

ACESSANDO O RACIOCÍNIO DOS ESTUDANTES ATRAVÉS DO TESTE EM TRÊS CAMADAS

ACCESSING STUDENT REASONING THROUGH THREE-TIERS TEST

Viviane Florentino de Melo¹, Wanderley Paulo Gonçalves Junior², Amanda
Amantes Neiva³

¹Universidade Federal da Bahia/Faculdade de Educação, melovivi221@hotmail.com

²Universidade Federal do Rio de Janeiro/Colégio de Aplicação, wpgjunior@gmail.com

³Universidade Federal da Bahia /Instituto de Física, amandaamantes@gmail.com

Resumo

Neste trabalho apresentamos um estudo sobre o uso do teste em três camadas como instrumento de pesquisas na área de ensino, em duas perspectivas: itens baseados em teoria e adaptação de testes consolidados. O objetivo foi elaborar itens a partir da teoria de perfil epistemológico e adaptar testes consolidados na literatura ao formato em três camadas. Esse último foi realizado para o inventário conceitual de força amplamente utilizado na área de ensino de Física. Testes em três camadas são uma ferramenta com três níveis de aferição do conhecimento, denominados como camadas. Na primeira camada o estudante é questionado sobre um conteúdo de maneira a explicitar, de forma dicotômica, se ele detém o conhecimento em questão. Na segunda camada o questionamento se refere a uma explicação sobre o conhecimento do item anterior, sendo que na terceira o estudante deve indicar o grau de confiança que sente em relação às suas repostas das duas primeiras camadas, demonstrando uma avaliação metacognitiva do seu conhecimento. Por meio do teste em três camadas é possível compensar as limitações inerentes a itens de múltipla escolha simples, que não acessam o raciocínio dos estudantes, e ainda assim se beneficiar de suas vantagens, como a correção objetiva e rápida. Como o trabalho é de cunho metodológico, o resultado diz respeito à consolidação de duas ferramentas de pesquisa. Consideramos que a construção validação de testes em três camadas é relevante apresenta grande potencial para pesquisas na área de ensino e para sala de aula, pois podem ser utilizados enquanto um instrumento de avaliação formativo.

Palavras-chave: teste em três camadas, perfil epistemológico, inventário conceitual de força, instrumentos de avaliação educacional, variáveis latentes.

Abstract

In this work, we present a study on the use of the three-tiers test as a research instrument in the teaching subject, under two points of view: theoretical based items, and adaptation of consolidated tests. The objective is to elaborate items from the epistemological profile theory, and to adapt consolidated tests in the literature to the three-tiers form. The later was accomplished to the conceptual inventory of force, largely used in the physics teaching field. Three-tiers test consist of a first layer in which the students are asked about the content in question, a second one in which they must indicate an alternative that justifies their answers to the previous item and a third one in which they must indicate the degree of confidence they feel in relation to their responses to the first two tiers. Through this instrument it is possible to compensate for the limitations inherent to simple multiple-choice items, which do not access the

students' reasoning, and still benefit from its advantages, objective and quick correction. As the work is of a methodological nature, the result concerns the consolidation of two research tools. Therefore, the construction and validation of three-tiers test has a lot of potential for research in the teaching area, in addition to being able to be used as an instrument of formative assessment in the classrooms.

Keywords: three-tiers test, force concept inventory, epistemological profile, educational assessment tool, latent variables.

Introdução

Pesquisas na área de Educação e Ensino, que se debruçam sobre questões relacionadas ao aprendizado dos estudantes, lidam recorrentemente com a dificuldade de construir instrumentos capazes de acessar esse constructo teórico. Isso porque não há um consenso na área sobre o que é aprendizagem, tampouco, como acessá-la. Algumas teorias, de cunho cognitivo (FISCHER, 2008; COMMONS e PEKKER, 2008), consideram a aprendizagem como a evolução do entendimento sobre determinada temática ou assunto. Desse modo, a partir delas, para se acessar o aprendizado deve-se utilizar instrumentos capazes de captar a evolução do entendimento dos sujeitos.

