



INSTITUTO DE MATEMÁTICA

Universidade Federal do Rio de Janeiro



UFRJ

**A Avaliação, o Erro e o Feedback:
um estudo sobre a correção de questões de
Matemática**

Rafael Filipe Novôa Vaz

Rio de Janeiro, Brasil
05 de novembro de 2021

Avaliação, o Erro e o Feedback: um estudo sobre a correção de questões de Matemática

Rafael Filipe Novôa Vaz

Tese de Doutorado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Ensino e História da Matemática e da Física.

Universidade Federal do Rio de Janeiro
Instituto de Matemática
Programa de Pós-Graduação em Matemática

Orientadora: Lilian Nasser

Rio de Janeiro, Brasil
05 de novembro de 2021

Tese de Doutorado apresentada ao Programa de Pós-graduação em Matemática do Instituto de Matemática da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Ensino e História da Matemática e da Física.

Trabalho aprovado por:

Orientadora – UFRJ

Dra. Lilian Nasser

Membro Externo – IPS Portugal

Dr. Jorge Manuel Bento Pinto

Membro Externo – UEL

Dra. Regina Luzia Corio de Buriasco

Membro Externo – UFJF

Dr. Marco Aurélio Kistemann Junior

Membro Externo – UERJ

Dra. Maria Isabel Ramalho Ortigão

Membro Externo – UFF

Dr. Carlos Augusto Aguilar Junior

Membro Interno – UFRJ

Dr. Agnaldo da Conceição Esquinca

Rio de Janeiro, Brasil
05 de novembro de 2021

*Ao eterno amigo Paulo Alves Moreira
e a todas as outras vítimas da Pandemia da Covid-19.*

Agradecimentos

Agradeço aos colegas do Grupo de Pesquisa em Avaliação em Matemática da UFRJ e aos colegas do PEMAT pela convivência e por compartilhar pesquisas e conhecimentos. Sigamos juntos!

Agradeço à duas grandes amigas e doutoras, Viviane e Amanda. Vocês são duas grandes referências para mim.

Agradeço aos amigos David Braga, Fabio Ferreira, Fabricio Cordeiro, Flavio Silvino e Poncio Mineiro por toda ajuda e incentivo.

Agradeço aos membros da banca por participarem desse importante momento e contribuírem para a minha trajetória acadêmica.

Agradeço à Lilian, minha orientadora, parceira e amiga, por todo apoio e ajuda na escrita e na pesquisa.

Agradeço à minha família, em especial à Rozeli, Chico e Luciene por toda ajuda e afeto.

Agradeço à Luciana, minha esposa, por todo apoio e incentivo na pesquisa e escrita da tese em tempos tão difíceis.

Agradeço aos meus filhos, Ruan e Maria Júlia, e ao meu afilhado Daniel, por me inspirarem a melhorar, a lutar e a não desistir. Essa inspiração foi traduzida em fotos que iniciam os capítulos desta tese, numa singela tentativa de eternizar nossos momentos e a minha gratidão.

RESUMO

Embora o exame tenha sido um instrumento criado pela burocracia chinesa para eleger membros de castas inferiores, o modelo de avaliação das aprendizagens, caracterizado pela aplicação de exames, se consolidou na Europa e na América do Norte no final do século XIX e início do século XX. Posteriormente, a pedagogia do século XX substituiu a palavra exame pelas palavras teste e prova, atribuindo características científicas a esse instrumento avaliativo. Desde o início do século XX, diversos estudos, principalmente na França, tentaram elaborar estratégias para eliminar características interpretativas e subjetivas dos corretores na correção de itens discursivos. Nas fases anteriores à pesquisa de tese, investigamos a existência de discrepâncias em atividades de multicorreção de questões discursivas de Matemática. Neste trabalho, investigamos como os corretores interpretam o erro na correção de itens, justificam suas pontuações e elaboram *feedbacks*. Este estudo envolveu 51 corretores, entre licenciandos em final de curso e professores formados e atuantes em uma atividade de multicorreção de seis soluções distintas para um problema de matemática elementar. Os resultados foram analisados em duas perspectivas teórico-metodológicas. A pragmática, relacionada à pesquisa científica tradicional, que busca encontrar verdades a partir da generalização dos resultados. E a narrativa, que tenta valorizar as particularidades de cada respondente, identificar aspectos relevantes e singulares das respostas, produzindo conhecimento a partir das analogias criadas. Os resultados indicaram que os corretores raramente descreveram critérios claros e concisos de correção, que valorizaram mais a resposta final do que a interpretação ou os procedimentos desenvolvidos pelo estudante para a solução do problema e, ainda, que interpretaram o erro como um não-saber. Alguns dos corretores construíram critérios de correção que valorizaram a interpretação dos estudantes. Alguns, também, foram capazes de elaborar *feedbacks* escritos com grande potencial pedagógico, que denominamos de *feedbacks formativos*, pois ajudam o estudante a identificar o que errou, por que errou e como proceder para resolver corretamente a questão. A análise desses resultados, à luz do referencial teórico utilizado, nos permite compreender que a ressignificação do erro em Matemática, de um não-saber para um saber em construção, é a chave para a consolidação de uma cultura avaliativa mais voltada à aprendizagem, seja na correção de itens discursivos, seja na construção de *feedbacks*.

Palavras-chave: Avaliação em Matemática. Avaliação Formativa. Feedback. O erro.

ABSTRACT

Although the exam was an instrument created by the Chinese bureaucracy to elect members of lower castes, the learning assessment model, characterized by the application of exams, was consolidated in Europe and North America in the late 19th and early 20th centuries. . Subsequently, the pedagogy of the 20th century replaces the word exam with the words test and proof, attributing scientific characteristics to this evaluative instrument. Since the beginning of the 20th century, several studies, mainly in France, have tried to develop strategies to eliminate interpretive and subjective characteristics of the correctors in the correction of discursive items. In the phases prior to the thesis research, we investigated the existence of discrepancies in the multi-correction activities of discursive questions in Mathematics. In this work, we investigate how the brokers interpret the error in the correction of items, justify their scores and elaborate feedbacks. This study involved 51 correctors, including undergraduates at the end of the course and teachers trained and active in a multi-correction activity of six different solutions to an elementary mathematics problem. The results were analyzed in two theoretical-methodological perspectives. Pragmatics, related to traditional scientific research, which seeks to find truths from the generalization of results. And the narrative, which tries to value the particularities of each respondent, identify relevant and unique aspects of the answers, producing knowledge from the created analogies. The results indicated that the correctors rarely described clear and concise correction criteria, that they valued the final answer more than the interpretation or the procedures developed by the student to solve the problem, and that they interpreted the error as not knowing. Some of the correctors built correction criteria that valued the students' interpretation. Some were also able to elaborate written feedbacks with great pedagogical potential, which we call formative feedbacks, as they help the student to identify what was wrong, why he made a mistake and how to proceed correctly to resolve the issue. The analysis of these results, in the light of the theoretical framework used, allows us to understand that the resignification of the error in Mathematics, from a non-knowledge to a knowledge under construction, is the key to the consolidation of an evaluative culture more focused on learning, be it in the correction of discursive items, or in the construction of feedbacks.

Keywords: Assessment in Mathematics. Formative Assessment. Feedback. The mistake.

SUMÁRIO

A Travessia	10
Eu, o Viajante.....	10
A motivação da pesquisa.....	11
A questão de pesquisa	13
O mapa do percurso	15
Capítulo 1 – Os Primeiros Passos	17
1.1 O que carregamos conosco sobre a avaliação	17
1.2 Olhando para trás	22
1.3 Atravessando o mito da medida	24
Capítulo 2 - Planejando a Coleta	30
2.1 O coletor de conchas	31
2.2 O modo pragmático de interpretar as conchas	33
2.3 O modo narrativo de analisar as conchas	36
2.4 Nossa forma de coletar e interpretar conchas	39
Capítulo 3 – Observando o <i>Feedback</i> do Além Mar	42
3.1 Percorrendo as funções da avaliação	42
3.2 O <i>feedback</i> e a avaliação	47
3.3 O <i>feedback</i> formativo	52
Capítulo 4 – Aprendendo com os Erros do Caminho	56
4.1 Critérios para a correção de problemas de Matemática	58
4.2 O erro como um tropeço	63
4.3 O Erro como um trampolim	65
4.4 O erro e a avaliação.....	68
4.5 Práticas avaliativas insubordinadas.....	69
Capítulo 5 – Coletando conchas, encontrando pérolas	76
5.1 Um panorama das pontuações atribuídas às soluções.....	77
5.2 O estabelecimento de critérios para a correção.....	79
5.3 A posição em que o erro ocorre	82
5.4 A desconexão com a matriz curricular	86
5.5 A ênfase na resposta final e a desvalorização de elementos relacionados à interpretação do problema.....	88

5.6 A Construção de Feedbacks Formativos.....	89
Capítulo 6 – Algumas Reflexões sobre a Travessia	95
6.1 A Dupla Diversificação Avaliativa	95
6.2 A utilização formativa de testes somativos.....	98
Capítulo 7 – Considerações Finais	103
7.1 Reflexões sobre a coleta.....	103
7.2 Reflexões sobre os resultados	105
7.3 Sobre as pérolas	110
7.4 Últimos passos e reflexões da travessia	111
Referências	113
Apêndice 1 – A folha de correção	119
Apêndice 2 – As quatro gerações da avaliação ao longo do século XX.....	122
A Geração da Medida	122
A Geração da Descrição.....	124
A Geração do Juízo de Valor	126
Problematizando as Três Primeiras Gerações	128
A Geração da Avaliação como Negociação e Construção.....	129
Referências do apêndice 2.....	132

A TRAVESSIA



*Viajar! Perder países!
Ser outro constantemente,
Por a alma não ter raízes
De viver de ver somente!
Não pertencer nem a mim!
Ir em frente, ir a seguir
A ausência de ter um fim,
E a ânsia de o conseguir!
Viajar assim é viagem.
Mas faço-o sem ter de meu
Mais que o sonho da passagem.
O resto é só terra e céu.*
Fernando Pessoa

Eu, o Viajante

Minha trajetória acadêmica é composta de altos e baixos, como de muitos estudantes que precisam conciliar estudo, trabalho e vida social. Concluí a graduação na UERJ (Universidade do Estado do Rio de Janeiro) nos primeiros meses de 2004, efeito da greve ocorrida em 2003. Em 2004 iniciei minha trajetória profissional, pelo menos a oficial, no município de Itaboraí, e no Centro Educacional da Lagoa, trabalhando por 10 anos nestes locais. Trabalhei no PEJA, o Programa de Educação de Jovens e Adultos da prefeitura do Rio de Janeiro, e na FAETEC (Fundação de Apoio a Escola Técnica), antes de ingressar, em 2015, no IFRJ (Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro) no regime de dedicação exclusiva. No Instituto, passei a atuar no Ensino Médio Técnico Profissionalizante, na Licenciatura em Matemática e, mais recentemente, na especialização em Educação e Diversidades.

Em 2011, ingressei no PEMAT para cursar o mestrado. Como acontece frequentemente com muitos estudantes, foi no mestrado que dei meus primeiros passos na pesquisa. A dissertação intitulada *Metodologia Didática de Análise de Soluções aplicada no ensino de Frações*, que foi inicialmente orientada pela professora Elizabeth Belfort e, posteriormente, pela professora Lilian Nasser, envolveu a criação de um método de ensino que utilizava os erros dos estudantes para promover a aprendizagem. O esboço teórico envolvia a Análise de Erros e o Ensino de Frações. Nós optamos por utilizar o termo Análise de Soluções, pois explorávamos não somente as soluções incorretas, mas também as incompletas e corretas com resoluções alternativas.

A ‘metodologia’ consistia na utilização das soluções dos estudantes para a formulação de novas atividades pedagógicas baseadas nesses erros. Utilizando a pesquisa-ação, desenvolvi e utilizei esse método em duas turmas do sétimo ano, revisando um conteúdo estudado no ano anterior: frações. Os resultados obtidos nesta pesquisa indicaram que a ressignificação dos erros cometidos pelos estudantes poderia promover grandes ganhos pedagógicos. O erro não foi interpretado como algo a ser execrado, pelo contrário, regulou o ensino se transformando em um trampolim para a aprendizagem (VAZ, 2013).

Em 2018, ingressei no doutorado no PEMAT com o objetivo de investigar a avaliação das aprendizagens. O projeto de tese utilizado na seleção e nos primeiros anos de doutorado não fazia qualquer associação com a dissertação. A ideia inicial era compreender por que os corretores atribuíam pontuações tão distintas nas suas correções. Os primeiros estudos comprovaram as discrepâncias nas pontuações dos corretores, indicando um alto grau de subjetividade presente na correção de questões de Matemática (VAZ; NASSER, 2018, 2019, 2021). Essa subjetividade está associada, também, à interpretação do erro dada pelos professores. Como na pesquisa de mestrado, o erro se tornou um elemento central do meu estudo, desta vez, na tese de doutoramento.

A motivação da pesquisa

Ao longo dos mais de quatro anos de pesquisa de doutorado, experimentei diversas sensações e opiniões sobre os exames escolares. Os resultados dos primeiros estudos de multicorreção que realizei, apresentados no capítulo 1, mostraram que os critérios utilizados pelos corretores variavam significativamente. As notas finais de um

exame com questões abertas eram, frequentemente, impactadas pelas concepções desses corretores. O reconhecimento de que esse instrumento¹ apresentava baixa confiabilidade fez com que, particularmente, desenvolvesse aversão aos exames escolares. Meu sentimento era de descrença.

Este incômodo não era relacionado às questões discursivas, exclusivamente. A utilização de questões objetivas, de múltipla escolha, era uma alternativa problemática. Se por um lado, o exame objetivo possui alta confiabilidade, por outro, tende a ter baixa validade. Afinal, as questões de múltipla escolha são corrigidas de modo menos interpretativo do que as questões discursivas, além disso, é mais difícil avaliar através de um exame de múltipla escolha certas habilidades de demonstração de teoremas ou a capacidade de argumentação. Uma analogia possível seria aquela que compara o exame escolar a um cobertor curto. Se cobrirmos o rosto, descobriremos os pés, e vice versa.

Se esse instrumento apresenta tantos problemas, por que ainda o usamos? Nesse sentido, a história da avaliação, apresentada no apêndice, pode nos ajudar a entender como as perspectivas positivistas do início do século passado influenciam o modo como avaliamos. Ajudam a compreender algumas razões históricas da utilização de exames nas avaliações, que nos fazem crer na sensação de justiça relacionada ao exame. De modo geral, acreditamos que um exame é imparcial, preciso e eficaz. A cultura avaliativa escolar é sustentada no discurso de objetividade. Comumente questionamos a dificuldade de um exame e a ‘mão pesada’ do professor que o corrige, mas raramente questionamos este modelo, que está culturalmente internalizado.

Há, certamente, outros aspectos que justificam essa utilização. A prova², como é amplamente conhecido o exame escolar, é pragmática, funcional, bem adaptada à escola. Sua praticidade também é notória. Uma prova pode ser aplicada por outro professor ou algum funcionário da escola. Os alunos são ‘avaliados’ em poucas horas, poucos minutos. Hoje, o professor aplica provas on-line. Existem aplicativos que podem corrigir exames com questões objetivas ou com respostas curtas, atribuindo uma pontuação. Na pandemia,

¹A palavra instrumento, referindo-se às provas e testes, será utilizado com o significado de um documento utilizado para “recolher informações sobre os trabalhos dos alunos ou comunicar lhes indicações a esse respeito” (BARLOW, 2006, p. 134).

²O termo prova será utilizado neste texto para se referir ao exame escolar.

passamos a utilizar ainda mais programas para avaliarem nossos estudantes. É pratico demais! Entretanto, seus efeitos a médio e longo prazo são desconhecidos, precisam ser estudados.

Do reconhecimento de sua praticidade e de suas limitações, enquanto instrumento avaliativo, surgiu a tentativa da compreensão dos exames. No início da pesquisa, a questão motivadora era ‘que discrepâncias ocorrem entre corretores’, agora, o questionamento é ‘que elementos estão presentes nessas discrepâncias’. A compreensão de como os docentes corrigem as produções dos seus estudantes, de como lidam com os erros e de que tipo de *feedbacks*³ escritos elaboram para seus estudantes passou a nortear a construção da questão de pesquisa.

A questão de pesquisa

Nas pesquisas anteriores à tese, os objetos principais das investigações eram as pontuações atribuídas às questões e à prova. A ideia era entender “que” vieses ocorriam na correção de questões/provas e “por que” isso acontecia. Estudei o impacto da ordem das questões na correção dos professores (efeito halo) e a discrepância das pontuações de corretores licenciandos e professores (VAZ; NASSER, 2018, 2019).

Nesta tese, a pontuação assume outro papel, se configurando como um elemento complementar de análise. Afinal, ‘pontuar questões’ nada mais é do que ‘atribuir uma quantidade’ associada à qualidade de uma produção. As ideias envolvidas nesta tese são: compreender “como” os professores interpretam as soluções de seus estudantes e “como” essa interpretação pode impactar o ensino e a aprendizagem. A análise será realizada a partir de três aspectos identificados nas correções, de professores ou licenciandos, de seis resoluções diferentes atribuídas a alunos do 6º ano do Ensino Fundamental:

- (1) A pontuação dada ao item (quantitativo)
- (2) A justificativa escrita da pontuação (qualitativo)
- (3) O *feedback* escrito (qualitativo)

³ A palavra *feedback* significa reação a um estímulo; efeito retroativo; informação que o emissor obtém da reação do receptor à sua mensagem, e que serve para avaliar os resultados da transmissão. Posteriormente, no capítulo 3, será apresentada uma definição de *feedback* voltada à educação.

Utilizaremos a análise das justificativas (2) e a pontuação (1) para compreender como os docentes interpretam o erro dos seus estudantes, e associaremos esses aspectos ao *feedback* produzido (3) para compreender os possíveis efeitos pedagógicos dessa interpretação. A Figura 1 mostra a relação entre os elementos analisados.

Figura 1 – Elementos analisados na tese



Fonte: elaborado pelo autor

Nossa **hipótese** é que o modo como os docentes interpretam o erro dos estudantes esteja no cerne da avaliação escolar. A partir do estudo de como os docentes interpretam os erros, poderemos compreender como esses profissionais corrigem as produções dos seus estudantes e quais são os impactos dessa interpretação para o ensino e aprendizagem

Nesta pesquisa, procurei dialogar com a literatura com o objetivo de

1. refletir sobre o papel do erro na avaliação pedagógica de Matemática;
2. refletir sobre as potencialidades pedagógicas do *feedback* no ensino de Matemática.

E, a partir de uma atividade investigativa de multicorreção, pretendo compreender

1. como licenciandos e professores interpretam diferentes tipos de erro,
2. como essa interpretação se relaciona com o *feedback* escrito e
3. como esses resultados se relacionam com a avaliação.

Sucintamente, podemos resumir a questão de pesquisa em:

Quais são os aspectos relacionados à avaliação que emergem da interpretação dos erros dos estudantes na correção de itens, e como estes aspectos influenciam a avaliação em Matemática?

O mapa do percurso

Como será descrito no **capítulo 1**, nossos estudos iniciais constataram características de aleatoriedade na correção de questões discursivas de Matemática, especificamente, na atribuição de pontuações para as questões. Naturalmente, perguntas que surgiram a partir desses resultados despertaram nossa curiosidade e interesse de pesquisa. Alguns desses questionamentos, relevantes para esta Tese, são discutidos nos capítulos 3 e 4.

O **capítulo 2** apresenta ao leitor reflexões teóricas sobre duas perspectivas possíveis para a pesquisa em Educação (Matemática): Narrativa e Pragmática. A Análise de Conteúdo, representante da perspectiva pragmática e a Análise Narrativa são as teorias metodológicas utilizadas para nortear o tratamento dos dados coletados nesta tese. Esse capítulo apresenta, também, o instrumento utilizado para coletar pontuações, justificativas e feedbacks, ambos escritos, de professores e licenciandos.

Um dos questionamentos corresponde à qualidade do *feedback*, elemento central, de acordo com alguns estudos, para a avaliação formativa e para a promoção da aprendizagem. No **capítulo 3** discutiremos as razões que fazem da avaliação formativa uma ideia poderosa para o Ensino, com diversas implicações para a aprendizagem e quais as características de um bom *feedback*.

O segundo aspecto analisado se refere à interpretação do erro em Matemática. O erro pode ser visto como um desconhecimento ou como um lapso, pode ser interpretado como algo que deva ser eliminado ou algo inerente ao processo de aprendizagem. Lidar com o erro é uma das atribuições do trabalho docente. O professor lida com o erro nas aulas e nas avaliações, conseqüentemente, o modo como o professor interpreta o erro e (não) o utiliza para a promoção da aprendizagem impacta o ensino, a avaliação e a própria aprendizagem. Tais questões são discutidas no **capítulo 4**.

Os resultados obtidos na investigação realizada com 51 respondentes (professores e licenciandos em final de curso) são analisados e discutidos no **capítulo 5**. As justificativas das pontuações e os *feedbacks*, elaborados pelos respondentes durante a tarefa de correção de questões, serão analisados com o objetivo de identificar possíveis aspectos relacionados à avaliação.

O **capítulo 6** apresenta algumas reflexões teóricas sobre a avaliação desenvolvidas na pesquisa: A Dupla Diversificação Avaliativa e a Utilização Formativa de Testes Somativos. São sugestões que podem contribuir para a (re)construção de práticas avaliativas mais direcionadas à aprendizagem, em uma tentativa de valorizar diferentes habilidades e potencialidades dos estudantes.

Nos capítulos finais estão as **referências** utilizadas nesta pesquisa e alguns **apêndices**, dentre eles, um capítulo referente à **história da avaliação** que apresenta modificações nos modos de compreender a avaliação ao longo do século XX, permitindo compreender as origens e razões envolvidas em algumas das práticas e concepções avaliativas contemporâneas.

CAPÍTULO 1 – OS PRIMEIROS PASSOS



*Há um tempo em que é preciso abandonar as roupas usadas,
que já tem a forma do nosso corpo,
e esquecer os nossos caminhos,
que nos levam sempre aos mesmos lugares.
É o tempo da travessia:
e, se não ousarmos fazê-la,
teremos ficado, para sempre,
à margem de nós mesmos.*
Fernando Teixeira de Andrade

1.1 O que carregamos conosco sobre a avaliação

O que é avaliar? Avaliar é um verbo transitivo bitransitivo no sentido de estabelecer a valia, o valor ou o preço de. Ao avaliar uma peça de antiquário, atribuímos um preço a esta peça, atribuímos também valor a esta peça. Valor não é sinônimo de preço. Algo pode custar mais do que vale. O preço é impessoal, o valor não. O preço é objetivo, o valor é subjetivo. E o que é avaliar no contexto escolar? Se a avaliação significa, na etimologia da palavra, ‘dar valor’ a alguma coisa, esse valor é atribuído a partir de um modelo de referência e do conhecimento do objeto que se avalia. Ao avaliar estudantes, atribuímos a eles um valor a partir de suas produções. Trata-se de uma tentativa de inferir um valor sobre suas aprendizagens, uma visão técnica da avaliação, um julgamento. No entanto, esta visão da avaliação desconsidera a complexidade da avaliação como elemento pedagógico.

Na escola, a avaliação é programada, realizada em um espaço artificial, associada ao sentimento de dever cumprido, de obrigação e julgamento. Comumente, é associada com a ideia de medir a aprendizagem (LUCKESI, 1998; MORGAN, 2000; HADJI, 2001). Na prática da aferição do aproveitamento escolar, os professores realizam,

basicamente, três procedimentos sucessivos: “medida do aproveitamento escolar; transformação da medida em nota ou conceito e utilização dos resultados identificados” (LUCKESI, 1998, p. 72).

A avaliação escolar é um processo de vai e vem de seleção e de produção de informação, de evidências, entre a produção do estudante e o modelo de referência do professor. Como ilustrado na Figura 2, o modelo de referência pode ser influenciado por diversos fatores, tais como expectativas do professor em relação ao desempenho e à aprendizagem dos estudantes, a experiência deste profissional e a escala de notação desenvolvida para a pontuação (NOIZET; CAVERNI, 1985; SANTOS; PINTO, 2018).

Figura 2 – Avaliação a partir de um modelo de referência



Fonte: SANTOS; PINTO (2018, p. 505)

A avaliação é construída a partir de um processo que inclui diversas etapas, não necessariamente desenvolvidas em uma sequência única: planejar o que fazer e para quê, recolher a informação, produzir um julgamento a partir da interpretação da informação recolhida e agir em conformidade (SANTOS, 2019). De modo que “toda avaliação pressupõe um julgamento e uma tomada de decisão sobre o que é relevante para determinado fim, uma recolha de informação, a sua interpretação e o desenvolvimento de uma ação decorrente dela” (SANTOS, 2011, p. 155). Buriasco (1999), em sua tese de doutorado, afirma que avaliar pressupõe definir princípios em função de

objetivos que se pretendem alcançar; estabelecer instrumentos para a ação e escolher caminhos para essa ação; verificar constantemente a caminhada, de forma crítica, levando em conta todos os elementos envolvidos no processo. Sendo assim, ela não possui uma finalidade em si, mas sim subsidia o curso de uma ação que visa construir um resultado previamente definido (BURIASCO, 1999, p. 72)

Possivelmente, o instrumento de avaliação mais utilizado em Matemática no Brasil, e provavelmente no mundo, é um exame escrito, comumente chamado de prova, teste ou avaliação. Há um problema na utilização de termos etimologicamente tão distintos. Para Barriga (2002), a substituição do termo exame por teste, pela pedagogia

do século XX, teve o intuito de atribuir a esse instrumento um caráter científico e a associação, reducionista, da avaliação ao exame teve por objetivo fornecer uma conotação acadêmica ao instrumento.

Neste texto, o termo avaliação será utilizado em seu sentido mais amplo, como um processo de análise da aprendizagem que se diferencia e transcende à utilização do instrumento. A avaliação também não pode ser resumida pelo estabelecimento de notas ou conceitos. A aplicação de instrumentos avaliativos, a análise da produção dos estudantes nas aulas e nos exames, o estabelecimento de notas, todos esses são componentes de um processo, que denominamos de avaliação das aprendizagens.

O termo *prova*, por sua vez, será utilizado nesta tese referindo-se ao teste escrito comumente aplicado nas escolas brasileiras. As questões que compõem esse instrumento podem ser de múltipla escolha, dissertativas, verdadeiro ou falso, preenchimento de lacunas, etc. Geralmente, as provas são individuais, escritas e sem consulta. Quando nos referirmos a uma exceção, utilizaremos um adjetivo adicional, por exemplo, a *prova em fases* ou *prova com cola*.

