

## Medidas Comportamentais de Eficácia: Contribuições na Avaliação do Ensino de Operações Aritméticas

Marcelo Henrique Oliveira Henklain<sup>1</sup>  
Universidade Federal de Roraima  
João dos Santos Carmo  
Universidade Federal de São Carlos  
Jair Lopes Júnior  
Universidade Estadual Paulista – Bauru

**RESUMO** - Investigaram-se medidas de eficiência (razão entre acertos e tempo de solução) e erro (diferença entre a resposta emitida e a resposta correta) com estudantes do Ensino Fundamental. O Experimento 1 compreendeu pré-teste, ensino de equivalência entre diferentes formas de apresentação de problemas, pós-teste 1, ensino do algoritmo de adição, pós-teste 2, ensino do algoritmo de subtração, pós-teste 3 e teste de generalização. O Experimento 2 foi semelhante ao anterior até o pós-teste 1, seguido de treino de solução de problemas na forma de balança, pós-teste 2, teste de generalização 1, ensino dos algoritmos de adição e subtração, pós-teste e teste de generalização 2. Medir eficiência e erro contribui para a avaliação da eficácia de programas de ensino de solução de problemas.

**Palavras-chave:** avaliação de programa educacional, solução de problemas, ensino da matemática

## Behavioral Measures of Efficacy: Contributions to the Evaluation of Arithmetical Operations Teaching

**ABSTRACT** - We investigated measures of efficiency (ratio of correct responses and resolution time) and error (distance between emitted response and the correct response) in a group of Elementary School students. Experiment 1 comprised pre-test, teaching of equivalence between different forms of problem presentation, post-test 1, teaching of addition algorithm, post-test 2, teaching of subtraction algorithm, post-test 3, and generalization test. Experiment 2 was similar to Experiment 1 until post-test 1, followed by training of scale problem-solving, post-test 2, generalization test 1, teaching of addition and subtraction algorithms, post-test 3, and generalization test 2. The results of this study suggest that measuring efficiency and error contribute to evaluate the effectiveness of educational programs in teaching problem solving.

**Keywords:** educational program evaluation, problem solving, mathematics education

Estudos apontam sistematicamente dificuldades no ensino do comportamento de resolver problemas de adição e subtração (Fossa & Sá, 2008; Magina, Santana, Cazorla, & Campos, 2010; Neef, Nelles, Iwata, & Page, 2003; Vasconcelos, 1998). Para um ensino eficaz, o professor precisa conhecer o repertório comportamental que define a solução de problemas, bem como o que determina a dificuldade do problema e as habilidades necessárias para resolvê-lo (Haydu, Pullin, Iégas, & Costa, 2010; Resnick, Wang, & Kaplan, 1973). A forma de apresentação do problema e as estruturas semântica e sintática, em especial a posição da incógnita, são algumas das variáveis que interferem na dificuldade do problema (Haydu, Costa, & Pullin, 2006; Henklain & Carmo, 2013a; Henklain & Carmo, 2013b).

A expressão *forma de apresentação* refere-se à topografia do estímulo, isto é, se o problema é apresentado, por exemplo, no formato de sentença escrita (Pedro tinha um bombom e ganhou mais dois. Quantos bombons ele possui

agora?), no formato de coleções (\* + \*\*) ou no formato de algarismos (1 + 2). *Estrutura semântica* refere-se ao significado do problema, que pode ser de *transformação*, nos casos em que uma condição inicial sofre uma mudança, de *combinação*, quando dois conjuntos são somados ou um conjunto é retirado de outro, e de *comparação*, quando um conjunto é comparado com outro. *Estrutura sintática* refere-se a aspectos do problema tais como o número de palavras e a ordem cronológica de apresentação das informações do problema; alterar a posição da incógnita é um tipo de mudança na estrutura sintática do problema. *Posição da incógnita* relaciona-se à localização do valor desconhecido do problema. Em um problema de três termos,  $a + b = c$ , existem três posições possíveis da incógnita.

Há consenso em considerar difíceis os problemas na forma escrita com estruturas semânticas de comparação e de combinação, que envolvem relações estáticas entre os elementos do problema, e com incógnita nas posições iniciais a e b (Fossa & Sá, 2008; Henklain & Carmo, 2013a; Henklain & Carmo, 2013b; Hiebert, 1982; Iégas, 2003; Magina et al., 2010; Neef et al., 2003; Resnick & Rosenthal,

<sup>1</sup> Endereço para correspondência: Rua do Cajueiro, n. 133, Caçari I, Boa Vista, Roraima, Brasil. CEP 69307-510. E-mail: marcelo\_henklain@hotmail.com

1974). Problemas escritos tendem a ser mais difíceis porque exigem habilidades matemáticas e linguísticas (Fayol, 1992; Verschaffel & De Corte, 1997). Estima-se que tais dificuldades são acentuadas quando os problemas apresentam estruturas semânticas que envolvam relações estáticas porque se torna menos evidente a relação entre a situação-problema e os comportamentos de solução necessários, uma vez que nada é adicionado ou retirado dos conjuntos (Carpenter & Moser, 1983; Nunes & Bryant, 1996). Problemas com incógnitas nas posições a e b são difíceis porque a criança se orienta pelas ações descritas no enunciado do problema para solucioná-lo, mas essa estratégia não pode ser adotada quando a incógnita está nas posições iniciais (Bryant, 2013; Haydu et al., 2010). Embora seja complexo ensinar o comportamento de resolver problemas, esse é um aprendizado crucial porque constitui pré-requisito para comportamentos matemáticos mais complexos (Costa, Galvão, & Ferreira, 2008; Magina et al., 2010; Nunes & Bryant, 1996).

Alguns estudos buscaram reduzir dificuldades na solução de problemas de adição e subtração (Haydu et al., 2006; Henklain & Carmo, 2013a, 2013b; Iéguas, 2003) e avaliaram a eficácia dos seus procedimentos de ensino (intervenções) com base na comparação entre as porcentagens de acertos no pré-teste e nos pós-testes. Assim, avaliaram, com base nas medidas de acerto no início e ao final da pesquisa, se os procedimentos empregados haviam conseguido ensinar a resolver problemas de matemática ou reduzir a dificuldade dos mesmos, o que foi inferido com base no aumento na porcentagem de acertos ao resolver esses problemas. Parcela expressiva dos relatos de pesquisa infere a eficácia de uma intervenção apenas com base na diferença entre a porcentagem de acertos registrados antes e após a exposição dos participantes aos procedimentos de ensino e de avaliação adotados. Embora admitindo a função informativa de tal medida, cumpre reconhecer que a mesma expressa a quantificação de acertos (e/ou erros), prescindindo de registros acerca de outras dimensões operantes, tais como a velocidade de realização da tarefa e a comparação da diferença entre o erro e a resposta correta. Estima-se, no presente artigo, que tais medidas poderiam aprimorar a avaliação da qualidade de uma intervenção pela ampliação da visibilidade a dimensões relevantes dos repertórios operantes envolvidos.