Nesse contexto, o constructo teórico entendimento se caracteriza por ser uma variável latente, a qual não se acessa diretamente, sendo evidenciada por meio de outras variáveis que podem ser observadas e mensuradas (BOND e FOX, 2007). É considerado, portanto, um traço latente o entendimento, a atitude, o conhecimento, as emoções ou qualquer outro atributo que não pode ser aferido diretamente e que só pode ser dimensionado a partir de uma inferência sobre o que se observa do indivíduo. A lógica que rege o modelo de variáveis latentes pressupõe que mudanças nessas variáveis resultam em mudanças nas variáveis observáveis (GOLINO e GOMES, 2015). A ligação entre as variáveis latentes e àquelas que podem ser observadas ocorre por meio de uma teoria que determina o que significam comportamentos específicos relacionados às variáveis.

Assim sendo, em um teste de conhecimento, a variável latente a ser acessada pode ser definida como o entendimento do estudante sobre determinado conteúdo. Já as variáveis observáveis referem-se as respostas dadas pelo sujeito aos itens do referido teste. No entanto, essa conexão não é estabelecida de forma trivial, fato que demanda a construção cuidadosa de instrumentos de medida confiáveis e fidedignos (GOLAFSHANI, 2003; RAYMUNDO, 2009).

Os testes de conhecimento, quando bem construídos e devidamente validados, são instrumentos eficientes para acessar o entendimento dos estudantes. Em pesquisas de natureza qualitativa são comuns testes baseados em itens discursivos, visto que por meio desses pode-se inferir sobre o raciocínio dos sujeitos. Todavia, a utilização de itens discursivos apresenta limitações, tais como: a dificuldade de escrita dos alunos, a quantidade de respostas e a subjetividade inerente às análises.

Em pesquisas de natureza quantitativa ou mista (quali-quantitativa) a quantidade amostral necessária, na maioria das vezes, inviabiliza o uso de itens discursivos, o que a isenta das limitações inerentes a esse tipo de item, mas também a impossibilita de inferir sobre o raciocínio dos estudantes. Geralmente em pesquisas dessa natureza os instrumentos de coleta são formados por itens dicotômicos, de múltipla escolha ou

verdadeiro ou falso. Toda coleta de dados carrega consigo limitações, de modo que o pesquisador deve procurar métodos condizentes com a teoria adotada e refletir não apenas sobre as potencialidades, mas também acerca das limitações de sua metodologia e pesquisa.

Neste trabalho apresentamos uma alternativa para testes de conhecimento compostos por questões de múltipla escolha que possibilita ao pesquisador ter acesso ao raciocínio dos indivíduos: o teste em três camadas.

Teste em três camadas

O teste em três camadas surgiu a partir da preocupação de se criar instrumentos de avaliação que apresentassem um caráter formativo, de modo a incentivar os estudantes a questionar, refletir e entender os conceitos científicos. Seu objetivo é fazer com que os estudantes, por meio de questões em camadas, possam extrapolar a simples memorização de fatos básicos, o que normalmente é o foco dos testes de múltipla escolha tradicionais.

Nesse contexto, a ideia do desenvolvimento de testes em três camadas surge a partir dos resultados positivos obtidos no uso de testes de múltipla escolha que solicitavam aos estudantes uma justificativa para sua resposta (KIRBULUT e GEBAN, 2014).

Na construção desse teste, a primeira camada de cada item é constituída por uma questão de múltipla escolha que contempla o conteúdo que se deseja avaliar. As alternativas dessa camada devem ser construídas a partir de concepções espontâneas ou não científicas de áreas de conteúdo bem definidas e delimitadas (CETIN-DINDAR e GEBAN, 2011; TREAGUST, 1998, 2012).

A segunda camada é composta por um conjunto de explicações plausíveis que servem de justificativa para as alternativas apresentadas na primeira camada. Cada alternativa da primeira camada deve ter uma explicação correspondente na segunda camada (TREAGUST, 1998, 2012). A terceira camada, por fim, tem como objetivo avaliar o grau de confiança do estudante ao realizar sua marcação da primeira e segunda camadas.

