

# Análise do processo de avaliação da aprendizagem de ações educativas de profissionais de enfermagem

ANALYSIS OF THE LEARNING EVALUATION PROCESS OF NURSING STAFF ACTIONS

ANÁLISIS DEL PROCESO DE EVALUACIÓN DE APRENDIZAJE DE ACCIONES EDUCATIVAS DE PROFESIONALES DE ENFERMERÍA

Vera Lucia Mira<sup>1</sup>, Marina Peduzzi<sup>2</sup>, Marta Maria Melleiro<sup>3</sup>, Daisy Maria Rizatto Tronchin<sup>4</sup>, Maria de Fátima Fernandes Prado<sup>5</sup>, Patrícia Tavares dos Santos<sup>6</sup>, Enilda Maria de Sousa Lara<sup>7</sup>, Jaqueline Alcântara Marcelino da Silva<sup>8</sup>, Jairo Eduardo Borges-Andrade<sup>9</sup>

## RESUMO

O objetivo deste estudo foi analisar o processo de avaliação da aprendizagem realizado em treinamentos ministrados à equipe de enfermagem, no que se refere à eficácia dos treinamentos e à validade dos instrumentos. Trata-se de um estudo empírico, cujos achados apoiarão formulações teóricas para a construção de uma metodologia de avaliação. A análise consistiu em avaliações da aprendizagem em seis treinamentos, totalizando 993 avaliações. Com base nos resultados dos seis treinamentos avaliados, pôde-se perceber a necessidade de proposição de uma metodologia de avaliação, construindo critérios, instrumentos e indicadores para essa finalidade. Notou-se também que é preciso aprimorar a realização do diagnóstico relativo à necessidade de treinamento e às estratégias de ensino.

## DESCRIPTORIOS

Equipe de enfermagem  
Capacitação  
Avaliação da educação  
Administração de recursos humanos em hospitais

## ABSTRACT

The objective of this study was to analyze the learning evaluation process of nursing staff training programs, regarding its effectiveness and instrument validity. This is an empirical study whose findings support theoretical formulations for the construction of an evaluation methodology. The analysis consisted of learning evaluations related to six training programs, totaling 993 evaluations. The results of these six training programs showed the need to propose an evaluation methodology, defining criteria, instruments and indicators for this purpose. Furthermore, it is important to improve the diagnosis process on training needs and teaching strategies.

## DESCRIPTORS

Nursing, team  
Training  
Educational measurement  
Personnel administration, hospital

## RESUMEN

El objetivo de este estudio fue analizar el proceso de evaluación de aprendizaje realizado en entrenamientos brindados al equipo de enfermería, en lo que se refiere a la eficiencia de los mismos y a la validez de los instrumentos. Se trata de un estudio empírico, cuyas conclusiones apoyarán formulaciones teóricas para la construcción de una metodología de evaluación. Se constituyeron en material de análisis las evaluaciones de aprendizaje relativas a seis entrenamientos, sobre un total de 993 evaluaciones. De los resultados de los seis entrenamientos evaluados fue posible percibir la necesidad de la proposición de una metodología de evaluación, construyendo criterios, instrumentos e indicadores para dicha finalidad. Se notó también que es preciso mejorar la realización del diagnóstico de necesidad de entrenamiento y las estrategias de enseñanza.

## DESCRIPTORIOS

Grupo de enfermería  
Capacitación  
Evaluación educacional  
Administración de personal en hospitales

<sup>1</sup> Professora Associada do Departamento de Orientação Profissional da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. vimirag@usp.br <sup>2</sup> Professora Associada do Departamento de Orientação Profissional da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. marinape@usp.br <sup>3</sup> Professora Associada do Departamento de Orientação Profissional da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. melleiro@usp.br <sup>4</sup> Professora Doutora do Departamento de Orientação Profissional da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. daisyrt@usp.br <sup>5</sup> Professora Doutora do Departamento de Orientação Profissional da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. fatima@usp.br <sup>6</sup> Enfermeira. Mestranda do Programa de Pós-Graduação em Gerenciamento em Enfermagem da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. patriciatavaress@yahoo.com.br <sup>7</sup> Enfermeira. Doutoranda do Programa de Pós-Graduação em Gerenciamento em Enfermagem da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. jaqueline.alc@gmail.com <sup>8</sup> Nutricionista. Doutoranda do Programa de Pós-Graduação em Gerenciamento em Enfermagem da Escola de Enfermagem da Universidade de São Paulo. São Paulo, SP, Brasil. enildalara@usp.br <sup>9</sup> Professor Titular do Instituto de Psicologia Social e do Trabalho da Universidade de Brasília. Brasília, DF, Brasil. jairo@unb.br

## INTRODUÇÃO

O desenvolvimento de ações educativas de profissionais de saúde e enfermagem é fundamental para assegurar e incrementar a qualidade da atenção à saúde e é tema de debate no contexto da política pública de saúde, no país. A reflexão construída com base nessas discussões resultou na implantação da Política Nacional de Educação Permanente em Saúde-PNEPS<sup>(1-2)</sup> que preconiza a articulação da formação e desenvolvimento dos trabalhadores ao perfil de necessidades de saúde de usuários/população e aos processos de mudança do modelo assistencial e gerencial na perspectiva da integralidade, da prática interprofissional e da efetividade da atenção à saúde.

