

Instrumento para validação de intervenções educacionais nas pesquisas em ensino de Ciências

Instrument for validation of educational interventions in research in science education

Lídia Cabral Moreira¹; Ana Paula Miranda Guimarães²; Amanda Amantes³

Universidade Federal da Bahia^{1,3}; Instituto Federal da Bahia;
proflidiacabral@gmail.com; apmguima@gmail.com;
amandaamantes@gmail.com

RESUMO

O objetivo do presente estudo é analisar a validação por juízes de uma intervenção educacional e discutir a importância da validação de intervenções nas pesquisas em ensino de Ciências. Inicialmente, planejamos e construímos uma intervenção para a abordagem de sistema nervoso contextualizado ao tema “drogas”, direcionada a estudantes do Ensino Médio. Posteriormente, elaboramos um barema para validação pareada da intervenção por juízes, a fim de avaliar se o desenho das aulas estava coerente com os objetivos de ensino e de desenvolvimento propostos. A validação da intervenção foi analisada de forma descritiva e a partir do coeficiente Kappa. A análise dos dados revela que a intervenção foi validada. Os procedimentos metodológicos apresentados nesse estudo, voltados para a validação por juízes de uma intervenção são importantes para conferir maior adequação e coerência à investigação e pode contribuir no sentido de se discutir com mais propriedade a eficácia das intervenções no contexto escolar.

Palavras-chave: Validação; Intervenção educacional; Pesquisa em ensino.

ABSTRACT

The objective of the present study is to analyze the validation by judges of an educational intervention and to discuss the importance of validation of interventions in research in teaching science. Initially, we planned and constructed an intervention for the approach of nervous system contextualized to the theme "drugs", directed to high school students. Subsequently, we elaborated a barema for paired validation of the intervention by judges, in order to evaluate if the design of the classes was coherent with the proposed teaching and development objectives. The validation of the intervention was analyzed in a descriptive way and from the Kappa coefficient. Data analysis reveals that the intervention has been validated. The etodological procedures presented in this study, aimed at the validation by judges of an intervention, are important to give greater adequacy and coherence to the research and can contribute to discuss more effectively the effectiveness of interventions in the school context.

Keywords: Validation; Educational intervention; Research in education.

Introdução

O debate sobre o rigor nas pesquisas educacionais, principalmente na área de Ciências, tem se intensificado (ANDRÉ, 2001; LONG; JOHNSON, 2000; SANTOS; GRECA, 2013) e um dos principais focos de discussão recai na questão sobre a coerência interna. Esse aspecto vem sendo pontuado como uma fragilidade que a área deve superar, principalmente, em relação à adequação das questões de pesquisa aos métodos de coleta, instrumentos e análise de dados (SHAFFER; SERLIN, 2004; GOLAFSHANI, 2003).

Em se tratando do aspecto metodológico, há um movimento crescente na área, principalmente no âmbito internacional, no sentido de conferir maior validade e confiabilidade aos instrumentos que são utilizados para coletar dados (CHO; TRENT, 2006; GOLAFSHANI, 2003; GORARD, 2002), pois a adequação de tais instrumentos, em relação ao referencial teórico adotado e aos objetivos da pesquisa, é essencial para garantir inferências e interpretações seguras sobre o nosso objeto de estudo. Santos e Greca (2013), ao realizarem um levantamento das pesquisas da área de ensino de Ciências na América Latina, no período de 2000 a 2009, apontaram que as pesquisas nessa área têm se preocupado mais com análises descritivas e explicações de fenômenos educacionais do que com análises sobre a eficácia e validade de tais avaliações. Não obstante a importância desse tipo de investigação, também é fundamental que as pesquisas apontem e discutam sobre os critérios de validade adotados nos estudos, para que possam gerar resultados consistentes e interpretações confiáveis.

A validade é um indicador de adequação de um instrumento de pesquisa ou de uma investigação científica e está associada com a precisão de determinada medida, ou seja, o grau em que um instrumento de medida realmente mede a variável que se propõe mensurar. No que se refere aos instrumentos, podemos interpretar a validade de diversas formas em termos do acesso ao atributo que se deseja avaliar. Portanto, pode-se interpretar a validade de instrumentos de coleta de dados (um questionário ou um teste, por exemplo). Nesse caso, a validade diz respeito à adequação dos instrumentos de coleta de dados em relação aos objetivos aos quais se propõe (BORSBSOOM; MELLENBERGH; HEERDEN, 2004; OLLAIK; ZILLER, 2012).