O conteúdo abordado no teste deve ser determinado a partir do conhecimento científico já existente e delimitado por uma lista de objetivos. As concepções espontâneas ou não científicas dos alunos utilizadas na construção da primeira camada devem ser obtidas por meio da literatura já existente. Na falta desta, pode-se usar as respostas dos estudantes em sala, entrevistas não estruturadas com estudantes que já tiveram instrução formal sobre o conteúdo avaliado e questões discursivas sobre o tema (TAN, GOH, *et al.*, 2002; CALEON e SUBRAMANIAM, 2010; ERYILMAZ, 2010; CETIN-DINDAR e GEBAN, 2011; ARSLAN, CIGDEMOGLU e MOSELEY, 2012; TREAGUST, 1998, 2012; KIRBULUT e GEBAN, 2014).

Nesse cenário, o teste em três camadas pode ser utilizado de forma eficiente para um grande número de respondentes, permitindo aos pesquisadores acessar o raciocínio dos estudantes, sem que seja necessária a realização de entrevistas para distinguir entre o que representa concepções espontâneas e o que é falta de conhecimento desses alunos (CALEON e SUBRAMANIAM, 2010; ERYILMAZ, 2010; KIRBULUT e GEBAN, 2014).

Elaboração do teste

Nesse trabalho trazemos a elaboração do teste em três camadas a partir de duas perspectivas – 1 - construção de itens baseados em teoria e, 2 - adaptação de testes consolidados. Apresentamos um exemplo de cada uma dessas perspectivas: um referente ao conceito científico de Densidade e outro ao de Força.

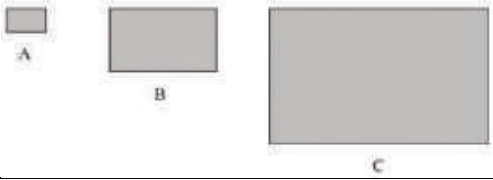
1 – Construção de itens baseados em teoria

A construção de teste em camadas a partir de uma teoria do conhecimento (Figura 1) foi feita tendo como base a teoria do perfil epistemológico de Gaston Bachelard (1979). A teoria do perfil epistemológico considera uma pluralidade de formas de entendimento dos conceitos científicos, segmentadas em zonas distintas, cada uma delas apresentando uma forma específica de compreensão da realidade.

Para a construção dos itens, inicialmente o conceito científico de densidade foi segmentado de acordo com as zonas do perfil epistemológico, para depois serem elaborados itens referentes à cada uma das respectivas zonas. Apresentamos aqui o exemplo de um item referente à zona do realismo ingênuo – primeira zona de acordo com a teoria de Bachelard, que diz respeito ao modo cotidiano como as pessoas operam com os conceitos científicos.

Uma associação muito comum nessa zona do perfil é a relação estabelecida entre densidade e tamanho ou forma dos objetos, que ocasiona afirmações cotidianas que associam erroneamente densidade à volume e/ou à massa. Visando acessar o entendimento dos estudantes acerca dessa associação, de modo a saber se eles estabelecem ou não essa relação e por quais razões, criamos a seguinte questão em três camadas.

Figura 1 – Item em três camadas construído a partir de uma teoria de conhecimento


Considere as três chapas ao lado, cada uma com suas respectivas dimensões, todas formadas pelo mesmo material, à mesma temperatura. <div style="text-align: right;">  </div>	
1ª Camada	2ª Camada
Com relação à densidade delas, podemos afirmar que:	Qual das seguintes explicações é o motivo de você ter marcado a resposta para a pergunta anterior?
(A) A densidade da A é a menor. (B) A densidade da A é a maior. (C) A densidade das três é a mesma. (D) A densidade da B é a média da densidade da A e da C. (E) A densidade da C é o dobro da de B.	(A) Como B tem volume e massa intermediária em relação a A e C, sua densidade é a média dos valores de A e C. (B) Como as três são formadas pelo mesmo material, a densidade delas é a mesma. (C) Como A é a menor, ela é a mais leve, logo a menos densa. (D) Como a C tem aproximadamente o dobro do tamanho da B, sua densidade também segue essa relação. (E) Como A é a menor, a massa dela é mais concentrada, por isso ela é a mais densa.
3ª Camada	
Quanto seguro(a) você se sente para dar as respostas das duas perguntas anteriores? “muito seguro (a)” [MS] / “seguro(a)” [S] / “neutro(a)” [N] / “inseguro(a)” [I] / “muito inseguro(a)” [MI]	
(A) MS	(B) S
(C) N	(D) I
(E) MI	

