuais acima de 0,30, indicando a ausência de dependência local significativa. Assim, ficou evidente a unidimensionalidade da medida, sendo adequado proceder às análises subsequentes.

TABELA 1
Estimativas de dificuldade, ajuste ao modelo de Rasch e análises de DIF para os 24 itens

<i>Item</i>	<i>Medida</i>	<i>Infit</i>	<i>Outfit</i>	χ^2 <i>Mantel-Haenszel</i>
22. Envolveu-se em acidente de trânsito?	2,66	1,01	0,50	0,33
8. Teve problemas com a lei?	2,15	0,91	0,57	3,17
2. Deixou de entregar um trabalho?	2,00	0,87	0,31	0,41
13. Teve necessidade de atendimento médico?	1,67	1,14	1,05	0,13
18. Teve pertences roubados?	1,56	1,09	0,90	0,65
16. Danificou ou destruiu propriedade privada ou patrimônio público?	1,26	0,93	0,70	7,80*
7. Envolveu-se em brigas?	1,13	0,94	0,82	5,35
24. Teve baixo desempenho em uma avaliação acadêmica?	1,09	1,00	1,84	0,74
3. Excedeu o limite máximo de velocidade permitido com um carro ou uma moto?	0,79	0,98	1,01	13,41*
6. Usou drogas ilícitas?	0,76	0,89	0,55	0,06
17. Teve um comportamento sexual que não acharia apropriado se não estivesse sob o efeito do álcool?	0,76	1,00	0,85	3,99
5. Foi infiel em um relacionamento afetivo?	0,48	1,03	0,79	4,70
1. Esqueceu-se de tarefas ou compromissos?	0,39	0,96	0,72	0,03
10. Fez sexo sem usar as proteções que você normalmente utiliza?	0,15	0,90	0,80	0,15
9. Discutiu ou teve atrito com familiares, amigos, amigas, parceira ou parceiro?	-0,56	1,02	1,08	0,03
15. Machucou-se de forma não-intencional?	-0,64	0,96	0,98	0,11
21. Faltou à aula ou ao trabalho?	-0,66	1,02	1,19	2,50
11. Usou tabaco (cigarro, cigarrilha, narguilé etc.)?	-0,95	1,14	1,29	0,64
12. Perdeu ou extraviou pertences?	-0,98	0,91	0,73	5,31
14. Gastou dinheiro de forma inconsequente?	-1,39	0,94	0,90	1,04
20. Esqueceu-se de coisas que fez sob o efeito do álcool?	-1,55	1,03	0,91	0,42
4. Não conseguiu estudar como desejava?	-1,73	1,08	1,17	4,46
19. Sentiu enjoo ou mal-estar?	-4,04	1,03	1,81	0,44
23. Sentiu dores de cabeça?	-4,33	1,11	1,93	3,50
Média	0,00	1,00	0,97	-
DP	1,74	0,08	0,41	-

Nota: Em negrito, itens com pouco ajuste ou com DIF significativo.

* $p < 0,01$.

Em um primeiro momento, foi avaliado o ajuste dos itens ao modelo de Rasch (ver Tabela 1). Os resultados apontaram quatro itens com valores fora do intervalo adequado de 0,50 a 1,50 para *infit* e *outfit*: “Deixou de entregar um trabalho?” (*infit*=0,87, *outfit*=0,31), “Foi mal em uma avaliação?” (*infit*=1,00, *outfit*=1,84), “Sentiu enjoo ou mal-estar?” (*infit*=1,03, *outfit*=1,81), e “Sentiu dores de cabeça?” (*infit*=1,11, *outfit*=1,93). Para a medida final, esses dois primeiros itens foram excluídos. Quanto aos dois últimos itens, ambos foram mantidos em função de uma explicação teórica para a falta de ajuste. Especificamente, o *outfit* observado acima de 1,50 é sugestivo de que muitos indivíduos com um nível de patologia mais alto do que o necessário para uma resposta positiva não responderam positivamente, conforme o esperado pelo modelo (Linacre, 1991). Isso pode ter ocorrido em virtude de ambos os itens se referirem a eventos fisiológicos que ocorrem após a ingestão de uma grande quantidade de etanol, mas que diminuem em intensidade na medida em que os indivíduos desenvolvem tolerância, um reconhecido sintoma de dependência de álcool (Hoeppner et al., 2011). Em virtude disso, é teoricamente aceitável que indivíduos com um nível mais severo de envolvimento com o álcool não endossem itens característicos de baixa tolerância. Analogamente, Kahler et al. (2004) decidiram não eliminar o item “*Sick or vomiting after drinking*” do YAAPST, que apresentou *infit*=1,05 e *outfit*=1,82. As informações sobre os demais itens também podem ser apreciadas na Tabela 1.

