

O Uso de Medidas Intelectuais na Análise do Monitoramento Metacognitivo de Crianças

Marília Zampieri¹
Patrícia Waltz Schelini
Universidade Federal de São Carlos

RESUMO – O monitoramento metacognitivo é uma das habilidades compreendidas pela metacognição. É avaliado utilizando-se instrumentos de autorrelato ou pela formulação de julgamentos. O objetivo do presente estudo foi investigar o monitoramento metacognitivo de crianças durante a realização de três subtestes de uma bateria de inteligência, cujo referencial é o Modelo Cattell-Horn-Carroll. Os participantes responderam aos subtestes e foram solicitados a emitir estimativas sobre seu desempenho. Os resultados indicaram que a amostra já apresentava habilidades de monitoramento cognitivo e algumas medidas de monitoramento mostraram-se significativamente melhores para o subteste Desempenho em Matemática. Os dados são relevantes para confirmar, na população nacional, as informações da literatura internacional, e também para discutir a importância do incentivo e estímulo ao treinamento das habilidades metacognitivas.

Palavras-chave: metacognição, monitoramento metacognitivo, medidas metacognitivas, medidas intelectuais, modelo CHC

The Use of Intellectual Measures in the Analysis of Metacognitive Monitoring of Children

ABSTRACT – Metacognitive monitoring is one component of metacognition. It is assessed either through self-report instruments or judgments. The aim of the present study was to investigate metacognitive monitoring of children during the execution of three subtests of an intelligence battery, based on the Cattell-Horn-Carroll intelligence model. Participants answered the subtests and were asked to estimate their performance. Results indicate that participants already had metacognitive abilities in their repertoire, and some monitoring measures were significantly better for the subtest Performance on Mathematics. The data of this study are relevant in confirming information from the international literature using a Brazilian sample, and in discussing the importance of incentives for metacognitive training.

Keywords: metacognition, metacognitive monitoring, metacognitive measures, intellectual measures, CHC model

A metacognição corresponde a uma das ramificações no estudo da cognição e pode ser definida como o conhecimento, cognição e compreensão do indivíduo acerca dos seus próprios fenômenos cognitivos (Flavell, 1979; Jou & Sperb, 2006; Schwartz & Perfect, 2002). O conhecimento sobre as próprias habilidades é útil para compreender e avaliar o rendimento durante a realização de tarefas (Schraw, 1998).

Dois modelos são importantes no estudo da metacognição: o modelo de Monitoramento Metacognitivo de Flavell (Flavell, 1979) e o modelo de Nelson e Narens (1994). Ambos consideram a importância da habilidade designada como monitoramento metacognitivo, que corresponde à capacidade de observar, refletir e experienciar o andamento dos processos cognitivos e, com isso, julgar ou caracterizar o funcionamento cognitivo (Schwartz & Perfect, 2002; Son & Schwartz, 2002).

O monitoramento metacognitivo permite que o indivíduo faça estimativas sobre seu desempenho cognitivo e sobre o grau de ajustamento desse desempenho em relação à demanda apresentada pelas tarefas. As informações obtidas por meio

do monitoramento metacognitivo fundamentam as decisões sobre escolha e mudança de estratégias, que correspondem ao controle metacognitivo, outra habilidade importante compreendida pela metacognição (Schwartz & Perfect, 2002; Son & Schwartz, 2002).

O produto do monitoramento metacognitivo é chamado de *julgamento* (Efklides, 2006; Nelson & Narens, 1994; Son & Schwartz, 2002). Os julgamentos podem ser emitidos pelos indivíduos em três momentos distintos: antes, durante ou após a realização de uma tarefa. Julgamentos feitos antes da realização das tarefas (prospectivos) correspondem às estimativas sobre a realização da tarefa. Durante a realização de uma tarefa, o indivíduo pode estimar se seu desempenho é satisfatório e adequado para atingir os objetivos pretendidos. Os julgamentos feitos antes ou durante as tarefas são chamados de: *ease-of-learning* (facilidade de aprendizagem), *judgments of learning* (julgamentos de aprendizagem), *feeling-of-knowing* (sentimento de conhecer). Os julgamentos feitos após a conclusão da tarefa são chamados de julgamentos de confiança e são designados como julgamentos retrospectivos. Os julgamentos de confiança são assim chamados por refletirem a estimativa do indivíduo sobre a probabilidade de acerto de uma tarefa já realizada, ou seja, a probabilidade de ter alcançado o objetivo da tarefa (Huff & Nietfeld, 2009; Son & Schwartz, 2002).