Fonte: autores do artigo

2 – Adaptação de testes consolidados

Outra forma de elaborar teste em três camadas é utilizar questionários já consolidados, como o “Force Concept Inventory” (FCI), traduzido e validado por Fernandes e Talim (2009). O inventário contém 30 questões de múltipla escolha (com cinco alternativas cada) que acessam conhecimento intuitivo e/ou científico dos alunos em relação aos conceitos/ da Mecânica Newtoniana, mais especificamente ao conceito de Força. Na segunda camada de cada item, buscou-se utilizar concepções intuitivas e não científicas dos estudantes para construção alternativas que justificassem a marcação feita na primeira camada.

Figura 2 – Item em três camadas construído a partir do FCI

<p>Na figura, o estudante “a” tem uma massa de 95 kg e o estudante “b” tem uma massa de 77 kg. Os estudantes estão sentados de frente um para o outro em cadeiras de escritório idênticas. O estudante “a” coloca os seus pés nos joelhos do estudante “b”, conforme mostrado. O estudante “a” subitamente empurra com os pés, fazendo com que ambas as cadeiras se movam.</p>				
1ª camada		2ª camada		
Durante o impulso e enquanto os estudantes ainda estão em contato um com o outro:		Qual das seguintes explicações é o motivo de você ter marcado a resposta para a pergunta anterior?		
(A) Nenhum estudante exerce uma força no outro. (B) O estudante “a” exerce uma força em “b”, mas “b” não exerce nenhuma força em “a”. (C) Cada estudante exerce uma força no outro, mas “b” exerce a força maior. (D) Cada estudante exerce uma força no outro, mas “a” exerce a força maior. (E) Cada estudante exerce a mesma força no outro.		(A) A força aplicada pelo corpo é inversamente proporcional a sua massa. (B) Não há forças atuando nos corpos porque nenhum dos dois ainda entrou em movimento. (C) Ainda que o estudante “a” seja responsável pelo impulso, ao exercer uma força no estudante “b”, ele também recebe uma força de mesmo valor. (D) Como o responsável pelo impulso é o estudante “a”, somente ele faz força sobre o estudante “b”. (E) Quanto maior a massa do corpo, maior a força aplicada por ele em suas interações.		
3ª camada				
Quanto seguro(a) você se sente para dar as respostas das duas perguntas anteriores?				
"muito seguro(a)" [MS] / "seguro(a)" [S] / "neutro(a)" [N] / "inseguro(a)" [I] / "muito inseguro(a)" [MI]				
(A) MS	(B) S	(C) N	(D) I	(E) MI

Fonte: autores do artigo

É importante ressaltar que o FCI é, até hoje, um dos instrumentos mais utilizado e testado para avaliar as concepções Newtonianas sobre o movimento (HESTENES, WELLS e SWACKHAMER, 1992; MCDERMOTT e REDISH, 1999; LUANGRATH e PETTERSSON, 2007; LASRY, ROSENFELD, *et al.*, 2011). No entanto, sua aplicação sem a segunda camada não permite que identifiquemos se ao marcar a alternativa correta o estudante utilizou o raciocínio correto e nem diferenciar o raciocínio equivocado (como a utilização de uma concepção intuitiva não científica, por exemplo) da falta de conhecimento (ERYILMAZ, 2010).

O item escolhido para exemplificar a construção de um item a partir de um teste consolidado (Figura 2) aborda a Terceira Lei de Newton (Ação e Reação). A segunda camada apresenta explicações que permitem a identificação de concepções espontâneas dos estudantes tais como a de que o corpo que tem maior massa ou que “realiza a ação” exerce mais força sobre o outro.