Admite-se, portanto, a necessidade de investigação de medidas eficazes e estáveis de avaliação de sucesso de procedimentos de ensino. Particularmente importante é a oferta de medidas de eficácia de procedimentos que visem reduzir dificuldades na solução de problemas aritméticos, tendo em vista o contexto das avaliações de larga escala que apontam para o baixo rendimento nesses repertórios (Araújo & Luzio, 2005), além das dificuldades encontradas por professores no ensino dos mesmos (Magina et al., 2010).

O objetivo do presente estudo foi investigar duas medidas alternativas da avaliação da eficácia de procedimentos redutores de dificuldades na solução de problemas: a comparação de eficiência e de erros, antes e após a exposição dos participantes a tais procedimentos. Determinou-se como *eficiência* o cálculo de divisão entre a porcentagem de acertos em um conjunto de problemas matemáticos e o tempo necessário para resolvê-los (em segundos). Por seu

turno, *erro* designa a diferença entre a solução correta dos problemas e os diferentes tipos de soluções apresentadas. Solução do problema, neste artigo, significa o resultado de uma operação e não o processo de solução. Portanto, a medida de erro expressa uma quantificação da diferença entre a solução considerada correta e a solução apresentada pelo participante. Assim, quando há coincidência entre resultado correto e apresentado, o erro é igual a zero. Considere ainda que não ocorrerão resultados negativos no cálculo do erro. Por exemplo: se a solução correta for 5 e a solução apresentada pelo participante for 4 ou 6, o erro será, nos dois casos, igual a 1.

Para demonstrar a utilidade dessas medidas na avaliação da eficácia de procedimentos de ensino, serão utilizados dados publicados em artigos anteriores dos dois primeiros autores deste estudo (Henklain & Carmo, 2013a, 2013b). Neste texto, o primeiro artigo será denominado de Experimento 1 e o segundo, de Experimento 2. Serão descritos detalhes dos métodos empregados nos dois experimentos, o que permitirá julgar a qualidade do delineamento e dos dados produzidos.

O Experimento 1 teve o objetivo principal de avaliar o efeito da formação de um conjunto de classes de equivalência entre diferentes formas de apresentação de problemas de adição (operação com algarismos, problema escrito com estrutura de transformação, problemas na forma de balança e coleção) sobre o desempenho de solução de problemas de adição e subtração em diferentes formas de apresentação, estruturas semânticas e com incógnitas nas três posições. O aprendizado de relações de equivalência ocorre quando estímulos dissimilares são arbitrariamente relacionados por uma determinada comunidade verbal. Tais relações são chamadas de simbólicas (Matos, 1999). Analogamente, comportamento simbólico é a resposta controlada por estímulos arbitrariamente relacionados (símbolos e seus referentes) e substituíveis entre si, de tal maneira que símbolo e referente podem exercer a mesma função no controle de repertórios específicos (McIlvane, Goulart, Brino, Galvão, & Barros, 2005). Conceitos matemáticos podem ser interpretados como relações simbólicas (Carmo, 2002). E comportamentos matemáticos podem ser operacionalizados como redes de relações simbólicas, sendo denominados de comportamentos simbólicos que envolvam, por exemplo, relações entre numerais, quantidades, outros símbolos e referentes matemáticos.

Na Análise do Comportamento, as primeiras análises sobre relações simbólicas são encontradas em Keller e Schoenfeld (1973). Entretanto, foram os trabalhos relatados por Sidman em 1994 que propiciaram a investigação sistemática desse fenômeno com base no paradigma de equivalência de estímulos que estabelece o procedimento para o ensino de relações arbitrárias entre estímulos e os critérios para atestar a formação de uma relação simbólica ou classe entre estímulos equivalentes (Sidman & Tailby, 1982). O procedimento usualmente utilizado para o estabelecimento de classes de equivalência é o de *escolha de acordo com o modelo* (MTS, do inglês *matching to sample*), que consiste na apresentação de tentativas em que o participante escolhe entre dois ou mais estímulos de comparação condicionalmente a um estímulo-modelo. Estímulos relacionados condicionalmente, no contexto desse procedimento, podem se tornar equivalentes.

Para atestar que são equivalentes, é preciso demonstrar a emergência das relações de simetria, transitividade e simetria da transitividade (de Rose, 1993).

Uma classe de estímulos equivalentes constitui-se em uma rede de relações, algumas ensinadas e as demais emergentes (de Rose, 1993), de tal modo que funções adquiridas por um membro da classe podem ser transferidas aos demais membros e esta pode ser expandida, bastando relacionar um estímulo novo a algum dos estímulos que compõem a classe (Sidman & Tailby, 1982). Investigar como diferentes estímulos são substituíveis entre si e como essa rede de relações se expande fornece indícios sobre a aquisição de comportamento simbólico sofisticado, como o comportamento matemático. Tecnologias educacionais para ensino de leitura e matemática, derivadas do estudo de classes de equivalência, são eficazes e econômicas, porque envolvem o ensino de algumas relações que produzem a emergência de outras (Carmo & Galvão, 1999; de Rose, 2005; Escobal, Rossit, & Goyos, 2010; Green, 2010; Prado & de Rose, 1999).

Foi estabelecido o objetivo adicional de avaliar, após essa história experimental de formação de classes de equivalência, o efeito do ensino de algoritmos para solução de problemas de adição e subtração sobre o desempenho de solução de problemas. O Experimento 2 teve por objetivos avaliar se a formação de dois conjuntos de classes de equivalência (adição e subtração) entre três diferentes formas de apresentação de problemas (algarismo, problema escrito com estrutura de transformação e balança) produziria melhora na solução de problemas e avaliar, após essa história, o efeito do treino com a balança e do ensino de algoritmos sobre esse desempenho.