¹ Endereço para correspondência: LADHECO - Laboratório de Desenvolvimento Humano e Cognição. Rodovia Washington Luis, Km 235. São Carlos – SP. CEP 13565-905.
E-mail: ma.zampieri@uol.com.br; pws@ufscar.br

O monitoramento metacognitivo pode ser estudado e avaliado por meio da solicitação dos julgamentos ou por meio de instrumentos de autorrelato. Isto porque o uso de estratégias metacognitivas para o exercício do monitoramento não é um comportamento observável (Garret, Alman, Gardner, & Born, 2007). Um critério de categorização da avaliação do monitoramento metacognitivo é o momento de sua realização: a avaliação *on-line* compreende a avaliação concomitante à execução das tarefas com demanda cognitiva, e avaliação *off-line* corresponde às avaliações feitas em um momento diferente da realização da atividade cujo monitoramento se deseja investigar (Desoete, 2008).

Diante das limitações que cada tipo de medida pode apresentar, torna-se prudente escolher a modalidade de avaliação de acordo com a tarefa a ser realizada, minimizando a interferência da avaliação sobre sua execução. Uma alternativa para contornar as limitações é considerar a possibilidade de realizar uma avaliação da metacognição por meio de medidas variadas (Desoete, 2008).

Para que o julgamento feito pelo indivíduo seja uma ferramenta útil na avaliação do funcionamento metacognitivo, é necessário compará-lo com o desempenho de fato apresentado pelo indivíduo (Pieschl, 2009). Algumas medidas que comparam os julgamentos com o desempenho real têm sido utilizadas: acurácia absoluta, acurácia relativa, viés, dispersão e discriminação (Schraw, 2009). Três destas medidas serão descritas em função de sua importância para o presente estudo.

A acurácia absoluta, ou índice de calibração (*calibration*), avalia a precisão dos julgamentos que o indivíduo faz sobre seu desempenho. É obtida elevando-se ao quadrado a diferença entre o julgamento de confiança de acerto e o desempenho apresentado pelo indivíduo. Este valor, portanto, varia de 0 a 1. Quanto menor a diferença entre o desempenho estimado e o real, maior a acurácia do indivíduo (Schraw, 2009).

O viés mede o grau em que um indivíduo subestima ou superestima seu desempenho, e é obtido calculando-se a diferença entre a confiança de acerto e o desempenho real na tarefa. Este índice pode variar de -1 a 1. Índices positivos indicam que existe superestimativa do desempenho – com alto índice de confiança, porém baixa *performance*, e índices negativos indicam que existe subestimativa – ou seja, baixa estimativa e alto desempenho. A precisão do julgamento depende do valor do viés: quanto mais distante de 0, maior a discrepância entre a estimativa e o desempenho real (Pieschl, 2009; Schraw, 2009).

O índice de discriminação indica se o indivíduo é mais preciso em estimar acertos ou erros e é obtido pela diferença entre os julgamentos de confiança para as tarefas realizadas corretamente e os julgamentos relativos às tarefas executadas de forma incorreta. Valores positivos indicam maior confiança nos itens respondidos corretamente, e valores negativos indicam maior confiança de acerto dos itens respondidos incorretamente. O valor obtido indica a magnitude da diferença de precisão entre os dois grupos de itens (Schraw, 2009).

Estudos têm apontado que, em geral, os julgamentos tendem a ser mais acurados quanto melhor for o desempenho dos indivíduos nas tarefas que monitoram (Son & Schwartz, 2002; Vadhan & Stander, 1993). Além disso, alguns aspectos da tarefa exercem influência sobre a qualidade do monito-

ramento metacognitivo (Maki & McGuire, 2002). As atividades cujo monitoramento metacognitivo é avaliado por meio dos julgamentos devem conter um conjunto de itens, e não um item único, para garantir a confiabilidade do monitoramento. Além disso, uma distribuição dos itens em relação ao grau de dificuldade parece adequada para se obter informações sobre o desempenho metacognitivo do indivíduo, minimizando o risco de realizar uma avaliação parcial ou incompleta. Essas características, em geral, são encontradas em medidas padronizadas para a avaliação da inteligência.

Diante disso, as atividades utilizadas no procedimento do presente estudo são subtestes de uma bateria de avaliação da inteligência de crianças. O Modelo das Capacidades de Cattell-Horn-Carroll (CHC), pertencente à corrente psicométrica, fundamentou a elaboração de tal bateria.

