



## Avaliação de habilidade técnica em laboratório de Química usando modelamento Rasch

Terezinha Ribeiro Alvim, Alexandre da Silva Ferry

Centro Federal de Educação Tecnológica de Minas Gerais, CEFET-MG, Belo Horizonte, Brasil.

### ARTICLE INFO

**Received:** 29 October 2017

**Accepted:** 9 May 2018

**Available on-line:** 9 May 2018

**Keywords:** technical skill, evaluation, Rasch models, school laboratory.

**E-mail:** talvim@deii.cefetmg.br.

ISSN 2007-9842

© 2018 Institute of Science Education.  
All rights reserved

### ABSTRACT

This research aimed to develop and validate an instrument to evaluate the acquisition of technical skills titration of students of a technical course in mid-level Chemistry. The acquisition of technical practical skills is one of the objectives listed by teachers and researchers of the laboratory in science teaching and, even then, research on the subject are rare. Moreover, the absence of objective criteria makes subjective assessment of performance of students in the laboratory to be a problem for teachers. The tool developed in this research is the checklist type and was drawn from the Titration procedure described in the literature and direct observation through video recording of students performing titrations in the laboratory at different times during the technical teaching. Data were analyzed using the Rasch model for dichotomous data in order to validate the instrument for measuring ability. The good fit of the data to the Rasch model indicated that the instrument proved to be valid to measure titration ability, fulfilling the basic requirements of a measuring instrument as having an interval scale, measuring only one attribute and be invariant to the occasion measure.

Esta pesquisa teve como objetivo a construção e validação de um instrumento para avaliar a aquisição da habilidade técnica de titulação dos estudantes de um curso técnico em Química de nível médio. A aquisição de habilidades práticas de natureza técnica é um dos objetivos, listado por professores e pesquisadores, do ensino de ciências em laboratório e, mesmo assim, pesquisas sobre o tema são raras. Além disso, a ausência de critérios objetivos torna subjetiva a avaliação do desempenho dos estudantes no laboratório sendo um problema para os professores. O instrumento desenvolvido nesta pesquisa, é do tipo checklist e foi elaborado a partir do procedimento de Titulação descrito na literatura especializada, e da observação direta, por meio de gravação em vídeo dos estudantes realizando titulações no laboratório em ocasiões diferentes durante o ensino da técnica. Os dados obtidos foram analisados usando o modelo Rasch para dados dicotômicos com o intuito de validar o instrumento de medida da habilidade. O bom ajuste dos dados ao modelo Rasch indicou que o instrumento se mostrou válido para mensurar a habilidade de titulação, cumprindo os requisitos básicos de um instrumento de medida como: possuir uma escala intervalar, medir apenas um atributo e ser invariante em relação à ocasião de medida.

### I. INTRODUÇÃO

Este trabalho relata o processo de construção e validação de um instrumento, do tipo checklist, para avaliar a aprendizagem da habilidade técnica de titulação em laboratórios escolares de Química. O instrumento constitui-se de uma lista de indicadores ou itens cuja resposta sim ou não para cada um possibilita mensurar o nível da habilidade em questão. O objetivo central dessa pesquisa foi obter medidas verdadeiras da habilidade técnica dos estudantes que pudessem ser usadas em outro estudo que pretendia analisar e modelar a evolução da habilidade com o tempo de prática (Alvim, 2011).

O desenvolvimento de habilidade técnica é um dos objetivos das aulas de Química em laboratório nos cursos técnicos e de graduação (Bruck, Towns e Bretz, 2010) e a avaliação de sua aprendizagem apresenta-se como um desafio para os professores. Prades e Espinar (2010) mencionam que a avaliação do desempenho no laboratório é mais um resultado da impressão que o professor forma do estudante do que, efetivamente, da sua aprendizagem, devido à ausência de critérios específicos para avaliar. Acreditamos que o uso de instrumentos com critérios objetivos pode resolver esse problema, permitindo que se avalie com confiabilidade o desempenho do estudante.