## Método

### Participantes

**Experimento 1.** Oito crianças com idade entre 8 e 11 anos, sete do sexo masculino e uma do sexo feminino. Uma estava cursando o 5º ano; duas, o 4º ano; e cinco, o 2º ano do Ensino Fundamental. Sete estudavam na mesma escola pública municipal do interior de São Paulo e uma, em escola estadual da mesma cidade.

**Experimento 2.** Oito crianças com idades entre 7 e 12 anos, sendo sete meninas e um menino. Dois participantes cursavam o 2º ano; quatro, o 3º; e dois, o 5º ano do Ensino Fundamental. Todas estudavam em uma escola municipal da mesma cidade em que ocorreu o Experimento 1.

Os responsáveis pelos participantes de ambos os experimentos informaram que seus filhos não utilizavam medicamento permanente ou possuíam qualquer tipo de limitação sensorial, motora ou intelectual. Não foi avaliada a presença de transtornos de aprendizagem na amostra selecionada. Todos eram experimentalmente ingênuos em relação a pesquisas sobre ensino-aprendizagem da matemática, mas possuíam experiência com tarefas de MTS. Além disso, também não eram inexperientes em relação à solução de problemas de adição e subtração com um ou dois dígitos e incógnita na posição c por já terem sido expostos a esse conteúdo na escola. Eles foram selecionados para

participar do estudo porque apresentaram baixa porcentagem de acertos nas provas a que foram expostos na sessão de pré-teste. O critério de participação em ambos os experimentos foi ser capaz de ler problemas escritos simples e apresentar desempenho abaixo de 70% em provas de matemática especialmente elaboradas para ambos os estudos.

### Local e Materiais

A coleta foi realizada em uma sala da biblioteca de uma universidade do interior de São Paulo, onde funciona um projeto de extensão que atende crianças com dificuldade em leitura/escrita. No Experimento 1, as provas com problemas escritos e na forma de balança foram elaboradas no *software* Excel® 2007 e as demais sessões foram programadas no *software* ProgMTS (Marcicano, Carmo, & Prado, 2011). No Experimento 2, todas as sessões foram programadas no *software* ProgMTS.

Experimento 1			
Balança (A)	Algarismo (B)	Coleção (C)	Problemas Escritos (D)
	$? + 2 = 5$		César tinha 7 aviões. Ganhou mais dois e ficou com cinco ao todo.
	$2 + ? = 5$		João tinha dois aviões. Ganhou mais ? e ficou com cinco ao todo.
	$3 + 2 = ?$		Vitor tinha três aviões. Ganhou mais dois e ficou com ? ao todo.
Experimento 2			
Balança (A)	Algarismo (B)	Problemas Escritos (C)	
	$? + 2 = 4$	João tinha ? pipas. Ganhou mais duas e ficou com quatro ao todo. Quantas pipas o João tinha?	
	$3 + ? = 5$	Edu tinha três bolas. Ganhou mais ? e ficou com cinco ao todo. Quantas bolas o Edu ganhou?	
	$5 + 4 = ?$	Ana tinha cinco bolas. Ganhou mais quatro e ficou com ? ao todo. Com quantas bolas a Ana ficou?	

Figura 1. Estímulos empregados nos experimentos 1 e 2

### Estímulos

Em todas as sessões foram empregados estímulos com valores entre um e nove, exceto nos testes de generalização. A Figura 1 exibe os tipos de estímulos empregados em cada experimento.

### Procedimento

A Tabela 1 exibe um resumo dos estímulos e delineamentos adotados nos Experimentos 1 e 2.

### Pré-teste e pós-testes.

**Experimento 1.** O pré-teste e os pós-testes foram compostos por quatro provas: (a) problemas escritos; (b) problemas na forma de algarismos; (c) problemas na forma de balança; e (d) problemas na forma de coleção. Havia 18 problemas escritos, sendo nove por operação (adição e subtração), três para cada estrutura semântica (transformação, combinação e comparação) e seis para cada posição da incógnita. Para as demais formas de apresentação, havia seis problemas de cada operação e um para cada posição da incógnita.

Tabela 1. *Resumo dos métodos dos Experimentos 1 e 2*

Método	Experimento 1	Experimento 2
Participantes	8 crianças entre 8 e 11 anos, do 2º ao 5º ano do Ensino Fundamental.	8 crianças entre 7 e 12 anos, do 2º ao 5º ano do Ensino Fundamental.
Local	Sala do Projeto de Extensão para crianças com dificuldades em leitura e escrita.	Sala do Projeto de Extensão para crianças com dificuldades em leitura e escrita.
Software	ProgMTS e Excel.	ProgMTS.
Estímulos	Problema escrito e nas formas de algarismo, coleção e balança.	Problema escrito e nas formas de algarismo e balança.
Delineamento	Pré-teste, Ensino das discriminações condicionais, Pós-teste 1, Ensino do algoritmo de Adição, Pós-teste 2, Ensino do Algoritmo de Subtração, Pós-teste 3 e Teste de generalização.	Pré-teste, Ensino das discriminações condicionais, Pós-teste 1, Treino com a balança, Pós-teste 2, Teste de Generalização 1, Ensino dos algoritmos de Adição e Subtração, Pós-teste 3, Teste de Generalização 2.

**Experimento 2.** Pré-teste e pós-testes foram compostos por três provas: (a) problemas escritos; (b) problemas na forma de algarismos; e (c) problemas na forma de balança. Havia 12 problemas escritos, sendo 6 por operação (adição e subtração), 6 para cada estrutura semântica (transformação e comparação) e 4 para cada posição da incógnita. Para as demais formas de apresentação, havia seis problemas, sendo três problemas de cada operação e dois de cada posição da incógnita. O que precisa ser considerado é que cada problema foi apresentado três vezes para cada participante. Portanto, a primeira prova teve, na prática, 36 problemas, 18 por operação, 18 por estrutura semântica e 12 por posição da incógnita. Nas demais provas, foram 18 problemas, nove por operação e seis por posição da incógnita.

**Formação das classes de equivalência.** As tentativas utilizadas nas contingências de treino e de testes de discriminações condicionais foram construídas no formato do procedimento de MTS simultâneo. Nas sessões, havia um estímulo-modelo, três estímulos comparação e o comando *Aponte o igual* apresentado por meio de estímulo sonoro. O que diferia treino e teste (sondas) era a presença de *feedbacks* sobre acertos e erros: no treino havia correção e no teste não.