No modelo CHC, a inteligência é decomposta em um conjunto de 10 fatores gerais, que se subdividem em 73 capacidades específicas. As capacidades gerais são: inteligência fluida, inteligência cristalizada, processamento visual, processamento auditivo, velocidade de processamento, memória de curto prazo, armazenamento e recuperação de longo prazo, conhecimento quantitativo, rapidez para a decisão correta e leitura-escrita (Schelini, 2002, 2006; Schelini & Wechsler, 2006). Três dessas capacidades serão abordadas no presente estudo.

A inteligência fluida (*Gf*) compreende a flexibilidade e capacidade de análise e adaptação diante de situações-problema não experienciadas previamente e pode ser considerada como o conjunto de capacidades de raciocínio sobre conteúdos abstratos (Alfonso, Flanagan, & Radwan, 2005). A inteligência cristalizada (*Gc*) é a capacidade exigida na resolução de problemas cotidianos, fazendo uso dos conhecimentos adquiridos pela experiência educacional e do contato com a cultura para ser desenvolvida (Alfonso et al., 2005). O fator conhecimento quantitativo (*Gq*) representa um conjunto adquirido de conhecimentos matemáticos, envolvendo a capacidade de usar informação quantitativa e de manipular símbolos numéricos (Primi, 2003; Schelini, 2006; Schelini & Wechsler, 2006).

Frente à importância do monitoramento metacognitivo para o ajuste e planejamento do funcionamento cognitivo dos indivíduos, somada à escassez de estudos nacionais nesta área, despertou-se o interesse e os propósitos do presente estudo. Seus objetivos foram: investigar a) o monitoramento metacognitivo de crianças diante da realização de três medidas de capacidades intelectuais (inteligência fluida, inteligência cristalizada e conhecimento quantitativo), e b) possíveis diferenças na precisão dos julgamentos de acordo com a natureza da tarefa.

Método

Participantes

Participaram do estudo 44 estudantes do 5º ano do Ensino Fundamental (4ª série) de uma escola da rede de ensino público municipal, localizada em um município do interior do estado de São Paulo. A média de idade da amostra foi de

10,4 anos ($dp = 0,89$). Quinze participantes eram do gênero masculino e 29 do gênero feminino. O critério de composição da amostra foi a série escolar, a fim de equiparar o grau de instrução e conhecimentos adquiridos pela amostra.

Local

Cada participante foi submetido individualmente ao procedimento, conduzido em uma sala disponibilizada pela instituição de ensino frequentada por eles, com cadeiras e mesa suficientes para acomodar a pesquisadora e o participante, e livre de ruídos ou de outras atividades que pudessem comprometer a condução do procedimento.

Instrumentos e materiais

Foram utilizados três subtestes da Bateria Multidimensional de Inteligência Infantil – BMI (Schelini, 2002). A BMI, destinada a crianças de 6 a 12 anos, é formada por nove testes desenvolvidos para avaliar algumas capacidades relativas ao Modelo das Capacidades Cognitivas de Cattell-Horn-Carroll. A única alteração a respeito da utilização dos subtestes em relação ao contexto usual de sua aplicação – ou seja, aplicação do conjunto de subtestes que compõem a BMI – foi a decisão de não utilizar os critérios de interrupção previstos pelo manual de aplicação da BMI, e sim os subteste completos. Essa decisão foi tomada a fim de garantir que todos os participantes realizassem o mesmo número de tarefas.

O subteste Desempenho em Matemática é formado por 37 itens e tem como objetivo avaliar a capacidade de conhecimento quantitativo (Gq) do Modelo CHC. O modo de apresentação dos itens aos participantes foi igual à forma de apresentação prevista pela aplicação padrão da BMI: os itens foram lidos oralmente pelo experimentador, com o uso do caderno de itens, e o participante teve acesso ao caderno com os itens impressos, para acompanhar a leitura. Lápis e papel também foram disponibilizados, caso desejasse utilizá-los para fazer as contas. Um exemplo de item do subteste é: “João tem dois carrinhos e o irmão dele tem três. Quantos carrinhos eles têm, juntos?”.

O subteste Vocabulário Geral é composto por 35 itens e tem o objetivo de avaliar a extensão do vocabulário e o conhecimento do significado das palavras, capacidades específicas incluídas na capacidade geral de inteligência cristalizada (Gc). Solicitou-se aos participantes que definissem oralmente cada uma das palavras lidas em voz alta pela pesquisadora. Um dos itens que compõe o subteste é “O que é *alimento*?”.