Podemos desmembrar o processo avaliativo através da *prova* em quatro etapas, como ilustrado na Figura 3:



Fonte: elaborado pelo autor

Por elaboração/planejamento compreendemos a etapa da concepção desse instrumento, que inclui a seleção das questões, do modelo de avaliação, dos tipos de questões, dos objetivos pedagógicos, dos critérios utilizados para a correção. O docente deve conhecer o Projeto Político Pedagógico da escola, assim como as normas, currículos e leis que regem a Educação (a LDB, as resoluções do CNE, as deliberações do CEE), os Parâmetros Curriculares Nacionais⁴ (PCN), os currículos locais e a Base Nacional

⁴ Os PCN - Parâmetros Curriculares Nacionais foram diretrizes elaboradas pelo Governo Federal, a partir de 1997, com o objetivo principal de orientar os educadores por meio da normatização de alguns fatores fundamentais concernentes a cada disciplina. (BRASIL, 1998)

Comum Curricular⁵ (BNCC). De acordo com as normas vigentes, o docente deve selecionar também qual será a modalidade da avaliação que será utilizada: formativa, somativa ou diagnóstica (Não se trata, necessariamente, de uma questão de escolha, talvez de princípio/perspectiva teórica). Definida a modalidade, vem a escolha do instrumento – uma prova, um seminário, uma lista de exercícios ou algo similar, quais atividades ou questões serão propostas e quais os critérios de execução. Todos esses pontos devem estar em consonância com os objetivos pedagógicos do docente e do currículo adotado pela instituição de ensino. Uma possível ordem é apresentada na figura 4.

Figura 4 - Fases de elaboração de um instrumento avaliativo



Fonte: elaborado pelo autor

Na literatura encontramos relatos que demonstram a preocupação de professores com a elaboração dos instrumentos avaliativos. Teixeira Jr. e Silva (2017) realizaram uma investigação com 16 professores de Matemática da rede pública sobre suas concepções de avaliação e suas práticas de aplicação de provas. “Os professores pesquisados mostram que o problema da avaliação em Matemática não está apenas no que se faz com os resultados obtidos após sua aplicação e, sim, também, com o planejamento”. Esse estudo identificou preocupações dos professores referentes ao tempo de planejamento de provas, a escolha dos conteúdos, além de “relacionar os valores culturais, sociais, cognitivos e políticos” (TEIXEIRA JR.; SILVA, 2017, p. 50).

A coleta de informações é caracterizada por um rito próprio que varia de acordo com o instrumento, com a instituição de ensino e com o contrato didático⁶. No caso específico da prova, a recolha ocorre quando alunos respondem às questões e entregam

⁵ A Base Nacional Comum Curricular (BNCC) “é um documento de caráter normativo que define o conjunto orgânico e progressivo de aprendizagens essenciais que todos os alunos devem desenvolver ao longo das etapas e modalidades da Educação Básica, de modo a que tenham assegurados seus direitos de aprendizagem e desenvolvimento, em conformidade com o que preceitua o Plano Nacional de Educação (PNE)”. (BRASIL, 2017, p. 5)

⁶ Chama-se contrato didático ao conjunto de comportamentos do professor, que são esperados pelo aluno, e o conjunto dos comportamentos do aluno, que são esperados pelo professor [...] esse contrato é o conjunto de regras que determinam explicitamente, por uma pequena parte, mas, sobretudo, implicitamente, em grande parte, o que cada elemento da relação didática deverá fazer e que será, de uma maneira ou de outra, válido para o outro elemento. (BROUSSEAU apud HENRY, 1991)

as suas respostas ao professor ou fiscal. A utilização dos instrumentos avaliativos na escola está, comumente, inserida em um intervalo de tempo delimitado e obedece a regras estabelecidas na instituição e pelo docente. Um rito comum no dia da prova é aquele em que os alunos são posicionados estrategicamente pelo professor, normalmente de acordo com a disposição mais comum das carteiras, enfileirados, quase que equidistantes. Em algumas situações os estudantes que apresentam melhores desempenhos são posicionados próximo ao professor. Os professores agem na tentativa de impedir qualquer comunicação no decorrer da resolução do instrumento e os alunos agem minimizando o efeito da separação imposta.

Ainda hoje, na escola, a execução do rito de avaliar – aplicar uma prova ou um teste escrito e converter as resoluções e respostas de cada estudante a um valor numérico – parece ser suficiente para fazer acreditar que se cumpriu o esperado desse mito: medir e classificar de maneira precisa os alunos. (BURIASCO; FERREIRA; CIANI, 2009, p. 70-71)

Após a aplicação se inicia a etapa da correção da prova. Essa etapa consiste na identificação dos erros e acertos cometidos pelos estudantes e o estabelecimento de uma pontuação ou conceito do instrumento a partir das soluções e respostas apresentadas. A etapa da correção foi contemplada em uma perspectiva nos estudos iniciais desta pesquisa, que será relatada na seção seguinte, e será abordada em uma nova perspectiva nesta tese. A perspectiva inicial estava relacionada a interpretar as divergências entre os corretores, a perspectiva atual relaciona-se à compreensão dos docentes em relação ao erro dos seus estudantes em uma atividade avaliativa.

Definimos a etapa da devolução como sendo a etapa em que o estudante recebe a prova corrigida, o retorno de um trabalho apresentado ou, simplesmente, quando os estudantes de um grupo recebem o retorno avaliativo da apresentação de um seminário. Neste momento, o estudante pode reconhecer a evolução de sua aprendizagem, pode identificar os erros que cometeu e, a partir desta identificação, reconhecer o que precisa ser melhorado. Além do erro, esse *feedback* dado ao aluno é o aspecto investigado nesta tese.

No entanto, o que se observa, comumente, é que o *feedback* dessa importante etapa da aprendizagem é resumido a uma entrega de “notas”. Em alguns casos, é seguido pela apresentação do gabarito. Ao receber a prova, os estudantes realizam uma conferência da correção do professor, sempre buscando um erro de correção em que o professor tenha

atribuído uma pontuação indevidamente menor, com o intuito de aumentar a pontuação obtida. Um processo similar acontece em outros instrumentos avaliativos em que a atribuição da nota é centralizada na figura do professor.

As duas últimas etapas da avaliação nessa perspectiva que defendemos (correção e devolução) serão investigadas nesta tese. Mais especificamente, investigaremos a correção de questões discursivas de Matemática em duas vertentes: a primeira diz respeito ao modo como os docentes interpretam o erro e a segunda, como, a partir daí, fornecem *feedback* escrito aos estudantes.

1.2 Olhando para trás

Influenciados por uma visão positivista, os professores de Matemática podem crer que a prova seja um instrumento de medida justo, imparcial e preciso. Há, então, uma real necessidade de estudos que promovam o rompimento da crença no Paradigma Psicométrico da Avaliação Escolar (FERNANDES, 2009). Segundo este paradigma, é possível determinar exatamente o que os alunos sabem e são capazes de fazer; as aprendizagens constituem uma realidade que pode ser avaliada de forma objetiva e neutra; a avaliação ocorre através de instrumentos cientificamente construídos; a avaliação deve centrar-se mais nos resultados do que nos processos. (FERNANDES, 2009)

Guba e Lincoln (2011) defendem que para compreender a avaliação hoje é necessário recorrer à sua evolução histórica, marcada por diversas influências sociais. As diversas concepções da avaliação escolar ou avaliação das aprendizagens se desenvolveram no decorrer da história, inseridas e atreladas ao desenvolvimento das ciências da Educação e da própria sociedade.

A ausência de uma história da avaliação dificulta este tipo de olhar em toda a sua extensão. A avaliação é uma prática social inscrita num sistema mais vasto de outras práticas que tiveram importantes funções em situações de reorganização social. As funções da avaliação e a sua evolução estão assim estreitamente articuladas com os próprios movimentos da sociedade. (PINTO, 2016, p. 8)

A avaliação na Pedagogia Tradicional, por exemplo, estava relacionada ao controle dos educandos, não só no que se refere à “aprendizagem dos conhecimentos e habilidades de raciocinar, debater e discutir, mas também no que se refere à doutrina e às condutas morais e de *civilidade* dentro da sociedade” (LUCKESI, 1992, p. 57-58, grifo do autor).

A Pedagogia Tradicional emergiu com o alvorecer da sociedade moderna, cristalizando-se com a Revolução Francesa e se estendendo até os nossos dias; a Pedagogia da Escola Nova emergiu em fins do século XIX, cristalizando-se na primeira metade do século XX; a Tecnopedagogia emergiu e se desenvolveu na história recente da sociedade. (LUCKESI, 1992, p. 55)

Para demonstrar seus conhecimentos, os estudantes necessitavam memorizar fatos e demonstrar seu domínio em testes orais e individuais ou, em alguns casos, exames escritos, dissertativos. O objetivo da escola era ensinar às crianças o que se reconhecia como certo

Os exames escolares foram utilizados durante centenas de anos para avaliar se os alunos haviam conseguido dominar o conteúdo de diversos cursos e assuntos nos quais haviam sido iniciados. Conteúdos apropriados eram definidos com base em sua autoridade, fossem eles sobre Aristóteles, a Bíblia ou, mais recentemente, as descobertas da ciência. (GUBA; LINCOLN, 2011, p. 28)

No entanto, a avaliação como conhecemos hoje está relacionada ao ensino de massas, que se iniciou há pouco mais de 100 anos. Segundo Pinto (2016, p. 3), ela foi “concebida como uma medida, que se destinava a informar vários processos administrativos de organização (por ex. organização de turmas) e gestão dos sistemas educativos (ex. retenção, transição de ano ou ciclos de estudos)”. Nesse sentido,

podemos dizer que a avaliação, ao longo do último século, assumiu funções que se prendem mais fortemente com pedidos de natureza social do que pedagógicos. O controle das qualificações profissionais, a rentabilidade dos recursos investidos na Educação e a garantia da segurança dos indivíduos e das comunidades através da regulação do exercício de certas profissões, são três razões sociais invocadas por Reuchelin (1974), que justificam e legitimam a avaliação, não só no passado, como no presente. (PINTO, 2016, p. 10)

As concepções avaliativas sofreram diversas mudanças desde o final do séc. XIX⁷, acompanhando as transformações sociais, econômicas e educacionais ocorridas nesse período. A avaliação percorreu um longo caminho para abandonar, mais no plano teórico do que prático, a ideia da medida e ser compreendida como “um processo de decisão compreensiva contextualizada, como um processo de construção social orientada para sustentar uma utilização pedagógica, ou seja, uma intervenção reguladora das aprendizagens” (PINTO, 2016, p. 3).

⁷ Para entender com maior profundidade, nos Apêndices, há um capítulo que trata da História da Avaliação.

1.3 Atravessando mito da medida

A avaliação é um componente essencial e indissociável do processo de ensino-aprendizagem. A avaliação é um processo de análise sobre a aprendizagem do estudante, mais próximo de uma leitura do que uma medida (HADJI, 2001). A avaliação com a concepção de medida, representada pelas provas, implica na aceitação de instrumento de aferição das aprendizagens, desconsiderando que possa ocorrer interferência do avaliador na correção e pontuação (MORAES, 2012). Para Luckesi (1998, p. 72), “a medida é uma forma de comparar grandezas, tomando uma como padrão e outra como objeto a ser medido, tendo como resultado a quantidade de vezes que a medida padrão cabe dentro do objeto medido”. A medida da aprendizagem do educando está relacionada ao número de acertos nos exames, que por sua vez, são transformados em pontos.

O padrão de medida, então, passa a ser os pontos. A cada acerto corresponderá um número de pontos, previamente estabelecido, que pode ser igual ou diferenciado para cada acerto. [...] Para coletar os dados e proceder à medida da aprendizagem do educando, os professores, em sala de aula, utilizam-se de instrumentos que variam desde a simples e ingênua observação até sofisticados testes, produzidos segundo normas e critérios técnicos de elaboração e padronização. (LUCKESI, 1998, p. 73)

A suposta objetividade da avaliação escolar, em que seria possível medir o conhecimento de alguém, está associada a uma filosofia positivista, que tem como pilares a neutralidade e a imparcialidade. Nessa concepção, a ideia de avaliação como uma “medida vem emprestar à avaliação uma legitimidade científica. O exame e os seus rituais, tido como o instrumento de medida, acabam por preencher por completo a própria noção de avaliação” (PINTO, 2016, p. 10).

Raphael (1998) aponta que um dos pontos positivos da prova individual e sem consulta é que ela está atrelada a um sentimento de justiça por oferecer um tratamento genérico a todos os estudantes. Tais concepções estão presentes nos três pressupostos básicos que sustentam o pensamento geral na avaliação:

primeiro, presume-se que os indivíduos possuem atributos (como conhecimento, compreensão, habilidade, etc.) que são detectáveis e mensuráveis. Em segundo lugar, o objetivo principal da avaliação é descobrir e medir esses atributos. Finalmente, a avaliação e seus resultados são considerados fundamentalmente benignos ou mesmo benéficos (embora efeitos colaterais infelizes possam ser reconhecidos e tentativas feitas para melhorá-los) (MORGAN, 2000, p. 225-226, tradução nossa).

Os professores associam a objetividade na avaliação escolar à clareza, à uniformidade nos critérios de avaliação e à neutralidade no campo da Matemática. Fischer (2008) cita a fala ambígua de um professor que diz privilegiar “a exatidão do conteúdo, mas valoriza, na correção, o desenvolvimento apresentado pelo aluno” (p.85) e conclui que a pesquisa realizada

mostrou que os professores de Matemática ainda refletem, em suas práticas avaliativas, muitas das concepções acerca do conhecimento matemático e de seu ensino construídas com base no discurso de modernidade. As características apontadas como constituintes do hábito desse professor, como a busca pela objetividade, a concepção positivista de rigor no trato dessa ciência e de seu ensino, um certo descrédito do fazer pedagógico e a adoção de uma postura pouco flexível, tem fortes marcas desse paradigma de ciência. (FISCHER, 2008, p.96)

A redução da avaliação a uma medida surgiu com o fortalecimento da Psicologia no final do século XIX, com o advento dos testes psicométricos. A ideia era construir instrumentos científicos capazes de medir a inteligência humana.

Os referidos testes sofreram influências do positivismo das ciências físico-naturais, sustentados nas exigências do rigor científico; de pesquisas tais como os trabalhos de Darwin, que se alicerçava no recurso da medição para delinear categorias dos indivíduos e suas diferenças e no desenvolvimento dos métodos estatísticos. Surge assim, em 1922, na França, a ciência “docimológica” ou a “docimologia”, que significa “medida por exame” (ou em latim *dakimê*), caracterizando a era dos testes em avaliação. (MACEDO; LIMA, 2013, p. 158)

O objetivo das investigações de multicorreção, desenvolvidas inicialmente na França, era comparar as distribuições de notas de cada um dos avaliadores para compreender as razões das discrepâncias cometidas pelos corretores e criar estratégias para evitá-las (NOIZET, CAVERNI, 1985, MERLE, 2018). Partindo da ideia de que o exame fornecia uma medida, os testes docimológicos buscavam caminhos científicos para eliminar vieses humanos na correção. Em uma perspectiva pragmática e objetiva “buscava-se o rigor metodológico visando evitar discrepâncias nos resultados, que muitas das vezes, nos casos dos testes de concursos, redundavam em injustiças díspares” (MACEDO; LIMA, 2013, p. 158).

Se por um lado, os estudos de docimologia permitiram avanços notáveis nos exames, principalmente, naqueles construídos com questões de múltipla escolha, vide a Teoria de Resposta ao Item, utilizada no Exame Nacional do Ensino Médio, por outro, no que diz respeito à correção de provas discursivas, a influência do corretor apresentou-se ao longo da história como algo inevitavelmente presente e indesejável.

O primeiro estudo sistemático de multitorreção, nomeado de inquerito internacional sobre os exames de acesso, foi realizado por Laugier e Weinberg em 1936. Tal estudo foi realizado a partir de dados do *baccalauréat*, exame que estudantes franceses fazem ao final do ensino médio para entrar na universidade. Seis avaliadores corrigiram 100 exercícios de seis disciplinas: Francês, Latim, Inglês, Matemática, Filosofia e Física. Laugier e Weinberg propuseram, nessa pesquisa, que as divergências observadas com a experiência de multitorreção eram frutos do acaso e que os erros cometidos pelos professores nas correções eram similares aos erros das ciências físicas, ou seja, variações aleatórias. A Tabela 1 ilustra os acordos e desacordos sobre a notas obtida em exames de diferentes disciplinas.

Tabela 1 – Porcentagem de acordos e de desacordos em 100 exercícios

Disciplinas	Excluídos por todos os avaliadores	Admitidos por todos os avaliadores	Admitidos por alguns e excluídos por outros
Francês	21	9	70
Latim	40	10	50
Inglês	37	16	47
Matemática	44	20	36
Filosofia	9	10	81
Física	37	13	50

Fonte: NOIZET; CAVERNI (1985, p. 42, apud LAUGIER; WEINBERG, 1936)

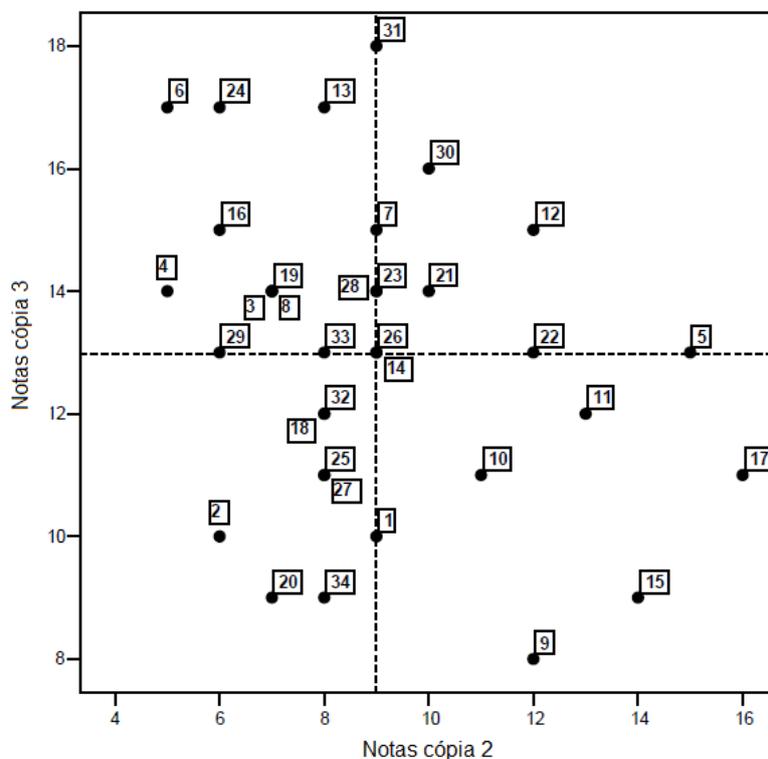
Em estudo mais recente, Suchaut (2008) reconhece a existência da subjetividade na correção de provas discursivas. A investigação foi realizada em duas universidades francesas, ao longo de dois anos, com cerca de trinta professores de ciências econômicas e sociais. Os professores foram convidados a corrigir 3 provas, comentando e justificando as notas. As três provas eram reproduções do *baccalauréat*, uma considerada boa (nota 15, de 20), e duas médias (notas 9 e 11). Os resultados de Suchaut (2008) mostram uma considerável dispersão da atribuição das notas em torno da média nas três cópias.

Para 4 cópias, a diferença máxima é de 10 pontos, e de 9 e 11 pontos para as outras duas cópias, o que é considerável. Os desvios padrão, que fornecem uma medida estatística da dispersão da pontuação média, também apresentam valores altos (de 1,8 a 2,7 pontos). Esses dados recentes confirmam claramente a incerteza das notas dos alunos, as marcas da mesma cópia podem variar bastante de um corretor para outro (SUCHAUT, 2008, p. 6, tradução nossa).

Outro aspecto que merece destaque no estudo de Suchaut (2008, p. 7) diz respeito à “instabilidade nas práticas de correção”, em que um corretor, ao mesmo tempo em que

era severo na correção de uma prova, era indulgente na outra. O gráfico apresentado na Figura 5 mostra pontuações das cópias 2 e 3 dadas pelos 34 corretores de uma das universidades. Quanto maior fosse a estabilidade dos corretores em relação à severidade e indulgência, maior seria o número de pontuações ocupadas no 1º e 3º quadrantes. Ao descartar os corretores cujos pontos estão em cima dos eixos indicativos das médias, o que se observa, no entanto, é uma distribuição oposta, visto que 11 corretores ocupam o interior dos 2º e 4º quadrantes e 8 ocupam os 1º e 3º quadrantes. Isto significa que não se trata de alguns corretores serem mais severos, enquanto outros, menos rigorosos, e sim, os mesmos corretores são mais severos em uma correção e menos em outra.

Figura 5 – Relação entre as notas atribuídas pelas cópias 2 e 3 por 34 corretores



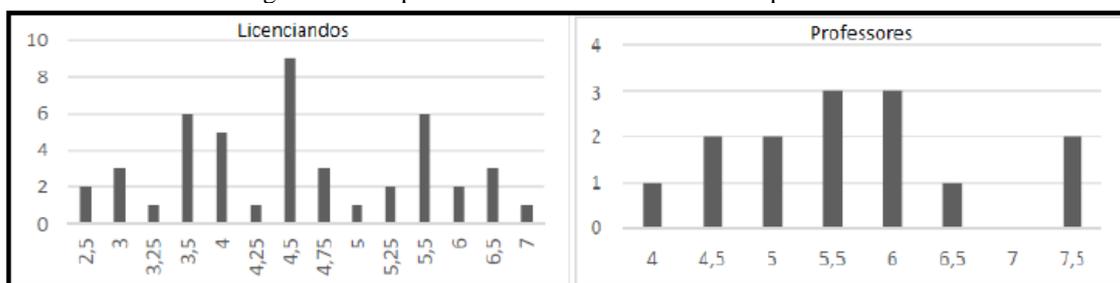
Fonte: SUCHAUT (2008, p. 10)

Suchaut (2008) concluiu que há grandes evidências estatísticas que apontam para a aleatoriedade de atribuição de notas na multicorreção e afirma que a experiência

apresentada e analisada neste texto apenas confirma através dados atuais, o que já sabemos há muito tempo sobre a incerteza da classificação. Nós só podemos lamentar que esta prática, tão difundida em todos os níveis do nosso sistema educacional, pode ter consequências em termos de desigualdades na certificação de aprendizado prévio escolar (SUCHAUT, 2008, p. 14, tradução nossa).

Em estudos preliminares à tese, realizamos uma investigação utilizando a multcorreção. A primeira fase desse estudo investigou o comportamento de licenciandos em Matemática nos últimos períodos da graduação (VAZ; NASSER, 2018) e na segunda fase, os corretores participantes foram professores formados e atuantes (VAZ; NASSER, 2019). O gráfico ilustrado pela Figura 6, apresenta à esquerda as notas atribuídas por 45 licenciandos e à direita, as notas de 13 professores em exercício. Ambos os grupos corrigiram a mesma prova discursiva, contendo 5 itens, cada um valendo 2 pontos. A imensa amplitude entre as notas atribuídas nos dois grupos é um indício que sugere o caráter subjetivo na correção das provas.

Figura 6 – Frequência de notas de licenciandos e professores.



Fonte: VAZ; NASSER (2019)

Se fosse escolhido, aleatoriamente, um desses corretores para corrigir uma prova com as características similares à que elaboramos nesse estudo, a nota atribuída à mesma prova dependeria do escolhido. A nota do aluno e, conseqüentemente, sua aprovação seriam incertas! Os resultados obtidos por Noizet e Caverni (1985), Suchaut (2008) e Vaz e Nasser (2019) confirmam o caráter interpretativo e subjetivo presente na correção de provas discursivas.

O julgamento do avaliador é infiltrado por elementos provenientes do contexto escolar e social, desde a carga afetiva e a dimensão emocional devido à presença efetiva dos alunos se, geralmente, ignora que se baseia em parte em uma representação construída do aluno e em convicções íntimas que nada têm de científicas (HADJI, 2001). Em consonância com Hadji (2001), compreendemos que avaliação seja uma leitura de algo que pode ser observado. Nós observamos o desempenho dos estudantes. Tal desempenho é influenciado por fatores externos e internos, inclusive, pela subjetividade de nós, professores corretores.

É necessário abandonar a ideia da medida. A possibilidade de quantificar o conhecimento aprendido de alguém é um mito. Precisamos reconhecer que o processo

avaliativo está imerso em subjetividades. Mesmo na prova chamada de objetiva, com questões de múltipla escolha, há subjetividade. Esta ocorre na correção, na escolha das questões e na confecção delas.

A objetividade é, antes de tudo, fruto de um trabalho coletivo e continuado, de explicitação pública e debate sobre as diferenças que se cruzam no espaço avaliativo. Quando as diferenças deixam de ser apenas divergências constatadas, fruto de más intenções ou interpretações, para se tomarem em diferenças institucionais assumidas e publicamente constatadas está-se no caminho da construção da objetividade. (PINTO, 2016, p. 7-8)

É necessário reconhecer que a avaliação pedagógica⁸ é imersa em complexidades e contradições, e neste sentido, se faz necessário que nós docentes tenhamos um olhar aberto e, necessariamente, plural (PINTO, 2016). A avaliação, como toda atividade humana, está sujeita ao erro, às diferentes interpretações, a uma multiplicidade de fatores que interferem e/ou influenciam o processo de correção e análise. Qualquer que seja a pretensão de atribuir mais objetividade ao processo avaliativo perpassa pelo reconhecimento das limitações inerentes às relações humanas e pedagógicas. Avaliar um estudante, partindo de uma norma, através da comparação com o desempenho de um grupo, apresenta algumas limitações. Os estudantes são diversos, plurais, cada um com suas idiosincrasias e particularidades. Os estudantes pensam de maneiras variadas, aprendem em momentos distintos. Reconhecer a multiplicidade de fatores sociais, econômicos, cognitivos, psicológicos torna qualquer processo avaliativo pedagógico mais complexo e problemático.

⁸ A Avaliação Pedagógica é uma terminologia utilizada nesta Tese para indicar a Avaliação Interna, aquela que é elaborada, aplicada e corrigida pelo corpo docente, geralmente professor regente.

CAPÍTULO 2 - PLANEJANDO A COLETA



*De tudo ficaram três coisas:
A certeza de que estamos começando;
A certeza de que é preciso continuar;
A certeza de que podemos ser interrompidos antes de terminar.
Façamos da interrupção um caminho novo.
Da queda, um passo de dança.
Do medo, uma escada.
Do sonho, uma ponte.
Da procura, um encontro.*
Fernando Sabino

A palavra *pesquisa* é originária do Latim *perquirere*, referindo-se à palavra *perquirir* no sentido de investigar, construída sobre o prefixo *per*, entendido como através ou por meio, e o verbo *quarere*, em relação a consultar, indagar ou procurar. (PESQUISA, 2020). A pesquisa

é um processo de estudo que consiste na busca disciplinada e metódica de saberes ou compreensões acerca de um fenômeno, problema ou questão da realidade ou presente na literatura o qual inquieta/instiga o pesquisador perante o que se sabe ou diz a respeito (FIORENTINI; LORENZATO, 2012, p. 60)

Os procedimentos metodológicos adotados por qualquer pesquisador são influenciados pela literatura revisitada, pelas suas convicções, certezas e incertezas, da sua forma de compreender a Matemática e a Educação Matemática, pelos seus pares, pelo seu grupo de pesquisa e pelo programa de Pós-Graduação ao qual pertence.

