

A Imprecisão Avaliativa na Correção de Provas Escritas Decorrente da Falta de Critérios Pré-Definidos

▮ Armando Terribili Filho *

▮ Andreia Melanda Chirinea **

Resumo

Pesquisa de campo realizada com dez professores que são alunos em um curso de pós-graduação stricto sensu em Educação, que visa evidenciar a relevância de se ter critérios bem definidos na avaliação de provas. A pesquisa constitui-se na simulação de uma prova realizada por dois alunos hipotéticos do ensino fundamental, que foi corrigida individualmente pelos professores. A primeira correção foi efetuada com critérios individuais; para a segunda correção, o grupo de professores discutiu, estabeleceu os critérios e efetuou nova correção individual. Os resultados apontam que a padronização de critérios reduziu as distorções de notas atribuídas. Todavia, os resultados ainda indicam que a avaliação exige uma reflexão crítica de professores e coordenadores educacionais, não somente quanto à correção de provas, e sim, no processo de avaliação e de seus instrumentos, elaboração de questões, evidenciando a importância do binômio “o que medir?” e “como medir?”.

Palavras-chave: Avaliação. Correção de provas. Critérios de correção.

1. Introdução

O Brasil possui na atualidade mais de 46 milhões de estudantes. Segundos dados do Censo da Educação Básica de 2014 (INEP, 2014a), a educação infantil (creche e pré-escola) tem 5.492.301 crianças matriculadas; o ensino fundamental, 23.446.790 e o ensino médio 7.031.624 estudantes. Deve-se acrescentar a esses números 2.985.304 de

* Pós-doutor e doutor em Educação pela UNESP de Marília (SP), mestre em Administração. Professor da Universidade Positivo e da ESIC de Curitiba (PR). E-mail: armando@impriamo.com.br.

** Mestre e doutoranda em Educação pela UNESP de Marília (SP). Professora da rede pública municipal de Bauru (SP) e professora da Universidade Sagrado Coração. E-mail: andreia.melanda@gmail.com.

matrículas no projeto de Educação de Jovens e Adultos (EJA), sendo 68% no ensino fundamental e 32% no médio. Na educação superior, conforme dados do Censo da Educação Superior de 2013, tem-se 7.526.681 matrículas na Educação Superior - cursos de graduação, sequenciais e pós-graduação *stricto sensu*. (INEP, 2014b).

Excetuando-se os alunos da educação infantil, os 41 milhões de alunos do ensino fundamental, médio e superior são submetidos frequentemente a processos avaliativos em sala de aula, a fim de identificar a efetividade da aprendizagem. É evidente que os processos avaliativos devem ser compostos por diversos e distintos instrumentos para que o aluno seja avaliado de forma holística, incluindo: provas escritas, trabalhos de pesquisa, seminários, estudos de caso, painéis, dinâmicas de grupo, dramatização, dentre outros.

Todavia, o que se constata é que a maior parte do sistema avaliativo se constituiu de provas escritas e trabalhos, quando então se calcula a média final do aluno, e, com base na média apurada (em geral, com peso maior para a prova escrita), o aluno é qualificado, sendo então, aprovado ou reprovado.

A prova escrita, além de se tornar o “grande instrumento” de avaliação apresenta inúmeros problemas, com destaque para a forma de medir o aprendizado dos estudantes, que tem como característica o processo “devolutivo” do aluno daquilo que lhe foi ensinado, ou seja, a valorização da reprodução, da memorização, limitando a análise e pouco estimulando a capacidade de dedução, de indução, de criação, de realização de análise sistêmica, de síntese e de reflexão crítica. De qualquer modo, o objetivo deste artigo é não é discutir os aspectos conceituais da prova escrita, mas mostrar o quão frágil é o processo de correção das provas.

Uma prova pode apresentar questões mal formuladas que causam dúvidas de interpretação ou até mesmo que contenham erros. Adicionalmente, o estado físico e mental do professor no momento da correção pode impactar o julgamento de valor que fará. Pior ainda, os critérios pessoais e subjetivos de avaliação do professor impactam o resultado da avaliação.

Com o objetivo de diagnosticar se o processo de correção de provas torna vulnerável a avaliação do estudante, foi realizada uma pesquisa de campo com a elaboração e correção de uma prova hipotética, com questões no nível do quarto ano

do ensino fundamental, tratando dos temas: língua portuguesa, matemática, geografia e história, cada questão com valor 2,5. Foi simulado que dois alunos (Maria Clara e João Pedro) haviam feito a prova. Essas provas foram então corrigidas pelos 10 professores que são alunos de um curso de pós-graduação *stricto sensu* em educação de uma instituição pública localizada em cidade do interior paulista. Para a prova de Maria Clara, a nota atribuída por um professor foi 2,0 e por outro professor, para a mesma prova, a nota foi 7,5, representando uma variação de 5,5 pontos! Na prova de João Pedro, a nota mínima atribuída foi 3,0 e a máxima 9,0, propiciando uma variação ainda maior, de 6,0 pontos! Não se pode ignorar que essa discrepância é preocupante, pois indica que os julgamentos efetuados estão muito distantes; pior, esses valores mínimos de 2,0 e 3,0 e máximo de 7,5 e 9,0 indicam, reprovação e aprovação de um aluno.

Para a segunda correção, para as mesmas provas, foram estipulados critérios claros e objetivos: (1) a resposta do aluno deve endereçar a questão proposta; (2) o gabarito não é binário; (3) erros de português devem ser considerados na correção; (4) aspectos pessoais do aluno devem ser ignorados; e (5) não se deve efetuar tipo algum de arredondamento.

Na correção efetuada, a variação da prova de Maria Clara diminuiu, porém, ainda ficou em patamar inadequado de 4,0 pontos. O mesmo ocorreu com a prova de João Pedro, com variação de 4,65.

Assim, se alunos do ensino fundamental realizam uma média de 6 provas mensais; os do ensino médio, 10 provas bimestrais e os alunos do ensino superior, 20 provas semestrais, poder-se-ia induzir que há um universo de injustiças cerca de 2 bilhões de provas escritas realizadas e corrigidas anualmente no país.

Ao se discorrer sobre as mudanças de enfoque na avaliação, a coleta de dados e a sistematização dos mesmos será possível explicitar melhor como a avaliação se constitui em julgamento de valor para o “bem” ou para “mal” dependendo dos critérios pré-estabelecidos pelo avaliador.

2. Avaliação: caminhos e sentidos

A avaliação escolar se constitui numa das etapas mais importantes do processo de ensino-aprendizagem e um dos mecanismos de sustentação da lógica organizacional do

trabalho e da cultura escolar. Suhr (2008) destaca que dentre as habilidades necessárias à competência docente está a de avaliar adequadamente o nível de aprendizagem dos estudantes, de modo a favorecer uma aprendizagem significativa.