O subteste Indução é destinado à avaliação da inteligência fluida (Gf), mais precisamente da capacidade específica de indução, por meio da identificação de regras para a formação de conceitos. Composto por 28 itens, o teste consiste na apresentação de grupos de figuras que variam em três dimensões (forma, tamanho e cor), e o participante deve identificar qual dessas dimensões explica a diferença entre dois grupos de figuras. Os itens se tornam progressivamente mais difíceis: inicialmente, a diferença é explicada pela variação de apenas um aspecto, depois pela variação simultânea de duas dimen-

sões em cada um dos desenhos que compõem os grupos de figuras e, por fim, pela variação de duas dimensões, porém não necessariamente na mesma figura. Antes da apresentação dos itens do teste, as instruções são acompanhadas da execução de itens de exemplo. Além dos materiais dos subtestes, foram utilizadas folhas de registro das respostas e dos julgamentos, bem como lápis e borracha.

Procedimento

Após o consentimento formal dos responsáveis pelos participantes, cada criança foi submetida individualmente ao procedimento. Os conjuntos de atividades foram apresentados em duas ocasiões: no primeiro dia, foram realizados os subtestes Desempenho em Matemática e Vocabulário Geral; no segundo dia, foi apresentado o subteste Indução. Essa divisão foi feita para equiparar a duração de cada encontro.

No primeiro encontro, antes de iniciar as atividades dos subtestes, foi realizada uma atividade de familiarização com a emissão de julgamentos (Garret, Mazzocco, & Baker, 2006). Foram feitas perguntas, das quais o participante soubesse a resposta, como “Quantos anos você tem?”, bem como perguntas com baixa probabilidade de que ele soubesse, tais como “Quanto pesa um carro?”. Isto porque cada resposta do participante era seguida pelas perguntas: “Você acha que acertou esta resposta?” e “De 0 a 100, qual você acha que é a chance de ter acertado?”. Este segundo conjunto de perguntas teve como objetivo mostrar ao participante que é natural não ter certeza sobre respostas a algumas perguntas (Garret et al., 2006).

Após a realização desse exercício de familiarização do participante com o procedimento e com a pesquisadora, era dado início à realização do primeiro conjunto de atividades. O participante recebia a instrução de que, ao todo, faria três conjuntos de atividades, e cada item da atividade seria seguido pelas mesmas perguntas feitas no exercício de familiarização. As respostas eram registradas pela pesquisadora.

Ao final de cada subteste, também era feita a seguinte pergunta: “Destes [37, 35 ou 28] exercícios que você acabou de fazer, quantos você acha que acertou?”. Terminada cada sessão de realização dos subtestes, o participante era reconduzido à sala de aula.

Resultados

Para análise dos resultados, foram utilizados testes não-paramétricos, uma vez que os dados correspondem a medidas de julgamento, referentes, assim, a uma escala ordinal; e, além disso, não apresentaram distribuição normal.

Na avaliação do monitoramento metacognitivo, é necessário considerar tanto o desempenho dos participantes nas atividades quanto os julgamentos emitidos por eles. A porcentagem média de acertos foi maior no subteste Indução ($M = 66,8$; $dp = 5,01$), seguida pelo subteste Desempenho em Matemática ($M = 63,3$; $dp = 6,78$). A menor porcentagem de acertos foi obtida no subteste Vocabulário Geral ($M = 61,3$; $dp = 4,49$). Entretanto, o teste estatístico ANOVA de Friedman apontou que a diferença de desempenho dos participantes

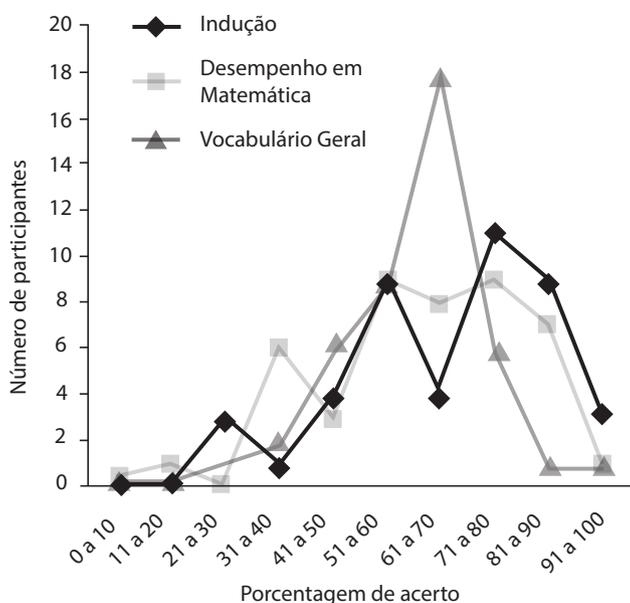


Figura 1. Distribuição da porcentagem de acertos por subteste da BMI.

nos subtestes não se mostrou estatisticamente significativa ($\chi^2 = 4,011$; $p = 0,135$; $gl = 2$).