Nesse sentido, cabe destacar que a proposta de educação permanente remete ao reconhecimento de limitações das ações educativas, tais como vêm sendo desenvolvidas, e aponta para um conjunto de mudanças a serem implementadas para a transformação das práticas de saúde, em especial, orientando-as segundo a lógica das necessidades de usuários de cada estabelecimento de saúde<sup>(3)</sup>.

Dentre as limitações, identifica-se a escassez de avaliação de atividades educativas de profissionais inseridos em serviços de saúde, que constitui um peculiar objeto de avaliação, reconhecido como necessário para o conhecimento e incremento das próprias ações educativas e de seus efeitos na qualidade da atenção à saúde<sup>(4-6)</sup>.

No Brasil, a avaliação de ações e programas educativos de trabalhadores de saúde ainda não se consolidou como uma tradição de pesquisa e constitui-se como uma temática relevante a ser investigada.

Para tanto, um longo caminho de acompanhamento das práticas e de pesquisa terá de ser percorrido, incluindo uma variedade de aspectos que permita a compreensão dos processos de levantamento e diagnóstico de necessidades e dos resultados esperados, das variáveis que afetam os resultados, do cumprimento dos objetivos de ações e programas educativos e sua efetividade para a atenção à saúde, bem como a definição de critérios e parâmetros de avaliação, de indicadores específicos de qualidade das ações e programas educativos de trabalhadores e de metodologia de avaliação.

O modelo de avaliação de ações educativas de profissionais pioneiro e consagrado internacionalmente é o de Donald Kirkpatrick que, nos anos 1950, publicou os artigos em que estabelecia quatro níveis de avaliação: avaliação de reação – medida da forma como o participante sentiu sua formação e suas reações pessoais frente à experiência de aprendizado; avaliação de aprendizagem – medida do aumento do conhecimento ou da capacidade intelectual, antes e depois da ação educativa; avaliação do compor-

tamento – medida em que os participantes aplicaram o aprendizado com mudança de seu comportamento, que pode ser verificada imediatamente ou alguns meses após a atividade educativa; e avaliação de resultado – medida do efeito sobre o serviço ou ambiente de trabalho resultante da melhoria do desempenho do participante<sup>(5)</sup>.

Esse modelo recebeu contribuições para seu aprimoramento por parte de Anthony Cradell Hamblin, que subdividiu o nível 4 da avaliação de resultados em dois subníveis: nível 4 – avaliação da organização, referido às mudanças no funcionamento das organizações e nível 5 – avaliação do valor final, relacionado às mudanças no alcance dos objetivos finais da organização<sup>(5)</sup>.

O modelo de Kirkpatrick também foi utilizado em estudos<sup>(7-8)</sup> sobre educação interprofissional (EIP) como base para a proposta de seis níveis de avaliação da efetividade da EIP: reação – percepção dos participantes sobre a experiência de aprendizagem e sobre a natureza da EIP; mudança de atitude/percepção – mudanças recíprocas de atitude/percepção entre os participantes e dos valores e aplicação do trabalho em equipe na atenção à saúde de grupos específicos; aquisição de conhecimentos e habilidades – aprendizagem de conhecimentos e habilidades associados a EIP; mudança de comportamento – transferência individual do aprendizado interprofissional para a prática do respectivo indivíduo; mudança na prática organizacional – identificação de mudança na atenção à saúde prestada pela organização; benefícios para pacientes/usuários, familiares e comunidade – identificação de melhora nas condições de saúde e bem-estar dos pacientes/clientes. A efetividade da EIP é concebida como a atenção à saúde que produz resultados positivos com custos aceitáveis e sem efeitos secundários inaceitáveis<sup>(8)</sup>.

No Brasil, foram desenvolvidos dois modelos de avaliação de ações educativas por autores do campo da psicologia social, fundamentados nos estudos de Kirkpatrick. O primeiro é o Modelo de Avaliação Integrado e Somativo (MAIS) de Borges-Andrade que acrescenta variáveis do ambiente e de processo e propõe a análise e interpretação integradas das informações, visando à formulação de políticas e estratégias organizacionais<sup>(9)</sup>. O segundo é o Modelo Integrado de Avaliação do Impacto do Treinamento no Trabalho (IMPACT), desenvolvido a partir do MAIS<sup>(10)</sup>, o qual analisa a relação entre os níveis de avaliação de reação, de aprendizagem e de impacto. Neste modelo, o impacto é avaliado pela transferência da aprendizagem e da influência do treinamento sobre o desempenho geral do participante.