**Experimento 1.** Foram ensinadas as relações Balança (A) e Algarismo (B), Balança (A) e Coleção (C), bem como Algarismo (B) e Problema escrito (D), sendo verificada a emergência das relações BA, CA, BD, DC, CD, AD, DA, BC e CB. Após esse treino, foi aplicado o Pós-teste 1.

**Experimento 2.** Antes de iniciar os treinos de discriminações condicionais, foram realizadas duas sessões

para ensinar os participantes sobre o funcionamento da balança (sessões de treino preparatório) com o intuito de reduzir dificuldades observadas no primeiro experimento com essa forma de apresentação de problemas<sup>1</sup>. Foram ensinados conjuntos de relações com o objetivo de formar dois conjuntos de classes de equivalência: uma de adição e outra de subtração. Isso foi diferente do Experimento 1, no qual foi formado apenas um conjunto de classes de equivalência entre problemas de adição. Para as duas classes de estímulos, foram ensinadas as relações Balança (A) e Algarismo (B) assim como Balança (A) e Problema escrito (C), sendo observada a emergência das relações BA, CA, BC e CB.

#### Treino de solução de problemas na forma de balança.

Essa fase ocorreu apenas no Experimento 2 e consistiu na exposição dos participantes a um conjunto de 40 problemas na forma de balança, 20 para cada operação, sendo sete com incógnita na posição *a*, sete na *b* e seis na *c*, com *feedback* caracterizado por uma voz sinalizando o acerto ou o erro e, no caso de respostas corretas, o vídeo de uma balança sendo equilibrada. Ao final dessa fase, foi aplicado o Pós-teste 2.

**Ensino de algoritmos.** Para os dois experimentos, nas tentativas das sessões de instrução e treino, havia um problema que consistia na exibição de uma tentativa de MTS com estímulo-modelo, nove estímulos-comparação (sendo que apenas um era a resposta correta) e o comando *Que número deve ser colocado no lugar desta interrogação para que esta conta fique correta?*

A instrução contendo algoritmo para solução de problemas de adição com incógnita nas posições *a* (exemplo:  $?+2=4$ ) e *b* (exemplo:  $2+?=4$ ) foi: *A conta é de mais ou de menos? Isso, é de mais. Não importa onde está a interrogação. Basta você ver quanto você tem (valor na posição a ou b) e quanto falta para chegar ao resultado (valor da posição c). Quanto temos aqui (posição a ou b)? [dois dedos eram levantados]. Quanto falta para chegar neste resultado? [mais dois dedos eram levantados; o valor da incógnita correspondia à quantidade de dedos levantados].* Após o ensino desse algoritmo, no caso do Experimento 1, foi aplicado o Pós-teste 2. No Experimento 2, foi ensinado também o algoritmo de subtração.

A instrução contendo o algoritmo para solução de problemas de subtração com incógnita em *a* (exemplo:  $?-3=2$ ) foi: *A conta é de mais ou de menos? Isso, é de menos. A interrogação está no começo ou no meio? Isso, no começo. Então você vai somar este número aqui (valor da posição b, no caso 3) [três dedos eram levantados] com este outro número aqui (valor da posição c, no caso 2) [mais dois dedos eram levantados e a soma de todos os dedos correspondia ao valor da incógnita na posição a]. Para os problemas de subtração com incógnita na posição b (exemplo:  $5-?=2$ ), foi adotada a seguinte regra: *A conta é de mais ou de menos? Isso, é de menos. A interrogação está no começo ou no meio? Isso, no meio. Então você vai fazer uma conta de menos com este número aqui (valor da posição a, no caso 5) [cinco dedos eram levantados] e este outro número aqui**

<sup>1</sup> Ocorreram erros nesse tipo de problema, possivelmente porque as instruções dadas não foram suficientes para instalar o comportamento de equilibrar a balança.

(valor da posição *c*, no caso 2) [dois dedos eram abaixados e os dedos levantados que haviam sobrado eram contados para fornecer o valor da incógnita em *b*]. Após essa fase, foi aplicado o Pós-teste 3.

### Teste de Generalização.

**Experimento 1.** Consistiu em um teste com 12 problemas, seis apresentados no monitor e seis ditados, seis por operação e dois para cada posição da incógnita.

**Experimento 2.** Consistiu em um teste com 24 problemas, 12 apresentados no monitor e 12 ditados, seis por operação e dois para cada posição da incógnita. Nesse experimento, o teste de generalização foi aplicado duas vezes, uma após o treino de solução de problemas na forma de balança e outra ao final de todas as fases do experimento.

## Resultados

### Experimento 1

Após o registro de resultados positivos nos testes de avaliação de relações de equivalência envolvendo as quatro formas de apresentação de problemas, houve aumento na porcentagem de acertos em todos os tipos de problemas, porém cinco participantes tiveram dificuldades com os problemas na forma de balança. Após o procedimento adicional de ensino de algoritmos para solução de problemas com incógnitas nas posições *a* e *b*, verificou-se que quatro participantes apresentaram aumento da porcentagem de acertos no Pós-teste 2 e seis, no Pós-teste 3. Esse resultado, embora positivo, indicou que mudanças de procedimento eram necessárias de modo que todos fossem beneficiados pela fase de ensino de algoritmos. Ao final desse experimento, no teste de generalização, verificou-se que os participantes acertaram todos os problemas.

### Experimento 2

Foram empregadas três formas de apresentação de problemas: algarismos, escrita e balança. Após a formação das classes, verificou-se, no Pós-teste 1, que a porcentagem de acertos aumentou para todos os participantes. No treino de solução de problemas sob a forma de balança, também foi verificada melhora de desempenho. No Teste de generalização 1, todos os participantes alcançaram porcentagens de acerto acima de 75%. Na fase de ensino conjunto dos algoritmos de adição e subtração, verificou-se aumento na porcentagem de acertos. Em seguida, foi aplicado o Teste de generalização 2, no qual todos alcançaram 100% de acerto. Nesse experimento, a cada pós-teste, observou-se melhora de desempenho.

A análise da eficácia dos dois experimentos foi fundamentada no aumento da porcentagem de acertos dos participantes nos diferentes tipos de problemas a que foram expostos antes e depois de cada procedimento de ensino.