Também podemos interpretar a validade de uma teoria de ensino ou de aprendizagem, a partir de estudos de intervenções educacionais, como discutido por Plomp (2009) e interpretar a validade dessas intervenções. Esse tipo de validação pode ser feita *a priori* (no caso deste estudo) ou a partir dos resultados da intervenção no contexto educacional. Nesses casos, a validade está relacionada ao alcance dos objetivos, seja de aprendizagem ou de ensino, a que a intervenção se propõe. Nesses termos, questões relacionadas à validade das intervenções demandam preocupação com as metodologias e referenciais teóricos utilizados no desenvolvimento da proposta e na análise de seus resultados, de maneira a garantir que a intervenção possua características para alcançar os objetivos de ensino e de aprendizagem esperados com a sua aplicação.

Contexto da pesquisa e a intervenção

Este estudo faz parte de uma pesquisa de Mestrado que teve por objetivo investigar características de uma intervenção educacional, relacionadas ao ensino e à aprendizagem de sistema nervoso contextualizado ao tema “drogas”, direcionada a estudantes do Ensino Médio. Inicialmente construímos e validamos a intervenção por juízes (Professores de

Biologia da rede pública de Educação Básica do Recôncavo e capital da Bahia). No segundo momento, analisamos os resultados de sua aplicação a partir do contexto de ensino e do desempenho dos alunos nos testes de conhecimentos e nas atividades realizadas em sala de aula. A intervenção foi aplicada a 25 alunos do 3º ano do Ensino Médio integrado ao técnico do Instituto Federal da Bahia, campus Camaçari, entre março e abril de 2015.

O referencial teórico e metodológico para o planejamento e avaliação da intervenção foi baseado na Metodologia da Problematização (MP), a partir do Arco de Charles Maguerez (BERBEL, 1999) e na *design research*, a partir da construção de princípios de planejamento (*design*). Para o planejamento da intervenção elaboramos alguns princípios de *design* que de forma resumida são apresentados: (1) Abordagem de ensino pautada na MP a partir da aplicação das etapas do Arco de Charles Maguerez; (2) Discussão acerca do sistema nervoso a partir de textos de divulgação científica; (3) Utilização de mídia audiovisual acerca de diferentes aspectos relacionados às drogas; (4) Enfoque no trabalho coletivo e cooperativo de aprendizagem.

Neste estudo apresentamos uma discussão sobre a validação da intervenção por juízes, antes de sua aplicação em sala de aula. O objetivo do presente estudo é analisar a validação por juízes de uma intervenção educacional para a área de Biologia e discutir a importância da validação de intervenções educacionais nas pesquisas em ensino de Ciências. A nossa aposta é de que os procedimentos metodológicos apresentados nesse estudo, voltados para a validação por juízes do primeiro protótipo da intervenção, são importantes para conferir maior adequação e coerência à investigação e pode contribuir no sentido de se discutir com mais propriedade a eficácia das intervenções no contexto escolar.

Procedimentos metodológicos

O instrumento de validação da intervenção

Cada aula e tarefas dos estudantes foram desenhadas de modo a contemplar as características da MP. Para cada semana de aula foi elaborada uma dinâmica específica com tarefas que tinham objetivos de naturezas diferentes – objetivos de ensino e de desenvolvimento do aluno. Assim, foi elaborado um barema de validação para analisar a adequação da intervenção, a partir da MP em termos de seus objetivos, ou seja, avaliar se a intervenção proposta tinha potencial para atingir tais objetivos estabelecidos.

O barema de validação foi distribuído a vinte e quatro professores de Biologia na capital e região do Recôncavo da Bahia, mas tivemos o retorno apenas de dez baremas avaliados, sendo um descartado por não atender ao que foi solicitado (as marcações foram indevidamente realizadas, demonstrando que o juiz pode não ter compreendido a estrutura do barema).