É por meio da avaliação que o professor é capaz de replanejar seu trabalho para melhor adequar os objetivos do curso e da aprendizagem. Para tanto o professor precisa atribuir um juízo de valor a cada etapa do trabalho executado pelo estudante. Embora seja um instrumento importante ao trabalho do professor, dependendo do juízo de valor atribuído e do tipo de avaliação realizada, ela pode legitimar o sucesso ou o fracasso escolar.

Hadji (1994) reconhece a dificuldade de realizar uma avaliação principalmente pelo julgamento de valor que é atribuído ao trabalho desenvolvido pelo aluno e avaliado pelo professor. O principal questionamento do autor é a ambiguidade do valor atribuído, assim como a definição das variáveis a serem avaliadas.

O maior impasse para a avaliação educacional é a subjetividade presente no ato de avaliar, e conseqüentemente de atribuir uma nota ou um valor para um determinado trabalho realizado pelo aluno. Neste sentido cabe a pergunta: Como julgar um trabalho que outra pessoa fez e qual medida deve ser a utilizada para atribuir um valor, uma nota ou um conceito?

Para o autor é preciso um distanciamento do sujeito avaliador e do objeto avaliado. Este distanciamento é importante na medida em que constrói fundamentos concretos para atribuição de valor que estão além da subjetividade apenas. Para concretizar esta perspectiva de avaliação, é necessário estabelecer critérios para avaliar, ou seja, definir com clareza o que e como avaliar. (HADJI, 1994).

Hadji (1994) considera que estes critérios precisam ser claros e precisos, sob a perspectiva do avaliador considerar o processo de “vir a ser”, ou seja, considerar neste processo, o real e o ideal. Segundo o autor, avaliar é verificar se houve um salto significativo entre o real para o ideal. E somente critérios pontuais definidos preliminarmente é que são capazes de direcionar o olhar do avaliador para questões significativas. Há, portanto, uma articulação entre a realidade, ou a situação concreta observada e o que é idealmente esperado. Sob este aspecto Hadji (1994) pondera que,

Articulação, em primeiro lugar, entre o referido e o referente, visto que avaliar consiste em atribuir um “valor” [...] a uma situação real à luz de uma situação desejada, ao confrontar assim o campo da realidade concreta com o das expectativas. (HADJI, 1994, p. 32).

Partindo-se do princípio que avaliar é um processo que tenciona ações reais para ações ideais, ou seja, a avaliação pressupõe uma relação entre aquilo que existe e foi diagnosticado e o que era esperado, ou seja, a consecução dos objetivos educacionais.

Durante muito tempo a avaliação escolar teve como pressuposto, o positivismo, e seu intuito era comprovar e comparar comportamentos. Segundo Sousa (1998), muitos teóricos como: Tyler, Bloom, Gagné, Mager e Popham viam na avaliação um processo de medir objetivos. Sousa (1998) afirma que para estes autores o processo avaliativo consiste em verificar o quanto o indivíduo conseguiu alcançar dos objetivos educacionais previamente estipulados. Medindo os objetivos poder-se-ia também medir os padrões de comportamento desejáveis no aluno, uma vez que a avaliação é o principal instrumento destinado a verificar as mudanças comportamentais, assim como o grau em que estas mudanças estão ocorrendo.

Nos anos de 1980, esta vertente psicológica de avaliação passa a ser questionada por diversos autores que se debruçaram sobre o tema, tentando desenvolver uma análise crítica da avaliação exercida no cotidiano escolar. Esta análise preconiza que a avaliação é uma atividade socialmente determinada, seja pela função da escola, pela própria sociedade ou pelo tipo de homem que se quer formar. Nestes casos os determinantes sociais definem a função da escola e o instrumento para legitimar esta função acaba sendo a própria avaliação. (SOUZA, 1998).

Esta influência impelia a necessidade de construir instrumentos avaliativos capazes não só de verificar o desempenho escolar, mas, sobretudo de compreender os processos de construção das desigualdades sociais no interior da escola, com isso

A tomada de consciência pelo professor dos efeitos do nível socioeconômico dos alunos sobre os resultados de aprendizagem não implica a realização de avaliações que mascarem este fato. É necessário identificar onde o aluno apresenta maiores dificuldades, como se mostram estas defasagens e garantir uma escola e um sistema de ensino que realmente ministra educação com justiça. (SOUZA, 1998, p. 168).

Além destas questões, outro argumento utilizado pela autora é o próprio questionamento que o professor deve fazer sobre sua prática avaliativa, entendendo que esse processo é permeado de insegurança, medos, ansiedade, desejos, e que muitas vezes estes sentimentos acabam interferindo no processo de avaliar.

Luckesi (1998) aponta algumas características para o processo de avaliação entendendo que num primeiro momento a avaliação se constitui num julgamento de valor sobre determinada realidade utilizando-se critérios pré-estabelecidos pelo avaliador. O valor julgado levará em consideração a satisfação de tais critérios, assim como a realidade, que é o objeto da avaliação.

Outra característica apontada por Luckesi (1998) é o sentido do julgamento para uma tomada de decisão. Neste sentido o avaliador terá a possibilidade de reconduzir o processo com o objetivo de aprimorar os indicadores.

Segundo o autor, estas características acabam por permear processos arbitrários da autoridade pedagógica. Isto porque na avaliação efetuada pela maioria das escolas o foco incide sobre a classificação, e não no diagnóstico. Desta forma, a arbitrariedade está em determinar posições dos estudantes indicando conceitos como insatisfatório, satisfatório ou plenamente satisfatório. Esta classificação impede que a avaliação cumpra com a função de reorientar a prática pedagógica de forma mais crítica, menos subjetiva, e mais criteriosa no que se refere ao julgamento de valor e na escolha do que é relevante julgar.

Penna Firme (1994) também concorda que muitas vezes a avaliação traz consequências perversas e inúteis para o processo educativo e que é preciso entender todo o movimento realizado no decorrer de cem anos de estudo sobre o tema, procurando entender os avanços e os retrocessos relacionados à avaliação.

Tanto Penna Firme (1994) quanto Fernandes (2009) utilizam das conceituações de Guba e Lincoln (1989) ao descreverem a trajetória dos estudos sobre avaliação educacional. Segundo os autores, em um primeiro momento, entre as décadas de 1920 e 1930, não havia distinção entre avaliação e medida e os estudos sobre avaliação educacional se debruçaram sobre instrumentos e testes que pudessem verificar o rendimento escolar. A avaliação neste período era vista como algo eminentemente técnico e sua função era essencialmente classificar os estudantes para então verificar seu desempenho.