Em acréscimo a essa medida de eficácia das intervenções utilizadas, estima-se necessário considerar outras medidas que possam fortalecer a conclusão de que a intervenção foi efetiva. Serão apresentadas, portanto, duas análises que demonstram outras formas de indicar a eficácia dos procedimentos de ensino adotados nesses experimentos.

A Figura 2 exibe a eficiência média dos oito participantes em cada prova a que foram expostos no pré-teste e nos pós-testes dos Experimentos 1 e 2. O cálculo da eficiência foi feito pela divisão da porcentagem de acertos em cada prova nas fases de pré-teste e pós-testes pela duração total de cada uma dessas provas em cada uma dessas fases.

Os tempos de duração das sessões do Experimento 2 foram todos obtidos automaticamente. No Experimento 1, por outro lado, apenas os tempos das provas com problemas na forma de algarismo e coleção foram obtidos automaticamente. O tempo de duração das provas com problemas na forma de balança e escrita foi obtido por meio do registro do tempo inicial e final da sessão. Porém, muitas provas foram aplicadas em um mesmo dia e, em alguns casos, não foi feito o registro do horário final de término das sessões. Portanto, os dados referentes a essas duas provas devem ser analisados com cuidado porque não são precisos; eles foram inseridos apenas para fornecer um parâmetro de qual pode ter sido a eficiência dos participantes nessas provas.

No Experimento 1, as provas com problemas nas formas de algarismo, coleção e balança tinham seis problemas e as

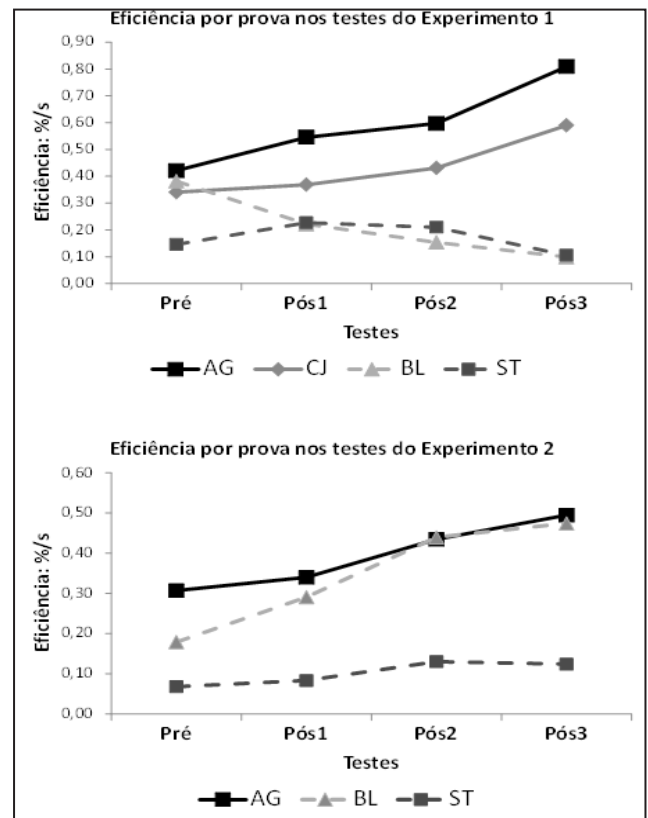


Figura 2. Eficiência (divisão da porcentagem de acertos por segundo) nas provas dos Experimentos 1 e 2. AG = Problemas na forma de algarismos; BL = Problemas na forma de balança; ST = Problemas escritos; CL = Problemas na forma de coleção

provas com problemas escritos tinham 18. No Experimento 2, as provas com problemas na forma de algarismo e balança tinham 18 problemas cada uma e a prova com problemas escritos, 36 problemas.

Ao comparar a eficiência nas diferentes formas de apresentação, é preciso considerar que, nos dois experimentos, a duração média das provas com problemas escritos foi maior. Ao comparar os dois experimentos, deve-se observar que, em termos absolutos, as provas do Experimento 2 tiveram maior duração porque tinham um número maior de problemas.

Esses dados devem ser analisados da seguinte forma: a eficiência é o resultado da divisão entre porcentagem de acertos e tempo para solucionar o problema. Portanto, será melhor a eficiência quanto maior for o numerador (porcentagem de acertos) e menor for o denominador (tempo para solucionar o problema). Portanto, apenas para exemplificar: em um cenário irreal no qual o participante alcance 1% de acerto e demore 60 segundos para resolver os problemas, a eficiência é de 0,0001. De outro lado, o participante que apresenta 100% de acertos e demora 1 segundo para solucionar todos os problemas, fica com eficiência igual a 1. Portanto, quanto maior o valor da eficiência, melhor o desempenho do participante, pois significa que ele acertou mais em menos tempo.

Nota-se na Figura 2 que, nos dois experimentos, quando comparada às demais provas, a eficiência na solução dos problemas na forma de algarismos (AG) já no pré-teste é elevada (Experimento 1: 0,42; Experimento 2: 0,31). Ao longo dos dois experimentos, essa eficiência aumenta de modo uniforme, alcançando, no Experimento 1, 0,81 de eficiência e, no Experimento 2, 0,49. No Experimento 1, a eficiência nas provas com problemas na forma de coleção (CJ) também cresceu de modo uniforme, mas com uma aceleração menor: no pré-teste, a eficiência foi de 0,34 e, no último pós-teste, foi de 0,59.

Com relação à eficiência nos problemas na forma de balança (BL), houve diferenças entre os dois experimentos. No Experimento 1, houve uma queda de eficiência (de 0,38 no pré-teste para 0,1 no último pós-teste), enquanto no Experimento 2 houve uma tendência de aumento uniforme (de 0,18 no pré-teste para 0,47 no último pós-teste). A eficiência na prova com problemas na forma de coleção do Experimento 1 é similar à eficiência na prova com problemas na forma de balança do Experimento 2, indicando que as dificuldades com a balança encontradas no Experimento 1 foram contornadas no segundo experimento.

Nos problemas escritos (ST), verificou-se, no Experimento 1, um crescimento inicial da eficiência seguido por uma queda (de 0,15 no pré-teste para 0,23 no Pós-teste 1 e 0,11 no Pós-teste 3). No Experimento 2, ocorreram aumentos entre o pré-teste e Pós-teste 1 (0,07 para 0,08) e entre os pós-testes 1 e 2 (0,08 para 0,13). Entre os pós-testes 2 e 3, houve uma pequena queda de eficiência (de 0,13 para 0,12). De todo modo, as diferenças encontradas foram sempre muito pequenas, em um padrão semelhante ao do Experimento 1.