Como visto na Tabela 1, o teste de DIF via Mantel-Haenszel evidenciou que apenas dois itens apresentaram estimativas de dificuldade significativamente menor ($p < 0,01$) para homens do que para mulheres. Esses itens foram: “Excedeu o limite máximo de velocidade permitido com um carro ou uma moto?” e “Danificou ou destruiu propriedade privada ou patrimônio público?”. Em virtude disso, para a versão final do instrumento, decidiu-se excluir esses dois itens com DIF, a fim de obter um mesmo instrumento padrão para ambos os sexos, como no estudo de Earleywine et al. (2008). Dessa forma, como visto na Tabela 1, houve 20 itens com ajuste adequado ao modelo de Rasch e sem DIF substancial. Para esse conjunto de itens, a fidedignidade para as pessoas foi de 0,73 e a fidedignidade para os itens foi de 0,99. A estimativa *logit* média de problemas decorrentes do beber (nível θ_n de habilidade) para a amostra foi de -2,40 (DP = 2,00), tendo cerca de 90% das pessoas apresentado estimativas abaixo da média da dificuldade δ_i dos itens (fixada em 0,00). A Figura 1 apresenta o Mapa dos Itens, com os 20 itens finais escalonados à direita e, à esquerda, “#s” representando a distribuição da amostra em termos da

medida *logit* de problemas decorrentes do beber (nível θ_n de habilidade).

A seguir, para fornecer uma interpretação intervalar dos escores brutos no instrumento, medidas *logit* correspondentes a cada escore bruto possível (soma de todos os itens endossados) foram calculadas. A relação entre os escores brutos e a posição de um indivíduo no contínuo da medida é tipicamente não linear (Bond e Fox, 2007; Embretson e Reise, 2000). Assim, a conversão dos escores brutos em medidas *logit* é um aspecto fundamental da análise de Rasch, uma vez que possibilita estimar a real posição de um indivíduo com um escore bruto x em um contínuo de problemas com bebidas. Esse contínuo, ao contrário dos escores brutos e dos escores percentis, leva em consideração as informações psicométricas de cada item, o que significa que estabelece o escalonamento dos mesmos a partir da saturação de cada item no construto. Dessa forma, são apresentados, na Tabela 2, escores brutos, percentis e descrições das frequências de ocorrência de cada valor possível de escore bruto, bem como escores *logit* e erros associados. Um escore *logit* padronizado também é apresentado, computado somando-se uma constante (6,21) aos escores *logit*, para evitar valores negativos. Essa escala pode ser empregada em estudos futuros como uma alternativa aos seus respectivos escores brutos, também satisfazendo o pressuposto da intervalaridade típico das estatísticas paramétricas.

DISCUSSÃO DOS RESULTADOS

Os resultados das análises sugerem que uma versão reduzida de 20 itens do instrumento original (com 24 itens) satisfaz as condições fundamentais do modelo de Rasch. Esses itens se mostraram representativos de um contínuo de problemas decorrentes do beber, variando de sintomas leves e frequentes (por exemplo, “Sentiu dores de cabeça?” e “Sentiu enjoo ou mal-estar?”) a sintomas mais graves e menos frequentes (por exemplo, “Envolveu-se em acidente de trânsito” e “Teve problemas com a lei?”). Os 20 itens do instrumento, assim, mostraram-se abrangentes de um amplo espectro de níveis de problemas com bebidas do contexto universitário (quase sete unidades *logit*). Esses itens, portanto, são úteis para diversos níveis de problemas decorrentes do beber característicos da população universitária brasileira. Além disso, tanto o ajuste dos itens ao modelo de Rasch (*infit* médio = 1,00), quanto a fidedignidade (0,99) se mostraram aceitáveis, sendo a fidedignidade para as pessoas um pouco mais baixa (0,73), em função da pouca variabilidade dos escores da amostra na medida.