A segunda geração de estudos sobre avaliação educacional (1930-1940) foi denominada descritiva, porque procurava descrever o que seria sucesso ou fracasso tendo como referência os objetivos estabelecidos. A avaliação ainda continuava a ser vista como algo técnico. No entanto, houve um avanço no que se refere principalmente na atuação do avaliador que utilizava padrões e critérios baseados no estabelecimento de objetivos educacionais. Percebe-se que a medida que os estudos sobre avaliação vão se consubstanciando de uma geração para outra, mais se avança no entendimento e na prática do que vem a ser avaliação educacional, ao menos no plano teórico.

Ainda segundo os autores, a terceira geração de estudos sobre avaliação educacional caracterizava-se com as mesmas limitações da fase anterior, principalmente com relação à falta de definição e clareza dos objetivos educacionais. Para solucionar esta problemática, de acordo com Penna Firme (1994) e Fernandes (2009), os especialistas Stake e Screven avaliaram a necessidade de estabelecer julgamento de valor, principal característica da terceira geração. Desta forma o avaliador teria o poder de julgar, tendo como referência tanto a mensuração como a descrição, fruto dos estudos das gerações anteriores. Segundo Penna Firme (1994, p. 108): “não bastaria medir e descrever, era preciso julgar sobre o conjunto de todas as dimensões do objeto, inclusive sobre os próprios objetivos”. Esta geração propiciou vários modelos de avaliação no intuito de conceber um julgamento de valor de forma mais clara e eficiente.

A quarta e última geração, já na década de 1990, ao identificar os problemas apresentados pelas gerações anteriores, procurou ancorar seus estudos na negociação da avaliação. Sob esta perspectiva diferentes atores buscam um consenso de forma negociada, interativa e pautado sob o paradigma do construtivismo. Neste cenário, Penna Firme (1994) reforça que,

Nessa perspectiva, quanto maior a participação na escolha das preocupações, das questões avaliativas e dos métodos, e na interpretação dos resultados, maior é o nível de negociação e mais aguçada é a capacidade de responder aos interessados no processo e no objeto da avaliação. É, pois, aspecto crucial de tal envolvimento, a identificação desses interessados (PENNA FIRME, 1994, p. 109).

Portanto, a quarta geração entende que para construir uma avaliação eficaz é preciso captar variáveis que estão além das prerrogativas de controle, pois estão envolvidas neste processo aspectos culturais, sociais, políticos e éticos. Por isso a

necessidade da negociação entre observador e observado. Muito embora a quarta geração de estudos sobre avaliação educacionais tenha contemplado a avaliação negociada, ainda hoje continuam a predominar práticas avaliativas que tem como objetivo a classificação, a seleção e a certificação dos estudantes, em detrimento de práticas de buscam melhor compreender competências adquiridas pelos alunos de forma mais clara.

2.1. A prova escrita

Camargo (2011) reuniu trabalhos desenvolvidos por especialistas em Educação, a fim de discutir o tema “avaliação escrita”, que ainda é o principal instrumento de aferição na escola, com o título “Professor posto à prova”. Deste artigo com base em livros, resultados de pesquisas e artigos de Cipriano Luckesi, Dirce Aparecida Foletto de Moraes, Vasco Moretto, Bernadete Gatti, Tadeu da Ponte e Jeffrey Karpicke, podem-se extrair dez elementos importantes da avaliação.

2.1.1. A origem da prova

Segundo Cipriano Luckesi, a prova surgiu na Europa no século XVI como um recurso de coleta de dados sobre o desempenho de alunos. Na atualidade, a prova é o mais comum instrumento usado em sala de aula, pois possibilita reunir informações sobre uma grande quantidade de alunos.

2.1.2. Instrumento de documentação e comprovação

Para Dirce Aparecida Foletto de Moraes, a prova documenta e registra o conhecimento, possibilitando a representação final por um valor numérico, retratando a aprendizagem do aluno avaliado. A professora afirma que “os educadores até buscam novos caminhos, mas, por desconhecerem as diferentes ferramentas ou por sentirem-se inseguros, garantem-se na prova como instrumento comprobatório”.

2.1.3. Comunicação família, escola e criança

A prova é um histórico e reconhecido meio de comunicação entre famílias, alunos e escola. Dirce Aparecida Foletto de Moraes ratifica essa situação, quando afirma que é mais frequente um pai indagar ao filho “como foi na prova?”, do que “o que você aprendeu?”. Adicionalmente, a prova se tornou um instrumento de status entre alunos, induzindo a existência de uma hierarquia em sala de aula.

2.1.4. A prova não é um mal

Se há críticas quanto à prova pelo seu uso inadequado (controle da classe e de poder do professor sobre os alunos), Camargo (2011) afirma que a prova deve ser adequadamente aplicada, exemplificando que uma prova escrita é útil quando se restringe às competências lógico-verbais.

2.1.5. Elaboração da prova e parâmetros de correção

Tadeu da Ponte destaca a importância da clareza quando da elaboração de cada questão, a objetividade e a adequação do vocabulário à faixa etária dos alunos. Para Vasco Moretto, uma característica típica de uma boa prova é sua capacidade de estimular a aprendizagem significativa, o que requer a contextualização do que está sendo perguntado. Vasco Moretto critica a falta de parâmetros claros para a correção.

2.1.6. Instrumento de reflexão para professores

Dirce Aparecida Foletto de Moraes afirma que "o feedback possibilita que ele [o professor] se situe em relação às aprendizagens, tenha condições de entender e não apenas constatar seus erros e acertos. É preciso refletir [...]”.

2.1.7. A “decoreba” pode ser importante

Pesquisa conduzida pelo psicólogo Jeffrey Karpicke da Universidade de Purdue (Indiana, EUA) com 200 jovens universitários apontou que a "decoreba" pode impactar positivamente o desempenho dos estudantes. Segundo o estudo, a memorização ajudou

os alunos a responderem questões que exigiam deduções mais complexas e cruzamento de informações, o que obrigaria o cérebro a reorganizar o assunto e priorizar determinados tópicos. Assim, a memorização, que é considerada um método ultrapassado por muitos educadores, não pode ser abandonado. O comentário de Tadeu da Ponte sobre isso é que "há, de fato, conhecimentos que precisam ser memorizados. Por outro lado, devemos procurar oferecer situações-problema que levem o aluno a se apropriar desses conteúdos, tornando-se capaz de pensar de forma diferente, conforme os contextos que se apresentam".

2.1.8. Professor repete o modelo

Camargo (2011) comenta que Gatti e Barreto (2009), quando analisaram os currículos de cursos de pedagogia de todo o país; identificaram que, em geral, o docente reproduz práticas que vivenciou em seu curso de graduação. Nas faculdades, o método mais comum de avaliação é a prova escrita, por isso, posteriormente aplicará este mesmo método de avaliação a seus alunos. O problema é que as consequências desse despreparo remetem diretamente a questões graves da educação brasileira, como o fracasso escolar.