Anteriormente, nosso instrumento de coleta envolvia a multiorreção de uma prova com diferentes itens, cada um deles associado a uma solução. O instrumento de coleta utilizado nesta tese envolve a correção de soluções distintas de um mesmo item. Esse instrumento, denominado *folha de correção*, apresentado na seção seguinte, solicita,

além da pontuação, comum em atividades deste tipo, a justificativa apresentada pelos professores em relação à pontuação e ao *feedback* escrito dado aos alunos.

2.1 O coletor de conchas

A pesquisa contou com a participação de 4 professores universitários de Matemática que atuam na licenciatura e na pós graduação, 30 estudantes de pós graduação e 17 licenciandos em Matemática (totalizando 51 participantes). Uma oficina sobre Avaliação Escolar, apresentada e dinamizada por mim, foi oferecida em dois momentos: no primeiro, em uma universidade pública, com professores universitários, mestrandos e licenciandos em Matemática; em seguida, uma oficina similar foi oferecida numa universidade privada para estudantes de pós-graduação em Educação Matemática com a participação de um professor. Ambos os estabelecimentos são localizados na cidade do Rio de Janeiro.

Durante essas oficinas, os participantes foram convidados a preencher um formulário⁹, denominado de *folha de correção*, que solicitava a correção de seis soluções distintas para uma mesma questão apresentada com gabarito, como ilustra a Figura 7.

Figura 7 – A questão investigada e seu gabarito

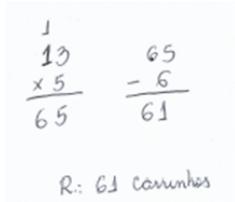
Questão	Gabarito
<i>Pedro possui 5 caixas onde guarda seus carrinhos. Em cada caixa havia 13 carrinhos, até que Pedro retirou 6 carrinhos de uma dessas caixas para brincar. Após essa retirada, quantos carrinhos ficaram nas 5 caixas ao todo?</i>	$5 \times 13 = 65$ $65 - 6 = 59$

Fonte: elaborado pelo autor

A proposta era supor que essas soluções tivessem sido produzidas por estudantes do 6º ano do ensino fundamental e que os entrevistados eram seus professores, corrigindo uma prova. Os participantes deveriam atribuir uma pontuação de 0 a 1, explicar o critério para a pontuação e oferecer um *feedback* escrito, nas lacunas correspondentes, para cada uma das seis soluções. Como exemplo, a Figura 8 mostra o quadro referente à primeira solução.

⁹A *folha de correção* encontra-se disponível no Apêndice 1

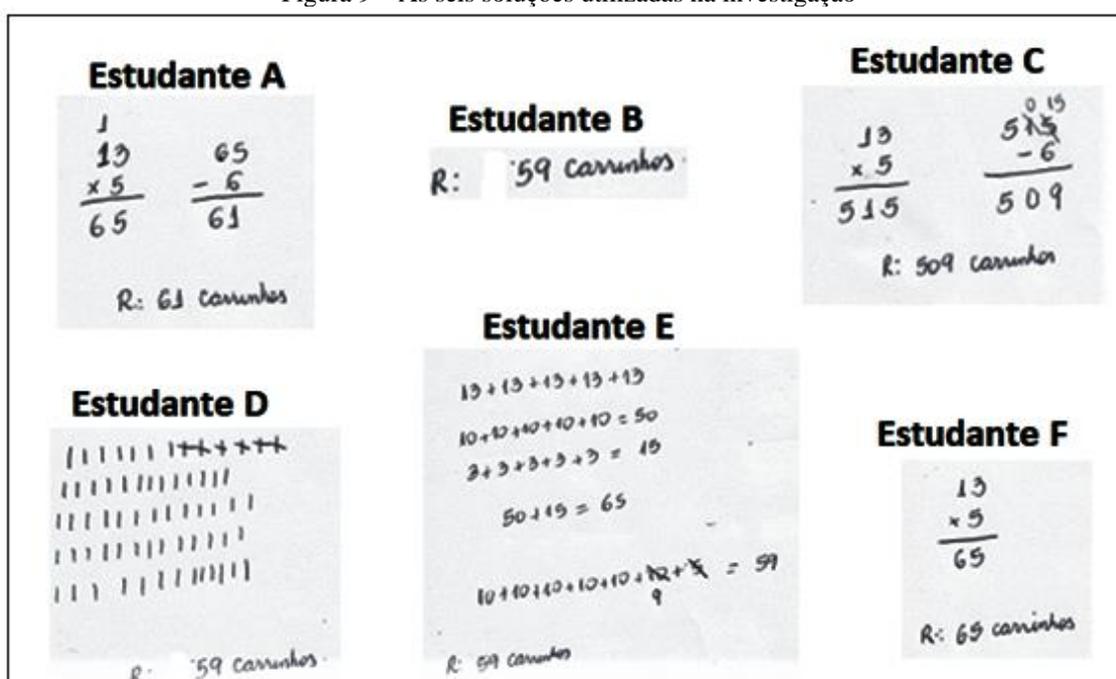
Figura 8 – Quadro referente à primeira solução

<p>Solução</p> 	<p>Pontuação na questão:</p> <hr/> <p>Critério/justificativa para a pontuação</p> <hr/> <hr/> <hr/> <hr/>
<p>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</p> <hr/> <hr/> <hr/>	

Fonte: elaborado pelo autor

As seis soluções dos estudantes fictícios foram elaboradas a partir da experiência do pesquisador (Figura 9). A ideia era diversificar as respostas entre soluções erradas, incompletas e alternativas. As soluções dos estudantes A e C apresentam uma das duas operações realizadas com algum erro. Na solução A o erro está na subtração (segunda operação) e na solução C, na multiplicação (primeira operação). Nesse caso, a ideia era observar se, e como, a posição do erro influenciaria a resposta. A solução do estudante B foi planejada para observar como os corretores interpretariam uma solução composta apenas da resposta final correta. Em relação às soluções dos estudantes D e E pretendíamos observar como os corretores iriam interpretar soluções alternativas. E por último, como os corretores interpretariam a solução incompleta do estudante F.

Figura 9 – As seis soluções utilizadas na investigação



Fonte: elaborado pelo autor

De acordo com Bruner (1987), há dois modos de conhecer e pensar, cada um com suas próprias maneiras distintas de ordenar a experiência, construir a realidade e compreender o mundo: o modo pragmático e o modo narrativo. O modo pragmático compreende o positivismo clássico, estando relacionado a uma forma de conhecer e pensar pertencente à tradição lógico-científica herdada (BOLÍVAR BOTÍA, 2002).

O instrumento de coleta apresentado nesta seção foi projetado em uma perspectiva pragmática, positivista. No entanto, poderia ser contraditório, mesmo incoerente em relação à fundamentação teórico-metodológica, utilizar exclusivamente uma abordagem pragmática em uma pesquisa que critica a visão positivista da avaliação escolar. Segundo Bolívar Botía (2002, p. 42), “estamos, portanto, em uma crise dos modos paradigmáticos de saber instituídos, onde se repensa o papel do sujeito investigador e a necessidade de incluir a subjetividade no processo de compreensão da realidade”.

Tanto na avaliação pedagógica, quanto na pesquisa em Educação Matemática, consideramos que a questão essencial não é abandonar concepções positivistas, pragmáticas, e sim, reconhecer a influência e interferência humana na elaboração e análise de avaliações e pesquisas. Trata-se de reconhecer e aprender a lidar com a subjetividade inerente ao processo de produção de conhecimento em relações humanas.

Em pesquisas educacionais, os modos pragmáticos e narrativos de produzir conhecimento são legítimos, complementares e irredutíveis (BRUNER, 1987; BOLÍVAR BOTÍA, 2002). Nesse sentido, analisamos as respostas dos professores utilizando algumas ideias metodológicas pragmáticas e narrativas. Por um lado, olhamos para o que é comum, semelhante e passível de uma generalidade por meio de processos de categorização e, por outro investigamos singularidades, maneiras idiossincráticas dos professores na correção de questões avaliativas de Matemática.

2.2 O modo pragmático de interpretar as conchas

O modo pragmático de produzir conhecimento tem seu foco no que é comum, semelhante e possível de ser agrupado. O pensamento pragmático “se concentra no que torna o item um membro de uma categoria. Não se concentra no que o torna diferente de outros membros da categoria” (POLKINGHORNE, 1995, p. 10). A ideia central é

estabelecer a categoria a que pertence cada uma das instâncias individuais, de incluir o particular no formal (categoria ou conceito), anulando qualquer diferença individual, que deve ser classificável. O modo paradigmático

concentra-se, especialmente, nos atributos que definem itens específicos como instâncias de uma categoria, e não no que diferencia alguns e outros membros de uma categoria (BOLÍVAR BOTÍA, 2002, p. 49, tradução nossa).

Nesse sentido, a Análise de Conteúdo pode ser classificada como um modo pragmático de produção de conhecimento. Diversas pesquisas realizadas no Brasil que investigam as respostas dos estudantes utilizam a Análise de Conteúdo com o intuito de “descrever as maneiras dos alunos lidarem com as questões em Matemática, buscando a partir disso, a construção de inferências sobre seus conhecimentos, em processos recursivos de construção de unidades de análises e categorizações” (VIOLA DOS SANTOS; DALTO, 2012, p. 2).

Silva e Buriasco (2005), por exemplo, utilizaram a Análise de Conteúdo para analisar a produção escrita de Matemática de estudantes do 5º ano do Ensino Fundamental. Para as autoras, a análise da produção escrita dos estudantes compreende um processo de "desconstrução" e "construção" para então realizar inferências: “o que o aluno fez? Por que fez assim? É isso mesmo que ele queria dizer? Realizar inferências serviu para que pudéssemos levantar hipóteses, estabelecer” (SILVA, BURIASCO, 2005, p. 503). As análises qualitativas

não se limitam a identificar um conjunto de categorias que fornecem identidade aos elementos específicos do banco de dados. Busca um segundo nível de análise que identifica as relações que existem entre as categorias estabelecidas. Esta análise busca mostrar como as categorias se vinculam (POLKINGHORNE, 1995, p. 10, tradução nossa).

Trevisan e Buriasco (2014) investigaram a análise de produção escrita de estudantes do segundo ano do Ensino Médio envolvendo a Prova em Fases. Os autores recorreram à Análise de Conteúdo, por considerá-la útil como estratégia metodológica para compreender as resoluções dos estudantes às questões propostas – interpretada pelos autores como um tipo de comunicação – além de seus significados imediatos.

A categorização utilizada nos trabalhos supracitados é considerada como um processo de

classificação de elementos constitutivos de um conjunto, por diferenciação e, seguidamente, por reagrupamento segundo o gênero (analogia), com os critérios previamente definidos. As categorias são rubricas ou classes, as quais reúnem um grupo de elementos (unidades de registro, no caso da análise de conteúdo) sob um título genérico, agrupamento esse efetuado em razão dos caracteres comuns destes elementos (BARDIN, 1977, p. 117).

Para Bardin (1977), o critério de classificação pode ser semântico, sintático, léxico ou expressivo. O semântico está relacionado aos significados dados pelos respondentes a algo em determinado contexto. O sintático está relacionado aos verbos e adjetivos utilizados, o léxico à “classificação das palavras segundo o seu sentido, com emparelhamento dos sinônimos e dos sentidos próximos e o expressivo, por exemplo, às categorias que classificam as diversas perturbações da linguagem” (BARDIN, 1977, p. 118).

A categorização é um processo estruturalista composto por duas etapas: a primeira delas é o inventário, em que se isola todos os elementos, a segunda é a classificação, em que os elementos são agrupados afim de organizar a mensagem (BARDIN, 1977). Bardin (1977, p. 119) afirma que “a análise de conteúdo assenta implicitamente na crença de que a categorização (passagem de dados brutos a dados organizados) não introduz desvios (por excesso ou por recusa) no material, mas que dá a conhecer índices invisíveis, ao nível dos dados brutos”. Neste sentido, é importante destacar que no

processo de fragmentação de um texto necessariamente se perde parte da informação do material analisado. A leitura feita representará sempre uma perspectiva do pesquisador. Entretanto, na medida em que se tem consciência de que não existe uma leitura objetiva e completa de um texto, esta perda de informação pode ser justificada pelo aprofundamento em compreensão que a análise possibilita. (MORAES, 1999, p. 6)

A categorização pode ser desenvolvida utilizando dois processos distintos: o de caixas e o de milha. No processo de caixas, o sistema de categorias é definido antes da coleta, a partir de elementos teóricos e hipotéticos. No procedimento por milha o sistema de categorias é desenvolvido durante a análise de dados: “o sistema de categorias não é fornecido antes, resultando da classificação analógica e progressiva dos instrumentos” (BARDIN, 1977, p. 119).

A Análise de Conteúdo parte do pressuposto que há nas produções escritas um conhecimento velado e escondido, e a partir de descrições sucessivas, realiza-se várias leituras para elaborar critérios e, com isso construir categorias. Daí, elabora-se algumas inferências com o objetivo de desvelar, descortinar verdades por meio da produção escrita (VIOLA DOS SANTOS; DALTO, 2012). O raciocínio paradigmático produz redes cognitivas de conceitos que permitem às pessoas construir experiências familiares, enfatizando os elementos comuns que aparecem continuamente. Ao fornecer um conhecimento familiar e descontextualizado do mundo, eles nos permitem gerenciar a

singularidade e a diversidade de cada experiência como se fosse a mesma das experiências anteriores (POLKINGHORNE, 1995).

2.30 modo narrativo de analisar as conchas

Enquanto na Análise de Conteúdo busca-se o comum, o semelhante, na análise Narrativa busca-se o singular, o particular. Enquanto o pensamento pragmático move-se do comum para o geral, o efeito cumulativo do pensamento narrativo é uma coleção de casos individuais nos quais o pensamento se move de caso a caso (POLKINGHORNE, 1995).

Nossa experiência e nossa memória são organizadas a partir de narrativas (BRUNER, 1987, 1991), de modo que a narrativa não apenas expressa dimensões importantes da experiência vivida, mas, mais radicalmente, medeia a própria experiência e configura a construção social da realidade (BOLÍVAR BOTÍA, 2002). Nós percebemos e interpretamos a realidade cotidiana e

de um modo muito particular, damos sentido às situações por meio do nosso universo de crenças, elaborado a partir das vivências, valores e papéis culturais inerentes ao grupo social a que pertencemos. As representações nos permitem decodificar e interpretar as situações que vivemos. [...] Os nossos filtros interpretativos nos permitem apropriarmo-nos dessa realidade e agir sobre ela utilizando, por vezes, modelos que antecipam o comportamento dos outros. E assim vamos construindo um percurso individual feito de cruzamentos de histórias que vivemos ou que ouvimos contar (GALVÃO, 2005, p. 325).

A preocupação no modo narrativo não é identificar cada caso em uma categoria geral. Uma coleção de experiências históricas fornece uma base para a compreensão de novos episódios de ação por meio de analogia.

A coleção de histórias é pesquisada para encontrar uma que seja semelhante em alguns aspectos à nova. A preocupação não é identificar o novo episódio como uma instância de um tipo geral, mas como semelhante a um episódio específico lembrado. O novo episódio é considerado semelhante, mas não igual ao episódio selecionado anteriormente. Assim, a compreensão da nova ação pode basear-se na compreensão anterior, ao mesmo tempo que se abre para os elementos específicos e únicos que tornam o novo episódio diferente de todos os anteriores. A compreensão analógica reconhece a improvisação e a mudança que constituem a variabilidade flexível do comportamento humano. (POLKINGHORNE, 1995, p. 11, tradução nossa)

Enquanto o pensamento pragmático é construído e orientado pela argumentação e por princípios logicamente estruturados, o pensamento narrativo se destaca por explicitar subjetividades (LIMA; GERALDI; GERALDI, 2015). As explicações dadas pelos professores contêm crenças e valores e a análise dessas explicações é pessoal e

situacional. Qualquer significado atribuído a essas explicações é fluido e contextual, não sendo nem fixo, nem universal (GALVÃO, 2005). Por conseguinte, “qualquer abordagem metodológica é, por natureza, incompleta, parcial e historicamente contingente” (RIESSMAN, 1993 apud GALVÃO, 2005, p. 332). O objetivo da análise narrativa é

produzir histórias como resultado da pesquisa. Os elementos de dados necessários para essa produção são descrições diacrônicas de eventos e acontecimentos. A análise narrativa compõe esses elementos em uma história. O pesquisador começa com perguntas como "Como isso aconteceu?" ou "Por que isso aconteceu?" e busca informações que contribuam para a construção de uma história que forneça uma resposta explicativa às questões. As informações podem vir de várias fontes, incluindo entrevistas, periódicos, documentos públicos e pessoais e observações. (POLKINGHORNE, 1995, p. 15, tradução nossa)

O trabalho com narrativas em Educação Matemática possibilita a produção de um novo olhar, de novas aberturas, do surgimento do novo em meio às diferenças. “Olhar para essas diferenças é uma postura ética, estética e política como formas de luta e resistência frente aos discursos hegemônicos, cujos compromissos são com a verdade no âmbito moral” (SILVA, 2020, p. 329).

A pesquisa narrativa constitui-se em um campo multifacetado e multidisciplinar, composto de diversas abordagens possíveis para o trabalho analítico (GALVÃO, 2005). Narrativas podem ser consideradas como abordagem metodológica, lente teórica ou uma combinação dos dois enfoques (ZACCARELLI; GODOY, 2013). Podemos considerar três possibilidades para a produção de textos narrativos:

(i) histórias contadas por participantes de pesquisa (que são, em si mesmos, interpretativos); (ii) *relatos interpretativos desenvolvidos por um investigador, baseados em entrevistas e observação de trabalho de campo* (uma história sobre histórias); e, ainda, (iii) narrativas que o leitor constrói após se engajar com as dos participantes e do investigador (RIESSMAN, 2008 apud ZACCARELLI; GODOY, 2013, p. 27, *grifo nosso*).

Nesta tese, utilizaremos a segunda possibilidade destacada por Zaccarelli e Godoy (2013), pois produziremos um relato interpretativo desenvolvido com base nas respostas fornecidas pelos docentes e licenciandos na *folha de correção*. Na Análise Narrativa, a produção de um relato ocorre a partir “de alguns traços, marcas e especificidades” contidas no texto analisado ou nos dados coletados, não havendo a necessidade de analisar todos os aspectos e também não tendo a pretensão “de olhar a totalidade dos possíveis modos de produzir significados” (VIOLA DOS SANTOS; DALTO, 2012, p. 13).

A análise narrativa pressupõe a exploração não só do que é dito, mas também de como é dito. Olha-se para o conteúdo e para a forma, podendo examinar-se o modo figurativo como a linguagem é usada. Metáforas, analogias, semelhanças e outros tipos de imagens, fornecem indicação sobre um significado diferente do que é dito. (GALVÃO, 2005, p. 335)

Prestar atenção aos professores, como indivíduos e como grupo, ouvindo suas vozes e as histórias que contam a respeito do seu trabalho e das suas vidas é uma condição necessária para compreender o ensino (ELBAZ-LUWISCH, 2002). Analisando as ideias de Bruner (1987) a respeito da análise pragmática e a análise narrativa, Polkinghorne (1995) propõe uma diferenciação entre Análise Narrativa, que emprega o raciocínio narrativo e Análise de Narrativas, que utiliza o raciocínio pragmático

Análise de narrativas: os dados recolhidos são histórias ou narrativas contadas pelos sujeitos da investigação; estes dados da narrativa são tratados de uma forma inteiramente paradigmática, analisados por temas comuns, imagens, enredos, etc., com o objetivo de se identificarem ideias gerais ou de se clarificarem conceitos a partir das narrativas recolhidas. *Análise narrativa:* são recolhidos dados experimentais sem que o objetivo do investigador é organizar os dados, de modo a criar uma narrativa com um enredo que unifique esses dados. A história daí resultante serve como uma explicação narrativa da história de vida ou do fenômeno que está a ser examinado. (ELBAZ-LUWISCH, 2002, p. 3, grifo nosso).

Na análise narrativa, os pesquisadores coletam dados, realizam entrevistas e os sintetizam ou configuram por meio de um enredo em uma história. Assim, na análise das narrativas as histórias são transformadas em elementos comuns e na análise da narrativa os elementos comuns são transformados em as histórias (POLKINGHORNE, 1995). Em relação à investigação em Educação Matemática, os textos narrativos como instrumento de pesquisa “possibilitam identificar, compreender e analisar como se dá o processo de produção de conhecimento através da investigação/exploração Matemática” (PASSOS; GALVÃO, 2011, p. 87).

A pesquisa narrativa, assim como a pesquisa convencional, na maioria das vezes envolve dois desempenhos: (a) a coleta de evidências e (b) a análise ou interpretação das evidências (POLKINGHORNE, 2007). Entretanto, a Análise Narrativa não tem por objetivo revelar *a verdade*, elemento único e imutável, e sim, construir uma interpretação possível de *uma verdade*, elemento situado em um contexto social e histórico. Não se trata de comprovar, e sim, de compreender a realidade. E nessa perspectiva o próprio conceito de validade é ressignificado.

Os pesquisadores narrativos não pedem aos leitores que concedam validade às suas afirmações apenas quando alcançam um nível de quase certeza sobre uma

afirmação. Os leitores são solicitados a fazer julgamentos sobre se a evidência e o argumento os convencem ou não no nível de plausibilidade, credibilidade ou confiabilidade da afirmação. (POLKINGHORNE, 2007, p. 477, tradução nossa)

A validade em Análise Narrativa é obtida a partir dos argumentos e não, exclusivamente, pelo rigor metodológico utilizado para coletar e analisar os dados. “Os pesquisadores narrativos precisam argumentar pela aceitação da validade das evidências coletadas e pela validade da interpretação oferecida” (POLKINGHORNE, 2007, p. 478). Isso não significa que as perspectivas metodológicas nessa forma de fazer pesquisa não devam ser bem delineadas pelo pesquisador.

2.4 Nossa forma de coletar e interpretar conchas

A interação entre o investigador e os professores/licenciandos, participantes da pesquisa, foi uma limitação para a utilização da Análise Narrativa nesta tese. Isto porque a investigação em Análise Narrativa

pressupõe uma postura metodológica firmada na interação entre investigador e participante, um jogo de relações baseado na confiança mútua e na aceitação da importância da intervenção de cada um na coleta dos dados e na sua interpretação. (GALVÃO, 2005, p. 342)

Nossa coleta de conchas foi planejada para uma análise pragmática. Não houve nenhum tipo de interação entre o pesquisador e os respondentes, além da palestra proferida no dia da aplicação do questionário. A decisão de utilizar a Análise Narrativa surgiu durante a análise das respostas, após a qualificação. Consequentemente, não seria coerente com a teoria estudada utilizar exclusivamente a Análise Narrativa.

Comprendemos, em consonância com as ideias defendidas por Silva (2020, p. 20) em sua tese de doutorado, que “em toda e qualquer pesquisa não se usa um método, mas se cria um método, dadas as circunstâncias, as singularidades e as sensibilidades de cada pesquisador e sua relação com o objeto de estudo”. Metaforicamente, nossa forma de coletar e interpretar as conchas e descobrir pérolas é particular.

Utilizamos a perspectiva discutida e citada anteriormente de que os modos pragmáticos e narrativos de produzir conhecimento são legítimos, complementares e irreduzíveis em pesquisas educacionais (BRUNER, 1987; BOLÍVAR BOTÍA, 2002). Nesta pesquisa, por um lado, procuramos valorizar as particularidades de cada respondente, identificar aspectos relevantes e singulares das respostas, de alguma forma

tentando valorizar suas ‘falas’, características da Análise Narrativa. Por outro, procuramos obter generalizações a partir das interpretações dos 51 respondentes, característica da pesquisa pragmática.

As respostas foram analisadas sem qualquer distinção prévia dos respondentes, por acreditar que todos eles compõem um corpo docente representativo de uma realidade. Os alunos de graduação e da pós influenciam e são influenciados, inevitavelmente e constantemente, pelos seus professores (formadores). São, todos eles, aprendizes, professores e corretores. No entanto, havendo qualquer peculiaridade na resposta dos professores formadores ou dos estudantes, esta será destacada na análise. Nosso objetivo foi compreender essa realidade, e não comprovar qualquer hipótese sobre a avaliação.

Em ambas as universidades, a palestra foi dividida em três momentos. No primeiro momento temas relacionados à avaliação em geral eram apresentados e debatidos (avaliação formativa e somativa, a avaliação é um processo, a prova é um instrumento e etc.). No segundo momento, as folhas de correção foram entregues. Os participantes foram convidados a realizar a correção, identificando apenas sua formação (licenciando, pós-graduando ou professor universitário). A ideia era corrigir as produções de fictícios estudantes do sexto ano do Ensino Fundamental. Antes do recolhimento, o palestrante pediu que ninguém alterasse a correção realizada, e sem seguida, as questões foram debatidas. O terceiro momento se iniciou com o recolhimento das folhas com as respostas corrigidas e a discussão de tópicos relacionados à multcorreção.

Após o recolhimento das *folhas de correção*, os dados foram analisados da seguinte forma:

Etapa 1 – Leitura inicial

Nessa etapa, ocorreu a leitura das respostas a fim de encontrar elementos relevantes nos comentários dos corretores. Com inspiração no modelo de milhas no desenvolvimento de categorias (BARDIN, 1977), alguns pontos foram observados e analisados durante a leitura. Não foram estabelecidas categorias no sentido de Bardin (1979) – não são mutuamente excludentes, por exemplo –, são apenas aspectos identificados nas respostas.

Etapa 2 – Análise das respostas

Após a identificação de pontos/tópicos de discussão, houve novas leituras para encontrar particularidades nas respostas que representassem as idiossincrasias dos sujeitos envolvidos no estudo. Houve, também, um tratamento estatístico básico das pontuações para a construção de um panorama das pontuações, a fim de apoiar a discussão realizada. As pontuações contribuíram, dessa forma, para a análise das respostas encontradas nas justificativas e nos *feedbacks* dos corretores. Dessa análise emergiram alguns aspectos relevantes para compreender o processo de correção de questões de resposta aberta, a interpretação do erro em Matemática e outros pontos relacionados à avaliação. Esses tópicos serão discutidos no capítulo 5.

Nos capítulos seguintes, apresentaremos elementos da fundamentação teórica e da revisão de literatura sobre dois aspectos centrais nesta tese: o *feedback*, capítulo 3, e o erro, capítulo 4.