Especificamente, quanto aos resultados e ao impacto dos processos educativos na qualidade dos serviços, vale destacar não ser possível ao trabalhador operar automa-

No Brasil, a avaliação de ações e programas educativos de trabalhadores de saúde ainda não se consolidou como uma tradição de pesquisa e constitui-se como uma temática relevante a ser investigada.

ticamente uma transposição mecânica e direta da capacitação para a situação de trabalho, pois esta envolve um conjunto de aspectos que dizem respeito inclusive às características do processo de trabalho, às condições de trabalho e à estrutura organizacional, o que vem ao encontro das proposições dos modelos MAIS e IMPACT.

Além disso, há que se considerar não apenas as mudanças no desempenho profissional, de curto prazo, advindas da aplicação de novos conhecimentos, habilidades e atitudes, que representariam os resultados, mas também que essas mudanças sejam aplicadas em médio e longo prazos e acompanhadas de efeitos positivos sobre o desempenho global do trabalhador. Nesse sentido, entende-se o conceito de impacto da capacitação no trabalho como os efeitos que o treinamento exerce sobre o desempenho global, assim como sobre a motivação e as atitudes do egresso em relação ao seu trabalho, em médio e longo prazos<sup>(11-12)</sup>.

Dos modelos disponíveis, verificamos que os mais praticados na área da saúde são as avaliações de reação e de aprendizagem. Neste estudo, abordaremos a avaliação da aprendizagem que mostra resultados de eficácia da ação e do programa educativo ou, como assinalado anteriormente, se a ação educativa foi capaz de promover aquisição ou ampliação do conhecimento no participante.

Assim, o presente estudo busca subsidiar a construção de uma metodologia de avaliação com a análise, em particular, da avaliação de aprendizagem tendo em vista a tomada de decisão que favoreça a gestão e a definição de instrumentos gerenciais no tocante ao processo de planejamento de programas educativos em saúde.

## OBJETIVO

Analisar o processo de avaliação da aprendizagem realizado em treinamentos ministrados à equipe de enfermagem, no que se refere à eficácia dos treinamentos e à validade dos instrumentos.

## MÉTODO

Trata-se de uma pesquisa empírica, uma vez que os achados apoiarão formulações teóricas para a construção de uma metodologia de avaliação, proporcionando maior concretude à base fatural e aos dados empíricos, que dependem do referencial teórico e que agregarão impacto no sentido de facilitarem a aproximação prática<sup>(13)</sup>.

O estudo foi quantitativo correlacional e testou, conjuntamente, variáveis para verificar o quanto uma variável mudou em decorrência da mudança de outra<sup>(14)</sup>. Para tal, as variáveis da avaliação da aprendizagem foram medidas por meio de notas, antes e após as ações educativas, ora denominadas, genericamente, de treinamento (T) por seu caráter de aprendizado cognitivo e de habilidades.

O foco de análise é o processo de avaliação da aprendizagem. Nesse sentido, caracteriza-se como uma pesquisa de avaliação, pois tem como objetivo determinar o valor de um programa, recorrendo a procedimentos científicos para tal<sup>(14)</sup>.

A análise está pautada em dois pontos, especificamente, na eficácia do treinamento, isto é, se ele contribuiu para aumentar o conhecimento do treinando, e a validade do instrumento, ou seja, se o teste de conhecimento foi capaz de medir o que se propunha medir. Embora esses aspectos sejam interdependentes, essa divisão na análise objetiva, a princípio, discutir os achados, propondo intervenções necessárias em cada um deles e, em seguida, para o conjunto do processo de treinamento.

O estudo foi desenvolvido em dois hospitais do município de São Paulo, um de ensino da rede pública – Hospital Universitário da Universidade de São Paulo (HU-USP) – e outro filantrópico da rede privada (HPSP), ambos de grande porte e de assistência geral. As duas instituições possuem serviço estruturado de educação, exclusivo para a equipe de enfermagem, desenvolvendo programas periodicamente, de acordo com as necessidades identificadas.

Constituíram-se no material de análise as avaliações da aprendizagem relativas a seis treinamentos, três de cada hospital, num total de 486 avaliações no HU-USP e 507 no HPSP, a saber:

1. Treinamento *Prevenção e tratamento de úlcera por pressão* (TUP) – realizado no HU-USP de maio a julho de 2005, com 96 participantes na pesquisa.
2. Treinamento *Ressuscitação cárdio-pulmonar no adulto* (TRCPA) – realizado no HU-USP nos meses de julho e agosto de 2005, com 218 participantes na pesquisa.
3. Treinamento *Anotações de Enfermagem* (TAE) – realizado no HU-USP em 2005, com 172 participantes na pesquisa.
4. Treinamento *Cuidados com cateterismo vesical de demora* (TCVD) – realizado no HPSP em dezembro de 2006, com 76 participantes na pesquisa.
5. Treinamento *Vias de administração de medicamentos* (TAM) – realizado no HPSP em dezembro de 2007 e janeiro de 2008, com 410 participantes.
6. Treinamento *Curativos* (TC) – realizado no HPSP em abril de 2007, com 21 participantes.