2.1.9. A prova – clareza, objetividade e contexto

A prova, de acordo com a síntese de Camargo (2011), deve ser planejada com antecedência, devendo ser bem elaborada e bem apresentada. Cada pergunta deve ser clara, com objetividade e intencionalidade, adequadas ao nível dos alunos, coerentes com as aulas ministradas pelo professor e contextualizadas para provocar o raciocínio do aluno.

2.1.10. A prova – o que evitar

Segundo Camargo (2011), devem ser evitadas provas exaustivas e questões com "pegadinhas", questões com uso de negativas, questões que exijam somente memorização e questões que abordem os extremos dos conteúdos estudados, ou seja,

não deve ser muito geral, nem muito específica, solicitando a informação contida em uma "nota de rodapé".

Suhr (2008) apresenta novas possibilidades quanto a elaboração de provas quanto a questões objetivas e dissertativas. Dentre as objetivas, apresenta as questões de resposta curta, de lacuna, de certo-errado, de correlação ou combinação e de múltipla escolha. Para as questões dissertativas, o autor destaca que o "verbo" pode definir o que se pretende do aluno. Por exemplo: relacionar ou enumerar (exige apenas recordação do aluno), organizar (solicita aspectos de memorização, porém com critério cronológico ou de importância), selecionar (requer a escolha fundamentada em normas de apreciação), descrever (solicita a exposição de características de um fato, objeto ou fenômeno), analisar (requer a estruturação de argumentos, relacionando-os), definir (requer que se enuncie os atributos essenciais e específicos de um fato, objeto ou fenômeno), exemplificar (exige a aplicação do conhecimento adquirido, por meio de contribuição pessoal), explicar (ênfata a relação causa e efeito), comparar (requer a análise simultânea de fatos, objetos ou fenômenos), sintetizar (exige a exposição de forma concisa uma ideia ou assunto), esquematizar (requer a elaboração de síntese organizada em tópicos e subtópicos), interpretar (requer a análise do significado de algo) e criticar (exige julgamento e elaboração de análise crítica com base em padrões ou critérios).

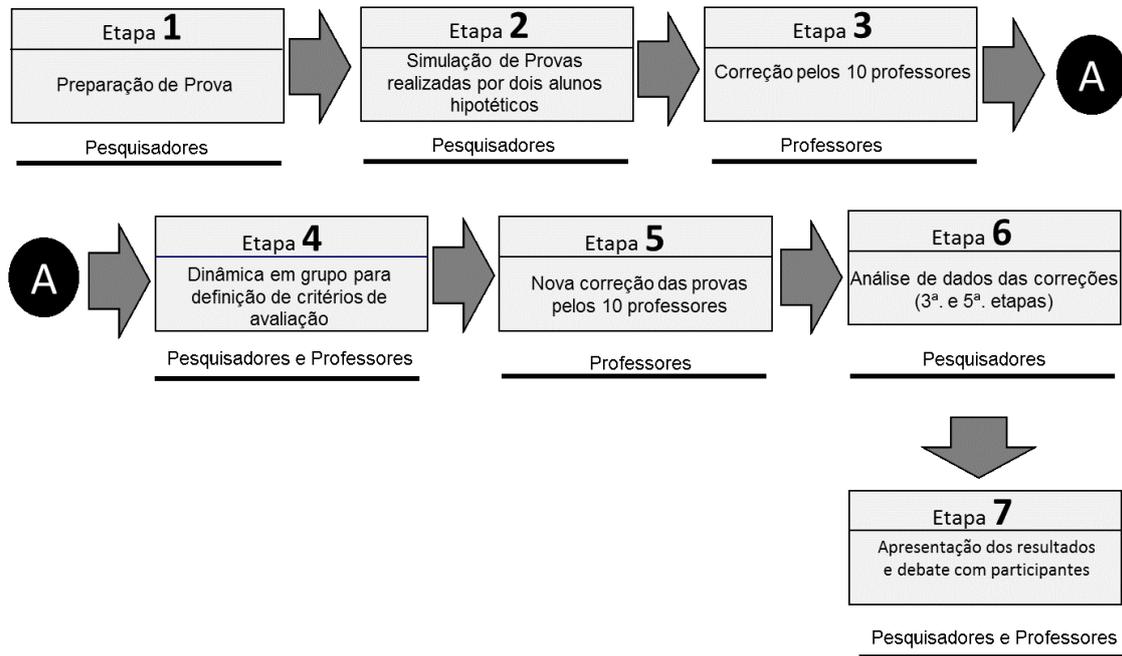
3. Coleta de dados

Com o objetivo de evidenciar que avaliação é um julgamento de valor pessoal em detrimento de práticas que buscam compreender de forma clara as competências adquiridas pelos alunos, foi realizada uma pesquisa em uma sala de aula com dez alunos de um curso de pós-graduação que atuam profissionalmente como professores.

São professores na faixa etária de 25 a 40 anos, que atuam há pelo menos três anos no ensino fundamental ou médio, com maior predominância nos anos iniciais do ensino fundamental (crianças de 6 a 10 anos).

Assim, após revisão bibliográfica do tema "avaliação", a pesquisa de campo foi dividida em sete etapas distintas e contínuas, conforme mostra a Figura 1. Na figura, cada etapa é representada por um retângulo devidamente identificado, sendo baixo de cada retângulo são apresentados os autores da referida etapa.

Figura 1 – Procedimentos metodológicos



Fonte: Os autores (2014).

A primeira etapa foi constituída da elaboração de uma prova contendo quatro questões abertas, no nível do quarto ano do ensino fundamental, tratando dos temas: língua portuguesa, matemática, geografia e história. As questões da prova, embora simples e com baixo nível de complexidade, possibilitavam intencionalmente mais de uma resposta correta, mesmo que não convencionais. As questões visavam exclusivamente verificar o rendimento escolar, que segundo Penna Firme (1994), constitui-se em uma característica dominante da primeira geração da avaliação (década de 1920). No Quadro 1, as quatro questões da prova.

Quadro 1 – Conteúdo da prova

Questão 1 – Utilizando as palavras: viagem, capital, hotel, passeio e cultura, escreva sobre uma viagem feita ao Estado de Santa Catarina.

Questão 2 – Hoje é aniversário de Daniela e Sérgio. Ele fez 13 anos e ela fez o dobro da idade dele. Quantas velinhas devo comprar para fazer a festa para eles?

Questão 3 – Dom Pedro II foi imperador do Brasil por quase 50 anos no século XIX. Pergunta-se: qual cidade ele fundou no Estado do Rio de Janeiro e quando se deu o final do Império no Brasil?