CAPÍTULO 3 – OBSERVANDO O FEEDBACK DO ALÉM MAR



*Assim moro em meu sonho:
como um peixe no mar.
O que sou é o que vejo.
Vejo e sou meu olhar.
Água é o meu próprio corpo,
simplesmente mais denso.
E meu corpo é minha alma,
e o que sinto é o que penso.
Assim vou no meu sonho.
Se outra fui, se perdeu.
É o mundo que me envolve?
Ou sou contorno seu?*
Cecília Meireles

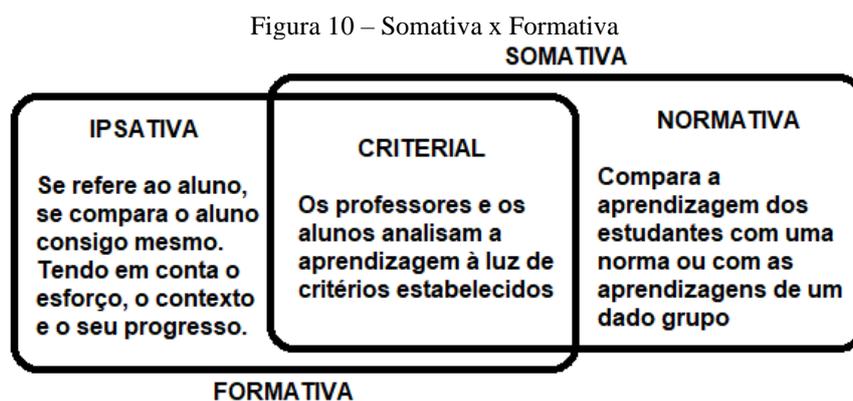
Para Santos (2019), o *feedback* é uma ação fundamentada e reativa tomada pelo professor. Na concepção somativa, o *feedback* corresponde “a todo comentário que informa sobre a qualidade do objeto sujeito à avaliação” (p. 173), e podemos incluir aí as notas ou conceitos atribuídos após a correção de uma prova, por exemplo. O *feedback*, em uma perspectiva formativa, consiste em toda informação que permite ao aluno identificar o que falta fazer e como fazer para alcançar o esperado.

3.1 Percorrendo as funções da avaliação

Para Hadji (2001), a avaliação em um contexto de ensino deveria ter por finalidade a contribuição para o êxito do ensino, ajudando os alunos a desenvolverem saberes e competências. Em relação à avaliação escolar, existem “duas ideias muito persistentes: uma que associa a avaliação à nota/classificação; outra aos testes ou exames, ou seja, aos instrumentos de recolha de informação” (SANTOS, PINTO, 2018, p. 504). Essas ideias sustentam a concepção somativa da avaliação. A avaliação somativa, normalmente

aplicada no fim de um período de ensino e aprendizagem, tem por finalidade certificar e fazer um balanço dos objetivos alcançados pelos estudantes (PINTO, 2019).

Fernandes (2009) considera que a avaliação somativa está mais diretamente associada à certificação, feita pelos professores e seria perfeitamente dispensável do ponto de vista da melhoria da aprendizagem. A avaliação somativa comumente “não está integrada ao ensino, não é utilizada sistematicamente para melhorar as aprendizagens, nem proporciona um *feedback* que oriente os alunos” (FERNANDES, 2009, p. 72-73). Tanto na avaliação formativa quanto na somativa, a produção dos estudantes é analisada à luz de critérios bem definidos, uma das suas diferenças, segundo o autor, está no referencial de comparação do desempenho. Na somativa, o desempenho de um estudante é comparado com o desempenho dos outros e na formativa com o desempenho dele mesmo. Tal perspectiva de Fernandes (2009) é apresentada na Figura 10.



Fonte: FERNANDES (2009)

No entanto, a principal diferenciação dessas modalidades de avaliação é a função (ou funções) para qual ela é pensada ou executada. São os propósitos para os quais a avaliação se destina que permitem diferenciar avaliações formativas de somativas, por exemplo, não o seu tipo. Ou seja, uma prova escrita tradicional de Matemática – compreendida como exame escrito, individual, sem consulta e com tempo delimitado – pode ter um caráter mais formativo que somativo (SANTOS, 2011).

Segundo Perrenoud (1999), a avaliação formativa esteve durante muito tempo associada à imagem da aplicação de um teste após um período de aprendizagem, seguido de ações de remediação para os alunos que não demonstrassem o domínio dos conhecimentos avaliados neste exame. Por essa concepção, os professores atuavam com a perspectiva formativa durante os testes em períodos específicos, limitados e

distanciados. Tal concepção evoluiu através do reconhecimento de ações de observação, intervenção e regulação mais frequentes e ampliadas.

Uma avaliação mais formativa não toma menos tempo, mas dá informações, identifica e explica erros, sugere interpretações quanto às estratégias e atitudes dos alunos e, portanto, alimenta diretamente a ação pedagógica, ao passo que o tempo e a energia gastos na avaliação tradicional desviam da intenção didática e das inovações. (PERRENOUD, 1999, p. 68)

Para Black e Wiliam (1998), a avaliação é formativa quando seus resultados são utilizados para adaptar o ensino e identificar as dificuldades dos estudantes. Estes dois pesquisadores realizaram um amplo estudo bibliográfico com pesquisadores de diversos países sobre a avaliação formativa. “Examinamos vários livros e, nos últimos nove anos, edições de mais de 160 periódicos, e estudamos revisões de pesquisas anteriores. Esse processo rendeu cerca de 580 artigos ou capítulos para estudar” (p. 140). Uma das conclusões foi que as práticas formativas trazem diversos benefícios para a aprendizagem, principalmente dos estudantes com menor desempenho acadêmico. A avaliação formativa é uma arma poderosa desde que seja comunicada de forma correta, alertam Black e Wiliam (1998).

Os alunos podem aceitar e trabalhar com essas mensagens, desde que não sejam ofuscadas por excesso de habilidade, competição e comparação com outras pessoas. Em resumo, a mensagem pode ser expressa da seguinte forma: o *feedback* para qualquer aluno deve ser sobre as qualidades particulares de seu trabalho, com conselhos sobre o que ele pode fazer para melhorar e deve evitar comparações com outros alunos. (BLACK; WILIAM, 1998, p. 143, tradução nossa)

Para Santos (2019), a avaliação somativa tem uma dimensão social, enquanto que a formativa, uma dimensão pedagógica. A primeira dirige-se às entidades externas e aos alunos e a segunda envolve os intervenientes diretos do processo de ensino e aprendizagem. A palavra observação estaria relacionada à construção de “uma representação realista das aprendizagens, de suas condições, de suas modalidades, de seus mecanismos, de seus resultados” (PERRENOUD, 1999, p. 104), já a palavra avaliação estaria muito vinculada à ideia de medida, de classificação e de contabilização do conhecimento. A observação será formativa quando seu foco estiver na orientação e otimização da aprendizagem e não na classificação, certificação e seleção.

A observação formativa pode ser instrumentada ou puramente intuitiva, aprofundada ou superficial, deliberada ou acidental, quantitativa ou qualitativa, longa ou curta, original ou banal, rigorosa ou aproximativa, pontual ou sistemática. Nenhuma informação é excluída a priori, nenhuma modalidade de percepção e de tratamento é descartada. (PERRENOUD, 1999, p. 104)

De acordo com Hadji (2001), “a avaliação só é formativa se for informativa” (p. 79). Para ser informativa, a avaliação deve responder perguntas que permitam conhecer o *objeto* avaliado, compreender do que o aluno é capaz, o que ele compreendeu e “se sabe, se sabe fazer, se sabe ser... (o quê?)” (p. 79).

O conceito de avaliação formativa foi inicialmente consolidado no desenvolvimento teórico do behaviorismo e, posteriormente, integrado em outras perspectivas teóricas relacionadas ao cognitivismo.

Os primeiros [behavioristas] usam-na mais frequentemente na análise de resultados, em um quadro de definição de objetivos muito específicos (comportamentais) e de tarefas que testam cada um desses objetivos, ao passo que os segundos [cognitivistas] utilizam-na mais na análise dos processos de aprendizagem dos alunos em um quadro de definição mais abrangente e integrada de objetivos e de tarefas que avaliam um leque mais amplo e integrado de saberes. (FERNANDES, 2009, p. 50)

Na perspectiva behaviorista associada a Pavlov, Watson e Skinner, o aprendizado era um produto do condicionamento.

O estímulo e resposta observáveis eram estudados cientificamente - não havia lugar para variáveis intervenientes invisíveis, como a 'mente'. A chave para o aprendizado era o experimentador ou professor manipular o ambiente para criar as respostas desejadas, principalmente por meio de reforço e *feedback* corretivo. (STOBART; HOPFENBECK, 2014, p. 35, tradução nossa)

As avaliações formativas de concepção behaviorista, normalmente constituídas por testes que *não valem nota* e utilizadas no final do período, prevalecem nos sistemas educacionais (BLACK; WILIAM, 1998; FERNANDES, 2009). As práticas avaliativas atuais, principalmente nas escolas dos EUA, sofrem forte influência das teorias behavioristas: “o uso generalizado de testes como *formativos* pode ser interpretado como oriundo da abordagem de *teste e correção* de Bloom” (STOBART; HOPFENBECK, 2014, p.35, grifo dos autores). Por outro lado, tal concepção tem sido fonte de controvérsia nos EUA, pois os educadores argumentam que a avaliação formativa não deveria ser interpretada como um teste, e sim, como um processo.

Há um problema relacionado à percepção de professores no que tange à utilização de avaliações formativas. Frequentemente percebem “a avaliação formativa como um complemento ao ensino, algo extra em vez de entendê-lo como parte central e integrada do ensino-aprendizagem e aprendizado” (DOLIN et al, 2018, p. 59). De certa forma, segundo esses pesquisadores, os professores estão certos, pois o aprendizado leva tempo,

os docentes possuem inúmeras demandas e, normalmente, atuam em um currículo sobrecarregado.

Para Fernandes (2009), as avaliações formativas behavioristas se caracterizam por serem pontuais, pouco ou nada interativas, por exigirem pouca participação dos alunos e por serem orientadas para a verificação da consecução de objetivos comportamentais e apresentarem reduzida exigência cognitiva.

Segundo Stobart e Hopfenbeck (2014, p. 36), as abordagens cognitivas rejeitaram a concepção behaviorista de estímulo-resposta da aprendizagem e “enfatizaram o papel dos indivíduos na construção de significado para si mesmos”. Na concepção cognitivista, “o foco está em *como a mente funciona* - como as informações são processadas e como os alunos *dão sentido* às informações”. Além disso, as teorias cognitivas trouxeram contribuições para a avaliação formativa principalmente em relação ao *feedback*, que passou a ser interpretado como um processo cognitivo muito mais complexo do que nas teorizações anteriores. A ênfase aqui está no desenvolvimento da aprendizagem autorregulada (STOBART; HOPFENBECK, 2014).

O papel limitado dado ao contexto social da aprendizagem nas abordagens cognitivas levou a um apelo, a um reconhecimento mais direto do social no processo de compreender as informações. As abordagens socioconstrutivistas se distinguem das socioculturais pelo reconhecimento da cognição individual e se assemelham no reconhecimento da importância da cultura e do contexto no ensino e na aprendizagem (STOBART; HOPFENBECK, 2014). Neste sentido, Fernandes (2009, p. 56) propôs a utilização do termo Avaliação Formativa Alternativa (AFA) para designar avaliações baseadas em “princípios que decorrem do cognitivismo, do construtivismo, da psicologia social e das teorias socioculturais e sociocognitivas”. A Avaliação Formativa Alternativa é

um processo eminentemente pedagógico, plenamente integrado ao ensino e à aprendizagem, deliberado, interativo, cuja a principal função é a de regular e de melhorar a aprendizagem dos alunos. Ou seja, de conseguir que os alunos aprendam melhor, com compreensão, utilizando o desenvolvimento de suas competências, nomeadamente as do domínio cognitivo e metacognitivo. (FERNANDES, 2009, p. 59)

Para Fernandes (2009, p. 83), a avaliação formativa alternativa (AFA) rompe o modelo imposto como padrão avaliativo, descrito como paradigma psicométrico de

avaliação. Pois, amparada na perspectiva construtivista, a avaliação das aprendizagens é “um processo complexo, subjetivo e que não nos permite apreender na totalidade o que os alunos sabem e são capazes de fazer”.

Um ponto de destaque na AFA é relevante para essa pesquisa de tese, o papel do *feedback*.

3.2 O *feedback* e a avaliação

Para Hattie e Timperley (2007), um *feedback* no ensino é algum tipo de informação oferecida por um agente ao estudante (por exemplo, professor, colegas de classe, o livro texto, os pais ou o próprio estudante), em relação aos aspectos de seu desempenho ou compreensão. Um professor ou um pai pode fornecer informações corretivas, um colega de classe pode fornecer uma estratégia alternativa, um livro pode fornecer informações para esclarecer ideias. Todas essas opções são exemplos de *feedbacks* existentes no processo de ensino-aprendizagem.

O feedback é a informação com a qual um aluno pode confirmar, adicionar, sobrescrever, ajustar ou reestruturar informações na memória, sejam essas informações, conhecimento de domínio, conhecimento metacognitivo, crenças sobre si mesmo e tarefas, ou táticas e estratégias competitivas (WINNE; BUTLER, 1994; apud HATTIE; TIMPERLEY, 2007, p. 82, tradução nossa)

O *feedback* assume, segundo Fernandes (2009), um papel de destaque na avaliação formativa. Se o *feedback* dado em uma avaliação não conduz a nenhuma ação para a melhoria do ensino ou da aprendizagem, esta não pode ser caracterizada como uma avaliação alternativa, pois um *feedback* inteligente é uma condição necessária para tal. Sete são os tópicos que permitiriam caracterizar a AFA:

(1) a avaliação é deliberadamente organizada para proporcionar um *feedback* inteligente e de elevada qualidade, tendo em vista melhorar a aprendizagem dos alunos; (2) o *feedback* é determinante para ativar os processos cognitivos e metacognitivos dos alunos, que, por sua vez, regulam e controlam os processos de aprendizagem, assim como para melhorar a sua motivação e autoestima; (3) a natureza da interação e da comunicação entre professores e alunos é absolutamente central porque os professores têm que estabelecer *pontes* entre o que se considera ser importante aprender e o complexo mundo dos alunos (...); (4) os alunos são deliberada, ativa e sistematicamente envolvidos no processo do ensino e da aprendizagem, responsabilizando-se pelas suas aprendizagens e tendo amplas oportunidades para elaborarem suas respostas e partilharem o que e como compreenderam; (5) as tarefas propostas aos alunos que, desejavelmente, são tanto de ensino, como de avaliação e aprendizagem, são criteriosamente selecionadas e diversificadas, representam os domínios estruturantes do currículo e ativam os processos mais complexos do pensamento (...); (6) as tarefas refletem estreita relação entre as didáticas específicas das disciplinas, que se constituem como elementos de referência

indispensáveis, e a avaliação, que desempenha um papel relevante na regulação dos processos da aprendizagem; (7) o ambiente da avaliação das salas induz uma cultura positiva de sucesso baseada no princípio de que todos os alunos podem aprender. (FERNANDES, 2009, p. 60-61).

O *feedback* também é apontado como um dos elementos diferenciadores das avaliações formativas e somativas por Pinto (2019) como ilustra o Quadro 1.

Quadro 1 – Modalidades de avaliação

	Somativa	Formativa
Momentos	no fim de um período de aprendizagem.	ao longo de todo processo de aprendizagem.
Objetivos	verificar o nível de qualidade de desempenhos, estabelecer um nível de competências a partir de um referencial.	identificar a natureza dos erros, desenvolver estratégias para que os estudantes superem suas dificuldades e ajudar o estudante a autorregular-se, ganhando autonomia.
Avaliadores	professores e entidades externas.	os pares e os próprios estudantes dividem com o professor a tarefa de avaliar.
Resultados	o <i>feedback</i> se resume às notas e às explicações do porquê dessas notas.	o <i>feedback</i> tem por finalidade orientar as aprendizagens.

Fonte: Adaptado de PINTO (2019)

Existem outros elementos que propiciam, ou podem proporcionar, uma avaliação mais formativa. Nesse sentido, Black e Wiliam (1998) apontam algumas sugestões. A primeira delas diz respeito à escolha das tarefas que serão utilizadas em sala de aula e aquelas destinadas para serem desenvolvidas em casa. O termo “escolha das tarefas” não diz respeito apenas à seleção de questões, mas também ao modo como os professores desenvolverão essas atividades em sala de aula. Para esses autores, os docentes devem proporcionar oportunidades para que os alunos expressem sua compreensão em todas as partes do ensino, pois assim, haverá a interação pela qual a avaliação formativa se desenvolve e fortalece o aprendizado.

As tarefas devem ser justificadas em termos dos objetivos de aprendizado a que servem e só podem funcionar bem se houver oportunidades para os alunos comunicarem sua compreensão em evolução no planejamento. Discussões, observação de atividades e marcação de trabalhos escritos podem ser usados para fornecer essas oportunidades, mas é importante observar ou ouvir atentamente a conversa, a escrita e as ações através das quais os alunos desenvolvem e exibem o estado da escola, o entendimento deles. (BLACK; WILIAM, 1998, p. 143)

Outro aspecto relevante é que, às vezes, por insegurança, o professor apresenta ao aluno a resposta esperada em um momento em que o professor poderia e deveria ajudá-lo na construção de suas próprias respostas. Tal modo de agir pode transmitir ao estudante a falsa impressão de que “o objetivo do exercício é descobrir - ou adivinhar - qual resposta o professor espera ver ou ouvir” (BLACK; WILIAM, 1998, p. 143). Fazer perguntas aos seus alunos faz parte da atuação do professor, no entanto, se o professor não oferece tempo adequando para que os estudantes trabalhem na resposta, o aprendizado fica comprometido (BLACK; WILIAM, 1998; WILIAM, 2011). Existem duas consequências dessa atitude:

Uma é que, como as únicas perguntas que podem produzir respostas em tão pouco tempo são questões de fato, elas predominam. A outra é que os alunos nem tentam pensar em uma resposta. Como eles sabem que a resposta, seguida de outra pergunta, surgirá em alguns segundos, não há sentido em tentar. Também é geralmente o caso de apenas alguns alunos de uma turma responderem às perguntas do professor. O resto, então, deixa para esses poucos, sabendo que eles não podem responder tão rapidamente e não estão dispostos a correr o risco de cometer erros em público. (BLACK; WILIAM, 1998, p. 143)

Para quebrar esse ciclo, os professores podem solicitar que os estudantes pensem nas possíveis respostas em pares ou pequenos grupos. Uma outra possibilidade seria oferecer algumas opções de respostas para que os estudantes pensem nas possibilidades. “O essencial é que qualquer diálogo evoque uma reflexão ponderada, na qual todos os alunos possam ser incentivados a participar, pois somente assim o processo formativo poderá começar a funcionar” (BLACK; WILIAM, 1998, p. 144).

A última sugestão apontada por Black e Wiliam (1998) se relaciona com o próprio instrumento avaliativo. Um bom teste pode ser a causa da aprendizagem, pode promover aprendizagem. Neste sentido, é melhor que os testes sejam curtos e frequentes. Outro ponto essencial para garantir a qualidade do *feedback* destacado pelos autores seria oferecer questões de qualidade. De acordo com Dolin e colaboradores (2018), somente oferecer conceitos ou notas aos estudantes como a única forma de *feedback*, indicando apenas o quão bom é o trabalho, não é consistente como objetivo de usar a avaliação para ajudar na aprendizagem. Se os estudantes

recebem apenas conceitos ou notas, eles não se beneficiam do *feedback*. O pior cenário é aquele em que alguns alunos que obtêm notas baixas dessa vez, mas também obtiveram notas baixas da última vez e esperam obter notas baixas da próxima vez. Esse ciclo de repetidas falhas se torna parte de uma crença

compartilhada entre esses alunos e os colegas de classe deles (BLACK; WILIAM, 1998, p. 144, tradução nossa).

Black e Wiliam (1998) constataram que o *feedback* melhora o aprendizado quando fornece a cada aluno orientações específicas sobre pontos fortes e fracos, de preferência sem quaisquer notas gerais e que para propósitos formativos, um teste no final de uma unidade ou módulo de ensino não faz sentido; é tarde demais para trabalhar com os resultados.

O *feedback* sobre testes, trabalhos de aula e trabalho de casa deve fornecer orientações a cada aluno sobre como melhorar, e cada aluno deve receber ajuda e uma oportunidade para trabalhar na superação de suas dificuldades e no desenvolvimento de sua aprendizagem. Dentro do contexto de ensino, consistente com a visão social construtivista da aprendizagem, o *feedback* pode ser dado de aluno para professor, de professor para aluno ou de aluno para aluno. Comentários de aluno para professor permitem que ele saiba como os alunos estão respondendo às atividades de aprendizagem e permitem que o professor saiba que ação tomar para ajustar as oportunidades e desafios oferecidos aos alunos.

As 5 estratégias chave para o pleno desenvolvimento de uma avaliação formativa são:

(1) Clarificar, compartilhar e compreender as intenções de aprendizagem e os critérios do sucesso; (2) Engendrar discussões eficazes em sala de aula, atividades e tarefas que promovam evidências de aprendizagem; (3) Promover *feedback* que conduza à aprendizagem; (4) Motivar os estudantes a atuarem como pesquisadores da sua aprendizagem e a de seus pares; (5) Fazer com que os estudantes se reconheçam como responsáveis pela própria aprendizagem. (WILIAM, 2011, p. 46, tradução nossa)

Entretanto, se os alunos estiverem envolvidos em autoavaliação ou avaliação por pares, as etapas precisam ser explícitas e os alunos precisam conhecer claramente a proposta. Também é necessário que os professores desenvolvam um *feedback* mais estruturado para que, com isso, a avaliação apresente certo grau de confiabilidade (DOLIN et al, 2018).

De certa forma, o ensino deve ser adaptado e adaptável para a melhoria da aprendizagem dos estudantes (BLACK; WILIAM, 1998, WILIAM, 2011). Para isso é necessário que se desenvolva

uma cultura de sala de aula, de questionamento e pensamento profundo, na qual os alunos aprendam com discussões compartilhadas com professores e colegas. O que emerge muito claramente aqui é a indivisibilidade das instruções e práticas de avaliação formativa (BLACK; WILIAM, 1998, p. 146, tradução nossa).

Então, o *feedback* não é apenas uma informação sobre o conteúdo, como em uma perspectiva behaviorista. O *feedback* é também sobre a compreensão das construções que os alunos fizeram da experiência da aprendizagem, ele pode ser aceito, modificado ou rejeitado (HATTIE; TIMPERLEY, 2007). Hattie (1999) criou um índice para verificar o quão eficaz é o *feedback* para a aprendizagem. Para isso, promoveu uma ampla pesquisa estatística a partir de 500 meta-análises envolvendo mais de 180 mil estudos. Quanto maior o índice, mais impactante para a aprendizagem dos estudantes é a atividade pedagógica. Nessa classificação, o *feedback* recebe pontuação menor apenas que a Instrução Direta¹⁰ e o Ensino Recíproco¹¹. A Tabela 2 apresenta uma comparação do *Feedback* com alguns fatores que contribuem para a aprendizagem.

Tabela 2 – Índices médios de contribuição para a aprendizagem

Maiores influências na aprendizagem	Índice médio do efeito na aprendizagem
Instrução direta	0,93
Ensino recíproco	0,86
Feedback	0,79
Capacidade cognitiva prévia	0,71
Influências socioeconômicas	0,44
Lição de casa	0,41
Uso de calculadoras	0,24
Tamanho da classe	0,12
Reprovação	-0,12

Fonte: HATTIE; TIMPERLEY (2007)

A tabela mostra que o *feedback* é claramente benéfico e poderoso pra contribuir com a aprendizagem dos estudantes.

¹⁰ É um termo para o ensino tradicional, expositivo, usando palestras ou demonstrações do material para os alunos.

¹¹ O ensino recíproco é uma atividade instrucional que assume a forma de um diálogo entre professores e alunos sobre segmentos de texto com o objetivo de construir o significado do texto. É composto por quatro estratégias: prever, questionar, clarificar e resumir. (WITTER, 2004).

3.3 O *feedback* formativo

Após uma avaliação o professor pode oferecer diversos tipos de *feedbacks*. Eles podem ser escritos ou orais, podem ser através de notas, conceitos ou comentários. Diferentes *feedbacks* podem proporcionar contribuições distintas à aprendizagem.

De acordo com Marzano, Pickering e Pollock (2001), o *feedback* precisa da natureza ‘corretiva’, oportuna, específica ao critério e ao aluno envolvido. Suas pesquisas indicam que simplesmente dizer a um aluno que sua resposta em um teste é certa ou errada não possui nenhum efeito positivo na aprendizagem. Para um melhor *feedback* possível, uma explicação sobre o que é preciso e o que é impreciso na solução dos estudantes é necessária.

Um processo [de *feedback*] que se baseia nos pontos fortes das crianças, em vez de se concentrar em suas fraquezas, é altamente positivo. O fornecimento de *feedback* descritivo pode evitar fazer julgamentos avaliativos sobre os alunos e também dará aos professores mais informações para compartilhar com esse aluno, seus pais e outro professor. [...] O objetivo do *feedback* descritivo é coletar informações sobre o que um aluno aprendeu e como ele aprendeu, bem como para que cada aluno esteja ciente de seus próprios processos de aprendizado (BARRY, 2008, p. 5, tradução nossa).

Nicol e Macfarlane-Dick (2006) afirmam que as boas práticas de *feedback* são aquelas que fortalecem a capacidade dos alunos de autorregular seu próprio desempenho. Esses autores listam sete princípios que caracterizam boas práticas de *feedback*.

(1) Ajudar a esclarecer o que é um bom desempenho (metas, critérios, padrões esperados); (2) facilitar o desenvolvimento da autoavaliação (reflexão) na aprendizagem; (3) fornecer informações de alta qualidade aos alunos sobre sua aprendizagem; (4) incentivar o diálogo entre professores e colegas em torno da aprendizagem; (5) incentivar crenças motivacionais positivas e autoestima; (6) oferecer oportunidades para fechar a lacuna entre o desempenho atual e o desejado; (7) fornecer informações aos professores que podem ser usadas para ajudar a moldar o ensino (NICOL; MACFARLANE-DICK, 2006, p. 205, tradução nossa)

Santos e Pinto (2018) discutem diferentes tipologias de *feedback* encontrados na literatura. Apresentam o *feedback avaliativo*, associado a um juízo de valor, centrado nas características dos estudantes e o *feedback descritivo*, relacionado à produção do estudante, classificações propostas por Gipps, em 1999. Em seguida, citando Jorro em uma publicação de 2000, distinguem dois tipos de escrita avaliativa: a *anotação como transmissão de informação*, que se traduz por juízos de valor ou por enunciados vagos e a *anotação como diálogo* que procura questionar, dar pistas e incentivar a reflexão por parte do aluno (SANTOS; PINTO, 2018).

Hattie (1999) estudou o impacto dos diferentes tipos de *feedback* para a aprendizagem dos estudantes, concluindo que os *feedbacks* relacionados às informações sobre uma tarefa e como resolvê-la de forma mais eficaz são os mais impactantes para a aprendizagem. Já os *feedbacks* relacionados aos elogios, às recompensas e às punições são ineficazes para a aprendizagem.

O *feedback* eficaz deve responder a três perguntas principais feitas por um professor e/ou por aluno: *Para onde vou?* (Quais são os objetivos?), *como estou indo?* (Que progresso está sendo feito em direção à meta?) e *para onde ir a seguir?* (O que as atividades precisam ser realizadas para progredir melhor?) Estas questões correspondem às noções de *feed up*, *feed back* e *feed forward* (HATTIE; TIMPERLEY, 2007, p. 86, tradução nossa).