Interpretação das respostas

Itens em três camadas permitem várias combinações de respostas por meio das quais é possível constatar a coerência do raciocínio do estudante. Considerando a primeira e a segunda camada são possíveis as seguintes situações (Quadro1):

Quadro 1 – Possibilidades de combinações de respostas entre a 1ª e 2ª camadas do teste

Combinação	1	2	3	4
1ª Camada	X	✓	X	✓
2ª Camada	X	X	✓	✓

A partir da relação entre a primeira e a segunda camada pode-se adicionar a terceira que versa sobre o grau de assertividade do estudante em suas respostas anteriores. Neste trabalho analisaremos duas possíveis combinações de respostas para essa camada – Seguro (para aqueles estudantes que marcaram a alternativa a e b) e Inseguro (para aqueles que marcaram d ou e).

A interpretação dessas combinações de respostas varia a depender de como as questões foram elaboradas. A seguir discorreremos acerca das interpretações que podem ser realizadas a respeito do entendimento dos estudantes, a depender de suas combinações de respostas para cada uma das perspectivas de construção de itens em três camadas abordadas neste trabalho.

1 – Itens baseados na teoria do perfil epistemológico

A questão apresentada como exemplo neste trabalho se refere à zona do realismo ingênuo (RI), sua função é evidenciar se o estudante iniciou a construção de seu perfil epistemológico para o conceito e se consegue discernir contextos para seu uso. Um acerto nesta questão, tanto na primeira quanto na segunda camadas (combinação 4) com segurança, pode ser interpretado como o início da construção de seu perfil epistemológico do conceito de densidade. O quadro a seguir (Quadro 2) sintetiza as possíveis interpretações de respostas.

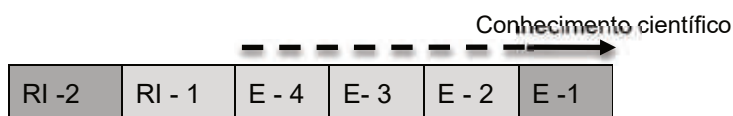
Quadro 2 – Interpretação das respostas dos itens baseados na teoria do perfil epistemológico

Combinação	Assertividade	Interpretação
1	Seguro	O estudante ainda não iniciou a construção de seu perfil epistemológico, opera na zona do realismo ingênuo (RI2)
	Inseguro	O estudante pode ter iniciado a construção de seu conhecimento sobre o conceito científico de densidade, visto que não está certo quanto a adequação de sua resposta. (RI1)
2	Seguro	Conhecimento em construção (E4). Apesar do estudante apresentar um entendimento que ultrapassa o realismo ingênuo, ele ainda não consegue justificar de forma adequada a sua resposta e não percebe isso.
	Inseguro	Conhecimento em construção (E3). Apesar do estudante apresentar um entendimento que ultrapassa o realismo ingênuo, ele ainda não consegue justificar de forma adequada a sua resposta, mas dispõe de algum nível de percepção quanto a isso.
3	Seguro	Conhecimento em construção (E4). Entretanto essa combinação de resposta é pouco esperada, sendo atribuída ao acerto ao acaso.
	Inseguro	Conhecimento em construção (E3). O estudante percebe de que está em dúvida quanto ao conhecimento.

4	Seguro	O estudante iniciou a construção de seu perfil epistemológico para o conceito de densidade (E1).
	Inseguro	Conhecimento em construção (E2). O estudante apresenta certo domínio do conceito, entretanto ainda não está seguro quanto a aplicação de seu conhecimento.

Para interpretação das combinações de respostas possíveis para essa questão adotamos a seguinte hierarquia:

Figura 3 – Hierarquia das possíveis interpretações



Essa hierarquia se baseia em dois extremos, nos quais estão situadas duas zonas do perfil epistemológico – RI -realismo ingênuo e E -empirismo. Entre essas duas zonas existem diferentes gradações que versam sobre o grau de conhecimento em construção dos estudantes.

2 – Itens adaptados de testes consolidados

No quadro a seguir apresentamos tanto as possíveis interpretações para as marcações da primeira e segunda camadas, quanto a articulação do conhecimento relacionada ao grau de confiança do estudante ao marcá-las.