Questão 4 – Assinale no mapa abaixo, onde está a capital do país e o Rio Amazonas.



Fonte: Os autores (2014).

Nota-se que na Questão 1, o aluno poderia criar uma frase bem estruturada, mas errar o nome da capital do Estado de Santa Catarina ou omitir o nome da cidade. Além disso, poderia mencionar que a cultura predominante na região é a alemã, embora tal conhecimento não fosse requisitado ao aluno. A formulação da questão vai ao encontro da dificuldade de se realizar uma avaliação registrada por Hadji (1994) quanto à definição das variáveis a serem avaliadas.

A Questão 2 propicia diversas interpretações, pois foi mal formulada. Sérgio com 13 anos e Daniela com 26, poder-se-ia comprar: (i) 39 velinhas, sendo 13 para Sérgio e 26 para Daniela, (ii) 26 velinhas, quantidade suficiente para ambos comemorarem de forma isolada, (iii) 4 velinhas, com os números 1, 3, 2 e 6, (iv) duas velinhas, uma para cada um, ou (v) apenas uma velinha, pois poderia se ter um único bolo para os dois aniversariantes contendo uma vela única para ambos.

A Questão 3 era mais objetiva, porém exigia do aluno apenas memorização; de qualquer modo, o termo “quando se deu o final do Império” pode induzir a diferentes respostas (ano específico, período ou século).

Finalmente, a Questão 4, também causa confusão. Quanto à posição do Rio Amazonas no mapa do Brasil, o aluno poderia colocar um “x” em qualquer parte da região norte do país ou desenhar o rio, mostrando onde nasce e onde desemboca. Quanto à capital do país, o aluno poderia marcar um “x” no local aproximado, ou fazer um grande círculo na região, impedindo uma avaliação mais precisa. Na concepção de Hadji (1994) deve-se definir com clareza o que e como avaliar; neste caso, tratava-se de uma questão sem esses propósitos.

Na segunda etapa, os pesquisadores simularam a realização da prova por dois alunos (Maria Clara e João Pedro) e acrescentaram dados pessoais hipotéticos desses dois alunos, como: idade, nível socioeconômico, relações familiares e condição do estudante na disciplina em termos de necessidade de nota na referida prova para obter a aprovação. Estes fatores foram definidos, a fim de possibilitar que avaliações, com base em valores pessoais dos professores viessem à tona e pudessem impactar a correção.

Na terceira etapa foram providenciadas cópias das duas provas. Os dez participantes da pesquisa receberam uma cópia de cada prova para ser corrigida. Deve-se ressaltar que os participantes não foram instruídos de como efetuar a correção da prova, nem informados dos critérios (certo, errado, meio-certo), das particularidades (notas com ou sem arredondamento), do processo de atribuição de valor para cada questão e nem mesmo dos passos seguintes da pesquisa em curso. Num primeiro momento, alguns participantes sentiram-se desconfortáveis com a proposta dos pesquisadores reforçando o posicionamento de Souza (1998) quando afirma que a avaliação é o principal instrumento para se verificar as mudanças comportamentais do aluno e o grau com que elas ocorrem, e a pesquisa não se propunha a efetuar esta validação (e nem tinha meios para isso), apenas em qualificar as respostas dadas e quantificar as pontuações. Todavia, por serem professores aceitaram realizar a tarefa e colaborar com a pesquisa em curso, ao considerarem que “a docência começa pela

pesquisa [...] a aula é consequência da pesquisa e é vitalizada pela pesquisa”. (DEMO, 2012, p. 86-87).

Após a conclusão da correção das provas, os materiais foram recolhidos. A etapa seguinte foi a realização de uma dinâmica com o grupo para definição dos critérios de correção das provas, atividade esta que foi coordenada pelos pesquisadores junto aos participantes. Foram então definidos cinco critérios, intencionalmente, não englobando gabarito de prova ou algo similar, a fim de manter a pessoalidade de “julgamento de valor”. Os cinco critérios definidos foram:

- (1) resposta deve ser à proposta da questão (mesmo que parcial);
- (2) o gabarito é “amplo” (não é binário);
- (3) erros de português devem reduzir nota (em até 20% de cada questão);
- (4) aspectos pessoais do aluno não podem ser considerados; e
- (5) não efetuar tipo algum de arredondamento.

A quinta etapa se constituiu em efetuar uma segunda correção das provas pelos participantes, tendo por base os critérios definidos. Após o recolhimento dos materiais, houve um breve debate sobre as percepções de cada um em cada momento da pesquisa.

A sexta etapa, ocorrida a posteriori, constituiu-se na análise dos dados coletados, quando foram identificados os valores mínimos e máximos obtidos em cada prova, calculadas as médias e os desvios padrão para que se pudessem avaliar os resultados obtidos nas Etapas 3 e 5 da pesquisa (primeira e segunda correções). Os resultados foram apresentados e debatidos com os participantes da pesquisa na sétima e última etapa dessa pesquisa.

4. Análise de Dados

Como foi dito anteriormente, para o desenvolvimento da pesquisa simulou-se a realização de duas provas por dois alunos: Maria Clara e João Pedro. As provas constavam de questões de português, matemática, história e geografia.

A primeira correção das duas provas ocorreu com base em critérios pessoais de cada um dos dez professores pesquisados, sem que houvesse qualquer debate prévio

acerca de critérios e padrões. Tão logo concluída a correção das provas, as mesmas foram recolhidas.

Em seguida, houve a definição dos critérios, tendo por base o debate entre os pesquisados que chegaram ao consenso de que: o que deveria ser avaliado é se as respostas dadas pelos alunos endereçam à proposição das questões; ademais, mesmo sem discutir as possíveis respostas corretas, o grupo decidiu que deveriam ser aceitas as respostas (com sentido correto), mesmo que fugissem a um padrão convencional (gabarito). Alguns itens operacionais foram: que os erros de língua portuguesa poderiam reduzir em até 20% o valor atribuído a cada questão, que não haveria nenhum tipo de arredondamento (seja em nível de questão ou de valor final da prova), e finalmente, que os aspectos pessoais do aluno não poderiam ser considerados na avaliação.

4.1. A prova de Maria Clara

Na prova, constavam alguns dados pessoais da aluna, como: idade e nível socioeconômico (10 anos de idade; de classe média), conforme Anexo A.