O processo de mensuração caracteriza-se como um problema nas ciências humanas (Wright, 1997), uma vez que os atributos a serem medidos não são diretamente visíveis. Medir a habilidade de uma pessoa não é um procedimento direto como, por exemplo, medir a sua altura. A dificuldade está no fato de que a habilidade é um traço latente e de estrutura complexa e, portanto, não é diretamente observável como a altura da pessoa o é. A partir de indicadores desse atributo, sua mensuração pode ser realizada usando os modelos Rasch (Bond e Fox, 2007; Mead, 2008; Andrich, 1988), modelos matemáticos probabilísticos, que, de acordo com Wright e Linacre (1989), além de especificarem exatamente como converter contagens em medidas lineares também permitem descobrir até que ponto uma conversão particular é bem sucedida o bastante para ser útil.

Em nossa pesquisa usamos o modelo Rasch para dados dicotômicos que diz que a probabilidade de uma pessoa  $n$  responder corretamente a um item  $i$  pode ser expressa como:

$$(1) \quad P(X_{ni} = 1) = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}}$$

Na equação (1),  $B_n$  é o parâmetro da pessoa e  $D_i$  é o parâmetro do item. Rasch (1977) associou aos parâmetros qualidades da pessoa e do item: o parâmetro da pessoa indica o “grau de habilidade” da pessoa e o parâmetro do item o “grau de dificuldade” do item.

A chance real (odd, em inglês) de uma pessoa responder corretamente a um item  $i$ , é dada por:

$$\frac{P}{1-P} = e^{(B_n - D_i)} \quad (2) \quad \ln\left(\frac{P}{1-P}\right) = B_n - D_i \quad (3)$$

A equação (3) mostra que o logaritmo da chance real (log odds)<sup>3</sup> de sucesso da pessoa sobre o item depende da diferença entre a habilidade da pessoa e a dificuldade do item. Quanto maior a diferença, maior a chance de a pessoa acertar o item. Considere a situação em que a pessoa  $n_1$  tem uma habilidade,  $B_1$ , maior do que a da pessoa  $n_2$ ,  $B_2$ . Então, a diferença entre as chances reais de sucesso sobre o mesmo item de dificuldade  $D_i$ , das pessoas  $n_1$  e  $n_2$ , é dada por:

$$(B_{n_1} - D_i) - (B_{n_2} - D_i) = B_{n_1} - B_{n_2} \quad (4)$$

A equação (4) mostra que a diferença entre as chances de acerto do mesmo item entre duas pessoas só depende da habilidade das duas pessoas e independe do parâmetro do item. O mesmo princípio pode ser aplicado à comparação de dois itens para uma mesma pessoa. Essa propriedade, a invariância dos parâmetros, está de acordo com um dos requerimentos da mensuração mencionado por Mead (2008, p. 6, tradução nossa): “A calibração dos agentes deve ser

independente dos objetos usados e a mensuração dos objetos deve ser independente dos agentes usados, sobre uma extensão útil”. Este princípio é exclusivo das medidas Rasch e é denominado por Rasch (1977) de objetividade específica.

As medidas das habilidades das pessoas e das dificuldades dos itens são construídas a partir das estimativas dos parâmetros  $B_n$  e  $D_i$  calculadas por um algoritmo, usando um programa de computador.

Um requisito básico dos dados para que possam ser analisados pelos modelos Rasch é a unidimensionalidade (Wright e Linacre, 1989). O instrumento usado para mensuração deve medir apenas um tipo de construto por vez.

Alcançar a unidimensionalidade configura-se como uma atividade problemática quando os construtos a serem medidos referem-se ao comportamento humano que é complexo e envolve múltiplas variáveis. Ao se tentar mensurar, por exemplo, uma habilidade técnica da pessoa, a proficiência que se está procurando capturar tem uma parte que envolve habilidade intelectual, outra parte que envolve habilidade perceptual-motora e outra parte que envolve motivação sendo, portanto, difícil obter-se unidimensionalidade irrestrita. Mesmo assim, segundo Andrich (1988) um teste deve aproximar-se o máximo possível do ideal de medida unidimensional se a obtenção de resultados generalizáveis é o que se busca.