A Figura 3 exibe, na parte superior, a diferença média de erros entre os pré-teste e pós-testes. A parte inferior da figura exibe essa mesma informação, mas destacando os erros médios por prova.

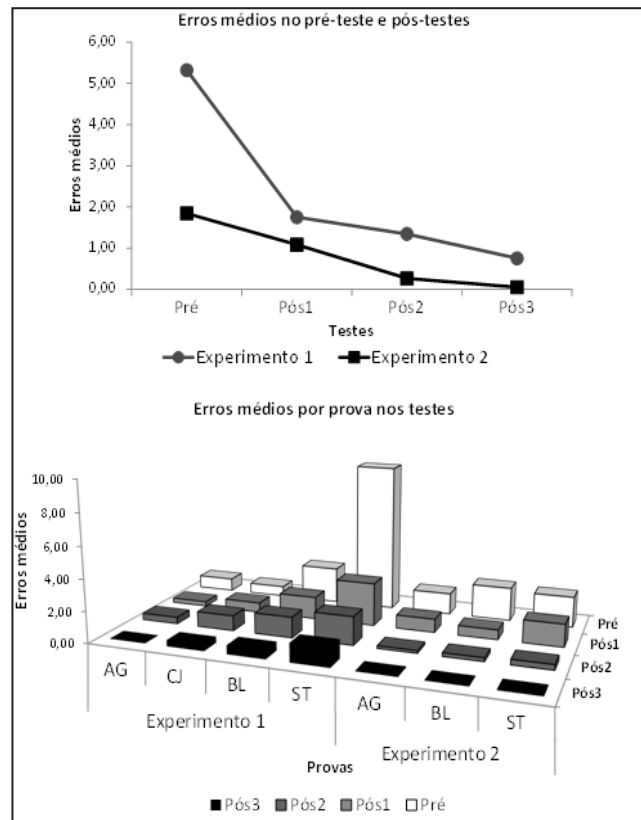


Figura 3. Diferenças entre os erros médios entre pré-teste e pós-testes (figura superior) e diferenças entre os erros médios entre pré-teste e pós-testes por prova (figura inferior)

O termo erro é utilizado neste texto em um sentido específico: significa a diferença matemática entre a resposta dada pelo participante e a resposta correta. Os erros médios são o resultado do cálculo da média aritmética dos erros de todos os participantes diante de cada condição apresentada na Figura 1 (por teste, figura superior, ou por prova em cada teste, figura inferior). Por exemplo, se a resposta correta era 10 e o participante respondeu 3, o erro foi de 7. Da mesma forma, se a resposta foi 17, o erro também foi 7.

Nota-se, na porção superior da Figura 3, que, no Experimento 1, os participantes apresentaram erros maiores que os participantes do Experimento 2, sendo que, nos dois experimentos, houve uma redução uniforme dos erros entre o pré-teste e o último pós-teste. Na porção inferior da Figura 3, é possível verificar que essa redução de erros aconteceu em todas as provas nos dois experimentos. Em ambos os experimentos, de um modo geral, os erros foram maiores em problemas escritos e na forma de balança.

Associadas com os dados de aumento da porcentagem de acertos, cuja diferença entre pré e pós-testes foi estatisticamente significativa (Experimento 1: teste de Friedman,  $X^2 = 12,3$ ,  $p = 0,003$ ; Experimento 2: teste de Friedman,  $X^2 = 24$ ,  $p = 0,001$ ), as medidas de eficiência e de erro fortaleceram a conclusão de que os procedimentos de ensino adotados foram eficazes. De acordo com os dados apresentados, houve aumento da porcentagem de acertos, mas ela não foi acompanhada por um aumento ainda maior no tempo de solução dos problemas, ou seja, os participantes não só passaram a acertar mais, como fizeram isso em um período

de tempo adequado. Da mesma forma, verificou-se que, em média, mesmo nos casos em que a resposta-solução não foi a correta, a diferença foi pequena e, ao longo do experimento, aproximou-se de zero, sugerindo maior atenção, engajamento e conhecimento de como resolver a tarefa.

## Discussão

O emprego de novas medidas em pesquisas experimentais pode ampliar as possibilidades de análises de processos e fenômenos psicológicos. Starkey e Cooper (1980), por exemplo, elegeram a duração do olhar em bebês ao investigar a capacidade inata de seres humanos para discriminar variações numéricas no ambiente. Barnes-Holmes et al. (2006), por seu turno, utilizaram o tempo de reação para identificar atitudes implícitas de uma pessoa em relação a um estímulo. No campo dos procedimentos de ensino de repertórios acadêmicos, novas medidas também são necessárias para se avaliar a ocorrência e a qualidade da aprendizagem.

No presente artigo foram propostas duas formas de avaliar a eficácia de uma intervenção, empregando os dados de porcentagens de acerto por tempo de solução das provas (eficiência) e diferença entre resposta dada e resposta correta. Esse tipo de análise possibilitou ampliar a visibilidade de medidas operantes relevantes em tarefas de solução de problemas a partir dos dados geralmente coletados na maioria das pesquisas em Psicologia.

Neste artigo, argumenta-se que a elevação da quantidade de acertos considerando o tempo de solução expressa um parâmetro positivo para a avaliação da eficácia de um procedimento de ensino. Essa relação entre quantidade de acertos e tempo de solução dos problemas foi designada como eficiência, sugerindo que, além do acerto, mostra-se também importante considerar a variação da quantidade de soluções corretas nos intervalos de tempo disponíveis. A análise de eficiência exposta neste artigo documentou estratégias que estendem a visibilidade de propriedades dos resultados que, a princípio, expressavam somente um aumento nos percentuais de acerto após a exposição aos procedimentos de ensino e de avaliação. As medidas de eficiência obtidas evidenciam que o aumento nos percentuais de acerto foi acompanhado pelo crescimento da taxa de respostas-soluções corretas no tempo disponível para a solução.

De modo complementar, foi considerada a diferença entre a resposta dada e a resposta-solução esperada. A diminuição dessa diferença parece indicar que não só os participantes passaram a acertar mais após a exposição aos procedimentos de ensino como também, quando erraram, o erro não foi do mesmo tipo observado ao início do experimento. Admite-se que a redução do erro ou da diferença entre a resposta emitida e a resposta esperada expressa a eficácia dos procedimentos de ensino adotados.