Quadro 3 – Interpretação das respostas dos itens construídos a partir de testes consolidados

Combinação	Interpretação	Assertividade	Nível de Articulação
1	O estudante apresenta uma concepção não científica (concepção ingênua) do fenômeno apresentado, possuindo coerência de raciocínio.	Seguro	Conhecimento mais articulado
		Inseguro	Conhecimento menos articulado
2	O estudante possui algum conhecimento científico, no entanto não apresenta coerência de raciocínio	Seguro	Conhecimento mais articulado
		Inseguro	Conhecimento menos articulado
3	O estudante não possui conhecimento científico nem coerência de raciocínio, acertando ao acaso	Seguro	Não há articulação de conhecimento
		Inseguro	
4	O estudante apresenta um conhecimento científico sobre o tema inquirido, havendo coerência de raciocínio.	Seguro	Conhecimento mais articulado
		Inseguro	Conhecimento menos articulado

Essa interpretação nos permite: i) identificar quando os estudantes dão uma resposta cientificamente correta a uma determinada questão, utilizando para isso um raciocínio intuitivo e não científico ; ii) constatar que errar a resposta da questão não significa necessariamente possuir um raciocínio não científico (PESMAN e ERYILMAZ, 2010); iii) avaliar em que medida o estudante apresenta um conhecimento mais explicativo e em que medida ele apresenta um raciocínio causal. Por fim, ainda é possível diferenciar o uso de uma concepção intuitiva para escolha da resposta de uma situação de falta de conhecimento.

Considerações Finais

O uso de testes em três camadas em pesquisas da área de ensino traz vantagens tanto em relação aos testes de múltipla escolha, quanto aos itens discursivos. Testes de múltipla escolha, apesar de poderem ser aplicados facilmente para amostras grandes, não acessam o raciocínio dos estudantes; itens discursivos, apesar de permitirem o acesso à dimensão explicativa, apresentam um carácter subjetivo e demandam muito tempo para correção. Nesse cenário, testes em três camadas se apresentam como instrumentos em potencial para superar algumas das limitações desses dois formatos de testes.

Ademais, os testes em três camadas podem contribuir significativamente para um processo de avaliação formativo, pois apresentam uma possibilidade de diagnóstico mais detalhado em relação aos obstáculos do entendimento, ao nível de articulação do raciocínio e à metacognição. Ao apontar a justificativa para a sua resposta, o aluno permite ao professor identificar possíveis falhas de raciocínio conceitual que possam ter perpetuado no decorrer do processo de ensino, assim como fornecer informações importantes sobre o grau de coerência de seu pensamento acerca do fenômeno. Além disso, ao informar o grau de assertividade com o qual respondeu à questão, o estudante é obrigado a fazer uma metarreflexão sobre seu próprio conhecimento.

Neste trabalho apresentamos duas perspectivas para o uso do teste em três camadas. Consideramos que esse instrumento apresenta muito potencial para ampliar as informações acerca da aprendizagem dos estudantes por meio de testes de conhecimento, se constituindo como uma ferramenta relevante não só para a pesquisa, mas também para avaliação docente. Nessa última perspectiva, tais ferramentas possibilitam que os professores adequem suas aulas e metodologias para sanar possíveis concepções errôneas. Ademais, a construção e validação de testes em camadas pode contribuir fortemente para dar suporte à construção de uma base de instrumentos metodológicos, coerentes e objetivos, que possam subsidiar toda forma de acesso e avaliação de conhecimentos científicos de estudantes, em diferentes níveis de escolaridade. Essa certamente é uma contribuição importante para a pesquisa na área, pois possibilita interpretações mais robustas sobre os fenômenos educacionais, pautadas em dados mais fidedignos obtidos por instrumentos validados.

Referências

- ARSLAN, H. O.; CIGDEMOGLU, C.; MOSELEY, C. A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. **International Journal of Science Education**, v. 34, n. 11, p. 1667-1686, July 2012.
- BACHELARD, G. A. Filosofia do Não.. In: _____ **Os Pensadores**. São Paulo: Abril Cultural, 1979. p. 01-87.
- BOND, T. G.; FOX, C. M. **Applying the Rasch Model: Fundamental Measurement in the Human Sciences**. 2ª. ed. Mahwah (NJ): Lawrence Erlbaum Associates Publishers, 2007.
- CALEON, I.; SUBRAMANIAM, R. Development and Application of a Three-Tier Diagnostic Test to Assess Secondary Student's Understanding of Waves. **International Journal of Science Education**, v. 32, n.7, p. 939-961, 2010. ISSN DOI: 10.1080/09500690902890130.
- CETIN-DINDAR, A.; GEBAN, O. Development of a three-tier test to assess high school students' understanding of acids and bases. **Procedia - Social and Behavioral Sciences**, Istanbul (Turkey), v. 15, p. 600-604, 2011.