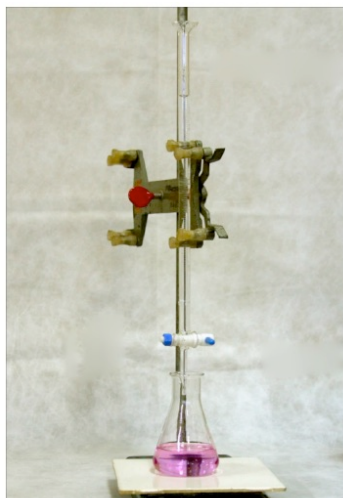
## II. METODOLOGÍA

### II.1. Participantes

Trinta e três (33) estudantes, do gênero masculino e feminino na faixa etária de dezesseis (16) a dezenove (19) anos participaram da pesquisa, todos pertencentes à turma da 2ª série do curso técnico de nível médio em Química de uma instituição federal de educação tecnológica.

### II.2. A tarefa e a lista de verificação

A técnica da Titulação é utilizada para determinar as quantidades de diversas espécies químicas em soluções, como, por exemplo, ácidos e bases. Resumidamente, o procedimento consiste em, por meio de uma bureta, adicionar lentamente a solução padrão, cuja concentração exata é conhecida, em um recipiente (geralmente um erlenmeyer) contendo a solução problema cuja concentração deseja-se determinar e um indicador do final da reação, até que se complete a reação. A Figura 1 mostra uma montagem típica de uma titulação ácido base. A partir dos volumes de ambas as soluções que reagem e da concentração da solução padrão, calcula-se a concentração da solução problema.



### FIGURA 1. Montagem de uma titulação.

A Titulação é um procedimento complexo composto de várias subtarefas que podem ser agrupadas em quatro (4) etapas de execução: preparação da bureta, preparação do erlenmeyer, titulação em si e cálculo do resultado. Os itens do instrumento (checklist), usados para avaliar a habilidade técnica são descritos como as subtarefas que constituem os passos do procedimento a ser seguido em cada uma das etapas da titulação que são respondidos dicotomicamente: fez ou não fez da forma como está descrito. Eles foram construídos com base nos procedimentos descritos na literatura especializada (Jeffery et al, 1992; Harris, 2008; Skoog et al, 2006; Ohlweiler, 1982) e da observação dos estudantes realizando a técnica.

### II.3. A coleta e a análise dos dados

Realizamos a coleta dos dados durante as aulas de Laboratório, cujo tema era titulação, por meio de gravações em vídeo dos estudantes executando a técnica em três ocasiões distintas: na primeira, segunda e terceira semana de aula prática sobre o assunto. As filmagens tiveram início duas semanas antes das aulas sobre titulação para que os estudantes se acostumassem com a presença das filmadoras dentro do laboratório.

Os vídeos foram analisados por ocasião (dia de aula), por estudante e por subtarefa, de modo que a resposta a cada item era registrada para todos os estudantes. Então, o instrumento era preenchido anotando-se, o escore de cada estudante para cada item como 0 (não fez) ou 1 (fez), para cada semana de aula de modo que obtivemos os escores empíricos da habilidade de um mesmo estudante para cada ocasião em que ele executou a técnica. A matriz resultante das respostas dos estudantes aos itens nas três ocasiões foi então analisada pelo programa WINSTEPS® (Linacre, 2010) usando o tratamento Rasch.

## III. RESULTADOS

### III.1. Construção do Instrumento para Avaliação da Habilidade Técnica de Titulação

O Instrumento para avaliação da habilidade técnica foi construído a partir do procedimento previamente descrito e da análise crítica dos vídeos que mostrou a necessidade de introduzir e modificar alguns itens para obtenção de um sistema de indicadores mais condizente com a realidade do ambiente de laboratório de ensino. Por exemplo, nas descrições das subtarefas antecipadas pela teoria não havia menção ao uso de vidraria limpa ou não, uma vez que essa é uma condição que está implícita para a execução de qualquer técnica de análise química. Mas, ao ver que na realização de algumas subtarefas alguns estudantes usaram vidraria contaminada ou suja, decidimos introduzir essa observação nos indicadores referentes àquelas subtarefas.