Como instrumentos de coleta de dados das avaliações da aprendizagem, foram elaborados testes, pelas enfermeiras instrutoras dos treinamentos, com base no conteúdo teórico ministrado, contendo questões de verificação do conhecimento específico do treinamento, e foram aplicados imediatamente antes e após o programa, a fim de verificar se a diferença ou o aumento pretendido no conhecimento poderia ser atribuído ao T. Os testes pré e pós-treinamento foram idênticos para que não houvesse

diferença no grau de dificuldade das provas e para que fosse possível a comparação do desempenho dos treinados, por meio da nota.

As avaliações foram aplicadas pelos instrutores dos T, devidamente orientados para tal, de modo a garantir a homogeneidade da coleta, bem como para conscientizar os participantes da finalidade da avaliação e utilização dos resultados.

Os dados obtidos foram analisados por meio de estatística descritiva e inferencial e o teste Shapiro Wilk foi usado para verificar a distribuição de todas as variáveis a fim de nortear a escolha do teste inferencial a ser empregado.

Para identificar a existência de diferença estatisticamente significativa entre as notas pré e pós-treinamento, foi feita uma comparação, desses dois momentos, por meio da aplicação do teste não-paramétrico de Wilcoxon, que é utilizado, em dados pareados, para testar a hipótese de não haver diferença estatisticamente significativa; geralmente, envolve uma variável medida no mesmo indivíduo em dois momentos diferentes. Entre as medições, há uma intervenção nos sujeitos, com o objetivo de verificar se a intervenção afetou as respostas. Calcula-se, então, para cada indivíduo, a diferença entre as medições final e inicial<sup>(15)</sup>.

Para verificação do poder de discriminação do instrumento de avaliação da aprendizagem, foi empregado o teste Qui-quadrado, com intervalo de erro de  $p < 0,05$  no TAE do HU. Nos demais, para verificar se as questões eram capazes de medir o conhecimento que se propunham medir, foi aplicado o teste exato de Fisher comparando as proporções das questões erradas nos dois momentos.

O projeto de pesquisa foi aprovado pela Câmara de Pesquisa do HU-USP e pelos Comitês de Ética em Pesquisa das duas instituições (CEP 555/05). Os participantes foram orientados a respeito dos objetivos das avaliações, para que participassem, espontaneamente, sem constrangimento ou temor de que pudessem sofrer sanções, dependendo do resultado de seus testes. Por essa razão, foi solicitada autorização dos mesmos para utilização dos dados das avaliações no estudo, conforme Termo de Compromisso para realização da pesquisa, observando os princípios que norteiam a Resolução 196/96.

## RESULTADOS

Como o objetivo desse estudo está focado na análise global dos dados de aprendizagem, não será realizada a apreciação dos aspectos técnicos dos treinamentos. Por essa mesma razão, os conteúdos ministrados e os de cada questão não serão apresentados e discutidos. Assim, os resultados estão apresentados em quadros síntese, por treinamento, no que se refere à eficácia medida pela nota nos testes de aprendizagem e à capacidade de discriminação de cada questão.

Conforme Tabela 1, a respeito das avaliações dos treinamentos ministrados no HUUSP, observamos que apresentaram alteração significativa da nota e da capacidade de discriminação as questões: 2, 5, 8a, 8d, 10b e 10d do TRCPA e as 1a, 1b, 2c, 3a, 3c e 3d do TAE.

**Tabela 1** – Comparação dos resultados das notas nos momentos pré e pós-treinamento e capacidade de discriminação da questão entre os momentos, nos treinamentos realizados no HUUSP, São Paulo – 2007