Esses juízes são professores atuantes de Biologia e pesquisadores da área de ensino de Biologia das redes estaduais e federais de ensino básico ou universidades de diferentes cidades da Bahia. O barema continha a seguinte estruturação: (1) uma descrição inicial, apresentando os objetivos de pesquisa; (2) informação ao professor de seu livre arbítrio em contribuir com os dados da pesquisa; (3) uma descrição de cada uma das etapas do Arco da MP; (4) uma orientação de como se deveria proceder a avaliação de cada etapa. Para cada natureza de objetivo foram oferecidas três opções, na qual apenas uma consistia no objetivo pretendido pelos pesquisadores a ser atingido em cada aula. Na coluna de sugestões os avaliadores poderiam informar quando não julgavam nenhum dos objetivos propostos apropriados, bem como sugerir mudanças na redação dos objetivos, mudança ou alterações na

dinâmica das aulas ou sugerir outras atividades. No quadro 1 pode ser observado o exemplo de como o barema foi construído.

Quadro 1. Exemplo de uma aula no barema de validação apresentado aos professores de Biologia para avaliação da intervenção educacional construída.

Aulas	Tarefa dos alunos	Das alternativas abaixo, qual você julga que mais se aproxima do objetivo dessa atividade?	Das alternativas abaixo, qual você julga que mais se aproxima do objetivo de desenvolvimento?	Comentário/sugestões
01: Socialização e problematização do tema: “Qual a relação do uso de drogas com o sistema nervoso?”, a partir dos questionamentos e da observação dos estudantes.	Exposição mural e discussão oral do que foi observado e registrado do cotidiano. (individual).	<p>a) Levantar os conhecimentos prévios dos alunos sobre o tema “Sistema nervoso”.</p> <p>b) Proporcionar maior contato com as inovações científicas na área de estudo sobre drogas.</p> <p>c) Fazer com que os alunos identifiquem elementos do cotidiano que se relacionem com Sistema Nervoso.</p>	<p>a) Estimular o desenvolvimento da autonomia para o estudo de questões relevantes do cotidiano.</p> <p>b) Despertar o interesse nos estudantes para o estudo de assuntos biológicos.</p> <p>c) Favorecer a aprendizagem de conceitos sobre drogas e seus danos ao sistema nervoso.</p>	

Fonte: Elaborado pelos autores.

Nesse barema, os juízes analisaram a dinâmica das aulas, a tarefa dos alunos e dentre os objetivos propostos julgaram qual o mais coerente com a atividade (Objetivo de ensino) e com o objetivo docente (Objetivo de desenvolvimento). No objetivo da atividade a avaliação deveria ser feita levando em consideração o que se pretende mobilizar no estudante, em termos de aprendizagem de conteúdos, em determinada aula. Ao mesmo tempo, no objetivo de desenvolvimento, deveriam avaliar o que o docente pretende que o aluno desenvolva com a atividade proposta, nos termos de habilidades, a partir do que é orientado nos documentos oficiais de ensino nacionais, por exemplo, os Parâmetros Curriculares Nacionais (BRASIL, 2000).

Metodologia de análise dos baremas

Como forma de determinar a validade da intervenção, os dados coletados através dos baremas foram tratados a partir de análise quantitativa e qualitativa. Inicialmente, procedeu-se a análise descritiva a fim de identificar como ocorria a concordância entre os juízes na marcação das alternativas dos objetivos estabelecidos no instrumento de validação. Isso foi feito pela análise de frequências de marcações, procedimento utilizado para dados de natureza dicotômica, que gera um percentual denominado IPC (Índice Percentual de Concordância).

Na análise quantitativa, utilizamos procedimentos estatísticos como forma de avaliar a validade da intervenção em relação ao que ela se propõe em termos de ensino e desenvolvimento. Existem vários enfoques estatísticos que podem ser utilizados para se determinar a confiabilidade entre avaliadores, e a escolha do método depende do tipo de sistema de classificação que será empregado e da natureza dos dados (PERROCA; GAIDZINSKI, 2003). Sendo assim, utilizamos o parâmetro de concordância do coeficiente Kappa (k) proposto por Cohen (1960). Esse coeficiente é um dos mais utilizados para

classificar dados em categorias nominais (FONSECA; SILVA; SILVA, 2007) e adequado para o nosso objetivo de avaliar a validade de uma intervenção na perspectiva de juízes. O coeficiente k pode ser definido como uma medida de associação usada para descrever e testar o grau de concordância (confiabilidade e precisão) numa determinada classificação (PERROCA; GAIDZINSKI, 2003). Esta medida de concordância assume valor máximo igual a 1, que representa total concordância ou ainda pode assumir valores próximos e até abaixo de 0, os quais indicam nenhuma concordância (ver Tabela 1).