Na Figura 1, é apresentada a distribuição da porcentagem de acertos dos participantes em cada subteste. As maiores diferenças entre as distribuições são encontradas na faixa que compreende 61 a 70% de acertos. Além disso, no subteste Vocabulário Geral houve menor número de participantes que atingiram os percentuais mais elevados de acerto.

Para cálculo dos índices de acurácia absoluta, viés e discriminação foram utilizadas pontuações para o desempenho dos participantes: a cada item dos subtestes respondido corretamente atribuiu-se um ponto; erros e omissões de resposta recebiam zero pontos. Além disso, as estimativas de acerto emitidas pelos participantes, em resposta às perguntas feitas pela pesquisadora, foram convertidas para uma escala de 0 a 1. A partir disso, eram calculados os índices para cada item dos subtestes, a partir das expressões apresentadas anteriormente (Schraw, 2009). Para a análise dos dados, foi utilizada a média aritmética de cada índice, por participante. Na Tabela 1 é apresentada a estatística descritiva desses índices, bem como da confiança de acerto, obtidos em cada subteste.

Observando-se a Tabela 1, é possível constatar que houve uma grande faixa de variação nas médias de confiança entre os participantes, com o valor mínimo de acertos bem próximo de 0 e o valor máximo chegando a 1 no subteste Indução. Apesar disso, a variância de confiança para cada subteste foi baixa. As médias de confiança indicaram que os indivíduos estavam relativamente confiantes em relação ao seu desempenho nos subtestes. A menor média de confiança foi encontrada para o subteste Vocabulário Geral e a maior média foi para o subteste Indução. A diferença das médias de confiança foi estatisticamente significativa entre os subtestes ($\chi^2 = 17,383$; $p < 0,001$; $gl = 2$).

Em relação à acurácia absoluta, o subteste Indução foi o que recebeu maior média e o subteste Desempenho em Matemática foi o que apresentou menor acurácia absoluta média. Essa diferença foi estatisticamente significativa ($\chi^2 =$

13,040; $p = 0,001$; $gl = 2$). Os valores em torno de 0,3 para os subtestes são indicativos de que a amostra conseguiu estimar com certa precisão seu desempenho. Esse índice pode variar de 0 a 1; quanto mais próximo de 1, menor a precisão dos julgamentos acerca do desempenho apresentado.

Observando-se os índices de discriminação apresentados na Tabela 1, em todos os subtestes a amostra apresentou mais precisão para prever corretamente seus acertos do que seus erros; isto é indicado pelos índices serem todos positivos. Porém, ao observar o valor dos índices, é possível constatar que tal habilidade ainda não é apresentada de forma tão consistente pelos participantes, uma vez que o valor máximo de discriminação é 1. Não foi encontrada diferença estatisticamente significativa entre os subtestes para este índice ($\chi^2 = 2,621$; $p = 0,270$; $gl = 2$).

Em relação ao viés, a média para o subteste Vocabulário Geral é um valor negativo, enquanto as médias para os outros dois subtestes são valores positivos. Isso significa que houve uma leve tendência a subestimar o desempenho, enquanto a tendência para os outros subtestes foi de superestimar. Ainda que os valores observados na Tabela 1 sejam próximos a 0 para os três subtestes, o teste de Friedman apontou diferença significativa para os índices de viés entre os três subtestes ($\chi^2 = 10,701$; $p = 0,005$; $gl = 2$).

Os participantes também foram levados a estimar o número de acertos em cada subteste. As estimativas foram convertidas em porcentagem de acertos, a fim de que fosse possível comparar os subtestes. A média mais alta da porcentagem estimada de acertos foi encontrada para o subteste Indução ($M = 72,07$; $dp = 21,23$), seguida pela média para o subteste Desempenho em Matemática ($M = 70,45$; $dp = 22,01$). A estimativa mais baixa foi para o subteste Vocabulário Geral ($M = 56,75$; $dp = 23,15$). A diferença na porcentagem de acertos estimados foi estatisticamente significativa ($\chi^2 = 13,063$; $p = 0,001$; $gl = 2$).