4.1.1. Primeira correção

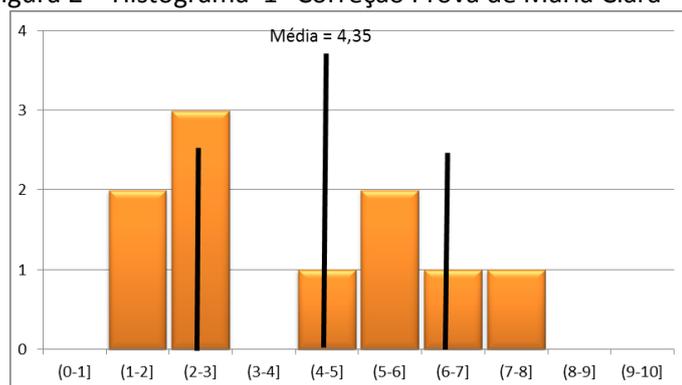
Quadro 2 – Resumo da 1ª correção da Prova de Maria Clara

Nota Mínima	2,0
Nota Máxima	7,5
Nota Média	4,35
Varição entre Nota Mínima e Máxima	5,5
Desvio Padrão	2,01

Fonte: Os autores (2014).

Deve-se destacar que para uma mesma prova, um professor atribuiu nota 2,0 (dois), enquanto outro atribuiu nota 7,5 (sete e meio), representando uma variação de 5,5 pontos. Na Figura 2 é apresentado o histograma das notas atribuídas na primeira correção para a prova de Maria Clara, incluindo a representação gráfica da média e o desvio padrão (dispersão em torno da média).

Figura 2 – Histograma 1ª Correção Prova de Maria Clara



Fonte: Os autores (2014).

4.1.2. Segunda correção

Após a definição consensuada dos critérios de correção da prova, a prova de Maria Clara foi corrigida pelos mesmos dez professores, com obtenção dos resultados apresentados no Quadro 3.

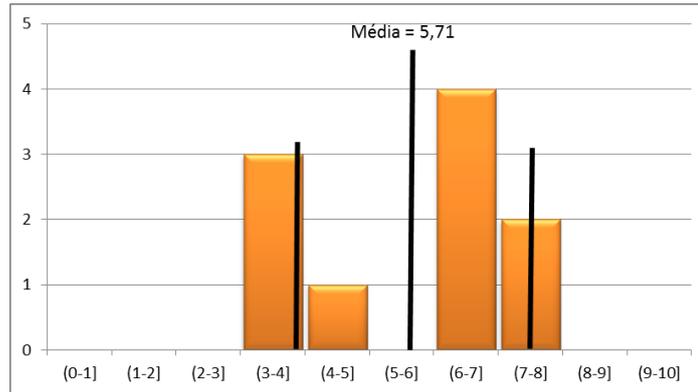
Quadro 3 – Resumo da 2ª correção da Prova de Maria Clara

Nota Mínima	3,5
Nota Máxima	7,5
Nota Média	5,71
Varição entre Nota Mínima e Máxima	4,0
Desvio Padrão	1,74

Fonte: Os autores (2014).

Deve-se destacar que para uma mesma prova, um professor atribuiu nota 3,5 (três e meio), enquanto outro, nota 7,5 (sete e meio), representando uma variação de 4,0 pontos. Na Figura 3 é apresentado o histograma das notas atribuídas na segunda correção para a prova de Maria Clara, incluindo a representação gráfica da média e o desvio padrão (dispersão em torno da média).

Figura 3 – Histograma 2ª Correção Prova de Maria Clara



Fonte: Os autores (2014).

4.1.3. Comparativo – Prova de Maria Clara

No Quadro 4 é apresentada uma síntese comparativa entre a primeira e segunda correções da prova realizada por Maria Clara

Quadro 4 – Resumo das correções da Prova de Maria Clara

Item	Correção 1	Correção 2	Variação	Comentários
Nota Mínima	2,0	3,5	1,5	Nota mínima subiu de 2,0 para 3,5.
Nota Máxima	7,5	7,5	0,0	Nota máxima não teve alteração.
Nota Média	4,35	5,71	1,36	A média subiu de 4,35 para 5,71.
Variação entre Nota Mínima e Máxima	5,5	4,0	-1,5	A variação entre a nota mínima e máxima caiu de 5,5 para 4,0.
Desvio Padrão	2,01	1,74	-0,27	O desvio padrão (dispersão em torno da média) diminuiu de 2,01 para 1,74.

Fonte: Os autores (2014).

Como se pode observar, a segunda correção foi mais consistente que a primeira. Do ponto de vista matemático, isto é evidenciado pela variabilidade dos dados, pois:

- 1) A variação entre a nota mínima e máxima (amplitude) foi reduzida em 1,5, passando de 5,5 para 4,0; e
- 2) O desvio padrão, que mede a dispersão em torno da média, foi reduzido em 0,27, passando de 2,01 para 1,74.

Adicionalmente, como benefícios acessórios que foram obtidos com a homogeneização do processo de correção, obteve-se uma elevação da média da nota em 1,36, passando de 4,35 para 5,71. Outra informação relevante, é que se fosse considerada a nota 5,0 (cinco) como sendo a mínima para aprovação, a aluna Maria Clara pela primeira correção realizada pelos professores, seria aprovada por apenas quatro deles; pela segunda correção, por seis professores.

4.2. A prova de João Pedro

Na prova, constavam alguns dados pessoais do aluno, como idade, relação familiar e desempenho escolar (11 anos, órfão e precisa obter nota mínima 5,0 para ser aprovado), conforme Anexo B.

De forma análoga à prova de Maria Clara, a prova de João Pedro foi corrigida por dez professores, conforme critérios individuais e sem prévio debate. Em seguida, após estabelecer de forma consensuada os critérios, foi realizada a segunda correção da prova. No Quadro 5 é apresentada uma síntese comparativa entre a primeira e segunda correções da prova realizada por João Pedro.

Quadro 5 – Resumo das correções da Prova de João Pedro

Item	Correção 1	Correção 2	Varição	Comentários
Nota Mínima	3,0	4,0	1,0	Nota mínima subiu de 3,0 para 4,0.
Nota Máxima	9,0	8,65	-0,35	Nota máxima foi reduzida de 9,0 para 8,65.
Nota Média	5,73	6,51	0,78	A média subiu de 5,73 para 6,51.
Varição entre Nota Mínima e Máxima	6,0	4,65	-1,35	A variação entre a nota mínima e máxima caiu de 6,0 para 4,65.
Desvio Padrão	1,96	1,39	-0,57	O desvio padrão (dispersão em torno da média) diminuiu de 1,96 para 1,39.

Fonte: Os autores (2014).

Como se pode observar, a segunda correção foi mais consistente que a primeira. Do ponto de vista matemático, isto é evidenciado pela variabilidade dos dados, pois:

1) A variação entre a nota mínima e máxima (amplitude) foi reduzida em 1,35, passando de 6,0 para 4,65; e

2) O desvio padrão, que mede a dispersão em torno da média, foi reduzido em 0,57, passando de 1,96 para 1,39.