COMMONS, M. L.; PEKKER, A. Presenting the formal theory of hierarchical complexity. **World Futures: The Journal of New Paradigm Research**, p. 375-382, 2008. ISSN <https://doi.org/10.1080/02604020802301204>.

ERYILMAZ, A. Development and application of three-tier heat and temperature test: Sample of bachelor and graduate students. **Eurasian Journal of Educational Research**, Eğitim Arastirmalari, v. 40, p. 53-76, 2010.

FERNANDES, S.; TALIM, S. L. **Tradução e Validação do Teste "Force Concept Inventory"**. Anais do XVIII Simpósio Nacional de Física. Vitória: [s.n.]. 2009.

FISCHER, K. W. Dynamic cycles of cognitive and brain development: measuring growth in mind, brain and education. In: BATTRO, A. M.; FISCHER, K. W.; LÉNA, P. J. **The Educated Brain: Essays in Neuroeducation**. [S.l.]: Cambridge University Press, 2008. p. 127-150.

GOLAFSHANI, N. Understanding Reliability and Validity in Qualitative Research. **The Qualitative Report**, Toronto, v. 8, n. 4, p. 597-607, dezembro 2003.

GOLINO, H. F. et al. **Psicometria Contemporânea: Compreendendo os Modelos Rasch**. São Paulo: Casapsi Livraria e Editora Ltda, 2015.

HESTENES, D.; WELLS, M.; SWACKHAMER. Force Concept Inventory. **Physics Teacher**, v. 30, 1992.

KIRBULUT, ; GEBAN,. Using Three-Tier Diagnostic Test to Assess Students' Misconceptions of States of Matter. **Eurasia Journal of Mathematics, Science & Technology Education**, Sanliurfa (TURKEY), v. 10, p. 509-521, July 2014. ISSN doi: 10.12973/eurasia.2014.1128a.

LASRY, N. et al. The puzzling reliability of the Force Concept Inventory. **American Journal of Physics**, v. 79, n.9, p. 908-912, September 2011. ISSN DOI: 10.1119/1.3602073.

LUANGRATH, P.; PETTERSSON, S. Adapting the force Concept Inventory to Lao Context (in Lao Language). **Scientific Journal of National University of Laos**, Laos, 2007. 117-128.

MCDERMOTT, L. C.; REDISH, E. F. Resource Letter: PER-1: Physics Education Research, v. 67, n. 9, p. 754-767, September 1999. ISSN doi: 10.1119/1.19122.

PESMAN, H.; ERYILMAZ, A. Development of a Three-Tier Test to Assess Misconceptions About Simple Electric Circuits. **The Journal of Educational Research**, v. 103, p. 208–222, 2010. ISSN ISSN: 0022-0671.

RAYMUNDO, V. P. Construção e validação de instrumentos: um desafio para a psicolinguística. **Letras de hoje**, Porto Alegre, v. 44, n. 3, p. 86-93, Julho/Setembro 2009.

TAN, K. C. D. et al. Development and Application of a Two-Tier Multiple Choice Diagnostic Instrument to Assess High School Students' Understanding of Inorganic Chemistry Qualitative Analysis. **Journal of Research in Science Teaching**, v. 39, n.4, p. 283-301, 2002.

TREAGUST, D. F. Development and use of diagnostic tests to evaluate students' misconceptions in science. **International Journal of Science Education**, v. 10, n. 2, p. 159-169, 1998. ISSN DOI: 10.1080/0950069880100204.

TREAGUST, D. F. **Diagnostic assessment in science as a means to improving teaching, learning and retention**. UniServe Science Assessment Symposium Proceedings. Bentley (Austrália): [s.n.]. 2012. p. 1-9.