Hattie e Timperley (2007) propõem quatro categorias de *feedbacks*. A primeira seria sobre uma *tarefa* ou um produto, apontando se o trabalho está correto ou incorreto. Este nível de *feedback* pode incluir orientações para adquirir mais informações diferentes ou corretas. Por exemplo: *Essa resolução está incorreta!*

A segunda seria um *feedback* direcionado ao *processo*, um *feedback* mais direto visando aos processos de aprendizagem. *Você errou a adição de frações, pois não igualou corretamente os denominadores.*

O terceiro tipo é direcionado à *autorregulação*,

incluindo maior habilidade em autoavaliação ou confiança para se envolver mais em uma tarefa. [...]. Esse *feedback* pode ter grandes influências na autoeficácia, proficiências de autorregulação e auto crenças sobre os alunos como alunos, de modo que os alunos sejam incentivados ou informados sobre como continuar a tarefa melhor e com mais facilidade (HATTIE; TIMPERLEY, 2007, p. 90).

A quarta e última categoria é o *feedback* destinado ao *self*, caracterizado por elogios e críticas, que como citado anteriormente, contribuem muito pouco para a aprendizagem. Santos e Pinto (2018)¹² apresentam a classificação de *feedbacks* descrita no Quadro 2.

Quadro 2 – Classificação de *feedbacks*

<i>Feedback</i>	Tipo	Comentário
Pouca atenção! Não estudaste!	Avaliativo, reprovação; Transmissão de informação; Dirigido <i>ao self</i> .	O professor recorre ao seu poder, fazendo uma inferência não assente em evidência

¹² Santos e Pinto (2018) utilizam para propor a classificação como referência os estudos de Gipps (1999), Jorro (2000), Hattie e Timperley (2007).

Tens de estudar mais.	Avaliativo, castigo ou incentivo; Transmissão de informação; Dirigido <i>ao self</i> .	Estudar o quê? Não fornece suficientes pistas para o aluno prosseguir.
Bem feito!	Avaliativo, aprovação Transmissão de informação Dirigido à tarefa	Não preciso de fazer mais nada? O que está bem feito?
O teu trabalho está cheio de erros de cálculo! Corrige-os	Descritivo, aperfeiçoamento específico; Transmissão de informação; Dirigido à tarefa	Mas quais são? Não esclarece o suficiente para o aluno prosseguir
Se em vez destes valores tivesses outros chegarias à mesma conclusão? Experimenta e compara com os teus resultados. O que conclus?	Descritivo, construção do percurso a seguir; Anotação como diálogo; Dirigido ao processo	Dá pistas ao aluno como continuar/reformular
Relê o enunciado da tarefa. Vai anotando as diferentes informações. No final, compara-as com as que usaste. São as mesmas?	Descritivo, construção do percurso a seguir; Anotação como diálogo; Dirigido à autorregulação.	Dá pistas ao aluno como continuar/reformular

Fonte: SANTOS; PINTO (2018)

Esses critérios encontrados na literatura sobre quais seriamos melhores tipos de *feedback* foram utilizados neste trabalho para sustentar o conceito que defendo de *feedback formativo* (VAZ; NASSER, 2021).

O *feedback formativo* é aquele que oferece pistas de como continuar, é descritivo, dirigido à regulação, ou seja, que permite ao aluno identificar o que falta fazer e como fazer para alcançar o esperado. Propõe uma situação que leve o estudante a rever sua solução, identificando quais foram, exatamente, o(s) erro(s) do processo. Que o ajude, além de identificar o que errou, a compreender o porquê do erro e o que deve ser feito para corrigi-lo. O termo formativo será utilizado partindo da hipótese de que este tipo de *feedback* é aquele que apresenta características didáticas, podendo atribuir aos testes e aos demais instrumentos avaliativos uma função reguladora, mais formativa.

O *feedback não formativo*, por sua vez, se aproxima mais do *feedback* avaliativo, é dirigido ao *self* e não oferece pistas de como continuar. Oferece informações relacionadas a um julgamento, com informações gerais que não auxiliam diretamente na aprendizagem. Trata-se daquele que, no máximo e somente, identifica o erro, mas não oferece nenhuma ajuda para como superar a dificuldade. Foram enquadradas nessa categoria respostas elogios ou críticas não acompanhadas de nada mais.

O quadro 3 apresenta a relação entre a proposta utilizada nesta investigação de tese com algumas das categorizações presentes na literatura.

Quadro 3 – *Feedback* não formativos e *feedback* formativo

<i>feedback não formativo</i>		<i>feedback formativo</i>		
Avaliativo ou Descritivo		Descritivo		GIPPS, 1999
Anotação como transmissão de informação		Anotação como um diálogo		JORRO, 2000
Dirigido ao <i>self</i> ;	Dirigido à tarefa;	Dirigido ao Processo;	Dirigido à autorregulação	HATTIE; TIMPERLEY, 2007

Fonte: elaborado pelo autor

Como discutido nesta seção, de acordo com Gipps (1999), Jorro (2000), Hattie e Timperley (2007), Santos e Pinto (2018), alguns *feedbacks* podem oferecer melhores contribuições à regulação do ensino e das aprendizagens. Não estamos minimizando qualquer impacto que exista no oferecimento de elogios ou palavras incentivadoras dadas pelos professores aos seus estudantes. Um *feedback* dirigido ao *self* pode promover uma mudança de comportamento em um estudante de forma mais poderosa do que um comentário sobre o que errou. Sabemos o quão importante e impactante é um professor que inspira e motiva seus alunos. No entanto, compreendemos que o *feedback formativo* oferece uma maior probabilidade de que a regulação ocorra, e por conseguinte, a avaliação ganhe tons mais formativos.

Nós, humanos, comumente, vemos o mundo de forma binária. Ou preto, ou branco (SUMPTER, 2019). Nesse sentido, a categorização do feedback em formativo e não-formativo poderia oferecer limitações se fossem utilizados ou interpretados como *feedbacks* bons ou ruins. As categorias propostas devem ser pensadas em termos didáticos e probabilísticos. Um *feedback formativo* apresenta maiores e melhores características de regulação, com isso, provavelmente, oferecerá maiores contribuições à aprendizagem do que um *feedback* não-formativo. Retornaremos a essa discussão no capítulo 5.

No capítulo seguinte discutiremos um elemento central para a construção, não somente de um feedback formativo, como também, de uma avaliação reguladora, o erro.

CAPÍTULO 4—APRENDENDO COM OS ERROS DO CAMINHO



*Escolas que são asas não amam pássaros engaiolados.
O que elas amam são pássaros em voo.
Existem para dar aos pássaros coragem para voar.
Ensinar o voo, isso elas não podem fazer,
porque o voo já nasce dentro dos pássaros.
O voo não pode ser ensinado.
Só pode ser encorajado.*
Rubem Alves

As concepções, crenças, saberes acadêmicos e experienciais influenciam no modo como os professores corrigem e pontuam as questões e interpretam a avaliação. Essa interpretação está diretamente relacionada a como esses docentes compreendem o erro e a Matemática. Uma das questões propostas em fases anteriores a esta pesquisa apresentava a solução de um estudante fictício, do nono ano do ensino fundamental, que comete um erro ao realizar uma divisão na segunda parte da solução de uma questão. A Figura 11 ilustra a correção dessa resolução por dois professores.

Figura 11 – Comparativo entre duas correções

<p>4 – Dois terrenos retangulares A e B possuem a mesma área. O terreno A possui 30 metros de comprimento por 24 metros de largura. Calcule o perímetro do terreno B, considerando que este possui 15 metros de largura.</p> <p>Solução:</p> <table border="0"> <thead> <tr> <th>Terreno A</th> <th>Terreno B</th> </tr> </thead> <tbody> <tr> <td>$A = 30 \cdot 24$</td> <td>$A = 15 \cdot x$</td> </tr> <tr> <td>$A = 720 \text{ m}^2$</td> <td>$720 = 15x$</td> </tr> <tr> <td></td> <td>$x = \frac{720}{15} = 66$</td> </tr> </tbody> </table> <p>Perímetro = $15 + 15 + 66 + 66 = 162 \text{ m}$</p> <p style="text-align: right;">1.8</p>	Terreno A	Terreno B	$A = 30 \cdot 24$	$A = 15 \cdot x$	$A = 720 \text{ m}^2$	$720 = 15x$		$x = \frac{720}{15} = 66$	<p>4 – Dois terrenos retangulares A e B possuem a mesma área. O terreno A possui 30 metros de comprimento por 24 metros de largura. Calcule o perímetro do terreno B, considerando que este possui 15 metros de largura. (1.0)</p> <p>Solução:</p> <table border="0"> <thead> <tr> <th>Terreno A</th> <th>Terreno B</th> </tr> </thead> <tbody> <tr> <td>$A = 30 \cdot 24$</td> <td>$A = 15 \cdot x$</td> </tr> <tr> <td>$A = 720 \text{ m}^2$</td> <td>$720 = 15x$</td> </tr> <tr> <td></td> <td>$x = \frac{720}{15} = 66 \cdot 48$</td> </tr> </tbody> </table> <p>Perímetro = $15 + 15 + 66 + 66 = 162 \text{ m} = 126 \text{ m}$</p> <p style="text-align: right;">C</p>	Terreno A	Terreno B	$A = 30 \cdot 24$	$A = 15 \cdot x$	$A = 720 \text{ m}^2$	$720 = 15x$		$x = \frac{720}{15} = 66 \cdot 48$
Terreno A	Terreno B																
$A = 30 \cdot 24$	$A = 15 \cdot x$																
$A = 720 \text{ m}^2$	$720 = 15x$																
	$x = \frac{720}{15} = 66$																
Terreno A	Terreno B																
$A = 30 \cdot 24$	$A = 15 \cdot x$																
$A = 720 \text{ m}^2$	$720 = 15x$																
	$x = \frac{720}{15} = 66 \cdot 48$																

Fonte: VAZ; NASSER (2019)

A solução apresentada na figura 11 é um exemplo interessante, pois está incorreta sob o ponto de vista matemático. No entanto, essa análise desconsidera saberes presentes na resolução dessa questão. O estudante compreendeu o problema e corretamente

estabeleceu uma estratégia para resolvê-lo. Observe que na correção da direita, mesmo identificando que o estudante cometeu um único erro no procedimento (a divisão de 720 por 15), o professor atribui a metade da pontuação ao estudante (1,0), enquanto na solução da esquerda o professor desconta apenas 10% da pontuação na questão (1,8). Esta diferença de pontuação pode estar relacionada ao modo como o docente compreende o erro em Matemática, o ensino da Matemática e a própria Matemática. Se o docente interpreta a matemática como algo exato, puro associado à genialidade, qualquer erro é “impuro”, deve ser eliminado, combatido. Por outro lado, se o erro é compreendido como parte do processo de aprendizagem da matemática, uma interpretação mais cuidadosa pode ser feita, impactando a pontuação.

O erro em uma atividade escolar pode ser descrito como “um padrão de resposta que não atingiu determinado critério que permita a aprovação por parte do professor/ou a passagem para outra etapa de aprendizagem” (CARMO, 2010, p. 214). No entanto, quem estabelece padrão? O docente? A escola? O padrão de respostas pode ser estabelecido pela Matemática ou pelo currículo. Talvez, um professor não espere encontrar uma adição de parcelas iguais na solução de um problema que trata da multiplicação em uma prova aplicada nos anos finais do Ensino Fundamental.

A interpretação dos erros, em contrapartida, não é influenciada unicamente pelos padrões estabelecidos pela Matemática ou pelo currículo, e sim, por uma pluralidade de saberes docentes. Uma solução, à luz da Matemática, pode ser interpretada como correta ou incorreta, mas à luz da Educação, essas interpretações se tornam muito mais subjetivas.

Uma possível interpretação para a correção deste item é aquela que considera sua solução como sendo composta de quatro etapas: (1) a interpretação do enunciado e o planejamento da resolução, (2) o cálculo da área do terreno A, (3) o cálculo da medida do lado do terreno B e (4) o cálculo do perímetro. Este estudante comete um erro na terceira etapa, um erro operacional. Provavelmente, para o segundo corretor, o erro cometido invalida a quarta etapa, pois utiliza parcelas incorretas. Obviamente, essa interpretação não é oriunda ou sustentada por um saber matemático, é uma opção pedagógica influenciada por interpretações particulares do erro em Matemática e no Ensino de Matemática. Uma solução pode ser interpretada de maneiras distintas entre os docentes. Essas interpretações influenciam não somente como atividades avaliativas são corrigidas,

mas também, como essas atividades são interpretadas e utilizadas para a regulação do ensino e da aprendizagem.

Na seção 4.1 discutiremos critérios possíveis que poderiam ser utilizados para a correção de problemas de Matemática encontrados na obra de Abrantes (1995) e nos estudos mais recentes de Nunes Neto e Mendoza (2021).

Nas seções seguintes, apresentaremos uma discussão teórica sobre duas possíveis interpretações do erro à luz da pedagogia: o erro como um tropeço (4.2) e o erro como um trampolim (4.3).

Na seção 4.4 realizaremos uma breve discussão sobre a interpretação do erro na avaliação, para que, em seguida, na seção 4.5, apresentemos algumas possibilidades de instrumentos e métodos alternativos de avaliação

4.1 Critérios para a correção de problemas de Matemática

A maioria dos professores não explicita os critérios de avaliação aos estudantes (FERNANDES, 2009), pelo menos, pude identificar em minha prática docente, que o estabelecimento de critérios se restringe à definição de instrumentos que serão utilizados e suas respectivas pontuações. Critérios de correção são raramente apresentados aos estudantes antes da aplicação dos testes. Analogamente, na literatura há poucos trabalhos que desenvolvam critérios para a correção de questões de Matemática, uma exceção pode ser encontrada na obra de Abrantes (1995).

Charles, Lester e O’Daffer (1987, apud ABRANTES, 1995) apontam vários processos de avaliação que podem ser usados pelos professores: observar e questionar, autoavaliação, testes de múltipla escolha e avaliação holística. Em relação à avaliação holística, utilizada para a avaliação de resolução de problemas, os autores descrevem três modos de pontuar as soluções: classificação analítica; classificação holística focada e classificação por impressão geral. Os dois primeiros correspondem a um conjunto de critérios utilizados para classificar as soluções dos alunos, enquanto que no terceiro não há uma escala prévia para ser usada (ABRANTES, 1995).

Na classificação analítica, são consideradas três perspectivas: compreensão do problema, planejamento da solução e apresentação de uma resposta. Em cada categoria,

o autor estabelece uma escala de 0 a 2 para cada uma das diferentes etapas do problema, como ilustra o Quadro 4.

Quadro 4 – Escala de classificação analítica

Escala	0	1	2
Compreender o Problema	Completa incompreensão do problema	Parte do problema incompreendido ou mal interpretado	Completa compreensão do problema
Planejar a Solução	Nenhuma tentativa ou um plano totalmente inadequado	Plano parcialmente correto baseado na parte do problema que foi interpretada corretamente	Plano que poderia (poderá) conduzir a uma solução correta se fosse (for) implementado adequadamente.
Dar uma resposta	Nenhuma resposta ou uma resposta incorreta baseada num plano inadequado	Erro de cálculo ou resposta parcial para um problema com múltiplas respostas	Resposta correta e corretamente indicada

Fonte: ABRANTES (1995, p. 53-54)

Na escala de classificação holística focada, as pontuações são estabelecidas, como o próprio nome sugere, analisando a solução como um todo (Quadro 5).

Quadro 5 – Escala de classificação holística focada

Pontuação	Características
0	Em branco; Os dados foram apenas copiados do enunciado, mas não há qualquer trabalho com esses dados ou há algum trabalho, mas não parece haver compreensão do problema; <u>Apresentam simplesmente uma resposta incorreta.</u>
1	Há um começo de trabalho para além da simples cópia dos dados refletindo alguma compreensão, mas a estratégia usada não conduziria a uma solução correta; Uma estratégia inadequada ou começada e não desenvolvida, não há evidência de que o aluno tenha tentado outra. Parece que o aluno usou uma estratégia que não funcionou e desistiu; O aluno tentou alcançar um subobjetivo do problema, mas não conseguiu.
2	O aluno usou uma estratégia inadequada e chegou a uma resposta incorreta, mas mostra alguma compreensão do problema; Foi usada uma estratégia adequada, mas: (a) ela não foi suficientemente desenvolvida para chegar a uma solução (por exemplo, o aluno apenas considerou duas entradas numa tabela); ou (b) ela foi implementada incorretamente e por isso não conduziu a qualquer resposta ou conduziu a uma resposta incorreta; O aluno alcançou um subobjetivo do problema, mas não foi longe; Apresenta uma resposta correta, mas: (a) o trabalho é incompreensível, ou (b) não apresenta qualquer trabalho a não ser a solução.
3	O aluno implementou a estratégia que poderia conduzir a uma resposta correta, mas não compreendeu uma parte do problema ou ignorou uma condição; Foram usadas corretamente estratégias, mas: (a) o aluno apresenta uma resposta incorreta sem que se compreenda porquê; ou (b) foi dada corretamente a parte numérica da resposta, mas ela não está bem indicada; ou (c) falta apenas a resposta; Foi dada uma resposta correta e há evidências de terem sido selecionadas estratégias adequadas. Contudo, a implementação das estratégias não é totalmente clara.
4	O aluno cometeu um erro ao desenvolver uma estratégia adequada, mas esse erro não reflete falta de compreensão nem do problema, nem do modo de implementar a estratégia, parecendo ser apenas erro de cálculo ou cometido ao copiar o enunciado; <u>Estratégias adequadas foram selecionadas e implementadas</u>

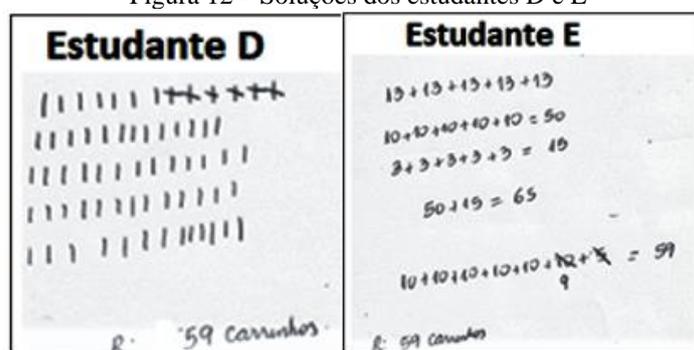
Fonte: ABRANTES (1995, p. 55-56)

As duas escalas apresentadas representam tentativas de implementar maior objetividade à correção de problemas de Matemática. Em ambas, os critérios mesmo sendo bem definidos, apresentam limitações, podendo, em algumas situações, depender da interpretação do professor (corretor).

Para exemplificar, faremos algumas ponderações sobre algumas soluções contidas na *folha de correção* apresentada na seção 2.1.

Como ilustrado na Figura 12, observamos que o estudante D planejou adequadamente a solução deste problema, mas não demonstrou conhecer os algoritmos envolvendo as quatro operações e apresentou uma estratégia pouco eficiente e/ou viável para resolver outros problemas similares com números maiores.

Figura 12 – Soluções dos estudantes D e E



Fonte: elaborado pelo autor

Nesse caso, utilizando a classificação analítica, poderíamos utilizar o critério “plano que poderia (poderá) conduzir a uma solução correta se fosse (for) implementado adequadamente” (ABRANTES, 1995, p. 54). Utilizando o processo de contagem através de ‘pauzinhos’ seria muito trabalhoso e, talvez, inviável utilizar essa estratégia para um problema com 1300 caixas, com 239 brinquedos em cada, por exemplo. Esse plano poderia produzir uma solução correta? Mesmo com esse critério bem definido, a interpretação docente seria necessária.

Nessa mesma perspectiva, há diferenças entre D e E. Como avaliá-los de forma objetiva, visto que o estudante E, talvez, tenha apresentado uma solução mais avançada no que se refere ao domínio dos números. No entanto, essa solução continua apresentando limitações para números bem maiores. A própria escolha do corretor de utilizar, ou não, essas limitações na avaliação, é particular.

Na holística focada, para pontuação 4, há a seguinte característica: “erro ao desenvolver uma estratégia adequada, mas esse erro não reflete falta de compreensão nem do problema, nem do modo de implementar a estratégia, parecendo ser apenas erro de cálculo ou cometido ao copiar o enunciado” (ABRANTES, 1995, p. 56). O erro cometido pelo estudante C, ilustrado pela Figura 13, por exemplo, seria somente um erro de cálculo ou um erro de conceito sobre o algoritmo da multiplicação? Esse erro reflete ou não a falta de compreensão do algoritmo? Novamente, há uma questão claramente interpretativa nessa análise. E esse fator interpretativo está presente no próprio critério na palavra “parecendo”.

Figura 13 – Solução do estudante C

The image shows a student's handwritten work. At the top, it says "Estudante C". Below that, there are two calculations. The first is a multiplication: 13 multiplied by 5, resulting in 515. The second is a subtraction: 515 minus 6, resulting in 509. Below the calculations, it says "R: 509 canudos".

Fonte: elaborado pelo autor

Dentre os critérios apresentados por Abrantes (1995), o holístico focado talvez seja aquele que apresenta mais vantagens em termos pedagógicos por analisar de maneira global a produção dos estudantes, no entanto, mesmo em relação a esse critério, há limitações.

Mais recentemente, Nunes Neto (2015) desenvolveu, em sua dissertação de mestrado, um modelo de correção de problemas em Matemática, que foi aperfeiçoado em seus estudos posteriores (NUNES NETO; MENDOZA, 2021). Neste modelo, há uma escala de pontuação de 1 a 4, que corresponde, respectivamente às quatro ações invariantes da Análise de Situações Problema (ASP): a) compreender o problema; b) construir o modelo matemático; c) solucionar o modelo matemático; e, d) interpretar a solução (Quadro 6).

Quadro 6 – Sistema de análise da Atividade de Situações Problema

Conceito/ Desempenho Quantitativo	Ações	Operações/Indicadores essenciais
		Qualitativo
1 Ponto – Fraco	Compreender o problema	a) Ler o problema e extrair os elementos desconhecidos. b) Estudar e compreender os elementos desconhecidos.

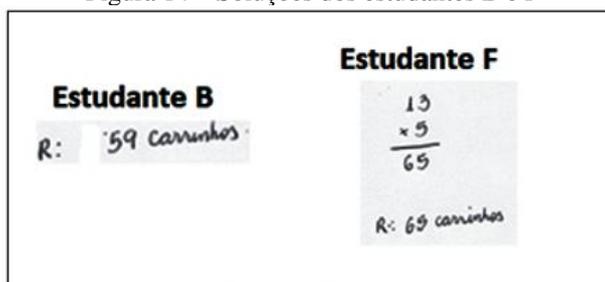
		c) Determinar os dados e as condições. d) Determinar os objetivos do problema.
2 Pontos – Regular	Construir o modelo matemático	a) Determinar as variáveis e incógnitas. b) Nomear as variáveis e incógnitas com suas unidades de medidas. c) Construir o modelo a partir das variáveis, incógnitas e condições. d) Realizar a análise das unidades de medida do modelo.
3 Pontos – Bom	Solucionar o modelo matemático	a) Selecionar os métodos matemáticos para solucionar o modelo. b) Solucionar o modelo.
4 Pontos – Ótimo	Interpretar a solução.	a) Interpretar o resultado obtido na solução do modelo. b) Extrair os resultados significativos que tenham relação com o(s) objetivo(s) do problema. c) Dar resposta ao(s) objetivo(s) do problema. d) Realizar um informe baseado nos objetivos do problema. e) Analisar a partir de novos dados e condições que tenham relação direta ou não com os objetivos do problema, a possibilidade de reformular o problema, construir novamente o modelo, solucionar.

Fonte: NUNES NETO; MENDOZA (2021, p. 119)

Vale observar que a tabela de Nunes Neto e Mendonza (2021) nos remete, de certa forma, às quatro fases de Polya (1978) para a resolução de um problema de Matemática de forma eficiente: (1) Compreender o problema; (2) Construir um plano de ação; (3) Executar o plano; (4) Validar a solução.

Utilizaremos o quadro proposto por Nunes Neto e Mendoza (2021) para analisar as soluções dos estudantes B e F ilustradas pela Figura 14.

Figura 14 – Soluções dos estudantes B e F



Fonte: elaborado pelo autor

O estudante B ‘dá uma resposta objetiva ao problema’, característica do conceito 4, mas não demonstra ‘selecionar métodos matemáticos para solucionar’ o problema, característica de um nível anterior, conceito 3. O estudante B, simplesmente, escreve a resposta. Como pontuar essa solução?

Em relação ao estudante E, podemos concluir que ele apresenta parte do método (a metade) para resolver o problema. Então ele constrói e soluciona parcialmente o modelo matemático que resolve o problema, características dos conceitos 2 e 3, respectivamente, indicando que este aluno tenha compreendido parte do problema (conceito 1). Como pontuar nesse caso? Metade do maior conceito? Novamente encontramos limitações na utilização desse modelo.

Consideramos que essas limitações indicam a dificuldade, ou até mesmo, a impossibilidade, de estabelecer critérios de correção isentos da necessidade da interpretação do professor ao avaliar a produção escrita dos estudantes. Refiro-me aqui às questões abertas, discursivas e/ou dissertativas. Isto não significa que estabelecer critérios seja algo ruim, pelo contrário, esses critérios podem proporcionar uma maior padronização e uma maior objetividade na correção e pontuação das soluções. No entanto, o caráter interpretativo é sempre indissociável do processo avaliativo.

O processo de correção de questões está relacionado ao modo que interpretamos os erros em Matemática. Nesse sentido, há pelo menos duas interpretações sobre os erros em Matemática: a primeira concebe o erro como um fracasso da aprendizagem e, por consequência, algo que deve ser evitado, combatido; a segunda concebe o erro como parte de um processo natural da aprendizagem e, conseqüentemente, algo que pode ser utilizado com fins pedagógicos, “identificando não apenas o quê, ou o quanto os alunos erram, mas passando a considerar o porquê e como erram, podemos contribuir para o desenvolvimento de ações para que os alunos possam superá-los” (RIBEIRO; GODOY; ROLKOUSKI, 2020, p. 114)

4.2 O erro como um tropeço

O erro pode ser interpretado como um tropeço, como uma barreira à aprendizagem. Algo a ser execrado. Os estudantes apagam as soluções erradas, substituindo-as por soluções copiadas do quadro. Esta forma de interpretar o erro é utilizado nas aulas e nos exames. Os erros são interpretados como uma ausência, uma falha ou incapacidade e não como uma incompletude ou um processo contínuo de construção de saberes.

Teixeira Jr. e Silva (2017) constataram que a maioria dos professores de Matemática corrige as provas apenas identificando os acertos e os erros, ao invés de

buscar informar os alunos sobre os problemas existentes nas resoluções apresentadas. Nessa perspectiva, “o erro, que deveria ser considerado como natural e parte do processo de aprendizagem, é visto como algo que deva ser combatido, aniquilado” (VAZ; NASSER; BELFORT, 2014).