Adicionalmente, como benefícios acessórios que foram obtidos com a homogeneização do processo de correção, obteve-se uma elevação da média da nota em 0,78, passando de 5,73 para 6,51. Outra informação relevante, é que se fosse considerada a nota 5,0 (cinco) como sendo a mínima para aprovação, o aluno João Pedro pela primeira correção realizada pelos professores, seria aprovado por apenas seis deles; pela segunda correção, por oito professores.

4.3. Discussão dos resultados

A discussão dos critérios a serem adotados na segunda correção da prova e a busca de consenso nos aspectos avaliados trouxeram um alargamento do processo avaliativo e uma padronização criteriosa e construtiva, que reduz a subjetividade.

Isto ficou evidente quando se identificou as reduções das oscilações entre nota mínima e máxima nas correções das mesmas provas. Na prova da aluna Maria Clara, a redução foi de 1,5 ponto, pois a variação na primeira correção era de 5,5 pontos, passando para 4,0 pontos na segunda. De modo análogo, nas correções da prova de João Pedro, obteve-se redução de 1,35 ponto na amplitude, pois passou de 6,0 para 4,65.

A redução do desvio padrão, que tem por consequência uma maior concentração em torno da média, foi outro indicativo de bons resultados. Na prova de Maria Clara a redução foi de 0,27 (passou de 2,01 para 1,74), e na prova de João Pedro, o resultado foi ainda melhor, por a redução foi de 0,57, passando de 1,96 para 1,39.

Como resultados indiretos, obteve-se elevação nas médias. Na primeira correção da prova de Maria Clara, a média era de 4,35 e na segunda, 5,71 (acréscimo de 1,36 ponto). No caso da prova de João Pedro, a média subiu em 0,78 ponto da primeira para a segunda correção, pois foram respectivamente, 5,73 e 6,51. Ademais, sendo considerado o valor 5,0 (cinco) como nota mínima para aprovação, Maria Clara seria aprovada por apenas quatro professores na primeira correção, e por seis na segunda. Já, João Pedro, na primeira correção seria aprovado por seis, e na segunda, por oito professores.

Outro ponto relevante foi “desconsiderar” as questões pessoais (sexo, idade, condições socioeconômicas, relação familiar e desempenho anterior) na avaliação, propiciando uma maior imparcialidade e senso de justiça; embora, nem sempre esse despojamento seja facilmente realizável, pois a racionalidade não é absoluta no ser humano, permeado pelas variações momentâneas de humor e emoções.

De difícil mensuração, os aspectos pessoais dos alunos podem ter impactado no processo de atribuição da nota da prova. Um comentário de um professor foi “fiquei com pena de João Pedro, pois precisava de 5,0 para ser aprovado, por isso, facilitei na correção para obtenção da nota mínima necessária para sua aprovação”.

Por fim, a elaboração de questões deve ser criteriosa e muito bem avaliada, evitando mal-entendidos e oferecendo a oportunidade para o estudante mostrar seu nível de aprendizagem. Camargo (2011) enfatiza que o planejamento de uma prova é fundamental para que seja bem elaborada. As perguntas devem ser claras, objetivas, adequadas ao nível dos alunos, coerentes com as aulas ministradas e bem contextualizadas.

5. Considerações Finais

A avaliação é um processo inerente à atividade educacional pelo qual se obtêm informações que permitem a tomada de decisões em vários níveis. A avaliação em sala de aula tem como objetivo principal verificar a aprendizagem dos alunos, muitas vezes, promovendo-os para o nível seguinte, ou retendo-os no mesmo nível, caso não tenham obtido o mínimo de conhecimento para a sua progressão. Ademais, por meio da avaliação dos alunos, podem-se identificar lacunas na aprendizagem, de modo pontual ou genérico, possibilitando ao docente a oportunidade de aperfeiçoar sua estratégia de ensino para determinados temas, ou mesmo, como aponta Luckesi (1998), permitindo aprimorar o processo de ensino-aprendizagem.

Neste cenário, torna-se imprescindível que a avaliação seja consistente com as estratégias de ensino-aprendizagem, devendo ser composta por múltiplos e diversificados instrumentos de avaliação e não somente pela tradicional prova escrita. Quando mais instrumentos existirem, quanto mais diversificados forem (trabalhos de pesquisa, seminários, painéis, debates, dramatização, discussões, experimentação,

diário de bordo, dinâmica de grupo, jogos, murais, visitas técnicas, prova oral), maior será o universo de informações do professor para fazer seu “juízo de valor” criterioso, imparcial e justo. Ademais, os múltiplos instrumentos reduzem a possibilidade de injustiças, sobretudo porque um aluno pode ter bom desempenho em um determinado tipo de atividade e em outro não. Suhr (2008) complementa a lista de técnicas e instrumentos para coleta de dados sobre a aprendizagem, como: autoavaliação, avaliação recíproca dos alunos, observação, arguição, produção de textos, defesa de projetos e trabalhos em grupos.

No caso específico da prova escrita, tão tradicional em nosso país, que tem sua existência perpetuada pelo sistema cartorial que ainda vigora no sistema educacional brasileiro, com apoio incondicional de professores, alunos e pais de alunos. A prova escrita é antes de qualquer coisa um “documento”. Documento que evidencia (ou se acredita evidenciar) o quanto um aluno sabe sobre um determinado assunto: é registro materializado que serve para se efetuar ou refutar contestações!

A avaliação de um aluno em uma determinada disciplina deveria ser holística, contínua, considerando além dos instrumentos formais, a participação do aluno, o interesse demonstrado, suas leituras, pesquisas e todos os trabalhos desenvolvidos.

Não se pode ter uma mesma prova avaliada por um professor com nota 2,0, e por outro, com nota 7,5 – como foi o caso de Maria Clara. Pior, um professor avaliando a prova de João Pedro com nota 3,0, e outro professor com nota 9,0. O estabelecimento de padrões de correção e documentados reduz as variações, ou seja, é condição necessária para a realização de uma avaliação mais criteriosa, porém, não é condição suficiente.

Os resultados da pesquisa indicam a necessidade de uma urgente reflexão crítica de professores e coordenadores educacionais no processo de avaliação e de seus instrumentos, tendo como prioridade o binômio “o que medir” e “como medir”.

Caso se queira continuar com as provas escritas, torna-se necessário que os professores sejam instruídos e capacitados em como elaborar provas para medir o conhecimento, como definir e divulgar os critérios de avaliação. Torna-se necessário também, que os professores tenham tempo para realizar as correções de provas e trabalhos, sendo remunerados por essas atividades de correção de prova, possibilitando

realizar o trabalho com critério e assertividade. Enfim, a avaliação na sala de aula por meio de provas escritas precisa ser revista em um contexto mais amplo, incluindo a formação e capacitação de professores.