Person	MAP	Item
		<more> <rare>
3	.	+
		Prob22. Envolveu-se em acidente de trânsito?
	.	
		Prob8. Teve problemas com a lei?
2	.	+
		S Prob13. Teve necessidade de atendimento médico?
		Prob18. Teve pertences roubados?
	.	
	.	T
		Prob7. Envolveu-se em brigas?
1	#	+
	#	
		Prob17. Teve um comportamento sexual que não acharia apropriado (...)?
		Prob6. Usou drogas ilícitas?
	.	
	.	Prob5. Foi infiel em um relacionamento afetivo?
	.#	
		Prob1. Esqueceu-se de tarefas ou compromissos?
	.##	
		Prob10. Fez sexo sem usar as proteções que você normalmente utiliza?
0	.	+
	.#	
	.##	M
		S
	####	
		Prob15. Machucou-se de forma não-intencional?
		Prob9. Discutiu ou teve atrito com familiares, amigos, amigas, parceira ou parceiro?
	.	
	.#####	
		Prob21. Faltou à aula ou ao trabalho?
-1	.	+
		Prob11. Usou tabaco (cigarro, cigarrilha, narguilé etc.)?
		Prob12. Perdeu ou extraviou pertences?
	.#####	
	.	
		Prob14. Gastou dinheiro de forma inconsequente?
	.#####	
		Prob20. Esqueceu-se de coisas que fez sob o efeito do álcool?
		Prob4. Não conseguiu estudar como desejava?
	.#####	
-2		M+S
	.#####	
	.#####	
-3		+
		S
	.#####	
		T
-4		+
		Prob19. Sentiu enjoo ou mal-estar?
		Prob23. Sentiu dores de cabeça?
	.#####	
-5	#####	+
		<less> <frequ>
EACH "#" IS 5. EACH "." IS 1 TO 4		

Figura 1. Mapa dos itens para o instrumento

TABELA 2
Escores brutos, frequências, percentis, escores *logit*, erros padrão e escores *logit* padronizados

Escore Bruto	<i>n</i>	%	% Acumulada	Escore Percentil	Escore Logit	Erro Padrão	Logit Padronizado
0	50	11,7	11,7	6	-6,21	1,92	0,00
1	38	8,9	20,6	16	-4,73	1,19	1,48
2	54	12,6	33,3	27	-3,63	0,95	2,58
3	48	11,2	44,5	39	-2,86	0,81	3,35
4	37	8,7	53,2	49	-2,28	0,71	3,93
5	26	6,1	59,3	56	-1,82	0,65	4,39
6	37	8,7	67,9	64	-1,43	0,61	4,78
7	27	6,3	74,2	71	-1,08	0,58	5,13
8	49	11,5	85,7	80	-0,75	0,56	5,46
9	13	3,0	88,8	87	-0,44	0,55	5,77
10	8	1,9	90,6	90	-0,14	0,55	6,07
11	12	2,8	93,4	92	0,16	0,55	6,37
12	9	2,1	95,6	94	0,47	0,55	6,68
13	10	2,3	97,9	97	0,78	0,56	6,99
14	4	0,9	98,8	98	1,11	0,58	7,32
15	2	0,5	99,3	99	1,46	0,61	7,67
16	1	0,2	99,5	99	1,84	0,64	8,05
17	1	0,2	99,8	99	2,29	0,70	8,50
18	0	0,0	99,8	99	2,86	0,81	9,07
19	0	0,0	99,8	99	3,71	1,08	9,92
20	1	0,2	100	99	5,02	1,87	11,23

A análise de Rasch também permitiu a construção de um sistema de avaliação dos problemas decorrentes do beber que considera a gravidade específica de cada item. É possível observar, a partir da Tabela 2, como os escores percentis facilmente subestimam a real diferença entre os indivíduos em termos do nível da severidade dos problemas decorrentes do beber. Isso se deve ao fato de que a relação entre os escores brutos (dos quais se derivam esses percentis) e o nível de severidade (*logit*) da medida de Rasch é não-linear (Bond e Fox, 2007; Embretson e Reise, 2000). Assim, por exemplo, suponha-se que dois indivíduos *x* e *y* são acompanhados durante um período de tempo por seus terapeutas e, em ambos os casos, constata-se que a quantidade de seus problemas decorrentes do beber (escore bruto) aumentou. Suponha-se também que ambos os indivíduos aumentaram seis pontos, sendo que o indivíduo *x* foi de seis para 12 pontos no instrumento e o indivíduo *y* foi de 13 para 19 (ver Tabela 2). Em termos de percentil, o indivíduo *x* foi quem teve um maior aumento na magnitude dos seus problemas com a bebida, indo do percentil 64 para o percentil 94, ou seja, 30 pontos, contra apenas dois