Questão	Momento	Média	p* Significância entre os momentos pré e pós	p** Capacidade de discriminação da questão
<b>TUP</b>				
1a	Pré	0,21	<0,001	0,289
	Pós	0,30		
1b	Pré	0,16	<0,001	0,478
	Pós	0,30		
1c	Pré	0,24	0,007	1,000
	Pós	0,28		
2	Pré	1,84	<0,001	0,448
	Pós	2,03		
3	Pré	1,51	<0,001	1,000
	Pós	2,20		
4	Pré	2,52	<0,001	0,453
	Pós	3,19		
Nota total	Pré	6,48	<0,001	----
	Pós	8,30		
<b>TRCPA</b>				
1	Pré	0,25	< 0,001	0,516
	Pós	0,69		
2	Pré	0,67	< 0,001	0,001
	Pós	0,91		
3	Pré	0,37	< 0,001	0,253
	Pós	0,81		
4	Pré	0,25	< 0,001	0,289
	Pós	0,67		
5	Pré	0,36	< 0,001	0,003
	Pós	0,60		
6	Pré	0,52	< 0,001	1,000
	Pós	0,98		
7	Pré	0,84	0,088	0,335
	Pós	0,90		
8a	Pré	0,19	0,016	0,042
	Pós	0,20		
8b	Pré	0,24	0,564	0,031
	Pós	0,24		
8c	Pré	0,19	0,257	0,021
	Pós	0,20		
8d	Pré	0,18	< 0,001	<0,001
	Pós	0,23		
9	Pré	0,94	0,527	0,277
	Pós	0,96		
10a	Pré	0,15	0,631	<0,00
	Pós	0,15		
10b	Pré	0,21	0,005	0,015
	Pós	0,24		
10c	Pré	0,16	< 0,001	0,711
	Pós	0,24		
10d	Pré	0,07	< 0,001	0,004
	Pós	0,12		
Total	Pré	5,16	< 0,001	---
	Pós	8,07		

Continua...

Questão	Momento	Média	p* Significância entre os momentos pré e pós	p** Capacidade de discriminação da questão
<b>TAE</b>				
1a	Pré	1,00	0,002	0,001
	Pós	1,08		
1b	Pré	0,99	< 0,001	0,0017
	Pós	0,80		
1c	Pré	0,95	0,050	0,001
	Pós	1,02		
1d	Pré	0,99	0,275	0,0015
	Pós	0,96		
2a	Pré	0,35	0,862	0,0024
	Pós	0,34		
2b	Pré	0,39	0,655	0,426
	Pós	0,40		
2c	Pré	0,30	0,020	0,001
	Pós	0,26		
3a	Pré	0,96	0,007	0,001
	Pós	1,04		
3b	Pré	0,93	0,050	0,001
	Pós	0,99		
3c	Pré	1,04	0,003	0,02
	Pós	1,09		
3d	Pré	0,58	0,008	0,001
	Pós	0,70		
Total	Pré	8,31	0,007	
	Pós	8,64		

Em todas as questões do TUP, houve alteração significativa nas notas pós, porém as questões não eram capazes de mensurar a diferença entre os momentos pré e pós. Além dessas, as questões 1, 3, 4, 6 e 10c do TRCPA apresentaram a mesma condição.

As questões 8b, 8c e 10a do TRCPA e as 1c, 1d, 2a e 3b do TAE não apresentaram significância estatística entre os momentos pré e pós-treinamento, sendo que as questões mostraram poder de discriminação.

As questões 7 e 9 do TRCPA e 2b do TAE não apresentaram alteração significativa nas notas e nem poder de diferenciação entre os momentos pré e pós-treinamento.

Os resultados das avaliações da aprendizagem dos três treinamentos ministrados no HPSP estão descritos nas Tabelas 2 e 3, nas quais verificamos que as questões 1 do TCVD e 1 do TAM não foram capazes de discriminar a diferença entre os dois momentos, nem tampouco houve diferença significativa entre os momentos.

Nas questões 3 do TC e 3 do TAM, houve diferença significativa entre os momentos, embora essas questões não tenham apresentado poder de discriminar tal diferença.

Já a questão 3b do TC, que tinha capacidade de discriminar a diferença entre os momentos, não apresentou diferença significativa.

As questões 2, 3 e 4 do TCVD e a 4 do TAM apresentaram diferença significativa entre os momentos, sendo que possuíam capacidade para mostrar essa diferença. Nas

questões 3c, 3d e 3e não houve nem diferença nem poder de discriminação.

**Tabela 2** – Distribuição do percentual de respostas certas nos momentos pré e pós-treinamento e da capacidade de discriminação da questão, nos treinamentos realizados no HPSP, São Paulo – 2008

Questão	Momento	Frequência Percentual	p* Significância entre os momentos pré e pós	p** Capacidade de discriminação da questão
<b>TCVD</b>				
1	Pré	88,31	1,000	1,000
	Pós	98,70		
2	Pré	42,47	0,001	0,001
	Pós	80,00		
<b>TC</b>				
4	Pré	94,12		---
	Pós	100,00		
5	Pré	17,65		0,669
	Pós	11,76		
<b>TAM</b>				
1	Pré	97,80	0,085	1,000
	Pós	99,02		
2	Pré	61,92	<0,001	0,00
	Pós	96,09		
3	Pré	91,71	0,036	0,057
	Pós	99,02		

**Tabela 3** – Comparação dos resultados das notas nos momentos pré e pós-treinamento e capacidade de discriminação da questão entre os momentos nos treinamentos realizados no HPSP, São Paulo – 2007