Obtivemos, então, um instrumento do tipo checklist contendo 26 itens (Quadro 1) para avaliar a habilidade técnica em titulação dos estudantes nas três ocasiões em que eles executaram a titulação como atividade de aula prática.

As respostas (1 para sim e 0 para não) dos 33 estudantes aos 26 itens foram registradas para cada ocasião e organizadas na forma de matriz. A soma dos resultados dos 26 itens para cada estudante constituiu seu escore empírico do desempenho na execução da técnica de titulação naquela ocasião.

Primeiro verificamos que o teste atende ao requisito da unidimensionalidade do modelo Rasch não havendo estrutura nos resíduos das respostas. E a dimensão Rasch do teste explica 34,5% da variância nos dados. À primeira vista parece um valor muito baixo, mas o modelo Rasch já prediz certa aleatoriedade nos dados, pois é um modelo probabilístico e, às vezes, nos surpreende o quão grande é esta fração do componente aleatório (Linacre, 2010, p. 317).

O máximo de variância explicada predita pelo modelo Rasch para dados dicotômicos para um desvio padrão igual a 2 logits para os parâmetros das pessoas e para os dos itens seria cerca de 50% (Linacre, 2010, p. 317). Os dados obtidos numa primeira análise apresentaram um desvio padrão para as pessoas igual a 0,75 e igual a 1,52 para os itens, correspondendo, aproximadamente, a um máximo de 45% de variância explicável pelo modelo Rasch. Portanto, explicar 34,5% corresponde a explicar 76% da variância explicável (45%).

**QUADRO 1. Instrumento de Avaliação da Habilidade Técnica de Titulação.**

Item	1ª fase - Preparação da bureta	Sim	Não
01	Fez ambiente na bureta com a solução básica.		
02	Verteu a solução básica contida em um béquer previamente limpo e seco, deste para a bureta.		
03	Durante o preenchimento a bureta estava fixada no suporte.		
04	Durante o preenchimento a bureta estava com a torneira fechada.		
05	Preencheu a bureta até um pouco acima do traço de aferição zero e deixou escoar até o zero preenchendo a ponta de primeira.		
06	Usou somente a mão esquerda, passando os quatro dedos por trás e o polegar pela frente da bureta, para abrir e fechar a torneira da bureta.		
07	Durante a aferição do menisco com o traço zero da bureta (escoamento da solução da bureta para o béquer) a ponta da bureta estava introduzida no béquer cerca de 1 cm.		
08	Aferiu o nível da solução em relação ao traço de aferição zero da bureta, colocando o traço de aferição nivelado com os olhos.		
09	Verificou se ficaram bolhas de ar no interior da bureta, incluindo a ponta.		
	2ª fase - Preparação do erlenmeyer	Sim	Não
10	Pipetou a solução ácida previamente transferida para um béquer limpo e seco.		
11	Fez ambiente na pipeta volumétrica com a solução ácida.		
12	Usando a pèra fez a sucção da solução ácida até um pouco acima do traço de aferição da pipeta volumétrica e fez escoar a solução até que a concavidade do menisco tangencie o traço.		
13	Aferiu o menisco da solução em relação ao traço de aferição da pipeta, colocando o traço nivelado com os olhos.		
14	Retirou a pipeta da solução do béquer e imediatamente introduziu sua ponta dentro do erlenmeyer limpo e deixou escoar, pressionando o botão de escoamento da pèra, todo o volume medido da solução dentro do erlenmeyer.		
15	Adicionou o indicador à solução do erlenmeyer imediatamente após a adição da solução ácida.		
16	Homogeneizou a solução após adicionar o indicador.		
	3ª fase - titulação	Sim	Não
17	Realizou a titulação de pé, de frente para a montagem durante todo o procedimento.		
18	Colocou a ponta da bureta introduzida cerca de 2 cm no erlenmeyer.		
19	Usou somente a mão esquerda, passando os quatro dedos por trás e o polegar pela frente da bureta, para abrir e fechar a torneira da bureta.		
20	Manteve a mão na torneira todo o tempo enquanto estava titulando.		
21	Livrou a solução básica da bureta lentamente, gota a gota e continuamente.		
22	Agitou o erlenmeyer com a mão direita de forma contínua à medida que livrava a solução básica da bureta sobre a solução.		
23	Durante todo o tempo da titulação manteve o olhar direcionado para o erlenmeyer.		
24	Obteve uma cor indicativa de viragem apropriada no ponto final da titulação.		
25	Leu o volume gasto na bureta corretamente sem apresentar dúvidas.		
	4ª fase - Cálculos para obtenção do resultado final	Sim	Não
26	Efetou os cálculos sozinho(a), sem ajuda.		