pontos percentis de aumento para o indivíduo *y* (de 97 para 99). Todavia, ao se analisarem as mudanças em termos de escore *logit*, há uma grande diferença com relação às conclusões tiradas a partir dos escores percentis. Nesse caso, enquanto o indivíduo *x* aumentou 1,90 *logit*, o indivíduo *y* aumentou, de fato, 2,93 logits. Dito de outro modo, o indivíduo *y* teve um agravamento maior do que o indivíduo *x*, o que pode também ser avaliado qualitativamente mediante o Mapa dos Itens. Consultando-se o Mapa dos Itens, é possível ver que o indivíduo *x* provavelmente passou de problemas como perder pertences ao beber (sexto item) a problemas como se esquecer de compromissos (décimo segundo item). Enquanto isso, o indivíduo *y* provavelmente passou de ser infiel ao beber (décimo terceiro problema) a algo mais grave, como ter problemas com a lei (décimo nono problema). Em resumo, apenas as diferenças em termos de escore *logit* refletem o verdadeiro aumento na gravidade dos problemas enfrentados pelos dois indivíduos. Isso ocorre porque os escores *logit* levam em conta o escalonamento dos itens em termos da saturação no construto (dificuldade δ_i), o que não ocorre com os escores brutos e com os escores percentis.

O instrumento construído permite, portanto, levar em consideração a dificuldade relativa de cada item no escore total individual, o que não teria sido possível mediante as análises estatísticas tradicionais em Psicometria. Torna-se possível ponderar o relato de consequências negativas do uso do álcool, considerando que um evento ameno e frequente (por exemplo, sentir dores de cabeça) não possui o mesmo significado, impacto ou saturação que outro mais grave e menos frequente (por exemplo, ter necessidade de atendimento médico). Assim, para a prática clínica, o relato de apenas um problema que conste no instrumento permite supor que o indivíduo, provavelmente, já apresentou todos os problemas anteriores, uma vez que é exatamente isso que significa o ajuste dos itens ao modelo de Rasch (Bond e Fox, 2007). Além disso, o fato de a medida ter seus itens dispostos em um contínuo possibilita ainda estabelecer um padrão de quais sintomas são mais prováveis de se manifestar, caso um indivíduo aumente seu nível de problemas com o álcool. Tal propriedade pode, por exemplo, ter um impacto positivo na alocação de esforços e recursos em perspectivas de intervenção em saúde mental, ajudando a identificar indivíduos que merecem atenção diferenciada.

Vale ressaltar que, apesar de terem sido excluídos da versão final dois itens que apresentaram DIF (“Excedeu o limite máximo de velocidade permitido com um carro ou uma moto?” e “Danificou ou destruiu propriedade privada ou patrimônio público?”), isso não impede que esses itens sejam aplicados em situações futuras de pesquisa ou prática clínica. De fato, obter informações descritivas relacionadas ao conteúdo de ambos os itens é desejável e pode ser relevante do ponto de vista da avaliação clínica do nível do envolvimento do indivíduo com bebidas alcoólicas. Contudo, apenas se sugere que esses itens não sejam empregados para o cômputo dos escores dos indivíduos. Isso porque ambos os itens se mostraram enviesados para o sexo masculino, ou seja, foram indicativos de um menor nível de patologia para homens do que quando observados em mulheres. Nesse caso, equivaler o nível de dificuldade desses itens entre os sexos pode levar à subestimação do nível de problemas decorrentes do beber enfrentados por indivíduos do sexo feminino. Por esse motivo, recomenda-se a utilização da Tabela 2, baseada no instrumento com apenas 20 (sem esses dois itens com DIF), como referência para a conversão dos escores brutos em escores *logit* intervalares.