Questão	Momento	Média	p* Significância entre os momentos pré e pós	p** Capacidade de discriminação da questão
<b>TCVD</b>				
3	Pré	2,89	0,01	0,025
	Pós	3,38		
4	Pré	1,25	<0,001	0,017
	Pós	1,44		
Total	Pré	7,24	<0,001	---
	Pós	8,75		
<b>TC</b>				
1	Pré	1,20	0,08	0,537
	Pós	1,51		
2	Pré	1,85	0,52	0,118
	Pós	1,94		
3	Pré	1,13	0,05	---
	Pós	1,38		
Total	Pré	7,89	0,10	
	Pós	8,62		
<b>TAM</b>				
4	Pré	1,45	1	0,001
	Pós	1,59		
Total	Pré	8,6	<0,001	----
	Pós	9,7		

(\*) nível descritivo de probabilidade do teste não-paramétrico de Wilcoxon; (\*\*) nível descritivo de Fischer.

## DISCUSSÃO

Observamos que a situação não é a desejada, pois se esperava que as questões apresentassem poder de discriminar

diferenças, sobretudo na presença de alteração significativa nas notas. Esses resultados exigem uma análise da qualidade dos testes e do grau de dificuldade das questões ou, ainda, do nível do conhecimento prévio dos participantes.

O esperado era que as questões apresentassem poder para medir a diferença de nota entre os momentos pré e pós-treinamento. No entanto, verificamos que algumas questões não foram capazes de mensurar essa diferença.

Resgatando os principais resultados relativos à aprendizagem, observamos que:

- No HU-USP, os três treinamentos apresentaram nota final com aumento estatisticamente significativo. No TUP, as sete questões apresentaram aumento significativo, mas as questões não possuíam poder de discriminação. No TRCPA, das 16 questões, 11 apresentaram aumento significativo; dessas, seis com poder de diferenciação e cinco sem; em cinco não houve diferença significativa. Nas 11 questões do TAE, seis apresentaram alteração significativa, sendo que em quatro houve aumento e em duas houve um decréscimo das notas; cinco não tiveram alteração significativa; exceto uma questão, que não apresentou diferença, todas as outras possuíam capacidade para discriminar os momentos.
- No HPSP, apenas o TC não apresentou aumento significativo na nota final. No TCVD, das quatro questões, três apresentaram aumento significativo por questões capazes de diferenciação; a outra não sofreu alteração e também não foi capaz de discriminar. Das nove questões do TC, apenas uma mostrou aumento significativo, porém sem capacidade de discriminar. Das demais, apenas uma possuía poder de discriminação. Ressaltamos que o baixo número de participantes (17) pode ter prejudicado o tratamento estatístico. No TAM, das duas com aumento significativo, uma tinha capacidade de diferenciar os momentos e a outra não, o mesmo ocorreu com as duas que não se alteraram.

Considerando, inicialmente, a alteração significativa das notas nas avaliações da aprendizagem, podemos afirmar que, de modo geral, os treinamentos influenciaram nas notas obtidas pelo participante no momento pós-treinamento em relação ao pré.

Percebemos que as questões dos testes que se referiam a conteúdos novos ou não habituais nos cursos de formação profissional foram essenciais para avaliar a aquisição de conhecimento, pois foram observados menores índices de acertos no teste pré e maior índice de aumento no pós, a exemplo do TRCPA, principalmente, entre os técnicos e auxiliares de enfermagem, fato esse que encaixa à conclusão de que conhecimentos novos foram adquiridos e as dúvidas foram sanadas.

Para confirmação objetiva desse aspecto será necessária, ainda, a verificação, em cada treinamento, de relação entre a variável nota e a variável conhecimento do

assunto, que pode ser obtida na avaliação da reação ou da satisfação.

As questões com alto índice de acerto, tanto no momento pré quanto no pós, mostram que os treinandos já possuíam informações sobre o tema antes do treinamento, como no TAE.

Por um lado, esse resultado mostra um aspecto importante que se refere à apropriação do saber por meio de uma superação constante e a possibilidade de ampliação do conhecimento quando há algum conhecimento prévio sobre o assunto<sup>(16)</sup>. Isso também reforça o entendimento de que informações conhecidas ajudam a fixar o conteúdo novo e permitem relacionar as novas informações com o conhecimento que o treinando traz consigo.

Por outro lado, sugere que o conteúdo ministrado estava aquém das necessidades dos participantes ou os treinandos não eram o público certo a ser treinado ou os instrutores não souberam elaborar o conteúdo na medida da necessidade dos treinandos. Compreendemos esse fato como um equívoco no levantamento de necessidades e na escolha dos participantes.