Discussão

Em relação ao desempenho na resolução dos itens dos subtestes, a maior porcentagem de acertos ocorreu no subteste Indução, ainda que a diferença não tenha sido estatisticamente significativa. Esse subteste avalia a inteligência fluida (Gf), uma capacidade mais dependente da maturação e desenvolvimento cognitivo do que contato com as informações (Primi, 2002). Em contrapartida, as capacidades inteligência cristalizada e conhecimento quantitativo, avaliadas, respectivamente, pelos subtestes Vocabulário Geral e Desempenho em Matemática, dependem de um conjunto de conhecimentos adquiridos, ou seja, são dependentes do acúmulo de experiências e informações. Isso pode explicar a maior taxa de erros nesses dois subtestes, possivelmente em função de os participantes não terem tido, até então, contato com o conhecimento requerido para responder aos itens. Isso porque a BMI, que contém os subtestes realizados, é destinada a crianças de até 12 anos, idade que pode corresponder a uma série escolar mais avançada e, portanto, com acesso a mais informações.

Os julgamentos de confiança, que são um produto direto do monitoramento metacognitivo, podem ser discutidos com

Tabela 1. Confiança de acerto e medidas de monitoramento metacognitivo por subteste da BMI

Índice de monitoramento metacognitivo	Medida estatística	Subteste		
		Vocabulário Geral	Desempenho em Matemática	Indução
Confiança	Média	0,64	0,72	0,76
	Desvio-padrão	0,22	0,24	0,25
	Mínimo	0,06	0,08	0,14
	Máximo	0,99	0,99	1,00
	Variância	0,05	0,06	0,06
Acurácia Absoluta	Média	0,29	0,25	0,30
	Desvio-padrão	0,12	0,12	0,16
	Mínimo	0,04	0,02	0,06
	Máximo	0,64	0,56	0,76
	Variância	0,01	0,01	0,02
Discriminação	Média	0,28	0,34	0,29
	Desvio-padrão	0,17	0,27	0,31
	Mínimo	-0,08	-0,32	-0,49
	Máximo	0,80	0,89	0,86
	Variância	0,03	0,07	0,09
Viés	Média	-0,06	0,04	0,08
	Desvio-padrão	0,28	0,23	0,30
	Mínimo	-0,73	-0,61	-0,86
	Máximo	0,43	0,48	0,70
	Variância	0,08	0,05	0,09

base em dois tipos de informação coletados: as estimativas de erro ou acerto de cada item do subteste e a formulação de uma estimativa de probabilidade (ou chance) de acerto. Partindo da média de confiança por subteste, o índice mais alto de confiança foi obtido para o subteste com melhor desempenho da amostra: Indução.

A porcentagem de acertos estimada também acompanhou o desempenho real no subteste: porcentagens mais altas foram estimadas para os subtestes em que foi observada mais alta porcentagem de acertos. Essas duas informações a respeito da confiança de acerto, seja por item respondido, seja pelo total de acertos estimado, são indícios da habilidade de monitoramento metacognitivo da amostra. Esses resultados corroboram os de estudos como o de Vadhan e Stander (1993), em que os melhores desempenhos cognitivos foram acompanhados de melhor monitoramento metacognitivo. Os autores discutem que o maior preparo dos indivíduos para realizar uma determinada tarefa oferece condições para que a avaliação do desempenho seja mais precisa.

Em relação à acurácia absoluta, as médias por subteste podem ser consideradas valores baixos; por serem valores mais próximos de 0 do que de 1; as médias de acurácia

absoluta indicam que a amostra, em termos gerais, conseguiu emitir estimativas de acerto que se assemelharam ao desempenho apresentado pelos participantes. No entanto, para essa medida, o melhor índice de acurácia foi obtido no subteste Desempenho em Matemática, que não obteve o melhor desempenho da amostra. O mesmo observou-se para o índice de discriminação.