CONSIDERAÇÕES FINAIS

Uma limitação do presente estudo é que a média da medida *logit* para as pessoas (-2,04) coincidiu com um

gap ou espaço vazio em que não há itens-alvo específicos para indivíduos com essa magnitude de problemas. Houve apenas itens mais leves (“Sentiu enjoo ou mal-estar) ou mais severos (“Não conseguiu estudar como desejava?”) do que a média da presente amostra. Assim, desenvolvimentos futuros do instrumento devem considerar a inclusão de outros itens representando sintomas leves do contínuo de severidade de problemas com o álcool. A inclusão e a reformulação de itens são, no entanto, práticas recorrentes no desenvolvimento de instrumentos psicométricos com os procedimentos da Teoria de Resposta ao Item (Embretson e Reise, 2000).

Embora o tamanho amostral utilizado ($N=401$) seja suficiente para a obtenção de estimativas paramétricas estáveis para os itens (Nunes e Primi, 2005), são ainda necessários estudos sobre DIF avaliando outras variáveis potencialmente influentes. Nesse sentido, o tipo de universidade – se particular ou pública –, bem como outras variáveis sociodemográficas poderão ser investigadas em termos de DIF futuramente. Essas novas informações ajudarão a refinar o instrumento, reduzindo potenciais vieses e o erro de medida em diversas situações práticas. Na ausência de DIF, espera-se que os parâmetros dos itens se mantenham semelhantes para novas amostras, tendo em vista a característica do modelo de Rasch da independência das estimativas paramétricas com relação à amostra empregada (Bond e Fox, 2007; Embretson e Reise, 2000).

Estudos futuros também devem estimar a capacidade preditiva do instrumento, buscando assim evidências de validade de critério para o instrumento. Esses estudos poderão se beneficiar de uma medida intervalar de escores *logit*, mais representativos da severidade de problemas com o beber do que escores brutos. Enfim, apesar de ter apresentado adequadas propriedades psicométricas e ser um instrumento promissor no contexto nacional, novos estudos ainda poderão ajudar a refinar a medida.

REFERÊNCIAS

- Andrade, A.G., Duarte, P.C. & Oliveira, L.G. (2010). *I Levantamento Nacional sobre o uso de álcool, tabaco e outras drogas entre universitários das 27 capitais brasileiras*. Brasília: Secretaria Nacional de Políticas contra Drogas (SENAD).
- Andrich, D. (1988). *Rasch Models for Measurement*. Sage Publications, Inc.
- Babor, T.F. & Grant, M.G. (1989). From clinical research to secondary prevention: international collaboration in the development of the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Health and Research World*, 13, 371-374.
- Bond, T., Fox, C. (2007). *Applying the Rasch model* (2nd ed). Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