Examinando os temas dos treinamentos, destacamos que, com exceção do treinamento de parada cardíaco-respiratória, que é mais complexo, e do de tratamento de úlcera por pressão pela oferta de novos produtos, os demais temas – anotações, curativos, administração de medicamentos e cateterismo vesical – relacionam-se a atividades cotidianas da enfermagem. Dessa forma, era esperado que os treinandos apresentassem considerável conhecimento antes do treinamento.

A esse exemplo, observamos que dos três treinamentos analisados no HU, o maior aumento da nota média do pré para o pós foi no TRCPA e o mais discreto no TAE. O que se pôde deduzir, ao relacionar nota e tema, é que o TRCPA foi o que apresentou conteúdo mais complexo e novo aos participantes, enquanto que o TAE foi o de conteúdo mais comum à prática e mais habitual na formação profissional.

No HPSP, o TC não apresentou nota final maior no pós-treinamento, embora com *n* pequeno, confirma a conjectura de que os temas abordados e os conteúdos ministrados eram de domínio dos treinandos.

Contudo, houve em cada treinamento um aproveitamento do conteúdo transmitido, apesar do prévio conhecimento, pois encontramos no TUP aumento significativo em todas as questões do teste; no TRCPA em 11 das 16 e no TAE em quatro de 11; em duas houve um decréscimo.

No HPSP, no entanto, o aproveitamento foi mais discreto, houve aumento significativo em uma das quatro questões do TCVD; em uma das nove do TC e em duas do TAM.

Na perspectiva do instrumento – teste de conhecimentos –, o ideal é que todas as questões tenham capa-

cidade de medir o conhecimento de modo comparativo entre os momentos. Entretanto, isso ocorreu em pouco mais que a metade das questões, considerando os seis treinamentos.

Sob o ponto de vista da avaliação da aprendizagem, o fato do instrumento não ter apresentado capacidade de discriminação enfraquece a comprovação da eficácia do treinamento, pois, embora tenha havido significância estatística entre os dois momentos, houve, também, problemas na elaboração das questões.

Esse resultado exige uma apreciação detalhada dos testes, da construção à aplicação, relacionando conteúdo e objetivos da atividade instrucional e analisando de modo qualitativo, cada uma das questões.

É preciso averiguar se o conteúdo medido no teste de conhecimento foi compatível com o conteúdo apresentado no treinamento, porquanto uma das limitações da avaliação da aprendizagem é a não-correspondência entre o conteúdo exigido no teste e o desenvolvido pelo instrutor durante o treinamento<sup>(17)</sup>.

Igualmente, para responder aos testes, os treinandos fazem o exercício mental de reconhecer a aplicabilidade das questões, compreendendo-as como se estivessem agindo. O instrutor deverá proceder ao mesmo raciocínio para formular os testes, de modo que as questões tenham sentido prático.

Além disso, a aprendizagem ocorre quando o indivíduo passa a dominar inteiramente o objeto do aprendizado, sendo esse domínio traduzido na aquisição de novas atitudes e habilidades<sup>(18)</sup>.

Sabendo que a aprendizagem é uma das condições necessárias à transferência do aprendido ao trabalho, mas que não é suficiente para tal, recomenda-se o aprofundamento das técnicas de avaliação, bem como a verificação de existência de relações entre as variáveis da aprendizagem com as de satisfação e de impacto.

## REFERÊNCIAS

1. Brasil. Ministério da Saúde. Portaria n. 198/GM, de 13 de fevereiro de 2004. Institui a Política Nacional de Educação Permanente em Saúde como estratégia do Sistema Único de Saúde para a formação e o desenvolvimento de trabalhadores para o setor e dá outras providências. Brasília; 2004.
2. Brasil. Ministério da Saúde. Portaria n. 1.996/GM, de 20 de agosto de 2007. Dispõe sobre as diretrizes para a implementação da Política Nacional de Educação Permanente em Saúde e dá outras providências [Internet]. Brasília; 2007 [citado 2011 ago. 30]. Disponível em: [http://portal.saude.gov.br/portal/arquivos/pdf/Portaria\\_N\\_1996\\_GMMS.pdf](http://portal.saude.gov.br/portal/arquivos/pdf/Portaria_N_1996_GMMS.pdf)
3. Peduzzi M, Guerra DAD, Braga CP, Lucena FS, Silva JAM. Atividades educativas de trabalhadores na atenção primária: concepções de educação permanente e de educação continuada em saúde presentes no cotidiano de Unidades Básicas de Saúde em São Paulo. *Interface Comunic Saúde Educ.* 2009;13(30):121-34.
4. Organización Panamericana de la Salud (OPAS). Capacitación del personal de los servicios de salud. Quito; 2002. (Proyectos relacionados com los Procesos de Reforma Sectorial, 137)
5. Mira VL. Avaliação de programas de treinamento e desenvolvimento da equipe de enfermagem de dois hospitais do Município de São Paulo [tese livre-docência]. São Paulo: Escola de Enfermagem, Universidade de São Paulo; 2010.