Valor do Kappa	Nível de concordância
<0	Nenhuma concordância
0 – 0.19	Concordância fraca
0.20-0.39	Concordância razoável
0.40-0.59	Concordância moderada
0.60-0.79	Concordância substancial
0.80-1.00	Concordância quase perfeita

Tabela 1. Valores de coeficiente k e seus níveis de concordância. Fonte: Landis e Koch (1977).

Se o índice de concordância assumir coeficientes entre 0.4 e 1 pode-se interpretar que a intervenção está validada (LANDIS; KOCH, 1977). Para obter essas medidas, foi utilizado o programa *SPSS (Statistical Package for the Social Sciences)* versão 20.

Resultados e Discussões

Todas as análises foram feitas sob as perspectivas das duas naturezas distintas de objetivos estabelecidos para a intervenção (objetivos de ensino e objetivo de desenvolvimento). No Gráfico 1 apresenta-se os dados sobre a concordância entre os juízes com os objetivos de ensino e o Gráfico 2 para concordância com os objetivos de desenvolvimento. Os números na vertical representam o total de juízes que avaliaram a intervenção, os números na horizontal dizem respeito às aulas elaboradas, sendo que para a terceira aula foram destinadas duas semanas de aula (4h/aulas), que correspondem aos números 3 e 4 nos gráficos. As letras correspondem às alternativas dos objetivos, contendo uma única opção correspondente ao objetivo pretendido pelas pesquisadoras.

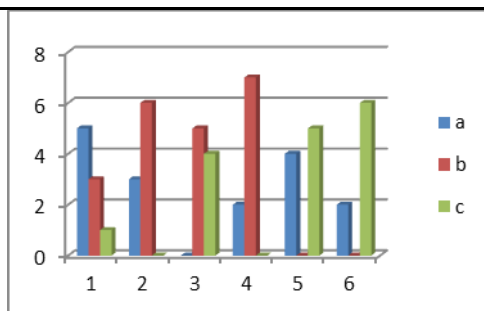


Gráfico 1 - Dados de concordância entre os juizes com relação aos objetivos de ensino. **Gráfico 2** - Dados de concordância entre os juizes com relação aos objetivos de desenvolvimento do aluno. Fonte: Dados da pesquisa.

A análise do gráfico 1 permite verificar que para a aula 1, sete professores marcaram alternativa *c* e 2 a alternativa *a*. Para a aula 2, houve 100% de concordância com relação a marcação da alternativa *a*. No entanto, observa-se uma dispersão para o objetivo da atividade da aula 3, com 5 professores que marcaram a opção *b* e 4 que optaram pela letra *c*, não havendo, portanto, muita concordância entre os juizes para o objetivo desta aula. Isso sugere que é preciso rever a configuração da aula ou a clareza do objetivo estabelecido. Para as aulas 4, 5 e 6 observa-se que houve pouca dispersão na marcação das alternativas, o que é um indicativo de alto nível de concordância entre os professores.

De maneira geral, houve alta concordância entre os professores sobre os objetivos da atividade. Esses objetivos estão relacionados aos objetivos de ensino docente, ou seja, o que o professor pretende, em determinada aula e com determinada atividade, em termos de aprendizagem de conteúdos. Esses são objetivos geralmente estabelecidos pela coordenação escolar no plano de ensino anual, ou pelo próprio professor no seu planejamento de conteúdos da unidade. Eles são orientados, usualmente, por parâmetros bem direcionados, uma vez que contemplam conteúdos determinados pela grade curricular e se remetem a uma aceção consensual sobre o que deve ser ensinado. Acreditamos que esse fato torna mais fácil estabelecer consenso sobre essa natureza de objetivo, os quais exigem critérios mais específicos de objetividade na sua elaboração e com pouca interferência da subjetividade do professor, uma vez que estão subsidiados pelo parâmetro do conteúdo.