### III.2. Construção da escala Rasch de medida da habilidade

Verificadas as evidências de ajuste ao modelo, e obtidas as medidas estimadas das habilidades dos estudantes e das dificuldades dos itens, realizamos a análise diferencial dos itens e do teste graficamente para verificar a invariância dos itens em relação às três ocasiões e a diferentes grupos de desempenho. Primeiramente, para cada ocasião, dividimos os

estudantes em dois grupos, um de baixo e outro de alto desempenho, de acordo com o escore total no teste. Então, construímos os gráficos das curvas características do teste e dos itens com seus respectivos intervalos de confiança.

Nesses gráficos foram “plotadas” a localização de cada grupo de desempenho como sendo a média das medidas estimadas das habilidades das pessoas do grupo para cada ocasião e a probabilidade de acerto foi calculada a partir do escore médio empírico do grupo. Os intervalos de confiança foram calculados num nível de confiança de 83% como proposto por Goldstein e Healy (1995) para que o critério de qualquer sobreposição entre eles fosse suficiente para afirmar que as medidas são estatisticamente equivalentes. Dessa forma uma simples inspeção visual dos gráficos permitiu avaliar o ajuste dos resultados observados aos previstos pelo modelo. Se havia superposição entre o intervalo de confiança dos pontos referentes aos grupos e o intervalo de confiança da curva, considerava-se que os resultados empíricos observados se ajustavam aos previstos pelo modelo e que aquele item funcionava igualmente para os dois grupos e para as três ocasiões.

A Figura 2 mostra a curva para o teste e as Figuras 3 e 4 mostram, respectivamente, as curvas de um item bem ajustado e de um com um ajuste não tão bom, mas aceitável. A curva do teste indicou um bom ajuste do teste como um todo. Os grupos de alto desempenho nas três ocasiões estão abaixo da curva do item, mas todos estão dentro do intervalo de confiança, isto é, as barras estão superpondo bem, portanto o teste funcionou igualmente para esses grupos.

O teste também funcionou igualmente para os grupos de baixo desempenho nas três ocasiões. Três (3) itens apresentaram problemas de ajuste em relação às ocasiões e foram então separados por ocasião.

Uma possível explicação pode ser o fato de que a tarefa não foi exatamente igual nas três ocasiões apesar de ser sempre uma titulação ácido base. Na primeira ocasião a tarefa foi a padronização de solução de hidróxido de sódio com o padrão hidrogeno ftalato de potássio; na segunda foi a padronização de solução de ácido clorídrico com a solução de hidróxido de sódio padronizada e na terceira ocasião foi a determinação da alcalinidade da soda cáustica comercial.