Esses resultados, apesar de não conclusivos, vão ao encontro da análise na qual se detectou que o nível de avaliação de reação e satisfação são fracamente relacionados com a aprendizagem; essas variáveis estão fortemente correlacionadas com impacto<sup>(11)</sup>. Corroboram, ainda, o estudo a respeito de ações educativas da área da saúde que evidenciou a predominância de estratégias de ensino tradicionais e incipiente experiências de avaliação<sup>(19)</sup>.

A avaliação do impacto é indicadora de mudança do comportamento no cargo e da efetividade das ações de treinamento no nível individual<sup>(12)</sup>, da qual presume-se a necessidade de ampliação das técnicas de avaliação, uma vez que o objetivo final do treinamento é provocar mudança no ambiente de trabalho, logo, é preciso avaliar o impacto.

## CONCLUSÃO

No tocante à aprendizagem, os resultados apontam para a eficácia dos treinamentos, embora a significância de aquisição ou aumento do conhecimento pretendido tenha sido enfraquecida pelos instrumentos de medida que apresentaram problemas de discriminação da variável nota entre os momentos pré e pós-treinamento.

É preciso analisar, junto aos instrutores e treinandos, a construção de cada questão, comparando o comportamento das notas aos objetivos propostos, ao conhecimento prévio dos participantes e ao conteúdo do treinamento que, além de explicar os equívocos do instrumento, permitirá avaliar ao mesmo tempo o ensino, levando a uma apreciação completa do processo ensino-aprendizagem e não à aprendizagem isoladamente.

Dos resultados dos seis treinamentos avaliados, pôde-se perceber a necessidade de proposição de uma metodologia de avaliação, construindo critérios, instrumentos e indicadores para essa finalidade. Notou-se, também, que é preciso aprimorar a realização do diagnóstico de necessidades de treinamento e as estratégias de ensino.

6. Otrenti E. Avaliação de processos educativos formais para profissionais da área da saúde: revisão integrativa de literatura [dissertação]. São Paulo: Escola de Enfermagem; Universidade de São Paulo; 2011.
7. Barr H, Koppel I, Reeves S, Hammick M, Freeth D. Effective interprofessional education: arguments, assumption & evidence. Oxford: Blackwell/CAIPE; 2005.
8. Freeth D, Hammick M, Reeves S, Koppel I, Barr H. Effective interprofessional education: development, delivery & evaluation. Oxford: Blackwell; 2005.
9. Borges-Andrade JE. Avaliação integrada e somativa em TD&E. In: Borges-Andrade JE, Abbad GS, Mourão L. Treinamento, desenvolvimento e educação em organizações e trabalho: fundamentos para a gestão de pessoas. Porto Alegre: Artmed; 2006. Parte III – Avaliação dos sistemas de TD&E; p. 343-58.
10. Abbad G. Um modelo integrado de avaliação de impacto de treinamento no trabalho [tese doutorado]. Brasília: Universidade de Brasília; 1999.
11. Abbad G, Gama AL, Borges-Andrade JE. Treinamento: análise do relacionamento da avaliação nos níveis de reação, aprendizagem e impacto no trabalho. Rev Adm Contemp. 2000;4(3):25-45.
12. Pilati R, Abbad G. Análise fatorial confirmatória da Escala de Impacto do Treinamento no Trabalho. Psicol Teoria Pesq. 2005;21(1):43-51.
13. Demo P. Metodologia do conhecimento científico. São Paulo: Atlas; 2008.
14. LoBiondo WG, Haber J. Pesquisa em enfermagem: métodos, avaliação crítica e utilização. 4ª ed. Rio de Janeiro: Guanabara Koogan; 2001.
15. Menezes RX, Azevedo RS. Bioestatística não paramétrica. In: Massad E, Menese RX, Silveira PSP, Ortega NRS. Métodos quantitativos em medicina. Barueri: Manole; 2004. p. 307-18.
16. Freire P. Educação e mudança. São Paulo: Paz e Terra; 2001.
17. Schaan MH. Avaliação sistemática de treinamento: guia prático. São Paulo: LTr; 2001.
18. Carvalho AVC, Nascimento LP. Administração de recursos humanos. São Paulo: Pioneira; 2000.
19. Tronchin DMR, Mira VL, Peduzzi M, Ciampone MHT, Melleiro MM, Silva JAM, et al. Permanent education of health professionals in public hospital organizations. Rev Esc Enferm USP [Internet]. 2009 [cited 2011 Ago 30];43(n.esp 2):1210-5. Available from: [http://www.scielo.br/pdf/reeusp/v43ns-pe2/en\\_a11v43s2.pdf](http://www.scielo.br/pdf/reeusp/v43ns-pe2/en_a11v43s2.pdf)