Em contrapartida, na interpretação do gráfico 2 observa-se uma maior dispersão para os dados da concordância para os objetivos de desenvolvimento. Pelo gráfico 2 verificamos que para o objetivo de desenvolvimento, os juizes não apresentaram alta concordância, como aconteceu para os objetivos de ensino. Por exemplo, para os objetivos da aula 1, houve marcação para as três opções de objetivos. Também na aula 3 não houve concordância entre eles para o objetivo de desenvolvimento. Na avaliação da aula 5, em que houve 90% de concordância entre os juizes para o objetivo de ensino, percebe-se que eles tomaram decisões divergentes para os objetivos de desenvolvimento, sendo que 4 optaram pelo objetivo descrito na letra *a* e 5 optaram pela letra *c*.

Os objetivos de desenvolvimento estão sendo entendidos neste estudo como as finalidades da aprendizagem, ou seja, o que o docente pretende que o estudante desenvolva em termos de habilidades e competências com a tarefa de ensino. Dito de outra maneira, é saber identificar quais características uma tarefa deve ter para mobilizar o desenvolvimento de habilidades, da autonomia e do pensamento crítico. Essas finalidades da aprendizagem também são estabelecidas nos documentos oficiais de educação no Brasil, como nos Parâmetros Curriculares Nacionais (BRASIL, 2000).

Essa discordância entre os juízes em relação a um dos objetivos da intervenção nos dá indícios de que eles apresentam pouca clareza (em relação à sua definição) para os professores em geral, pois não há um direcionamento específico para definição sobre os objetivos de desenvolvimento, o que gera interpretações divergentes sobre eles. Os objetivos de desenvolvimento são definidos de forma mais genérica pelos documentos oficiais, dando maior margem para interpretá-los de diferentes maneiras, tornando seu entendimento mais difuso. Dessa forma, sua interpretação não é tão consensual, e é intrínseca ao entendimento do professor do que representa, por exemplo, “senso crítico”; isso torna essa avaliação mais subjetiva quando comparada à avaliação de objetivos de aprendizagem, portanto, susceptível a maiores discordâncias.

Esse resultado é um forte indicador de que devemos nos preocupar com a clareza dos critérios que definem habilidades e competências no âmbito do ensino de Ciências para que as intervenções educacionais e as abordagens de ensino atinjam os objetivos de aprendizagem e de desenvolvimento do aluno.

Coeficiente k

Para interpretar a concordância dos juízes com o barema realizamos a análise de duas formas: primeiro, analisando a concordância com os objetivos de ensino e de desenvolvimento separadamente e depois com o barema de modo geral, para então, determinar a validade da intervenção.

Na tabela 2 apresentam-se os dados da análise de k para a concordância dos juízes com os objetivos da atividade e de desenvolvimento.

	Objetivo da atividade	Objetivo de desenvolvimento
	Valor de k	Valor de k
Prof. 1	1,000	,250
Prof. 2	,429	-,250
Prof. 3	,308	,063
Prof. 4	,700	,500
Prof. 5	,520	,250
Prof. 6	1,000	,500
Prof. 7	1,000	1,000
Prof. 8	,739	,500
Prof. 9	,739	,750

Tabela 2. Valores de coeficiente de concordância de cada juiz com os objetivos da atividade e de desenvolvimento. Fonte: Elaborado pelos autores a partir do programa SPSS.

Observa-se, a partir da Tabela 2 que para os objetivos da atividade houve três concordâncias quase perfeitas (prof. 1, 6, 7), três concordâncias substanciais (prof. 4, 8, 9), duas moderadas (prof. 2 e 5) e uma concordância razoável (prof. 3). Sendo assim, o instrumento tem bons

índices de concordância, o que nos leva a interpretar que ele está validado para esta natureza de objetivo, pois os níveis de concordância ficaram entre quase perfeitas, substanciais e moderadas, não se obtendo nenhum nível de concordância fraca. Com essa medida pode-se inferir, a partir da análise dos juízes, que a intervenção está adequada do ponto de vista dos objetivos estipulados para as atividades de ensino propostas.