É preciso destacar que essas propostas de análise não suplantam a avaliação de acertos e erros no início e ao final do experimento. São medidas complementares. Da mesma forma, cabe reconhecer que as análises expostas neste artigo, por si só, igualmente não representam garantia de uma avaliação completa da eficácia de uma intervenção.

Novas avaliações podem ser propostas, sendo que o próprio delineamento da intervenção é fundamental para a construção da inferência de que houve sucesso.

É preciso lembrar que as análises de acerto, bem como as análises de eficiência e erro, concentram-se nos resultados ou produtos do comportamento. Faz-se necessário avaliar também o processo de aprendizagem ao longo da intervenção. Por exemplo: quais seriam as possíveis regras que exerceram controle sobre o comportamento do aprendiz? O que significam, nessa perspectiva, os erros cometidos pelo aprendiz em termos de relações de controle previstas, tanto quanto das relações espúrias? Quão reforçadora foi a tarefa de resolver problemas, de modo que se torne mais provável que, no futuro, a criança continue emitindo esse comportamento? A avaliação do processo, portanto, também é um aspecto crucial para determinar a eficácia de uma intervenção. Esse tema fugiu ao escopo deste trabalho, mas merece tratamento detalhado e exige medidas específicas.

De um ponto de vista comportamental, ensinar é produzir aprendizagem, e aprender é o nome dado à mudança de comportamento (Kubo & Botomé, 2001). Portanto, se a pessoa “não sabia fazer algo” e alguém, por meio de um procedimento, fez com que passasse a “saber como fazer”, pode-se afirmar que houve mudança de comportamento (aprendizagem). Consequentemente, é lícito classificar o procedimento adotado como ensino. Logo, ensino não se define pelas ações do professor em relação ao aluno, mas pelos efeitos de suas ações sobre o comportamento do aluno.

Por esse motivo, o ensino formal deve ser planejado para que de fato seja eficaz no sentido skinneriano, ou seja, para que de fato produza mudança de comportamento no sentido de gerar consequências reforçadoras (Mazzo & Gongora, 2009). É preciso conhecer o que se espera que o aluno seja capaz de fazer após o término do procedimento ou programa educacional (Henklain & Carmo, 2013c). Quando se conhece o objetivo terminal a ser alcançado, avaliar a eficácia fica muito mais fácil.

Resnick et al. (1973), por exemplo, desenvolveram um programa para o ensino de comportamentos matemáticos. Isso foi feito por meio da análise dos comportamentos que compõem (a) o conceito de número, (b) a contagem, (c) a comparação de conjuntos, (d) a seriação e a ordenação e (e) a soma e a subtração, e da indicação da sequência de ensino e da hierarquia de aprendizagem que pode ser adotada com base em uma análise de quais são os comportamentos mais básicos e mais complexos na aprendizagem desses conceitos.

O presente estudo sugere que não está claro tudo o que se espera do aluno. Depois que as medidas de eficiência e erro foram apresentadas, ficou evidente que acertar a solução do problema e demorar demais para conseguir isso não é o que se espera dos alunos. Da mesma forma, pequenas diferenças entre a resposta fornecida e a correta podem ser toleradas, mas grandes diferenças são caracterizadas como erros “grotescos” e não são aceitas. Nesse sentido, Saville, Lambert e Robertson (2011) defendem que a educação seja baseada em evidências e argumentam que muitas propostas baseadas na psicologia operante de Skinner atendem a esse critério, mas não são postas em prática. O presente artigo demonstra a utilidade de outras formas de avaliar a eficácia do ensino e sugere que

as mesmas sejam aplicadas, por exemplo, na avaliação do ensino de matemática.

Planejamento e avaliação de eficácia são, portanto, processos essenciais ao ensino. A avaliação da eficácia de procedimentos de intervenção adotados para o estudo de processos comportamentais constitui-se em condição crucial para ampliar a visibilidade de dimensões operantes que se mostram sensíveis às características de tais procedimentos. A ampliação de tal visibilidade apresenta-se como condição necessária para o aprimoramento das condições experimentais comprometidas com a investigação de processos comportamentais relevantes no âmbito da solução de problemas, tanto para fomentar reflexões dos professores sobre procedimentos de ensino no contexto da alfabetização matemática.