Referências

- CAMARGO, P. de. Professor posto à prova. *Revista Educação*, São Paulo, v. 174, set. 2011. Disponível em: <<http://revistaeducacao.uol.com.br/textos/174/artigo236301-1.asp>>. Acesso em: 19 jul. 2015.
- DEMO, P. *Praticar ciência: metodologias do conhecimento científico*. São Paulo: Saraiva, 2012.
- FERNANDES, D. *Avaliar para aprender: fundamentos, práticas e políticas*. São Paulo: UNESP, 2009.
- GATTI, B. A.; BARRETO, E. S. de S. (Coord.). *Professores do Brasil: impasses e desafios*. Brasília, DF: UNESCO, 2009. Disponível em: <<http://unesdoc.unesco.org/images/0018/001846/184682por.pdf>>. Acesso em: 19 jul. 2015.
- GUBA, E. G.; LINCOLN, Y. S. *Fourth generation evaluation*. Newbury Park: Sage Publications, 1989.
- HADJI, C. *A avaliação, regras do jogo: das intenções aos instrumentos*. 4. ed. Porto: Porto, 1994.
- INEP. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. *Censo Escolar da Educação Básica 2014*. Brasília, DF: INEP, 2014a. Disponível em: <<http://portal.inep.gov.br/basica-censo>>. Acesso em: 22 mar. 2015.
- _____. *Censo da Educação Superior 2013*. Brasília, DF: INEP, 2014b. Disponível em: <http://download.inep.gov.br/educacao_superior/censo_superior/apresentacao/2014/coletiva_censo_superior_2013.pdf>. Acesso em: 22 mar. 2015.
- LUCKESI, C. C. *Avaliação da aprendizagem escolar: estudos e proposições*. 8. ed. São Paulo: Cortez, 1998.
- PENNA FIRME, T. Avaliação: tendências e tendenciosidade. *Ensaio*, Rio de Janeiro, v. 1, n. 2, p. 105-115, jan./mar. 1994.
- SOUZA, C. P. de. Descrição de uma trajetória na/da avaliação educacional. *Série Ideias*, São Paulo, n. 30, p. 161-174, 1998.
- SUHR, I. R. F. *Processo avaliativo no ensino superior*. Curitiba: IBPEX, 2008.

Recebido em: 22/03/2015

Aceito para publicação em: 15/06/2015

Inaccurate Evaluation in the Correction of Written Exams Resulting from Lack of Pre-Defined Criteria

Abstract

Field research conducted with ten teachers who are students in a stricto sensu postgraduate course in Education, which aims to highlight the importance of having well-defined criteria in test evaluation. The research was based on a careful simulation of a test applied to two hypothetical elementary school students, which were graded by individual teachers. The first correction applied individual criteria; for the second correction, the group of teachers discussed and established the criteria and made another individual grading. The results showed significant improvements in the assessment, since the distortion of marks awarded was reduced. However, the results also indicate that the evaluation requires a critical reflection of teachers and educational coordinators, not only for test grading, but in the process of evaluation and its instruments, development issues, highlighting the importance of the binomial "what to measure?" and "how to measure?".

Keywords: Review. Test grading. Grading criteria.

Las Imprecisiones Evaluativas en la Corrección Escrita Debido a la Falta de Criterios Definidos

Resumen

La presente es una investigación de campo realizada con diez profesores que son estudiantes de un curso stricto sensu de posgrado en Educación, y pretende exponer la importancia de tener criterios bien definidos al realizar la evaluación de pruebas. La investigación se basa en la simulación de una prueba hecha por dos estudiantes hipotéticos de la enseñanza fundamental, que fue corregida individualmente por los profesores. La primera corrección se hizo con criterios individuales; para la segunda corrección, el grupo de profesores discutió y estableció los criterios que serían usados, después corrigió nuevamente de forma individual. Los resultados muestran mejoras significativas en la evaluación, ya que la distorsión de la clasificación se redujo. Sin embargo, los resultados indican que la evaluación aún requiere una reflexión crítica por parte de los docentes y

coordinadores educativos, no sólo al corregir las pruebas, sino también en el proceso de evaluación y sus instrumentos, elaboración de planteos, lo que señala la importancia del binomio "¿qué medir?" y "¿cómo medir?".

Palabras clave: Evaluación. Corrección de pruebas. Criterios de corrección.

ANEXOS

Anexo A – Prova de Maria Clara

PROVA FINAL – 4ª. série do Ensino Fundamental

Nome: Maria Clara
 Idade: 10 anos
 Sexo: feminino de classe média.
 Particularidades: _____

Cada questão vale 2,5 pontos.

Questão 1 – Utilizando as palavras: viagem, capital, hotel, passeio e cultura, escreva sobre uma viagem feita ao Estado de Santa Catarina.

A viagem com meus pais pela praia de Florianópolis foi legal pois fizemos passeios e o hotel era legal.

Questão 2 – Hoje é aniversário de Daniela e Sérgio. Ele fez 13 anos e ela fez o dobro da idade dele. Quantas velinhas devo comprar para fazer a festa para eles.

4 velinhas. ~~velas~~ velhinhas.

Questão 3 – Dom Pedro II foi imperador do Brasil por quase 50 anos no século XIX. Pergunta-se: qual cidade ele fundou no Estado do Rio de Janeiro e quando se deu o final do Império no Brasil?

Teresopolis. Terminou com a chegada da república

Questão 4 – Assinale no mapa abaixo, onde está a capital do país e o Rio Amazonas.

 → não fica muito perto não

NOTA FINAL : _____

Anexo B – Prova de João Pedro

PROVA FINAL – 4ª. série do Ensino Fundamental

Nome: João Pedro
 Idade: 11 anos
 Sexo: masculino
 Particularidades: órfão e precisa de nota 5,0 para passar.

Cada questão vale 2,5 pontos.

Questão 1 – Utilizando as palavras: viagem, capital, hotel, passeio e cultura, escreva sobre uma viagem feita ao Estado de Santa Catarina.

Fui viajar com minha família para Blumenau que tem cultura alemã e foi um passeio muito ~~gostoso~~ gostoso e

Questão 2 – Hoje é aniversário de Daniela e Sérgio. Ele fez 13 anos e ela fez o dobro da idade dele. Quantas velinhas devo comprar para fazer a festa para eles.

Dois sendo uma para cada um

Questão 3 – Dom Pedro II foi imperador do Brasil por quase 50 anos no século XIX. Pergunta-se: qual cidade ele fundou no Estado do Rio de Janeiro e quando se deu o final do Império no Brasil?

A cidade foi Petrópolis e o império terminou no final do século XIX.

Questão 4 – Assinale no mapa abaixo, onde está a capital do país e o Rio Amazonas.



NOTA FINAL : _____

*legado
Ficou
no
hotel*