Com relação aos objetivos de desenvolvimento nota-se uma dispersão maior, explicada pelo baixo índice de concordância do prof. 2, com um valor k abaixo de 0, aquém do esperado; ou seja, para este juiz não houve concordância com o barema em termos de objetivo de desenvolvimento. Já para o objetivo da atividade, este mesmo juiz obteve uma concordância moderada ($k= 0,43$). Isso significa que para este professor a intervenção está adequada e válida para alcançar os objetivos de ensino, porém, não está adequada do ponto de vista do que se pretende que o aluno desenvolva. A mesma dispersão ocorre para o prof. 1 que para o objetivo da atividade obteve coeficiente de concordância ($k= 1,00$), enquanto para os objetivos de desenvolvimento obteve coeficiente de concordância de apenas ($k= 0,25$). O prof. 6 também obteve concordância quase perfeita para a primeira natureza de objetivo e concordância moderada ($k= 0,50$) com os objetivos de desenvolvimento. Esses resultados corroboram a análise descritiva, na qual foi encontrada maior concordância com os objetivos de ensino do que com os objetivos de desenvolvimento.

A partir dessa análise, os resultados sugerem que é preciso rever a dinâmica das aulas e as tarefas elaboradas, de forma que os objetivos de desenvolvimento colocados sejam de fato alcançados, também é preciso rever a redação desses objetivos, de modo a possuírem maior clareza para o professor que avalia.

Quando analisamos a concordância dos juízes com o barema de modo geral, considerando as duas naturezas de objetivos propostos (Tabela 3), obtivemos uma concordância quase perfeita, uma concordância fraca e uma razoável, e as demais concordâncias moderadas e substanciais. Obteve-se uma média de ($k= 0,56$) de concordância dos juízes com o barema, o que de acordo com a teoria de coeficiente k de Cohen (1960) quer dizer uma concordância moderada, o que indica validade da intervenção.

Concordância geral		
	Valor k	Índice. Sig.
Prof.1	,613	,003
Prof.2	,097	,636
Prof.3	,205	,284
Prof. 4	,596	,003
Prof. 5	,400	,023
Prof. 6	,753	,000
Prof.7	1,000	,000
Prof. 8	,621	,002
Prof. 9	,745	,000
Média	,559	
SD	0,266	

Tabela 3. Valores de coeficiente de concordância de cada juiz com o barema em geral (considerando as duas naturezas de objetivos propostas). Fonte: Elaborado pelos autores a partir do programa SPSS.

A medida obteve um desvio padrão alto ($SD= 0,27$) que pode ser explicado pelo baixo índice de concordância do prof. 2, que teve uma concordância fraca com o barema, ($k=0,097$). Este

juiz também obteve coeficiente de concordância menor que 0 para os objetivos de desenvolvimento.

Acreditamos que a explicação para a alta dispersão na concordância com os objetivos de desenvolvimento é devido à natureza difusa da aceção dos professores sobre o que significa desenvolver habilidades e competências dos estudantes a partir da aprendizagem de conteúdos científicos. Embora a subjetividade seja inerente a qualquer tipo de avaliação, a alta dispersão para com os objetivos de desenvolvimento também pode ser explicada por este ponto de vista, pois, assim como é difuso o entendimento do que seja desenvolvimento do aluno, também são diversos os aportes teóricos sobre ensino e aprendizagem que embasam a prática pedagógica dos professores. Esse embasamento teórico orienta como o professor concebe sua prática de ensino e como ele definirá os objetivos desse processo. No entanto, não descartamos a hipótese de falta de clareza do instrumento de validação, no que diz respeito à redação dos objetivos e ao desenho das aulas, o que sugere que é preciso rever esses elementos do barema e submetê-lo a uma nova validação por juízes.

Acreditamos que se não elaborarmos intervenções com potencial para atingir os objetivos de ensino e de desenvolvimento, também não garantimos um bom desempenho escolar dos estudantes. O desafio do ensino é sempre refletir como alcançar os objetivos de aprendizagem dos conteúdos científicos e os objetivos relacionados a determinadas competências e habilidades que pretendemos desenvolver nos estudantes, além de criar estratégias para atingi-los. O professor desempenha papel fundamental na elaboração de atividades de ensino, pois é por meio delas que o aluno poderá construir conhecimento sobre os processos científicos, desenvolver autonomia e habilidades que lhe permitam, sobretudo, refletir e atuar na sociedade de forma ativa. Por isso, a validação de intervenções se configura como uma prática importante nas pesquisas em ensino de ciências no contexto escolar, pois pode conferir maior rigor a investigação.