## Referências

- Araújo, C. H., & Luzio, N. (2005). *Avaliação da educação básica: Em busca da qualidade e equidade no Brasil*. Brasília: MEC/ INEP.
- Barnes-Holmes, D.T., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you really know what you believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, 32(7), 169-177.
- Bryant, P. (2013). Children's understanding and use of inversion in arithmetic. *Cuadernos de Investigación y Formación en Educación Matemática*, 8(11), 231-238.
- Carmo, J. S. (2002). *Comportamento conceitual numérico: Um modelo de rede de relações equivalentes* (Tese de doutorado não publicada). Programa de Pós-graduação em Educação, Universidade Federal de São Carlos, São Carlos, SP.
- Carmo, J. S., & Galvão, O. F. (1999). Aquisição do conceito de número em crianças pré-escolares através do ensino de relações condicionais e generalização. In J. S. Carmo, L. C. C. Silva, & R. M. E. Figueiredo (Orgs.), *Dificuldades de aprendizagem no ensino de leitura, escrita e conceitos matemáticos* (pp. 50-87). Belém: UNAMA.
- Carpenter, T. P., & Moser, J. M. (1983). The acquisition of addition and subtraction concepts. In R. Lesh & M. Landau (Eds.), *The acquisition of mathematical concepts and processes* (pp. 7-44). New York: Academic Press.
- Costa, A. L. M., Galvão, O. F., & Ferreira, B. P. (2008). ARIT - um software baseado em equivalência de estímulos dirigido a crianças com histórico de fracasso na aprendizagem de conceitos aritméticos. In Sociedade Brasileira de Computação (Ed.), *Anais do XIX Simpósio Brasileiro de Informática na Educação* (CD, pp. 125-134). Fortaleza: SBC.
- de Rose, J. C. (1993). Classes de estímulos: Implicações para uma análise comportamental da cognição. *Psicologia: Teoria e Pesquisa*, 9(2), 283-303.
- de Rose, J. C. (2005). Análise comportamental da aprendizagem de leitura e escrita. *Revista Brasileira de Análise do Comportamento*, 1(1), 29-50.
- Escobal, G., Rossit, R., & Goyos, A. C. N. (2010). Aquisição de conceito de número por pessoas com deficiência intelectual. *Psicologia em Estudo*, 15, 467-475.
- Fayol, M. (1992). From number to numbers in use: Solving arithmetic problems. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 283-306). New Jersey: Lawrence Erlbaum Associates.
- Fossa, J. A., & Sá, P. F. (2008). Uma distinção entre problemas aritméticos e algébricos. *Revista Educação em Questão*, 33(19), 253-278.
- Green, G. (2010). A tecnologia de controle de estímulo no ensino de equivalências número quantidade. In J. S. Carmo & P. S. T. Prado (Orgs.), *Relações simbólicas e aprendizagem da matemática* (1a ed., pp. 159-172). Santo André, SP: ESETec.
- Mazzo, I. M. B., & Gongora, M. A. N. (2009). Conceito skinneriano de comportamento eficaz. *Interação em Psicologia*, 13(2), 229-240.
- Haydu, V. B. H., Costa, L. P. C., & Pullin, E. M. M. P. (2006). Resolução de problemas aritméticos: Efeito de relações de equivalência entre três diferentes formas de apresentação dos problemas. *Psicologia: Teoria e Pesquisa*, 19(1), 44-52.
- Haydu, V. B., Pullin, E. M. M. P., Iégas, A. L. F., & Costa, L. P. (2010). Solucionar problemas aritméticos: Contribuições da análise do comportamento. In J. S. Carmo & P. S. T. do Prado (Orgs.), *Relações simbólicas e aprendizagem da matemática* (pp. 159-172). Santo André, SP: ESETec.
- Henklain, M. H. O. H., & Carmo, J. S. (2013a). Equivalência de estímulos e redução de dificuldades na solução de problemas de adição e subtração. *Psicologia: Teoria e Pesquisa*, 29(3), 341-350.
- Henklain, M. H. O. H., & Carmo, J. S. (2013b). Stimulus equivalence and increase of correct responses in addition and subtraction problems. *Paidéia*, 23(56), 349-358.
- Henklain, M. H. O., & Carmo, J. S. (2013c). Contribuições da análise do comportamento à educação: Um convite ao diálogo. *Cadernos de Pesquisa*, 43(149), 704-723.
- Hiebert, J. (1982). The position of the unknown set and children's solutions of verbal arithmetic. *Journal for Research in Mathematics Education*, 13(5), 341-349.
- Iégas, A. L. F. (2003). *Software para a resolução de problemas aritméticos: O modelo da balança* (Dissertação de mestrado não publicada). Programa de Educação, Universidade Estadual de Londrina, Londrina.
- Keller, F. S., & Schoenfeld, W. N. (1973). *Princípios de psicologia*. São Paulo: EPU.
- Kubo, O. M., & Botomé, S. P. (2001). Ensino-aprendizagem: Uma interação entre dois processos comportamentais. *Interação em Psicologia*, 5, 133-170.
- Magina, S. P. M., Santana, E. R. S., Cazorla, I. M., & Campos, T. M. M. (2010). As estratégias de resolução de problemas das estruturas aditivas nas quatro primeiras séries do Ensino Fundamental. *Zetetiké*, 18(34), 15-50.
- Marcicano, D. C., Carmo, J. S., & Prado, P. S. T. (2011). Software ProgMTS: Possibilidades de delineamento e condução de programas de ensino em análise experimental do comportamento. In Sociedade Brasileira de Psicologia (Ed.), *Resumos de Comunicações Científicas - 41ª Reunião Anual da Sociedade Brasileira de Psicologia: Formação e produção do conhecimento em psicologia* (CD - ISSN 2176-5243). Belém: SBP.



- Matos, M. A. (1999). Controle de estímulo condicional, formação de classes conceituais e comportamentos cognitivos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 1(2), 159-178.
- McIlvane, W. J., Galvão, O. F., Goulart, P. R. K., Brino, A. L. F., & Barros, R. S. (2005). Variáveis de procedimento na pesquisa sobre classes de equivalência: Contribuições para o estudo do comportamento simbólico. *Revista Brasileira de Análise do Comportamento*, 1(1), 15-27.
- Neef, N. A., Nelles, D. E., Iwata, B. A., & Page, T. J. (2003). Analysis of precurrent skills in solving mathematics story problems. *Journal of Applied Behavior Analysis*, 36(1), 21-33.
- Nunes, T., & Bryant, P. (1996). Giving meaning to addition and subtraction. In T. Nunes & P. Bryant (Eds.), *Children doing mathematics* (pp. 114-141). Oxford: Blackwell.
- Prado, P. S. T., & de Rose, J. C. (1999). Conceito de número: Uma contribuição da Análise Comportamental da Cognição. *Psicologia: Teoria e Pesquisa*, 15(3), 227-235.
- Resnick, L. B., & Rosenthal, D. J. A. (1974). Children's solution processes in arithmetic word problems. *Journal of Educational Psychology*, 66(6), 817-825.
- Resnick, L. B., Wang, M. C., & Kaplan, J. (1973). Task analysis in curriculum design: A hierarchically sequenced introductory mathematics curriculum. *Journal of Applied Behavior Analysis*, 6(4), 679-710.
- Starkey, P., & Cooper-Jr., R. G. (1980). Perception of numbers by human infants. *Science*, 210(28), 1033-1034.
- Saville, B. K., Lambert, T., Robertson, S. (2011). Interteaching: Bringing behavioral education into the 21st century. *The Psychological Record*, 61, 153-166.
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.
- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior*, 37(1), 5-22.
- Vasconcelos, L. (1998). Problemas de adição e subtração: modelos teóricos e práticas de ensino. In A. Schliemann & D. W. Carraher (Orgs.), *A compreensão de conceitos aritméticos: Ensino e pesquisa* (pp. 32-45). Campinas, SP: Papirus.
- Verschaffel, L., & De Corte, E. (1997). Word problems: A vehicle for promoting authentic mathematical understanding and problem solving in the primary school? In T. Nunes & P. Bryant (Eds.), *Learning and teaching mathematics: An international perspective* (pp. 69-97). Hove: Psychology Press

Recebido em 03.08.2014

Primeira decisão editorial em 27.02.2015

Versão final em 27.03.2015

Aceito em 27.03.2015 ■