Borasi (1985) observou como as frustrações causadas por uma postura negativa em relação ao erro podem ser maléficas para o desenvolvimento dos alunos e na sua disposição para aprender Matemática. Ao cometer erros, os estudantes experimentam diversos sentimentos, tais como desapontamento, frustração, vergonha e raiva (BORASI, 1985). Deste modo, um aluno pode desenvolver o que Sales, Carmo e Henklain (2020) denominam de Ansiedade Matemática, desestimulando os estudantes e criando obstáculos que irão atrapalhá-los no decorrer da sua trajetória escolar. A Ansiedade Matemática é um transtorno associado a

casos de extrema dificuldade em lidar com a Matemática e que se manifesta por meio das seguintes respostas: (a) autonômicas, como taquicardia, sudorese, náuseas; (b) cognitivas, como auto atribuições negativas e crenças irracionais em relação à Matemática; e (c) condutas de esquiva e fuga de situações em que se faz necessário algum tipo de habilidade Matemática (SALES; CARMO; HENKLAIN, 2020, p. 242)

Nessa perspectiva, o esforço do estudante é comumente ignorado, “como se o erro fosse um produto apenas do estudante e não da relação entre estudante e professor” (SALES; CARMO; HENKLAIN, 2020, p. 258). Então, adotando uma postura remediadora, o professor foca na correção do problema, sem se preocupar com suas causas ou eventuais potencialidades pedagógicas.

Quando o professor percebe que seus alunos estão errando muito nos problemas com frações e propõe, separadamente, exercícios de fixação, com treino de operações com frações, nem sempre está voltado para a superação de um possível problema. [...] Visto de forma simplificada, seu tratamento consiste em aplicar paliativos para eliminar seus efeitos. (PINTO, 2000, p.141-142)

Em um estudo desenvolvido em sua dissertação, Santos (2007) verificou que os trabalhos acadêmicos que abordam Análise de Erros tratam o erro dos estudantes como uma falta ou lacuna na aprendizagem, e não como parte de um processo. Na escola, mais especificamente no caderno escolar, os erros são apagados. Ao identificar o erro, o estudante apaga o que foi feito a lápis ou risca o que foi produzido à caneta, geralmente, de modo que o que foi errado fique ilegível, de modo que o estudante não consiga revisar esse erro posteriormente.

Caracterizamos os alunos pela falta, ou seja, ao “dizer que eles erraram, que não fizeram o certo, que suas interpretações estão equivocadas, que eles não utilizaram uma determinada estratégia ou procedimento [...]” (SANTOS; PASSOS; MARQUES, 2014, p. 88), mas normalmente, não permitimos que o erro seja utilizado pedagogicamente para a aprendizagem. Nossa solução geralmente é a repetição. Faça mais exercícios! Pratique mais! Dizemos aos nossos estudantes. No entanto, a repetição mecânica de exercícios pode não corresponder à ação pedagógica necessária para a compreensão das razões que geraram o erro.

4.3 O Erro como um trampolim

Os erros podem mostrar as dificuldades e limitações, e ajudar na melhoria dos métodos e resultados utilizados pelo professor e pela escola. Nessa concepção, o erro assume uma função pedagógica. Segundo os Parâmetros Curriculares Nacionais (PCN), desde que trabalhado consistentemente

pele professor, o erro é um elemento que pode permitir ao estudante tomar consciência do seu próprio processo de aprendizagem, da apropriação que faz de diferentes conteúdos, percebendo que há diferenças entre o senso comum e os conceitos científicos e que é necessário saber aplicar diferentes domínios de ideias em diferentes situações. (BRASIL, 1998, p. 32)

O erro, inserido no processo de ensino-aprendizagem, é visto como algo natural, como “um saber que o aluno possui, construído de alguma forma, e é necessário elaborar intervenções didáticas que desestabilizem as certezas, levando o estudante a um questionamento sobre suas respostas” (CURY, 2007, p.80). Borasi entende que, ao se adotar uma visão mais ampla sobre o conhecimento matemático, o professor pode reconhecer o lado positivo dos erros e, utilizando procedimentos didáticos adequados, o erro se torna “um trampolim para a aprendizagem em Matemática” (BORASI, 1985, p.1).

Ainda apontando os benefícios didáticos gerados pelo conhecimento dos erros que os alunos cometem, Cury (2012) elucida que:

para poder trabalhar com os erros e tomar decisões sobre eles, é preciso ter conhecimento do conteúdo envolvido e das fases da análise, tomando decisões que são específicas dos professores, porque levam em conta, ao mesmo tempo, o que o aluno sabe, o que não sabe e o que pode ser feito para ajudá-lo a reorganizar seu pensamento sobre o conteúdo em questão. (CURY, 2012, p.31)

A identificação e categorização dos erros cometidos pelos estudantes possibilitam adequar o planejamento de ensino para corrigir os erros conceituais e procedimentais.

Além de classificar a resposta como certa ou errada, é importante investigar como os conceitos e procedimentos adotados por um estudante se relacionam com os conhecimentos prévios. Carmo (2010) identificou alguns tipos de erros comuns cometidos por estudantes e analisou quais variáveis poderiam explicá-los. O pesquisador verificou que mesmo sendo constantemente delineado como pertencentes ao estudante, muitas dessas variáveis identificadas como prováveis geradoras dos erros estavam associadas ao modelo de ensino ou à postura pedagógica adotada pelo professor. Por isso, devemos "redimensionar nosso entendimento acerca da produção de erros" (CARMO, 2010, p. 226). Alguns erros podem não indicar falta de atenção ou descuido, e sim, "uma transferência equivocada de algo que é verdadeiro num domínio, mas falso em outro" (RIBEIRO; GODOY; ROLKOUSKI, 2020, p. 155).

Alguns erros são podem não ser representantes de um não saber, pelo contrário, podem representar um saber legítimo em outro domínio. No sétimo ano o estudante aprende a resolver uma equação com termos fracionários, obtendo frações equivalentes, eliminando o denominador, multiplicando-se ambos os lados da igualdade. Na resolução da adição de frações, no ano anterior, e de adição de monômios, no ano seguinte, essa simplificação não é correta. (vide Quadro 7)

Quadro 7 – Verdadeiros em um domínio, errados em outro.

	6º ano	7º ano	8º ano
Tema/conteúdo escolar	Adição de frações	Resolução de equação do primeiro grau	Adição de monômios
Ação realizada	Mantém-se os denominadores	Elimina-se os denominadores	Mantém-se os denominadores
Exemplo	$\frac{1}{2} + \frac{4}{3}$ $= \frac{3}{6} + \frac{8}{6}$ $= \frac{11}{6}$	$\frac{x}{2} + \frac{x}{3} = 1$ $\frac{3x}{6} + \frac{2x}{6} = \frac{6}{6}$ $3x + 2x = 6$	$\frac{x}{2} + \frac{x}{3}$ $= \frac{3x}{6} + \frac{2x}{6}$ $= 3x + 2x$

Fonte: elaborado pelo autor

Alguns desses erros podem ser justificados pela ênfase dada no ensino de Matemática aos procedimentos operatórios em detrimento dos conceitos. Em prol de agilidade ou praticidade, a adição dos dois membros de uma equação por +2 é substituída

pela ideia de ‘jogar para o outro lado trocando o sinal’, o que pode ocasionar erros gerados pela sobreposição de conhecimentos verdadeiros. Um exemplo é mostrado no Quadro 8.

Quadro 8 – Sobreposição de erros

Procedimento verdadeiro (A)	Procedimento verdadeiro (B)	Sobreposição dos procedimentos A e B
$-2 + x = 10$ $x = 10 + 2$	$+2x = 10$ $x = \frac{10}{+2}$	$-2x = 10$ $x = \frac{10}{+2}$

Fonte: elaborado pelo autor

Essa ênfase no operacional em detrimento do conceitual pode ser observada no ensino de frações (FAZIO; SIEGLER, 2011), no ensino da divisão de naturais (MEGID, 2012), tanto no ensino de expressões e equações, quanto no ensino de Cálculo.

A literatura apresenta outras fontes de erro: falhas do programa de ensino, a possível existência de lacunas entre o que o aluno já sabe e o que deve fazer, a linguagem utilizada pelo professor e a escrita dos comandos das questões (CARMO, 2000; GRIS; POLOMBARINI; CARMO, 2019). Santos, Passos e Marques (2014) apresentam uma interessante discussão sobre as maneiras de lidar com o erro a partir do exemplo ilustrado na Figura 15. Podemos observar, inicialmente, que o estudante não domina o algoritmo da adição. Para estes autores, seria possível categorizar esta solução como oriunda de “erros cometidos no domínio deficiente de pré-requisitos de habilidades, fatos e conceitos”¹³(SANTOS; PASSOS; MARQUES, 2014, p. 89).

Figura 15 – Erro no algoritmo da adição

$$\begin{array}{r}
 781 \\
 7283 \\
 + 1687 \\
 \hline
 15999
 \end{array}$$

Fonte: SANTOS; PASSOS; MARQUES (2014)

Santos (2009) defende a ideia de realizarmos a leitura das produções dos estudantes buscando compreender o modo como os estudantes interpretam os problemas e constroem suas soluções, ou, nas palavras do autor, realizar essa leitura a partir de suas ‘maneiras de lidar’. Essa leitura está relacionada à postura investigativa dos professores em relação à produção dos seus estudantes.

¹³ Os autores apresentam essa classificação tendo como sustentação teórica Radatz (1979)

Ao se adotar a análise da produção escrita como prática de investigação assume-se um olhar sobre as *maneiras de lidar*. Nessa perspectiva, o que existe é análise das *maneiras de lidar*, e, nela, o ‘erro’ é apenas um julgamento, uma das formas de caracterizar as ‘maneiras de lidar’ dos sujeitos. Portanto, o ‘erro’ não está na produção do estudante, mas é o resultado da caracterização que o professor faz dessa produção. O que se pode dizer que existe, de fato, são as maneiras com as quais os estudantes lidam com as questões, maneiras essas que nem sempre são tão acessíveis ou ‘visíveis’, de modo a serem, portanto, suscetíveis a algum julgamento (BURIASCO; FERREIRA; CIANI, 2009, p. 80).

Ao realizarmos uma análise mais detalhada do exemplo ilustrado na figura 15, podemos observar um conhecimento Matemático muito maior do que a análise superficial permite. A adição entre 3 e 7 resultou em 10, de fato. Mas, o estudante desmembrou o número 10 em duas parcelas 1 e 9. “Como ele só pode colocar no máximo 9 unidades no resultado, então ele coloca as 9 e ‘sobe’ uma unidade” (SANTOS; PASSOS; MARQUES, 2014, p. 90). Essa ‘maneira de lidar’ do estudante é utilizada em toda a operação.

Em consonância com as ideias de Buriasco (2000), Buriasco, Ferreira, Ciani, (2009), Santos, Passos e Marques (2014), acreditamos que utilizando uma postura investigativa na interpretação e correção das produções dos estudantes poderemos transformar os erros em ‘trampolins para a aprendizagem’, como defende Borasi (1985).

4.4 O erro e a avaliação

A avaliação é um elemento indissociável do Ensino e da Aprendizagem. Se é impossível pensar em Avaliação sem pensar em Ensino e em Aprendizagem, a recíproca também é verdadeira. Nesse sentido, a mudança na percepção e na significação dada aos erros cometidos pelos estudantes permite que o professor de Matemática atribua novos significados à avaliação. “O erro, muitas vezes mais que o acerto, revela o que a criança *sabe*, colocando este saber numa perspectiva processual, indicando também aquilo que ela *ainda não sabe*, portanto pode *vir a saber* (ESTEBAN, 2002, p. 21). Após o reconhecimento do erro como um saber diferente, algo em construção, interpretá-lo de forma estática, como geralmente ocorre nos exames, passa a não fazer mais sentido.

A dicotomia entre o erro e o acerto e entre o saber e o não saber, marcos da concepção classificatória de avaliação, são aspectos profundamente enraizados em nossa forma de ver o mundo. A compreensão de que toda resposta, seja ela *certa* ou *errada* do ponto de vista do avaliador, traz em si conhecimentos e desconhecimentos, leva à conclusão de que o que permite o movimento é o *não saber* tem tanto valor, para se tecer novos conhecimentos, quanto o *saber*: rompe com essa dicotomia e traz outros olhares para o processo avaliativo (ESTEBAN, 2002, p. 26)

Um dos olhares possíveis está na Avaliação Formativa. Para analisar as produções orais e escritas dos estudantes e elaborar *feedbacks* formativos é necessário, a priori, que o professor compreenda o erro como um elemento de grande potencial pedagógico, seja na sala de aula, seja em uma prova escolar. Nesse sentido, consideramos que a ressignificação do erro para o ensino – incluindo a ruptura das dicotomias erro e acerto e saber e não saber – torna-se um pré-requisito para o desenvolvimento de algumas vertentes da avaliação formativa e, especificamente, para a construção de *feedbacks* formativos.

Por outro lado, a ressignificação do erro pode impulsionar que novas interpretações sejam atribuídas aos instrumentos avaliativos e à própria avaliação. Nessa nova concepção, a avaliação passa a ser interpretada e utilizada como uma investigação (ESTEBAN, 2002; BURIASCO, 2004). Os docentes utilizam as soluções dos estudantes para compreender não só o que foi aprendido, mas como foi aprendido. Trata-se de uma rica possibilidade de investigação docente para

conhecer as estratégias que os alunos elaboram, os procedimentos que utilizam, os modos de interpretação que fazem do enunciado, as características dos problemas que constroem a partir da sua interpretação do enunciado original, ao resolverem uma questão. (VIOLA dos SANTOS; BURIASCO; CIANI, 2008).

Se o erro é parte de um processo, um saber em construção, na prova ou em outro instrumento avaliativo, conseqüentemente, deve ser reinterpretado. O errado não é mais riscado ou apagado, deve ser reconstruído, ressignificado e reutilizado, em prol da aprendizagem. Por sua vez, um bom feedback pode ser um elemento de ensino promissor, pois atua na zona de desenvolvimento proximal, auxiliando a transição do erro para o acerto, do saber em processo, para um saber mais avançado, mas ainda em processo.

Na seção seguinte apresentaremos algumas possibilidades práticas para o desenvolvimento de processos avaliativos mais formativos em que o erro é ressignificado.

4.5 Práticas avaliativas insubordinadas

A palavra insubordinação está relacionada à desobediência a uma regra ou superior, desafio à autoridade ou recusa em obedecer a ordens. No entanto, a insubordinação criativa “é legitimada pelo foco em práticas profissionais pautadas em fundamentos éticos” (D’AMBROSIO; LOPES, 2015, p. 3). Essas pesquisadoras defendem que os professores devem questionar a Matemática que se apresenta na escola,

devem apoiar a incerteza da Matemática como disciplina e enfatizar a humanidade combatendo, por exemplo, discursos discriminatórios sobre os alunos. D'Ambrosio (2015, p. 3) afirma que a escola por vezes “espera que os professores sigam as regras às cegas, sem se preocupar com os resultados de suas ações na Educação das crianças”.

Alrø e Skovsmose (2010) denunciam a existência do absolutismo burocrático no ambiente escolar. Na escola, “as coisas são como são por causa de regras e normas” (ALRØ; SKOVSMOSE, 2010, p. 26). Neste sentido, se o professor está impedido de modificar as regras relacionadas à escolha e utilização de exercícios (ou pelo menos encontra barreiras para tal mudança), provavelmente, encontrará dificuldades ainda maiores em modificar concepções relacionadas às práticas avaliativas. Desse modo, seria plausível supor que esse absolutismo atue nas concepções e, principalmente, nas práticas dos professores referentes às avaliações em Matemática.

Entendemos como avaliação insubordinada toda avaliação que rompe, ou tenta romper, o absolutismo burocrático presente nas avaliações e as concepções positivistas que sustentam o modelo avaliativo tradicional. Consideramos como uma avaliação insubordinada a avaliação que se constrói sobre a autonomia do professor, assumindo e redefinindo o seu papel de mediador da aprendizagem.

Nós, educadores matemáticos, precisamos ousar, em nossa produção de conhecimento, de forma autônoma, a partir de nossas visões de mundo, constituídas por crenças e concepções adquiridas ao longo de nossas vidas, mas lembrando-nos sempre de considerar, nos processos de interações sociais dos espaços formativos, a heterogeneidade e a diversidade (D'AMBROSIO; LOPES, 2015, p. 10-11).

Algumas ideias alternativas de avaliação são encontradas na literatura. Nasser e colaboradores (2019) apresentaram um relato em que os Cenários para Investigação¹⁴ eram utilizados no ensino e avaliação. Um professor avaliou seus estudantes utilizando atividades investigativas de Geometria Espacial realizadas em sala de aula

durante um período de dois meses em uma turma do segundo ano do Ensino Médio do Instituto Federal do Rio de Janeiro (IFRJ). A turma era composta por cerca de 20 alunos com um encontro semanal com 4 aulas consecutivas. Durante cerca de 5 encontros, os alunos foram divididos em grupos de 4 a 5 integrantes para desenvolver atividades de investigação. Usando o telefone celular disponível na Internet, sólidos geométricos acrílicos, papelão e material escolar regular, os alunos devem pesquisar informações que vão desde o

¹⁴ Ole Skovsmose (2000)

tamanho real da pirâmide até fórmulas de área e volume de sólidos (NASSER et al, 2019, p. 119)

Cada grupo produzia pequenos relatórios durante as aulas, que eram utilizados como um documento que substituía as provas tradicionais, comuns na disciplina e naquela instituição de ensino. Os autores relatam que tanto a abordagem pedagógica quanto o modelo de avaliação foram acordados e negociados com a turma na aula inicial da disciplina. “O resultado observado foi um aumento de assiduidade e participação. A interação foi maior do que nos períodos anteriores, os alunos assumiram o papel de protagonistas da própria aprendizagem” (NASSER et al, 2019, p. 120).

As diversas vertentes da prova tradicional estudadas por pesquisadores da área – destacamos aqui as pesquisas desenvolvidas na Universidade Estadual de Londrina (UEL) e na Universidade Tecnológica Federal do Paraná (UTFPR) – se configuram como excelentes alternativas em romper a cultura da prova tradicional. Antunes e Mendes (2018) apresentam as seguintes vertentes alternativas à prova tradicional:

(1) *Prova de ensaio*: um texto literário breve, que expõe ideias, críticas e reflexões a respeito de um determinado tema. Como um instrumento de avaliação em Matemática, os alunos podem ser convidados a escrever respondendo a um artigo de jornal, ou dar a sua opinião sobre alguma situação da vida cotidiana; (2) *Prova de levar para casa*: consiste em uma prova escrita elaborada pelo professor a respeito de um ou mais conteúdos que o aluno leva para resolver as tarefas em casa, com mais tempo, utilizando quaisquer materiais de apoio ou até mesmo pedindo ajuda a outras pessoas; (3) *Prova em duas fases*: um instrumento de avaliação que combina vários formatos de provas escritas. Inicialmente, o professor deve deixar claro para os alunos como será cada uma das fases. A primeira fase se dá na sala de aula, durante um período de tempo preestabelecido. O instrumento pode conter um número maior de tarefas, uma vez que é possível resolvê-las em duas fases; (4) *Prova de produção*: consiste em solicitar que os alunos elaborem suas provas escritas com as respectivas tarefas. O professor pode reorganizar a sala de aula em grupos menores para melhor acompanhar o encaminhamento. Em seguida, os grupos trocam as provas e cada um, agora com uma prova diferente, resolve as tarefas solicitadas; (5) *Prova em grupo*: consiste em uma prova escrita para a qual o professor elabora as tarefas de acordo com os conteúdos trabalhados em sala de aula. O professor informa os alunos de que podem fazer grupos compostos de n alunos, de acordo com a quantidade de alunos da turma, e estabelece o período de tempo para a sua aplicação. (ANTUNES; MENDES, 2018, p. 26-27)

Nessas cinco vertentes, Antunes e Mendes (2018) destacam o potencial didático existente na utilização de diferentes versões de uma prova. Nós destacamos também que estas práticas estão relacionadas à resignificação da avaliação pedagógica, pois ao propor provas em grupo, feitas em casa, em mais de um momento e utilizando questões abertas, o professor rompe paradigmas bem estabelecidos e conhecidos. O papel do erro, as

concepções sobre a ‘cola’, a ideia da avaliação como uma medida e a interação com os outros alunos são elementos pedagógicos que devem ser ressignificados para que essas práticas avaliativas sejam implementadas.

A prova com “cola”, estudada por Forster (2016) em sua dissertação de mestrado, em que os alunos devem elaborar a “cola” em uma $\frac{1}{4}$ de folha é uma alternativa interessante para ser utilizada. Para caber na pequena folha, os alunos devem estudar para poder sintetizar sua “cola”. Os professores podem, neste método, avaliar não só pela prova em si, como a partir da “cola” elaborada. Além de avaliar a aprendizagem dos estudantes, a “cola” permite que se avalie o ensino. O que foi valorizado e compreendido pelos estudantes pode ser identificado na “cola” mais claramente do que na prova, se configurando como uma excelente oportunidade para a regulação do ensino e da aprendizagem. Ao utilizar a cola, os estudantes podem analisar suas soluções durante a realização, podem identificar previamente e evitar lapsos que venham a se configurar como um erro. Se o erro é um saber em processo, há, nessa identificação, a construção de um novo saber.

A prova ou teste¹⁵ em duas fases foi desenvolvida(o) originalmente na Holanda, pelo *Projeto Hewet*, com alunos dos últimos anos do Ensino Médio, ideia que foi adaptada para alunos mais novos pelo Projeto MAT789 (ABRANTES, 1995). As questões propostas incluía(m):

(tipo 1) Perguntas de interpretação, justificações e problemas de resolução relativamente rápida, e (tipo 2) alguns problemas abertos requerendo investigação e respostas mais desenvolvidas. Em qualquer dos casos, não se pedia a reprodução de definições ou regras tanto mais que os alunos podiam consultar os cadernos e apontamentos. Mas pedia-se que as respostas fossem pormenorizadas e explicadas (ABRANTES, 1995, p. 24)

A ideia era que os estudantes resolvessem, na primeira fase, as questões do tipo 1. Os professores davam pistas e sugestões após corrigir essas questões. Na segunda fase, os estudantes faziam as questões do tipo 2. Uma possibilidade de formato para a prova em duas fases seria o seguinte:

(1) o professor elabora a prova e os alunos, em uma primeira fase, resolvem sem nenhuma indicação do professor, em tempo determinado; (2) o professor avalia as resoluções iniciais dos alunos e tece comentários pedindo

¹⁵ Abrantes (1995) utiliza *teste em duas fases*; Pires e Buriasco (2012, 2017), Trevisan (2013) e Mendes e Buriasco (2017) utilizam *prova em (duas) fases*.

justificativas e esclarecimentos; (3) na segunda fase, os alunos tentam responder as questões postas pelo professor, podendo dispor de um tempo maior que na primeira fase. Nessa etapa espera-se que os alunos melhorem as respostas dadas na primeira fase. (PIRES; BURIASCO, 2012, p. 8-9)

Trevisan (2013), em sua tese de doutorado, afirma que a segunda fase é uma etapa mais trabalhosa, pois se refere às questões do tipo ensaio, em que o estudante deve discorrer a respeito de um item ou um tema.

A primeira fase é encaminhada como uma prova escrita tradicional: os estudantes devem responder tantas questões quanto possível num espaço de tempo limitado. Espera-se que respondam principalmente as questões abertas da primeira parte. Depois de corrigida pelo professor, a prova é devolvida aos estudantes com indicação da nota parcial e apontamento dos erros mais graves. Numa segunda fase, e em posse dessas informações, o estudante repete o trabalho em casa, podendo responder as questões da maneira que escolher: independente umas das outras, ou na forma de um ensaio. Após o tempo combinado (algumas semanas, por exemplo), a prova é devolvida e novamente corrigida pelo professor (TREVISAN, 2013, p. 80)

A prova em duas fases apresenta algumas vantagens. Primeiramente, de caráter prático, trata-se de um instrumento similar àquele utilizado na prova tradicional, com pequenos ajustes nas questões, pois nesse modelo há a necessidade de itens de respostas mais abertas. A segunda vantagem está presente na segunda fase. De acordo com Trevisan (2013), a segunda fase não é vista como uma segunda chance, mas como um meio para que o estudante reflita sobre a sua produção na primeira fase. “As pistas e sugestões dadas pelo professor ao avaliar a primeira fase, variáveis consoantes as respostas, desempenhavam um papel crucial no trabalho subsequente dos alunos” (ABRANTES, 1995, p. 24). No momento em que os professores solicitam justificativas, tecem comentários, apontam caminhos, o fazem a partir dos erros dos estudantes, e se configuram como *feedbacks* que apresentam características mais formativas do que não formativas, pois são construídos com o intuito de direcionar para a aprendizagem.

A prova em fases, uma adaptação da Prova em Duas Fases idealizada por De Lange (1987), se configura como um excelente exemplo da utilização do *feedback formativo*. Com a inserção de n fases, adaptadas às singularidades de cada situação, avaliação e estudante, há em si, uma tendência a uma maior regulação das aprendizagens.

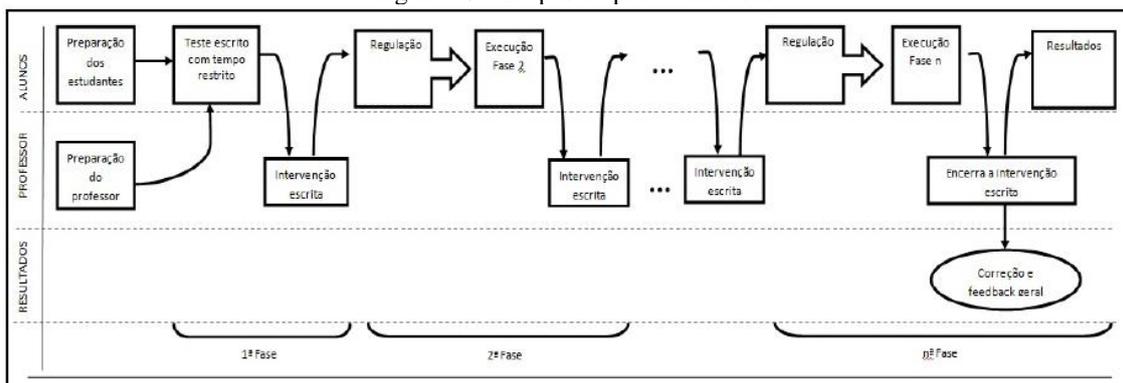
A característica formativa deste tipo de regulação se dá pelo *feedback* e pela adaptação de situações de ensino e aprendizagem feitas após a aplicação do instrumento de avaliação. Não há um número máximo limitante de vezes em que o instrumento pode ser reaplicado e novas adaptações possam ser feitas (PIRES; BURIASCO, 2017, p. 477)

De acordo com essas autoras, esse formato de avaliação apresenta a vantagem em permitir que o estudante volte a refletir sobre sua produção apoiado nas observações, orientações oferecidas pelos professores. Na prova em fases, os estudantes têm a oportunidade de

- (1) estabelecer um processo de comunicação por escrito: ao explicar o que fez, pode, ao mesmo tempo, mostrar o que compreendeu das considerações feitas pelo formador;
- (2) refletir sobre sua resposta inicial procurando reconstituir e criticar o seu próprio raciocínio, podendo descrever e explicar o que fez;
- (3) desenvolver a resolução feita inicialmente. (PIRES; BURIASCO, 2017, p. 478)

As intervenções realizadas pelos professores não correspondem a correções, não havendo necessariamente indicação de correto ou incorreto. Considerando a intervenção, “cada aluno pode interpretar e decidir o caminho a seguir nas fases seguintes, tanto na produção escrita nessa prova, como em seus estudos” (MENDES; BURIASCO, 2017, p. 14). A Figura 16 apresenta um infográfico que descreve as etapas da prova em n fases.