A familiaridade com a tarefa é uma das variáveis apontadas como facilitadora da emissão de julgamentos acurados (Maki & McGuire, 2002). Diante disso, pode-se supor que as tarefas compreendidas pelo subteste Desempenho em Matemática apresentam tal familiaridade, uma vez que os itens se assemelham muito aos exercícios escolares, tanto na sua forma de apresentação – os itens do subteste são redigidos na forma de problemas – quanto na sua resolução. Além disso, a execução de tarefas de matemática na rotina escolar permite que o aluno entre em contato com as correções e devolutivas destas tarefas. É possível considerar que o contato com informações a respeito do desempenho ofereça condições para que o indivíduo compare suas experiências metacognitivas e estimativas com seu desempenho real (Maki & McGuire, 2002). Além disso, receber *feedback* a respeito

dos próprios julgamentos de confiança pode aumentar a calibração das estimativas de confiança, tornando-as mais próximas do desempenho de fato apresentado pelo indivíduo (Dunlosky & Metcalfe, 2009). A partir dessas experiências, a realização futura de tarefas semelhantes talvez possa ser favorecida com a emissão de julgamentos de confiança mais acurados e, conseqüentemente, com o aprimoramento do monitoramento metacognitivo.

Os valores obtidos para o viés também indicam que a amostra conseguiu estimar com mais precisão seu desempenho no subteste Desempenho em Matemática. Em relação ao subteste Indução, o viés apresentou uma tendência sutilmente mais forte do que o Desempenho em Matemática para a emissão de julgamentos que superestimavam o desempenho. Já no subteste Vocabulário Geral, a média do viés apontou a característica de que, nesse subteste, houve mais estimativas que subestimaram o desempenho.

A confiabilidade das medidas de monitoramento metacognitivo pode ser afetada por algumas características das tarefas realizadas, como o número de itens e o grau de dificuldade (Maki & McGuire, 2002). A opção metodológica de se aplicar todos os itens dos subtestes, adotada no presente estudo, parece contemplar a condição de apresentar aos indivíduos um conjunto de tarefas com grau de dificuldade variado, bem como um número de itens não muito baixo, criando condições para que as medidas de monitoramento metacognitivo fossem confiáveis e representativas do funcionamento metacognitivo da amostra.

Comparando-se as três medidas de monitoramento metacognitivo entre os subtestes – acurácia absoluta, discriminação e viés – é possível afirmar que um dos principais resultados do presente estudo é a constatação de que o monitoramento metacognitivo apresentou diferença em relação ao domínio de aplicação. Mais do que isso, essa habilidade se apresentou mais precisa para a capacidade de raciocínio quantitativo do que para as demais capacidades.

Vale considerar que as habilidades metacognitivas investigadas podem ser aprimoradas à medida que os participantes tiverem progressos em seu desenvolvimento cognitivo. Há autores que defendem que o baixo desempenho metacognitivo pode ser melhor explicado pela presença de habilidades metacognitivas imaturas do que pela sua ausência (Garret et al., 2006).

Partindo desse argumento, torna-se relevante discutir a viabilidade e a importância do ensino e incentivo ao uso das habilidades metacognitivas, principalmente as de monitoramento. Diversos autores têm defendido que as habilidades metacognitivas não se desenvolvem espontaneamente (Cornoldi, 2010; Garret et al., 2006; Garret et al., 2007; Panaoura & Panaoura, 2006). Somam-se a tal argumento as informações já apresentadas de que o monitoramento metacognitivo adequado acompanha o bom desempenho cognitivo. Torna-se interessante, então, considerar a possibilidade de estimulação das habilidades metacognitivas como estratégia para otimização do desempenho dos indivíduos em tarefas com demanda cognitiva.

Além disso, uma vez que as habilidades metacognitivas são componentes da aprendizagem autorregulada (Jou & Sperb, 2006; Son & Schwartz, 2002), a promoção de condições favoráveis para o seu desenvolvimento pode tornar os

indivíduos mais autônomos em seu processo de aprendizagem. Por meio do monitoramento metacognitivo acurado e preciso, os indivíduos teriam melhores condições de tomar decisões corretas a respeito de como estudar, por exemplo (Mokhtari & Reichard, 2002).

Considerações finais

O presente estudo proporcionou a constatação de que a população contemplada já apresentava habilidades de monitoramento metacognitivo. A contribuição de pesquisas subseqüentes a esse respeito pode se dar por meio da expansão da faixa etária e da série escolar dos participantes.

Outra contribuição a ser destacada é a produção de dados para a população brasileira, escassos até o momento. Mais do que isso, os dados da amostra do presente estudo confirmaram evidências apontadas na literatura internacional a respeito do monitoramento metacognitivo.