- Caetano, R., Mills, B., Pinsky, I., Zaleski, M. & Laranjeira, R. (2012). The distribution of alcohol consumption and the prevention paradox in Brazil. *Addiction*, 107(1), 60-8. doi:10.1111/j.1360-0443.2011.03567.x
- Clements, R. (1999). Prevalence of alcohol-use disorders and alcohol-related problems in a college student sample. *Journal of American College Health*, 48(3), 111-8. doi: 10.1080/07448489909595682
- Draba, R.E. (1977). The identification and interpretation of item Bias (Research Memorandum No. 25). Chicago: The University of Chicago, Department of Education, Education Statistics Laboratory. [Online] <<http://www.rasch.org/memo25.htm>>. Acesso em: mar. 2012.
- Earleywine, M., LaBrie, J.W. & Pedersen, E.R. (2008). A brief Rutgers Alcohol Problem Index with less potential for bias. *Addictive Behaviors*, 33(9), 1249-53. doi:10.1016/j.addbeh.2008.05.006
- Embretson, S.E., Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Engs, R.C., Hanson, D.J. (1994). The Student Alcohol Questionnaire: An updated reliability of drinking patterns, problems knowledge and attitude subscales. *Psychological Reports*, 74, 12-14.
- Fidalgo, A.M. (2011). GMHDIF: A computer program for detecting DIF in dichotomous and polytomous items using generalized Mantel-Haenszel Statistics. *Applied Psychological Measurement*, 35, 247-249.
- Fidalgo, Á.M. & Scalón, J.D. (2012). Using Mantel-Haenszel methods for detecting differential item functioning. *Psicologia: Reflexão e Crítica*, 25(1), 60-68.
- Hoepfner, B.B., Kahler, C.W. & Jackson, K.M. (2011). Evaluating the validity and utility of scaling alcohol consumption indices alongside AUD symptoms in treatment-seeking adolescents. *Drug and Alcohol Dependence*, 115(3), 196-204. doi:10.1016/j.drugalcdep.2010.10.023
- Hurlbut, S.C. & Sher, K.J. (1992). Assessing alcohol problems in college students. *Journal of American College Health*, 41(2), 49-58. doi:10.1080/07448481.1992.10392818
- Kahler, C.W. Strong, D.R. & Read, J.P. (2005). Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: The brief Young Adult Alcohol Consequences Questionnaire. *Alcoholism: Clinical and Experimental Research*, 29(7), 1180-1189.
- Kahler, C.W., Strong, D.R., Read, J.P., Palfai, T.P. & Wood, M.D. (2004). Mapping the continuum of alcohol problems in college students: a Rasch model analysis. *Psychology of Addictive Behaviors*, 18(4), 322-333. doi:10.1037/0893-164X.18.4.322
- Larimer, M.E., Cronce, J.M. (2002). Identification, prevention, and treatment: A review of individual-focused strategies to reduce problematic alcohol consumption by college students. *Journal of Studies on Alcohol, Suppl.*(14), 148-163.
- Linacre, J.M. (1991). *A user's guide to Winsteps, Program Manual 3.72.0*. Chicago; Winsteps.com.
- Maddock, J.E., Laforge, R.G., Rossi, J.S. & O'Hare, T. (2001). The College Alcohol Problems Scale. *Addictive Behaviors*, 26(3), 385-398. doi:10.1016/S0306-4603(00)00116-7
- Nunes, C.H.S. da S. & Primi, R. (2005). Impacto do tamanho da amostra na calibração de itens e estimativa de escores por teoria de resposta ao item. *Avaliação Psicológica*, 4(2), 141-153.
- Perkins, H.W. (2002). Surveying the damage: A review of research on consequences of alcohol misuse in college populations. *Journal of Studies on Alcohol, Suppl.*(14), 91-100.
- Peuker, A.C., Fogaça, J., Bizarro, L. (2006). Expectativas e beber problemático entre universitários. *Psicologia: Teoria e Pesquisa*, 22(2), 193-200.
- Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Test*. Copenhagen, Dinamarca: Denmark's Paedagogiske Institut.
- Read, J.P., Kahler, C.W., Strong, D.R., Colder, C.G. (2006). Development and preliminary validation of the Young Adult Alcohol Consequences Questionnaire. *Journal of Studies on Alcohol*, 67(1), 169-177.
- Silveira, C.M., Silveira, C.C., Silva, J.G. da, Silveira, L.M., Andrade, A.G. de, & Andrade, L.H.S.G. de. (2008). Epidemiology of heavy drinking and heavy episodic drinking in Brazil: a systematic review of literature. *Revista de Psiquiatria Clínica*, 35, 31-38.
- Souza, F.G.M., Landim, R.M., Perdigão, F.B., Morais, R.M. & Carneiro Filho, B.A. (1999). Consumo de drogas e desempenho acadêmico entre estudantes de medicina no Ceará. *Revista de Psiquiatria Clínica*, 26(4), 188-194.
- White, H.R. & Labouvie, E.W. (1989). Towards the assessment of adolescent problem drinking. *Journal of Studies on Alcohol*, 50, 30-37.

O presente trabalho foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) em forma de bolsa de doutorado ao primeiro autor.

Recebido em: 05.05.2012 Aceito em: 06.09.2012.

Autores:

Nelson Hauck Filho – Doutorando do Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Sul – UFRGS.

Marco Antônio Pereira Teixeira – Professor do Programa de Pós-Graduação em Psicologia da Universidade Federal do Rio Grande do Sul – UFRGS.

Enviar correspondência para:

Nelson Hauck Filho
Universidade Federal do Rio Grande do Sul
Instituto de Psicologia
Rua Ramiro Barcelos, 2600, sala 117 – Bairro Santana.
CEP 90035-003, Porto Alegre, RS, Brasil
E-mail: hauck.nf@gmail.com
hauck_psychology@hotmail.com