Figura 16 – Etapas da prova em fases



Fonte: MENDES; BURIASCO (2017)

Mendes e Buriasco (2017) concluíram em seus estudos que as intervenções escritas geradas a partir do lidar com a Prova em Fases favoreceram aos alunos o desenvolvimento de uma regulação da aprendizagem com relação:

- (1) ao compromisso com as questões Matemáticas na identificação de erros;
- (2) à busca por compreender o enunciado da questão; busca por contradições em produções de fases anteriores; avaliação de sua produção;
- (3) à estratégia de resolução individual na elaboração de hipóteses; orientação por meio das produções de fases anteriores; identificação e retificação dos erros;
- (4) a articular ideias próprias na justificativa da resolução; no desenvolvimento de conceitos;
- (5) à eficiência Matemática na reflexão a respeito dos erros e de sua superação; comparação entre as produções de fases anteriores;
- (6) à autoavaliação na busca de validar suas respostas; revisão do processo de resolução (MENDES; BURIASCO, 2017, p. 18)

Como dito anteriormente, a ressignificação do erro é uma condição necessária para que a prova em fases seja utilizada. Resignificar o erro significa compreender que há saberes no erro e há não saberes no acerto. Um excelente exemplo dessa perspectiva pode ser observado na percepção de Trevisan (2013), ao identificar que, possivelmente, seus estudantes estivessem realizando uma operação para a resolução de um problema sem a devida compreensão do processo.

Percebemos que todos os estudantes haviam utilizado essa estratégia em suas resoluções: dividir 3780 por 360. Porém, instigava-nos saber se compreendiam o significado do algoritmo, interpretando o significado do quociente e do resto dessa divisão. Apresentamos ao lado da maioria das resoluções o questionamento: O que representa o resultado dessa divisão? (TREVISAN, 2013, p. 95)

Nesse item, os estudantes deveriam localizar no círculo trigonométrico o arco na primeira volta correspondente ao ângulo de 3780° . Ao realizar perguntas como “o que representa essa divisão?” o professor pode realizar uma leitura mais apurada sobre o que os estudantes realmente compreenderam. Trevisan (2013) observou que alguns alunos que haviam acertado não haviam compreendido, enquanto, que outros, mostraram que haviam compreendido ao reconhecer o que haviam errado.

CAPÍTULO 5 – COLETANDO CONCHAS, ENCONTRANDO PÉROLAS



*Todos crescemos e evoluímos.
Mas só uma concha que possui as duas metades
é capaz de gerar uma pérola.
Se não tivermos nossa outra metade,
somos só como conchas,
pequenas
e quebradiças*
Matheus de Souza.

Relembrando aspectos metodológicos, já discutidos no capítulo 2, as folhas de correção foram preenchidas no decorrer de duas oficinas, realizadas por mim, em duas universidades, uma pública e outra, particular. As oficinas foram realizadas a convite de professores dessas instituições. Na instituição particular, a palestra foi oferecida a professores de Matemática que cursavam uma pós-graduação *latu sensu* em Educação Matemática. Na universidade pública, foi destinada a um grupo de mestrandos e licenciandos nos últimos anos de formação. Ao todo, 51 corretores responderam a *folha de correção* apresentada no capítulo 2, incluindo 5 professores universitários que assistiram às palestras.

Analisaremos todas as respostas, sem qualquer distinção prévia dos respondentes, por acreditar que todos eles compõem um corpo docente representativo de uma realidade. Os alunos de graduação e da pós influenciam e são influenciados, inevitavelmente e constantemente, pelos seus professores (formadores). São, todos eles, aprendizes, professores e corretores. No entanto, havendo qualquer peculiaridade na resposta dos professores formadores ou dos estudantes, esta será destacada na análise. Com o apoio da Análise Narrativa, uma das perspectivas teóricas desta tese, nosso objetivo é compreender essa realidade, e não comprovar qualquer hipótese sobre a avaliação.

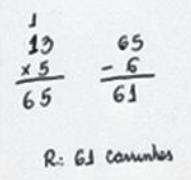
Em ambas as universidades, a palestra foi dividida em três momentos. No primeiro momento temas relacionados à avaliação em geral eram apresentados e debatidos (avaliação formativa e somativa, a avaliação é um processo, a prova é um instrumento e etc.). No segundo momento, as folhas de correção foram entregues. Os participantes foram convidados a realizar a correção, identificando apenas sua formação (licenciando, pós-graduando ou professor universitário). A ideia era corrigir as produções de fictícios estudantes do sexto ano do Ensino Fundamental. Antes do recolhimento, o palestrante pediu que ninguém alterasse a correção realizada, e sem seguida, as questões foram debatidas. O terceiro momento se iniciou com o recolhimento das folhas com as respostas corrigidas e a discussão de questões relacionadas à multicorreção.

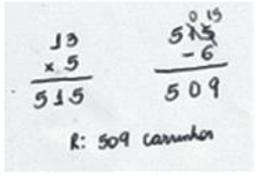
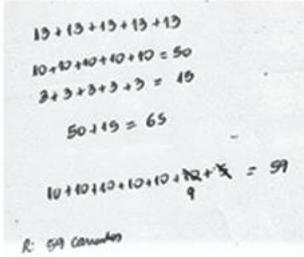
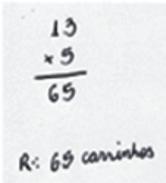
Na seção 5.1 apresentaremos um tratamento estatístico básico das pontuações para a construção de um panorama das pontuações, a fim de apoiar a discussão realizada nas seções seguintes. As pontuações contribuíram, dessa forma, para a análise das respostas encontradas nas justificativas e nos *feedbacks* dos corretores. Dessa análise emergiram alguns aspectos relevantes para compreender o processo de correção de questões de resposta aberta, a interpretação do erro em Matemática e outros pontos relacionados à avaliação. Esses aspectos correspondem às seções posteriores à seção 5.1.

5.1 Um panorama das pontuações atribuídas às soluções

Antes de discutirmos os critérios de correção utilizados, apresentaremos um panorama das pontuações dadas às resoluções apresentadas. Como discutido no capítulo 3, essas pontuações serão utilizadas para complementar a análise dos dados, oferecendo uma variável quantitativa às justificativas e *feedbacks*(qualitativas). O Quadro 9 apresenta para cada solução, a média, moda, as pontuações máxima e mínima e o desvio padrão.

Quadro 9 – Panorama das pontuações

Solução	Média	Moda	Menor pontuação	Maior pontuação	Desvio Padrão
<p>Estudante A</p> 	0,58	0,5	0	0,9	0,19

<p>Estudante B</p> <p>R: 59 carinhos</p>	0,54	0,5	0	1	0,38
<p>Estudante C</p>  <p>R: 509 carinhos</p>	0,34	0	0	0,9	0,28
<p>Estudante D</p>  <p>R: 59 carinhos</p>	0,95	1,0	0,2	1	0,16
<p>Estudante E</p>  <p>R: 99 carinhos</p>	0,95	1,0	0,2	1	0,14
<p>Estudante F</p>  <p>R: 65 carinhos</p>	0,40	0,5	0	0,7	0,18

Fonte: elaborado pelo autor

Uma análise inicial nos permite destacar alguns resultados que serão discutidos posteriormente. A **solução B**¹⁶ apresenta o maior desvio padrão em torno da média (0,38) se comparada com as demais, o que significa que os corretores oscilaram mais na atribuição de pontuação desta solução. A **solução C** apresenta a menor nota média (0,34), tendo zero como nota de maior frequência (moda). Nessa solução, o estudante fictício C demonstra compreender o problema e criar uma estratégia correta para resolvê-lo, além

¹⁶ Para dinamizar a leitura, utilizaremos somente o termo **solução X** para nos referir à **solução do estudante X**; $X \in \{A, B, C, D, E, F\}$.

de acertar a segunda operação (subtração). Podemos destacar também o baixo desvio padrão presente nas **soluções D e E**, indicando uma menor dispersão em torno das altas pontuações médias (0,95, em ambas) que tiveram.

5.2 O estabelecimento de critérios para a correção

Na *folha de correção* não havia nenhuma recomendação para que os corretores escrevessem os critérios utilizados para a pontuação das soluções no papel, no entanto, cinco corretores o fizeram. Obviamente, não escrever um critério geral para a correção não significa, necessariamente, a sua inexistência. No entanto, pelas divergências identificadas, é possível considerar que a maioria dos respondentes fez a correção sem o estabelecimento de um critério prévio e uniforme para a pontuação dos itens, realizando o que Abrantes (1995) denominou de classificação por impressão geral.

O corretor 27 estabelece um critério de pontuação de 0,5 para a multiplicação e 0,5 para a subtração, escrevendo o critério no gabarito fornecido. Ao longo da correção, utiliza o mesmo critério, não justificando apenas diferença das pontuações da **solução A** (0,8) e da **solução C** (0,7).

O corretor 44 escreveu *critério: 0,7 – produto, 0,3 – subtração*, se referindo às duas operações do problema. Este corretor também utilizou o mesmo critério para todas as soluções corrigidas, com exceção da **solução B**, em que estabeleceu 0,5. Ao justificar essa correção, escreveu *depende de quão bem eu conheço aluno* (...). Esse comentário será analisado posteriormente.

O corretor 2 escreveu na **solução A** os seguintes critérios: *0,5 para armar as duas contas ou indicá-las de outra maneira. 0,25 p/ a realização de cada conta. Vou considerar esse critério*. Como esse critério foi escrito no espaço destinado ao critério/justificativa da **solução A**, não ficou claro se este corretor tinha estabelecido esse critério com o intuito de utilizar nas demais soluções. Ao analisar as demais correções realizadas por ele, parece que algumas correções obedeceram a esse mesmo critério (A, D, E e F), enquanto nas soluções B e C foram utilizados outros critérios.

O corretor 2 atribuiu '0,75' de pontuação para a **solução A**, a pontuação máxima para as **soluções D e E**, e '0,5' para **solução F**. A coerência entre as pontuações das **soluções A e F** pode ser constatada se considerarmos que o '0,5' destinado à armação das

duas contas seja dividido em 2. Nesse caso, a pontuação de F foi ‘0,25’ pela armação e ‘0,25’ pela resolução da multiplicação (primeira conta).

Na **solução B**, o corretor 2 atribuiu ‘0,2’ como pontuação, escrevendo *não apresentou a resolução, acho importante avaliar o processo, não só o resultado*. A alta dispersão das pontuações para essa solução indica que não há consenso entre os corretores sobre a pontuação desse tipo de solução que contém somente a resposta final. Pragmaticamente, como o comando da questão não solicitava qualquer justificativa, a resposta não justificada poderia ter sido aceita. Soluções que apresentam somente a resposta mostram a necessidade do estabelecimento de critérios transparentes sobre a ‘regra do jogo’, entre o professor e seus alunos, uma espécie de Contrato Didático avaliativo.

O contrato didático é definido por Guy Brousseau como o conjunto de comportamentos do professor que são esperados pelo aluno e o conjunto dos comportamentos do aluno que são esperados pelo professor. Esse contrato é o conjunto de regras que determinam explicitamente em uma pequena parte, mas sobretudo implicitamente em grande parte, o que cada elemento da relação didática deverá fazer e que será, de uma maneira ou de outra, válido para o outro elemento. (BROUSSEAU, 1982, apud PESSOA, 2004, p. 1)

Diferentes professores e diferentes instituições de ensino podem esperar de seus alunos comportamentos/soluções distinto(a)s. Os critérios devem ser estabelecidos pelo professor e comunicado aos seus alunos de forma clara e transparente antes de qualquer etapa avaliativa. Na **solução C**, o corretor 2 não atribuiu pontuação, afirmando *esta, para mim, foi a mais difícil de avaliar*. Esta dificuldade será discutida na seção seguinte.

O corretor 14 estabeleceu um critério baseado em três perguntas: (1) *Respondeu corretamente à pergunta? Se sim, pontuação máxima. Se não, vou para a segunda pergunta.* (2) *Apresentou um raciocínio capaz de resolver a questão? Se sim, 0,5. Se não, vou para a terceira pergunta.* (3) *Apresentou uma parte do raciocínio capaz de resolver corretamente? Se sim, 0,25.*

Este corretor utilizou o critério em todas as soluções. As duas primeiras perguntas foram apresentadas no espaço destinado à justificativa da **solução A**, e a última, na **solução F**. Alguns pontos podem ser destacados sobre a correção realizada por esse corretor. O primeiro se refere à provável necessidade de realizar alguns ajustes no critério, algo que provavelmente ocorreu somente na correção da **solução F**. O segundo é que com um critério bem estabelecido, mesmo tendo realizado ajustes, o corretor 14 conseguiu

atribuir todas as pontuações de forma coerente: meio ponto para as **soluções A e C**, 1 (um) ponto para as **soluções B, D e E**, e para a **solução F**, 0,25.

Figura 17 – Solução F, corretor 14.

<p>Solução</p> $\begin{array}{r} 13 \\ + 5 \\ \hline 65 \end{array}$ <p>R: 65 carrinhos</p>	<p>Pontuação na questão: 0,25</p> <p>Critério/justificativa para a pontuação</p> <p>1^o Respondeu corretamente a questão? Não (não para 2^o)</p> <p>2^o Apresenta um raciocínio capaz de responder corretamente? Não (não para 3^o)</p> <p>3^o Apresenta um raciocínio capaz de responder corretamente? Não</p>
<p>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</p> <p>Faltou tirar os 6 carrinhos.</p>	

Fonte: elaborado pelo autor

Por último, o critério pragmático estabelecido pelo corretor 14 pode apresentar uma falha ao desconsiderar aspectos cognitivos observáveis nas soluções, ao pontuar com a nota máxima uma solução que utiliza contagem (**solução D**). Por outro lado, um critério claro e coerente pode reduzir vieses cometidos pelo corretor, aumentando a credibilidade do resultado e, sendo divulgado aos estudantes, pode contribuir com mais transparência ao exame.

O corretor 51, para a **solução B**, escreveu que a forma de atribuir a pontuação dependeria do estudante que realizaria a prova. *Zero: A prova exige que se tenha cálculos (avaliação para o aluno que só apresenta nota baixa). 1,0: avaliação para o aluno que apresenta boas notas.* O corretor 51 foi o único que escreveu algo que alguns colegas comentaram durante o debate das questões: depende de quem é o aluno. Por um lado, esse corretor compreende que o processo avaliativo transcenda a prova, pois utiliza de uma avaliação do estudante para atribuir a pontuação. No entanto, há nesse comentário um julgamento de capacidade e honestidade preocupante. Alunos considerados fracos, que geralmente possuem maiores e diversas limitações que impactam sua aprendizagem, encontram barreiras ainda maiores a serem superadas para conseguir um bom desempenho, devido ao julgamento dos seus professores. Um julgamento que atesta sua incapacidade de conseguir encontrar um bom resultado de forma honesta.

O corretor 11 estabeleceu o seguinte critério, no início da folha de correções, *interpretação (0,3), contas (0,3 + 0,3) e resposta (0,1)*. Interpretamos que as duas

parcelas que compõem o critério de pontuação ‘contas’ correspondem às duas operações utilizadas na solução. Este corretor atribui para as **soluções A e B** a mesma pontuação, 0,7. Na solução A, escreveu *errou uma das contas*, o que nos leva a concluir que descontou apenas 0,3 da segunda parte dos cálculos e pontuou 0,1 pela escrita da resposta, independentemente, de estar correta. Na **solução C**, o corretor escreve, com caneta azul na justificativa, *tiraria metade das contas* e, no espaço destinado ao feedback, *é preciso retornar em ambos os algoritmos*.

Como explicado anteriormente, após o preenchimento da *folha de correção*, houve um debate entre o pesquisador e os participantes sobre as soluções analisadas e corrigidas, como parte da oficina. Durante esse debate, ocorrido um pouco antes do recolhimento das folhas, os corretores foram solicitados a não alterarem suas respostas. O corretor 11 comenta, com caneta de cor vermelha no espaço da justificativa da **solução C**, *eu não tinha notado que a 2ª conta estava correta (o que foi comentado na discussão) por isso, eu aumentaria a nota*. O espaço destinado à pontuação estava sem rasuras e com uma única pontuação: 0,7. Não fica claro se a pontuação foi escrita antes ou depois do comentário em vermelho. O que podemos concluir é que esse corretor não havia observado, inicialmente, algo correto na **solução C**. Seria uma influência da posição em que o erro ocorre?

5.3A posição em que o erro ocorre

As **soluções A e C** apresentam desenvolvimentos parcialmente corretos. Elas foram planejadas para compreender o comportamento dos corretores ao se depararem com soluções que se diferenciam, principalmente, pelo lugar em que o erro ocorre, como ilustra a Figura 18. Ocorrendo do início, o erro ‘contamina’ as demais etapas da solução. Notem que na solução do estudante C, o minuendo (515) é bem diferente do minuendo esperado (65). Por outro lado, a subtração realizada está correta.

Figura 18 – A posição em que o erro ocorre

Estudante A	Estudante C
$\begin{array}{r} 13 \\ \times 5 \\ \hline 65 \end{array}$ $\begin{array}{r} 65 \\ - 6 \\ \hline 61 \end{array}$ <p>R: 61 carinhos</p>	$\begin{array}{r} 13 \\ \times 5 \\ \hline 515 \end{array}$ $\begin{array}{r} 515 \\ - 6 \\ \hline 509 \end{array}$ <p>R: 509 carinhos</p>
Operação esperada: 65 - 6	

Fonte: elaborado pelo autor

Alguns corretores atribuíram às **soluções A e C** a mesma pontuação, como é o caso do corretor 5. Estes corretores parecem ter estabelecido critérios de correção, nos quais, as pontuações em que os erros cometidos nas duas soluções possuem o mesmo peso para o estabelecimento da pontuação. No entanto, esses corretores são exceções.

Na **solução A**, o corretor 3 escreve que o estudante *acertou uma das contas*, atribuindo pontuação de 0,5. Na **solução C**, atribui zero e escreve *contas estão erradas*. Não é possível concluir se esse corretor não identifica o acerto ou desconsidera a segunda parte pela ‘contaminação’, ou seja, por utilizar no cálculo um número oriundo de um erro da primeira parte (515).

O corretor 8 identifica os erros e acertos das **soluções A e C**. Embora faça uma lista de itens que deveriam ser verificados na correção e que nessa lista não haja nenhum motivo aparente para o estabelecimento de pontuação distinta, esse corretor atribui pontuações distintas, como ilustra a figura 19:

Figura 19 – Pontuações e justificativas apresentadas pelo corretor 8 para as soluções A e C.

Pontuação na questão:	Solução A	Pontuação na questão:	Solução C
0,5		0,2	
Critério/justificativa para a pontuação		Critério/justificativa para a pontuação	
<ul style="list-style-type: none"> - Não tem as expressões matemáticas do problema - Errou a subtração - Acertou a multiplicação 		<ul style="list-style-type: none"> - Noção de tabuada: ok. - Subtração: ok. - O aluno não utilizou a multiplicação corretamente. 	

Fonte: elaborado pelo autor

A percepção aparente do corretor 1 é que o estudante A comete um erro (lapso) e o estudante C desconhece a operação de multiplicação, essa percepção talvez tenha influenciado a diferença de pontuações atribuídas às soluções: 0,9 para a **solução A** e 0,7 para a **solução B**. Ao explicar o critério adotado, este corretor afirma que o estudante A

identificou as operações necessárias e as realizou. Contudo, errou na subtração, em relação ao estudante B, afirma que aparentemente o estudante não sabe operar a multiplicação. Os corretores 9, 26, 42, 43 e 50 fizeram análises similares como ilustra o Quadro 9:

Quadro 9 - Comparativo das justificativas das soluções A e C

Corretor	Justificativa da pontuação na solução A	Justificativa da pontuação na solução C
9	<i>O aluno A acertou o raciocínio e cometeu um erro na subtração que pode ter sido falta de atenção ou não sabe subtração com recurso¹⁷</i>	<i>teve um raciocínio certo, porém não sabe fazer multiplicação com reserva (sistema de numeração)</i>
25	<i>Identificou corretamente as operações necessárias</i>	<i>O aluno ainda não compreendeu a utilização das operações.</i>
26	<i>Não se atentou quanto a solução do algoritmo da subtração.</i>	<i>apresenta dificuldade quanto à resolução de operações de multiplicação</i>
42	<i>Ele acertou metade do cálculo pedido.</i>	<i>Ele errou tudo! Não tem como aproveitar nada!</i>
43	<i>Apesar do raciocínio lógico estar correto, houve um erro de operação Matemática.</i>	<i>Raciocínio correto e execução toda errada.</i>
50	<i>Compreendeu o problema, porém apresentou falta de atenção.</i>	<i>Não sabe efetuar uma conta de multiplicação.</i>

Fonte: elaborado pelo autor

Uma possível explicação para as diferenças de pontuação e interpretação seria o efeito gerado pelo lugar em que o erro ocorre. Como o número utilizado na segunda operação depende do primeiro resultado, o impacto se torna maior caso o erro ocorra na primeira operação. A ‘conta’ esperada para a segunda parte da solução era ‘65 – 6’, isto ocorre na **solução A**, mas não na **solução C**, que apresenta a operação ‘contaminada’.

Considerando que a moda obtida na correção da **solução C** foi zero e analisando cuidadosamente algumas dessas respostas, é possível questionar se alguns desses corretores identificaram que a segunda operação apresentada pelo estudante C estava correta. O corretor 24, por exemplo, não identificou, pois escreveu como justificativa para

¹⁷ Provavelmente se referindo ao reagrupamento na subtração de naturais.

a **solução C**: *errou a questão por completo*. O corretor 42, também foi enfático ao afirmar que *ele errou tudo, não tem como aproveitar nada*. Ambos atribuíram zero à **solução C**.

Se a habilidade desenvolvida/avaliada na segunda etapa da solução deste problema é a subtração com recurso a uma ordem superior, a **solução C** atenderia a proposta. No entanto, essa ‘contaminação’ provocada pelo erro na multiplicação impacta na interpretação dos corretores. As justificativas do corretor 16 podem ser um bom exemplo dessa interpretação. Sobre a **solução A**, esse corretor justifica que *a primeira parte foi resolvida corretamente, porém houve um erro de cálculo na segunda parte, o que levou a uma resposta final que não corresponde ao problema*. Sobre a **solução C**, escreve que *houve um problema em resolver a primeira parte da questão, que acarretou na manutenção do erro durante sua continuação*. O corretor 16 atribuiu 0,6 à **solução A** e zero à **solução C**.

Outra possibilidade identificada nas respostas está relacionada à ordem de grandeza. Para a **solução A**, o corretor 1 escreve: *65 – 6 não é 61. Tente novamente. Experimente revisar suas contas ao final de cada atividade*. Novamente, o corretor deixa claro que neste caso o erro não significa falta de conhecimento, e sim, um lapso. Pois lapsos podem ser mais facilmente revistos e identificados. Para o estudante da **solução C** ele escreve o seguinte *feedback*: *observe que 13×5 é o mesmo que $13 + 13 + 13 + 13 + 13$ e este número não pode ser tão grande como 515. Tente refazer o exercício*. Talvez, para o corretor 1, o problema não está na posição do erro, e sim na ordem de grandeza. O resultado encontrado pelo estudante C na multiplicação (515) não está na ordem de grandeza esperada para aquela multiplicação (13×5).

O corretor 2 atribui zero à **solução C** escrevendo que *esta, para mim, foi a mais difícil de avaliar, o estudante não teve criticidade para o resultado encontrado*. O comentário foi escrito como uma observação acima da questão. Os espaços destinados ao critério e ao *feedback* não foram preenchidos por este corretor. Sobre essa discussão dois pontos podem ser destacados.

Primeiramente, corrigir questões escritas de Matemática, apesar de rotineira, não é uma tarefa fácil. A correção de questões discursivas é uma tarefa que faz parte da prática profissional de professores de Matemática, sobretudo em um sistema educacional que valoriza a utilização de provas escritas como instrumento avaliativo. O segundo aspecto

diz respeito ao trecho *não teve criticidade para encontrar o resultado*. Parece que novamente temos uma ponderação relacionada à ordem de grandeza. O corretor 10 que pontuou as **soluções A e C**, respectivamente em, 0,8 e 0,5, parece destacar também a ordem de grandeza como um fator relevante, ao afirmar que o resultado encontrado pelo estudante na **solução C** *é absurdo diante dos números do problema*.

Isso nos remete a outro questionamento: um estudante do sexto ano teria a habilidade de analisar a ordem de grandeza? Como discutiremos na seção a seguir, aspectos relacionados ao currículo parecem ser desconsiderados pelos corretores que participaram desta pesquisa.

5.4 A desconexão com a matriz curricular

Consideramos que o conhecimento sobre o currículo, e mais especificamente, sobre a matriz curricular é um pré-requisito para que um professor avalie seus alunos. Para avaliar um estudante o professor deve considerar o que os estudantes já estudaram, o que estão estudando e o que estudarão. As ações pedagógicas desenvolvidas e em desenvolvimento. As habilidades que foram, ou deveriam, ter sido desenvolvidas em anos anteriores.

As **soluções D e E** foram elaboradas com o objetivo de compreender como os corretores iriam interpretar e corrigir soluções alternativas. A ideia inicial era identificar se esses corretores ‘aceitariam’ soluções alternativas. Considerava-se a possibilidade de essas questões não serem bem pontuadas pelos corretores por se tratar de soluções distintas daquela esperada. O que se observou na pontuação é que essas soluções foram bem pontuadas, recebendo, ambas, 0,95 de média e 1,0 de moda.

Apenas no caso do corretor 16, é possível identificar um comentário indicando a não aceitação da **solução D** por ser alternativa: *a resposta está correta, mas a forma como foi encontrada não foi ensinada na sala de aula*. Como *feedback* para essa solução, esse corretor sugere que o estudante estude *o algoritmo da soma que facilitaria a resolução do problema*. Na **solução E**, esse corretor escreve *a solução foi exercida bem, mas poderia ser executada de forma mais polida*. Pelos comentários e pelas pontuações, 0,7 para a **solução D** e 1,0 para a **solução E**, podemos supor que para esse corretor a escrita do sistema de numeração indo-arábico era importante nessa avaliação.

Alguns comentários escritos nas justificativas pelos corretores destacavam que as **soluções D e E** não eram adequadas ao ano de escolarização (sexto ano).