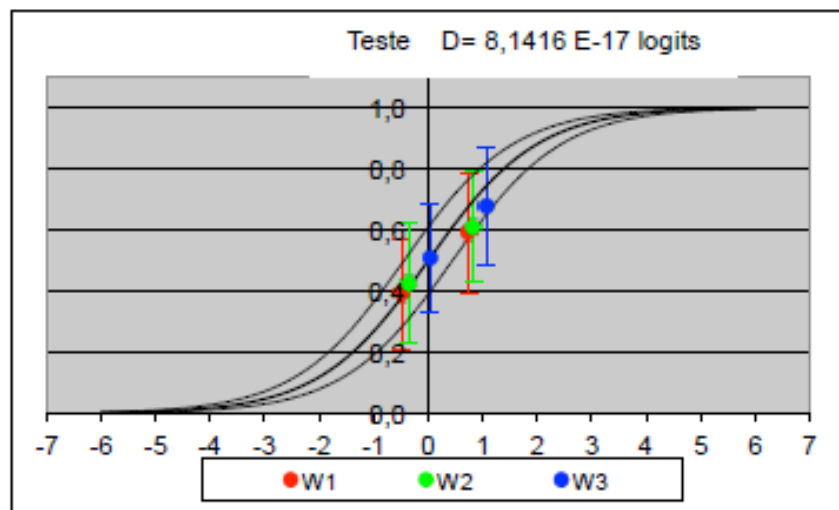


FIGURA 2. Curva do teste. W1, W2 e W3 indicam, respectivamente a 1ª, 2ª e 3ª ocasiões de medida.

Ao final, obteve-se uma escala com 30 itens que foi considerada invariante para as três ocasiões e para os dois grupos de desempenho. A variância explicada aumentou um pouco, indo para 37%. Os autovalores extraídos dos resíduos indicam que não há estrutura nos mesmos de modo que consideramos que nosso instrumento se aproxima bem da unidimensionalidade.

Além da análise gráfica para verificar o ajuste dos dados ao modelo Rasch também verificamos a estatística de ajuste INFIT (MNSQ - sigla para quadrados médios em inglês) que têm um valor esperado de 1 variando entre 0 e + (Golino et al, 2015). Valores maiores que 1 para este índice significam que as respostas ao item são inesperadas em relação ao esperado pelo modelo e valores menores que 1 significam que as respostas ao item são previsíveis em relação ao esperado pelo modelo. Em nossa análise o item 24b apresentou um valor de 1,26 para o INFIT (MNSQ) e o item 19a o valor de 0,75. Os demais apresentaram valores bem próximos de 1 (0,82 a 1,12). Segundo a interpretação recomendada por Linacre (2010) para essa estatística os valores encontrados indicaram que não houve distorção nem degradação das medidas para os itens (valores entre 0,5 e 1,5).

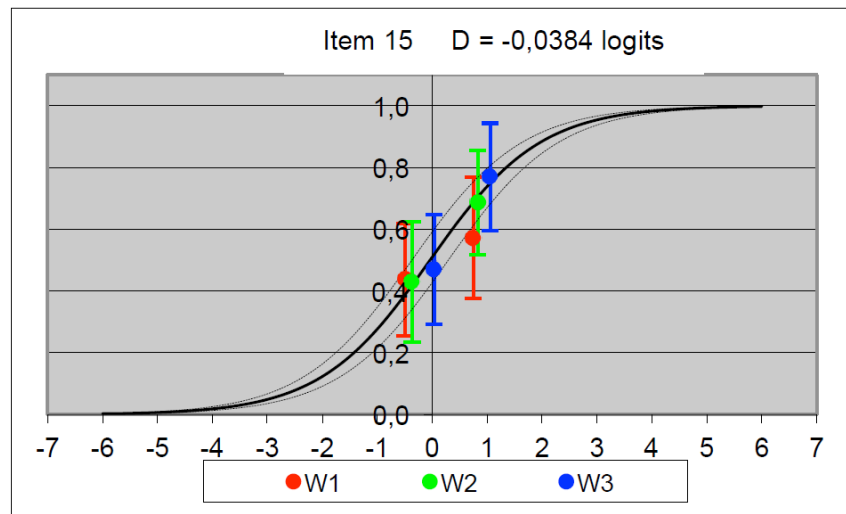


FIGURA 3. Curva de um item bem ajustado.