Diante disso, é possível considerar a conquista de avanços na produção de conhecimento a respeito da metacognição. Espera-se que o corpo de conhecimentos acumulados seja um fator a motivar a continuação da investigação neste campo, a fim de que os estudos investigativos produzam como benefícios a formulação de programas de intervenção e estimulação de habilidades metacognitivas. Sem dúvida, a população que pode se beneficiar de tais intervenções poderá se tornar um grupo de aprendizes mais autônomos e eficientes. Este pode ser considerado um dos objetivos atuais da educação.

Referências

- Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. P. Flanagan, & P.L. Harrison (Eds), *Contemporary intellectual assessment* (pp. 185-202). New York: The Guilford Press.
- Cornoldi, C. (2010). Metacognition, intelligence, and academic performance. In H. S. Waters, & W. Schneider (Eds), *Metacognition, strategy use, and instructions* (pp. 257- 277). New York: The Guilford Press.
- Desoete, A. (2008). Multi-method assessment of metacognitive skills in elementary school children: what you test is what you get. *Metacognition Learning*, 3(3), 189-206.
- Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Califórnia: SAGE Publications.
- Efklides, A. (2006). Metacognition and affect: what can metacognitive experiences tell us about the learning process? *Educational Research Review*, 1(1), 3-14.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: a new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911.
- Garret, J., Alman, M., Gardner, S., & Born, C. (2007). Assessing students' metacognitive skills. *American Journal of Pharmaceutical Education*, 71(1), 1-7.
- Garret, A. J., Mazzocco, M. M. M., & Baker, L. (2006). Development of the metacognitive skills of prediction and evaluation in children with or without Math disability. *Learning Disabilities Research & Practice*, 21(2), 77-88.

- Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*(2), 161-176.
- Jou, G. I. de, & Sperb, T. M. (2006). A metacognição como estratégia reguladora da aprendizagem. *Psicologia: Reflexão e Crítica, 19*(2), 177-185.
- Maki, R. H., & McGuire, M. J. (2002). Metacognition for text: findings and implications for education. In T. J. Perfect, & B. L. Schwartz (Eds), *Applied Metacognition* (pp. 39-67). Cambridge: University Press.
- Mokhtari, K., & Reichard, C. A. (2002). Assessing student's metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*(2), 249-259.
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1-25). Cambridge, MA: MIT Press.
- Panaoura, A., & Panaoura, G. (2006). Cognitive and metacognitive performance on Mathematics. In J. Novotná, H. Maraová, M. Krátká, & N. Stehliková (Eds.), *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education* (pp. 313-320). Praga: PME.
- Pieschl, S. (2009). Metacognitive calibration – an extended conceptualization and potential applications. *Metacognition Learning, 4*(1), 3-31.
- Primi, R. (2002). Inteligência fluida: definição fatorial, cognitiva e neuropsicológica. *Paidéia* (Ribeirão Preto), *12*(23), 57-75.
- Primi, R. (2003). Inteligência: Avanços nos modelos teóricos e nos instrumentos de medida. *Avaliação Psicológica, 2*(1), 67-77.
- Schelini, P. W. (2002). *Bateria Multidimensional de Inteligência Infantil: Proposta de instrumento*. (Tese de Doutorado não publicada). Pontifícia Universidade Católica de Campinas, Campinas - SP.
- Schelini, P. W. (2006). Teoria das inteligências fluida e cristalizada: Início e evolução. *Estudos de Psicologia, 11*(3), 232-332.
- Schelini, P. W., & Wechsler, S. (2006). Estudo da estrutura fatorial da Bateria Multidimensional de Inteligência Infantil. *Estudos de Psicologia, 23*(2), 105-112.
- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*(1), 113-125.
- Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition Learning, 4*(1), 33-45.
- Schwartz, B. L., & Perfect, T. J. (2002). Introduction: Toward an applied metacognition. In T. J. Perfect, & B. L. Schwartz (Eds), *Applied metacognition* (pp. 1-12). Cambridge: University Press.
- Son, L. K., & Schwartz, B. L. (2002). The relation between metacognitive monitoring and control. In T. J. Perfect, & B. L. Schwartz (Eds), *Applied metacognition* (pp.15-35). Cambridge: University Press.
- Vadhan, V., & Stander, P. (1993). Metacognitive ability and test performance among college students. *The Journal of Psychology, 128*(3), 307-309.

Recebido em 20.06.2012

Primeira decisão editorial em 20.11.2012

Versão final em 10.01.2013

Aceito em 14.01.2013 ■