Considerações Finais

Consideramos que, dada às limitações da pesquisa (tais como pouco número de juízes, primeira versão do barema e poucos dados para fazer uma análise estatística mais detalhada), a intervenção foi validada. Contudo, é preciso fazer uma ressalva em relação aos objetivos de desenvolvimento, que obtivemos maior dispersão, o que pode significar duas coisas: ou nosso instrumento não conseguiu dimensionar bem esses objetivos na primeira aplicação do barema ou não há realmente um consenso entre os professores. Nos dois casos temos como apontamento a dificuldade em definir os objetivos de desenvolvimento e esse resultado demanda maior aprofundamento na investigação, uma vez que tais concepções têm um impacto direto na própria prática docente dos professores e na interpretação dos resultados de pesquisas.

Diante dos resultados obtidos, esta pesquisa se insere numa área de estudos ainda pouco explorada nas pesquisas em ensino de Ciências e pretende contribuir com os debates sobre rigor e qualidade na elaboração de intervenções educacionais, assim como iniciar as discussões, em âmbito nacional, sobre a validação por juízes de propostas de ensino antes de sua aplicação nos contextos reais de sala de aula. Esperamos que os critérios metodológicos que foram empregados neste estudo para validação de uma intervenção educacional possam ser empregados em outras pesquisas de mesma natureza. Esperamos com esta pesquisa contribuir para diminuir a lacuna existente entre a construção de intervenções educacionais e o estudo da sua validade e eficiência para o ensino.

Agradecimentos e apoios

Agradecemos à Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB), que viabilizou o suporte financeiro para o desenvolvimento desta pesquisa.

Referências

- ANDRÉ, M. Pesquisa em Educação: Buscando Rigor e Qualidade. **Cadernos de Pesquisa**, n. 113, p. 51-64, julho, 2001.
- BERBEL, N. A. N. (Org.). **Metodologia da problematização**: fundamentos e aplicações. Londrina: EDUEL, 1999.
- BORSBOOM, D.; MELLENBERGH, G. J.; VAN HEERDEN, J. The concept of validity. **Psychological Review**, 111(4), 1061- 1071. 2004.
- BRASIL. Ministério da Educação, Secretaria de Educação Média e Tecnológica. **Parâmetros Curriculares Nacionais**: Ensino Médio. Brasília: Ministério de Educação, 2000.
- CHO, J.; TRENT, A. Validity in qualitative research revisited. **Qualitative Research**, v. 6, n. 3, p. 319-340, Aug. 2006.
- COHEN, J. A coefficient of agreement for nominal scales. **Educational and Psychological Measurement**, 20, 37-46. 1960.
- FONSECA, R.; SILVA, P.; SILVA, R. Acordo inter-juízes: O caso do coeficiente kappa. **Laboratório de Psicologia**, v. 5, n. 1, p. 81-90, 2007.
- GOLAFSHANI, N. Understanding reliability and validity in qualitative research. **The Qualitative Report**, Canadá, v. 8, n. 4, p. 597-607, 2003.
- GORARD, S. Can we overcome the methodological schism? Four models for combining qualitative and quantitative evidence. **Research Papers in Education**, v.17, n.4, p.345-361, 2002.
- LANDIS J. R.; KOCH G. G. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33:159-75.
- LONG, T; JOHNSON, M. Rigour, reliability and validity in qualitative research. **Clinical Effectiveness in Nursing**, n. 4, p. 30–37, 2000.
- OLLAIK, L. G.; ZILLER, H.M. Concepções de validade em pesquisas qualitativas. **Educação e Pesquisa**, São Paulo, v.38, n.1, 229-241, 2012.
- PERROCA, M. G.; GAIDZINSK, R. R. Avaliando a confiabilidade intervaliadores de um instrumento para classificação de pacientes - coeficiente Kappa. **Rev Esc Enferm USP**; 37(1): 72-80, 2003.
- PLOMP, T. Educational Design Research: an Introduction. In: PLOMP, T.; NIEVEEN, N. **An introduction to educational Design Research**. Enschede: SLO-Netherlands Institute for Curriculum Development. pp. 9-35. 2009.
- SANTOS, F. M. T.; GRECA, I. M. Metodologias de pesquisa no ensino de ciências na américa latina: como pesquisamos na década de 2000. **Ciência e Educação**, v. 19, n. 1, p. 15-33, 2013.
- SHAFFER, D. W.; SERLIN R. C. What Good are Statistics That don't Generalize? **Educational Researcher**, University of Wisconsin, Madison, v. 9, n. 33, p. 14-25, Dec. 2004.