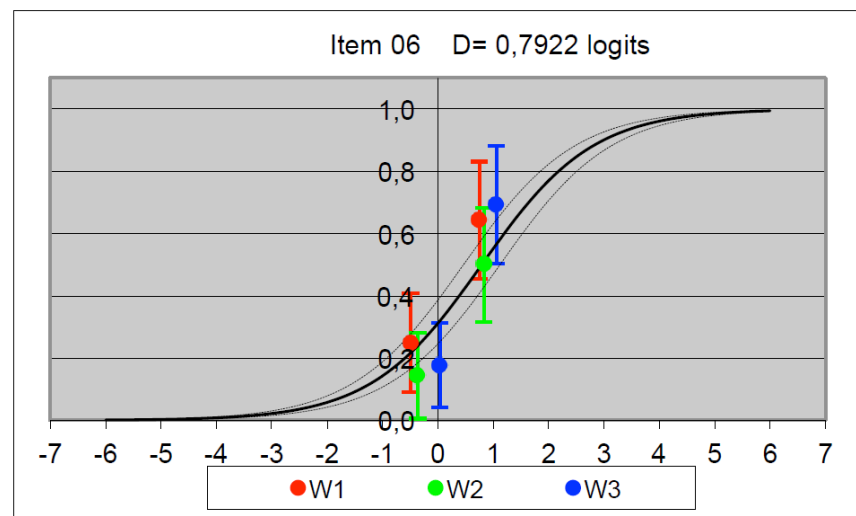


FIGURA 4. Curva de um item com ajuste aceitável.

#### IV. CONCLUSÕES

Consideramos que para dados obtidos em ambiente de ensino e de aprendizagem não controlado como é o laboratório escolar, nossos resultados apontam para um instrumento construído que é válido para mensurar a habilidade técnica de titulação em laboratórios de Química. Este tipo de instrumento pode ser usado para acompanhar o desenvolvimento da habilidade técnica dos estudantes durante o ensino bem como, para avaliar a aprendizagem da técnica.

## AGRADECIMENTOS

Ao CEFET-MG e ao Departamento de Química.

## REFERÊNCIAS

- Alvim, T. R. (2011) Desenvolvimento da Habilidade Técnica de Titulação em um Laboratório Escolar de Química. Tese (Doutorado em Educação) – Faculdade de Educação, Universidade Federal de Minas Gerais.
- Andrich, D. (1988) Rasch Models for Measurement. Sage Publications.
- Bond, T.G. e Fox, C. M. (2007) Applying the Rasch Model: Fundamental Measurement in the Human Sciences. London: LAWRENCE ERLBAUM ASSOCIATES.
- Bruck, L. B.; Towns, M.; Bretz, S. L. (2010) Faculty Perspectives of Undergraduate Chemistry Laboratory: Goals and Obstacles to Success. *Journal of Chemical Education*, 87, 1416.
- Goldstein, H.; Healy, M. J. R. (1995). The Graphical Presentation of a Collection of Means. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 158, 175-177.
- Harris, D. C. (2008) Análise Química Quantitativa. Rio de Janeiro: Livros Técnicos e Científicos Editora.
- Jeffery, G. H. et al. (Org.) (1992) VOGEL Análise Química Quantitativa. Rio de Janeiro: Livros Técnicos e Científicos S. A.
- Linacre, J. M. (2010) WINSTEPS® (Version 3.70.0) Rasch measurement computer program. Beaverton, Oregon.
- Mead, R. J. (2008) A Rasch primer: The measurement theory of Georg Rasch. Psychometrics services research memorandum 2008-001, Maple Grove: Data Recognition Corporation.
- Ohlweiler, O. A. (1983) Química Analítica Quantitativa, vol. 2, Rio de Janeiro: Livros Técnicos e Científicos Editora S. A.
- Prades, A.; Espinar S. R. (2010) Laboratory assessment in chemistry: an analysis of the adequacy of the assessment process. *Assessment & Evaluation in Higher Education*, 35, 449-461.
- Rasch, G. (1977) On Specific Objectivity: An Attempt at Formalizing the Request for Generality and Validity of Scientific Statements. Disponível em: <<http://www.rasch.org/memo18.htm>>.
- Skoog, D. A. et al. (2006) Fundamentos de Química Analítica. São Paulo: Pioneira Thomson Learning.



Wright, B. D. (1997) *Measurement for Social Science and Education: History of Social Science Measurement*, Disponível em: <http://www.rasch.org/memo62.htm>.

Wright, B. D.; Linacre, J. M. (1989) *Observations are Always Ordinal; Measurements, however, Must be Interval*. Disponível em: <http://www.rasch.org/memo44.html>.