Sobre a **solução D**, o corretor 2 escreveu *a resposta me surpreendeu, parece um aluno dos anos iniciais (1º ou 2º). Mas, ele resolveu a questão;* o corretor 9 comentou *acertou o raciocínio e mostrou que sabe o que fazer, mesmo sem usar os algoritmos da multiplicação e subtração* e o corretor 33 escreve *fez por contagem*. Ambos os corretores atribuíram a esta solução pontuação máxima. Para o corretor 8, em relação à **solução D**, *o aluno tem apenas a intuição da multiplicação, mas não sabe esta operação*, enquanto que, para o corretor 43, esse estudante apresenta *raciocínio correto, mas elementar*.

Sobre a **solução E**, o corretor 17 parece indicar que o desconto de 0,1 décimo ocorre devido à *falta do uso da ferramenta complexa da Matemática*, ou seja, a não utilização dos algoritmos esperados. O mesmo desconto foi dado pelo corretor 37 que escreve *o aluno teve um raciocínio perfeito, mas não domina as operações básicas*. Para o corretor 25, que atribuiu pontuação máxima a essa solução, *o aluno soube realizar a conta necessária, mas não conseguiu perceber/utilizar a operação de multiplicação na forma de conta. Mas, aparentemente, entende o conceito*.

Alguns corretores destacaram a inadequação do modo de resolver apresentado na **solução D** para a resolução de problemas envolvendo números maiores. O corretor 13 escreve que *o desenvolvimento e a resposta estão completamente corretos. Apesar de não ser ideal, pois seria praticamente impossível listar um número muito grande*. No feedback dado pelo corretor 37 ao estudante D também é possível identificar essa preocupação: *seu raciocínio está perfeito, porém você precisa aprender a realizar as 4 operações básicas, pois em situações que envolvam valores maiores, será complicado resolver dessa forma*.

Na análise das **soluções D e E**, pode-se destacar que, de algum modo, os docentes demonstraram alguma preocupação com o currículo, seja na indicação da ausência das operações básicas nas soluções, multiplicação e subtração, principalmente na solução D, seja pela preocupação com o desempenho desses estudantes em problemas futuros envolvendo números maiores. No entanto, essa preocupação foi limitada.

Diferentemente do que ocorreu nas outras soluções, que receberam pontuações bem menores, a preocupação ou a identificação da inadequação da solução com a escolaridade não foi transformada em descontos na pontuação. A percepção de que a

solução D, por exemplo, constatada e escrita nos comentários, refletia um procedimento envolvendo habilidades de séries iniciais, a contagem, não foi impedimento para a atribuição de pontuações altas. Para os corretores, cometer um erro operacional na utilização dos algoritmos adequados à escolaridade é mais grave, pelo menos merecedor de mais descontos na pontuação, do que não saber operar com esses algoritmos.

Isso pode refletir mais do que a ausência de critérios de correção, a valorização negativa do erro e a desconexão com o currículo.

5.5 A ênfase na resposta final e a desvalorização de elementos relacionados à interpretação do problema

Os corretores que participaram dessa pesquisa, de um modo geral, valorizaram mais a resposta final do que as habilidades Matemáticas envolvidas. As médias atribuídas às **soluções D e E** (0,95) foram superiores àquelas atribuídas às demais soluções. A média atribuída à **solução B** (0,54) é um pouco menor que a média atribuída à **solução A** (0,58) e maior que a atribuída à **solução C** (0,34).

As habilidades envolvendo a interpretação do enunciado e a compreensão dos algoritmos de subtração e multiplicação foram subvalorizadas em detrimento da resposta final. Se por um lado, os estudantes fictícios A, C, D e E interpretaram corretamente o problema, por outro, apenas os estudantes A e C demonstraram conhecer os dois algoritmos necessários, mesmo cometendo um erro operacional. O estudante D, por exemplo, resolveu esse problema com o artifício de contagem, através de “pauzinhos”, demonstrando um possível atraso em sua aprendizagem se comparado aos demais estudantes A e C.

A **solução E** pode ser interpretada de diferentes formas. O estudante pode não ter compreendido corretamente o problema, indicando problemas com leitura e interpretação de textos. Outra possibilidade é que tenha compreendido o problema, mas não tenha conseguido estabelecer uma estratégia completa para resolvê-lo e sem essa estratégia, optou em escrever como resposta final o único resultado que tinha alcançado. Pode não ter conseguido terminar a solução, faltou tempo. Os alunos, comumente, reclamam disso. Quaisquer que sejam as interpretações, nessa solução, o estudante não conseguiu mostrar a devida compreensão do problema através de uma estratégia correta. Mesmo assim, a

solução E atingiu a mesma moda que a **solução A**, ambas 0,5. Atingiu também média (0,4) e moda (0,5) superiores àquelas atribuídas à **solução C** (0,34 de média, 0 de moda).

De um modo geral, os corretores não utilizaram a compreensão/intepretação do enunciado e o estabelecimento de estratégias corretas como critérios relevantes na correção para o estabelecimento de pontuações. Erros de cálculos foram, por outro lado, fatores com grande impacto negativo, principalmente, porque esses erros afastavam a solução da possibilidade de atingir a resposta final, o aspecto mais valorizado pelos corretores. Tal situação, talvez, reflita a concepção positivista que os professores possuem da Matemática, como uma ciência pronta e acabada em que os erros devam ser execrados.

5.6A Construção de Feedbacks Formativos

Como discutido no capítulo 3, *feedbacks* como ‘estude mais’, ‘faça mais exercícios’ ou ‘tenha mais atenção aos algoritmos’ não apontam exatamente qual é o erro cometido no desenvolvimento, muito menos, indicam como proceder especificamente com esse erro. Do mesmo modo, elogios ou críticas dirigidas ao estudante e/ou ao trabalho, desprovidos de qualquer orientação direcionada à correção da tarefa, como por exemplo, ‘bom trabalho’, ‘parabéns!’ ou ‘você pode fazer melhor que isso’ podem oferecer aos estudantes contribuições motivacionais, psicológicas, mas não podem ser consideradas contribuições didáticas. Nos referiremos a esses *feedbacks* como não formativos.

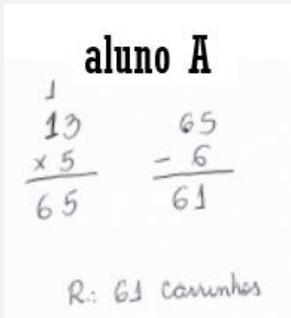
Por outro lado, um *feedback formativo* é aquele que apresenta características didáticas, que oferece pistas de como continuar, é descritivo, dirigido à regulação, ou seja, que permite ao aluno identificar o que falta fazer e como fazer para alcançar o esperado. Propõe uma situação que leve o estudante a rever sua solução, identificando quais foram, exatamente, o(s) erro(s) do processo. Que o ajude além de identificar o que errou, compreender o porquê do erro e o que deve ser feito para corrigi-lo.

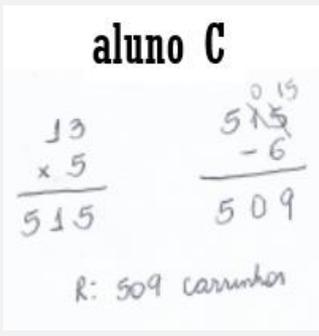
Os comentários como *observe quanto é $61+6$. O que você pode dizer sobre a resposta?* e (...) *observe que 13×5 é o mesmo que $13+13+13+13+13$ e este número não pode ser tão grande como 515*, dados pelos respondentes, permitem não somente a identificação do erro, mas oferecem ao estudante a possibilidade de pensar em uma *prova real* ou em caminhos que permitam refletir e identificar erros semelhantes em outras oportunidades. Consideramos também *feedbacks formativos* aqueles que identificaram o

caminho correto, como 5 unidades é menor que 6 unidades, pedimos 10 unidades emprestadas ao 6 que representa 6 dezenas (60). Daí, $15 - 6 = 9$. Logo, a resposta é 59 ou, simplesmente, como você sabe que ficaram 59 carrinhos? É preciso apresentar a resolução. Talvez, esse segundo grupo de respostas não ofereça um *feedback* tão bom quanto o anterior pois apenas indica o erro e/ou o caminho certo, diferentemente do primeiro grupo que traz algo agregado, uma técnica de testagem, de *prova real*.

Como os alunos eram do sexto ano do ensino fundamental, a resolução esperada utilizaria a multiplicação ($13 \times 5 = 65$), seguida da subtração ($65 - 6 = 59$). O Quadro 10 apresenta a solução dos estudantes A, B e C. A primeira coluna corresponde ao *feedback* não formativo. Trata-se de um *feedback* com, provavelmente, menores efeitos para a aprendizagem. Podendo ser avaliativo ou descritivo, transmite uma informação dirigida ao *self* ou à tarefa. A segunda coluna, por sua vez, corresponde ao *feedback* formativo, com características descritivas, mais dialógico e dirigido ao processo ou à autorregulação.

Quadro 10 - *Feedbacks* das questões A, B e C

Aluno/Solução	<i>Feedbacks</i> formativos	<i>Feedbacks</i> não formativos
<p>aluno A</p> 	<p>Observe quanto é $61 + 6$. O que você pode dizer sobre a resposta?</p> <p>Como 5 unidades é menor que 6 unidades, “pedimos” 10 unidades emprestadas ao 6 que representa 6 dezenas (60). Daí, $15 - 6 = 9$. Logo, a resposta é 59.</p>	<p>Estude mais a subtração;</p> <p>Rever conceitos de operação com número naturais;</p> <p>Mais atenção na resolução das questões;</p> <p>Se atente à subtração;</p> <p>Fazer mais exercícios semelhantes.</p>
<p>aluno B</p> 	<p>Como você sabe que ficaram 59 carrinhos? É preciso apresentar a resolução;</p> <p>Não dar resposta sem a justificativa;</p> <p>Cobrar do estudante a resolução. Caso saiba, coloque a conta. Nada na Matemática e na vida vem do nada.</p>	

<p style="text-align: center;">aluno C</p> 	<p><i>Observe que 13×5 é o mesmo que $13 + 13 + 13 + 13 + 13$ e este número não pode ser tão grande como 515;</i></p> <p><i>Apontar ao estudante que o número encontrado é absurdo em relação ao problema;</i></p> <p><i>Lembrar que o 1 não deveria ir para baixo e pedir para revisar isso;</i></p> <p><i>Ensinar a prova real para que o aluno confira as suas respostas.</i></p>	<p><i>Praticar operações básicas;</i></p> <p><i>Ter maior atenção aos algoritmos;</i></p> <p><i>É preciso retomar ambos os algoritmos;</i></p>
---	---	--

Fonte: elaborado pelo autor

Em relação à solução do aluno A, dos 51 entrevistados, 42 ofereceram *feedbacks não formativos*, 2 deixaram em branco e apenas 7 ofereceram *feedbacks formativos*. Dos formativos, em apenas um o professor não apresentou a resposta correta, respeitando a ideia de “timing”.

O maior desvio padrão em relação às notas atribuídas ocorreu na correção da solução do estudante B, o que pode ser justificado pela inexistência de um senso comum entre os professores em relação à resposta ‘seca’ (sem a presença de justificativas ou contas). Note que, neste caso, todos os *feedbacks* foram classificados como formativos. Mesmo dando nota máxima ou mínima, todos os professores alertaram aos estudantes sobre a necessidade de apresentar as contas. No entanto, somente dois professores propuseram modificar a nota mediante a apresentação de nova resolução. Alguns motivos podem justificar essa postura dos respondentes, o primeiro deles se refere à falta de autonomia da profissão, outra possibilidade está no modo engessado como a avaliação é comumente interpretada.

Para a solução do estudante C, 17 dos participantes ofereceram *feedbacks formativos*, 2 deixaram em branco e os demais, *feedbacks não formativos*.

As **soluções D e E** estão matematicamente corretas. Os corretores ofereceram 28 *feedbacks* classificados como formativos para a resposta do aluno D e 23 para a do aluno E. Essas duas soluções alternativas fornecem aos professores informações sobre o que o aluno não sabe. Não fica claro se o aluno E domina os algoritmos de multiplicação e subtração. Essa dúvida fica mais evidente quando observamos a **solução D** que utiliza o

método da contagem, refletindo um, possível, grau de aprendizagem (ou cognitivo) inferior desse estudante se comparado aos demais.

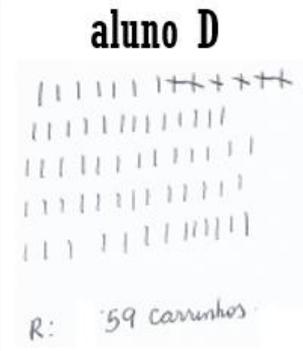
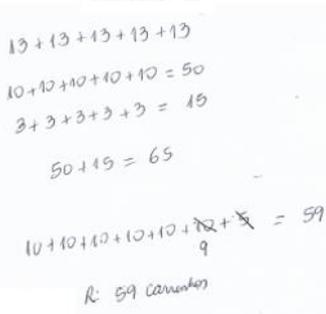
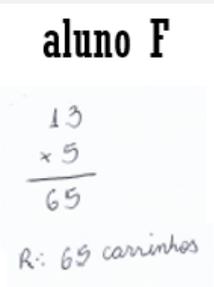
Note que diferentemente dos dois exemplos anteriores, o estudante F conhece o algoritmo da multiplicação. No entanto, há algumas possibilidades para a incompletude da questão: falta de tempo, esquecimento ou não compreensão do texto. Em uma perspectiva formativa, aquela que concebe a avaliação para as aprendizagens, um modo de ajudar esse estudante seria reconhecer qual é o motivo dessa incompletude. Desse modo, há um forte indício de uma prática formativa quando o professor convida o aluno a comparecer ao quadro para discutir com a turma sua solução alternativa ou, pelo menos, permite que ao estudante apresentar suas contas. O mesmo poderia ser aplicado ao estudante B.

Desse modo, *comentários como e se eu dissesse pra você que havia 500 caixas, cada uma com 13 carrinhos...? e tente refazer o problema com 7043 carrinhos em cada caixa* oferecem uma maior regulação da aprendizagem, pois permitem aos estudantes ao mesmo tempo refletir sobre a quão limitada foi sua estratégia inicial e, também, sobre a necessidade de um aprofundamento na resolução, o que no caso seria a apropriação de novos algoritmos.

Em relação a esses três estudantes (D, E e F), de modo análogo aos anteriores, consideramos que comentários como *releia o enunciado, praticar a resolução de problemas elementares envolvendo as quatro operações, precisa melhorar a multiplicação e ficou muito confuso os cálculos, porém chegou à resposta correta* não oferecem contribuições específicas para a aprendizagem do tópico em questão.

O Quadro 11 apresenta alguns dos *feedbacks* dados às soluções dos alunos D, E e F.

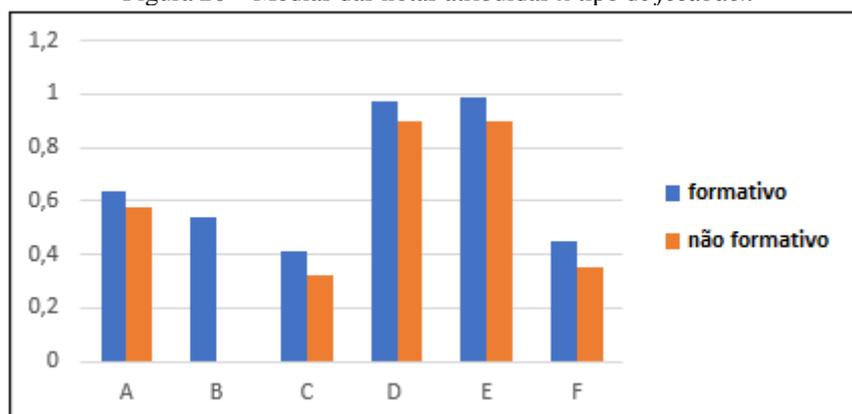
Quadro 11 – Feedbacks das questões D, E e F

Aluno/Solução	Feedbacks formativos	Feedbacks não formativos
<p>aluno D</p> 	<p>Tente refazer o problema com 73 carrinhos em cada caixa;</p> <p>... E se eu dissesse pra você que havia 500 caixas, cada uma com 13 carrinhos...?</p> <p>Considerando os algoritmos praticados em sala, você saberia utilizá-los e apresentar outra solução? ... nem sempre uma representação como essa é viável – por exemplo, quando os números são maiores.</p>	<p>Praticar resolução das quatro operações como ministrado em sala;</p> <p>Treinar contas;</p> <p>Parabéns!</p>
<p>aluno E</p> 	<p>Tente refazer o problema com 7043 carrinhos em cada caixa;</p> <p>Que tal ir ao quadro e compartilhar com a turma como você pensou? (este professor dividiu a resolução em partes correspondentes a 5 x 13 e 65 – 6);</p> <p>Parabéns! Agora tente fazer utilizando outro método. Isso irá facilitar em contas maiores.</p>	<p>Praticar resolução das quatro operações como ministrado em sala;</p> <p>Precisa melhorar a multiplicação;</p> <p>Ficou muito confuso os cálculos, porém chegou à resposta correta;</p> <p>O aluno compreende o sentido de unidade e dezena, e consegue estabelecer relação de igualdade ao estabelecer $13 = 10 + 3$.</p>
<p>aluno F</p> 	<p>O que representa 13×5 no problema? Você percebeu que carrinhos foram retirados?</p> <p>O cálculo foi correto, porém não foi contado a retirada de 6 carrinhos da caixa, sugiro que leia com mais atenção o enunciado.</p>	<p>Praticar a resolução de problemas elementares envolvendo as quatro operações;</p> <p>Ter mais atenção ao enunciado;</p> <p>Releia o enunciado;</p> <p>Estudasse mais a interpretação de enunciados.</p>

Fonte: elaborado pelo autor

O gráfico a seguir (figura 20) apresenta a comparação entre as médias oferecidas pelos participantes que forneceram *feedbacks formativos* e as médias relacionadas aos *feedbacks não formativos*.

Figura 20 – Médias das notas atribuídas x tipo de *feedback*



Fonte: elaborado pelo autor

Na figura 20, é possível observar que as notas médias atribuídas pelos professores e licenciandos que ofereceram *feedbacks formativos* foi ligeiramente superior àquela atribuída por professores e licenciandos com *feedbacks não formativos*. Considerando que esses dados sejam relevantes em termos estatísticos, duas interpretações são possíveis:

- (1) A primeira se refere ao zelo na execução da tarefa proposta na pesquisa. Os professores que atribuíram maior pontuação foram mais cuidadosos na correção e, por isso, ofereceram também, um *feedback* mais cuidadoso, formativo.
- (2) A segunda se refere à capacidade dos professores em oferecer melhores *feedbacks*. Professores com maiores aptidões didáticas, que podem oferecer melhores *feedbacks*, também tendem a valorizar melhor a produção dos seus estudantes, atribuindo, inclusive, maiores pontuações. Nesse caso, esse resultado pode ser explicado pelo modo como o professor compreende a Educação, a Matemática e o ensino de Matemática.

De um jeito ou de outro, esse resultado indica que a interpretação do erro como um processo natural da aprendizagem e de construção do conhecimento pode impactar positivamente a avaliação através de melhores *feedbacks* e uma maior valorização da produção dos estudantes.

CAPÍTULO 6 - ALGUMAS REFLEXÕES SOBRE A TRAVESSIA



*O fim de uma viagem é apenas o começo de outra.
É preciso ver o que não foi visto,
ver outra vez o que se viu já,
ver na primavera o que se vira no verão,
ver de dia o que se viu de noite,
com o sol onde primeiramente a chuva caía,
ver a seara verde,
o fruto maduro,
a pedra que mudou de lugar,
a sombra que aqui não estava.
É preciso voltar aos passos que foram dados,
para repetir e para traçar caminhos novos ao lado deles.
É preciso recomeçar a viagem.
Sempre.*

José Saramago

Ao longo da pesquisa e do estudo realizado no processo de doutoramento, pude refletir sobre diferentes aspectos relacionados à avaliação. Este capítulo apresenta duas dessas reflexões: A Dupla Diversificação Avaliativa e a Utilização Formativa de Testes Somativos. São sugestões que podem contribuir para a (re)construção de práticas avaliativas mais direcionadas à aprendizagem, em uma tentativa de valorizar diferentes habilidades e potencialidades dos estudantes.

6.1 A Dupla Diversificação Avaliativa

A avaliação é um processo complexo e dinâmico cujo objetivo central é compreender a realidade. A avaliação está mais próxima de um filme do que de uma foto. A aplicação de exames pontuais, como é praticada comumente nas universidades e nas escolas, geralmente não permite aos professores realizarem uma leitura precisa da realidade. Durante a pesquisa da tese, refletimos sobre a necessidade de aproximarmos a

foto ao filme. Um filme é, pelo menos antes da era digital, uma sequência de fotos de diferentes ações, obtidas, algumas vezes, de ângulos diferentes.

Essa foto gera uma nota. O que esse número, de fato, representa? Esse aluno tirou nota 6,0, está aprovado. Já esse outro, nota 5,0, reprovado! Acreditamos na nota! O que esses números representam? Acreditamos que o estudante que obteve a nota 6,0 teve um melhor desempenho que aquele da nota 5,0. Talvez, esses números representem somente a nossa crença. Uma crença legitimada por um mito. O mito de que é possível ‘medir’ a aprendizagem de alguém a partir do seu desempenho em um exame. Precisamos reconhecer que os critérios de seleção de itens e de correção de questões impactam esses resultados. Reconhecer que nossas aulas impactam esse resultado.

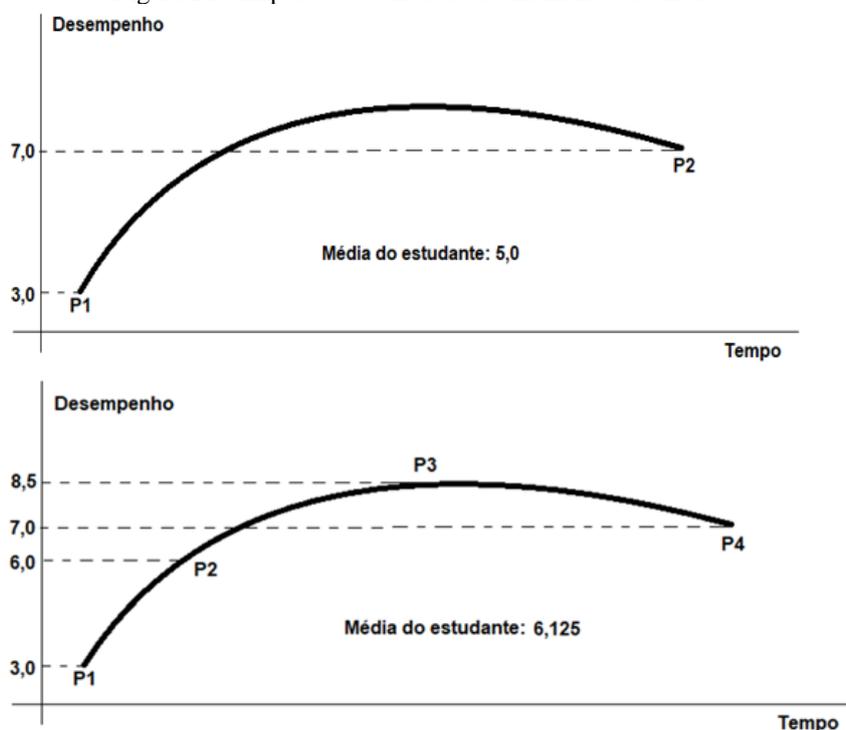
Partimos do pressuposto que o ensino, a aprendizagem e a avaliação são processos distintos, porém interdependentes e indissociáveis. Precisamos reconhecer que qualquer que seja o instrumento e método utilizados, a avaliação é um processo imperfeito. A avaliação é uma leitura, não uma medida. Uma leitura do que, e de como, os estudantes aprenderam. Do que, e por que, não aprenderam. Do que precisa ser feito para que a aprendizagem ocorra. A aprendizagem é o fim! O objetivo central do ensino e da avaliação.

Uma orientação pedagógica desenvolvida nesta tese é a *Dupla Diversificação Avaliativa*, que consiste em que, no desenvolvimento do processo avaliativo escolar, o professor diversifique momentos e instrumentos, a fim de proporcionar uma melhor leitura da aprendizagem dos estudantes.

Como discutido anteriormente, segundo Black e Wiliam (1998), um bom teste pode ser a causa da aprendizagem, pode promover aprendizagem e, neste sentido, segundo esses dois autores, é melhor que os testes sejam curtos e frequentes. Ao ampliar os momentos avaliativos, o professor tenderá a conseguir uma melhor leitura do desempenho dos estudantes. Na analogia do filme, quanto maior o número de fotos, melhor poderemos compreender uma cena. Entendemos que o desempenho dos estudantes varia de acordo com o tempo. O ser humano é assim! Alguns dias produzimos mais, em outros, menos. Quanto mais variamos os momentos, melhor poderemos compreender o desenvolvimento da aprendizagem dos estudantes, reduzindo, também, possíveis impactos negativos e pontuais na avaliação.

Tal ideia é ilustrada na Figura 21. Na imagem superior, utilizando duas provas P1 e P2 os professores conseguem uma leitura menos representativa da realidade do que na imagem inferior, em que são utilizados quatro instrumentos avaliativos.

Figura 21 - Impacto da variedade de momentos avaliativos.



Fonte: VAZ; NASSER; LIMA (2021)

De modo pragmático, quatro testes de 5 questões aplicados em quatro momentos distintos podem oferecer uma melhor leitura das aprendizagens do que dois testes de 10 questões cada.

Além da variedade de momentos, consideramos que a utilização de instrumentos pode promover ao professor melhores chances de compreender as limitações e potencialidades relacionadas às aprendizagens dos estudantes. Diferentes leituras em instrumentos diversificados podem proporcionar avaliações mais inclusivas por considerar multiplicidades de habilidades dos estudantes. A utilização de, além das provas tradicionais, seminários, portfólios, atividades em grupo, autoavaliação, provas com cola, provas em fases, avaliação por pares, ... pode contribuir para a construção de uma avaliação plural e, conseqüentemente, global.

Historicamente, a escola tem sido indiferente às diferenças.

Sob o prisma da igualdade, criamos mecanismos homogêneos para todos. Se acreditarmos que tratar os alunos de forma igual é oferecer as mesmas

oportunidades, as mesmas avaliações, o mesmo currículo, para quem apresenta diferenças, é certo que estaremos produzindo situações de desigualdades. (SANTIAGO; SANTOS; MELO, 2017, p. 639)

6.2 A utilização formativa de testes somativos

A utilização formativa em testes somativos foi proposta por Black e Wiliam. (WILIAM, 2011). Black e colaboradores (2003) argumentam que usar preparação pré-teste ou acompanhamento pós-teste como estratégias avaliativas aprimoram o desempenho dos alunos nas avaliações somativas e/ou promovem autorregulação. Para Fernandes (2019, p. 52), a utilização formativa de testes somativos parece ser aceitável, no entanto, a utilização de testes formativos para fins somativos parece, inicialmente, “mais problemática tendo em conta a sua natureza e sua inserção no processo pedagógico”. Lam (2013) problematiza a possibilidade de utilizar em uma mesma avaliação características somativas e formativas.

Avaliações somativas e formativas são tipicamente caracterizadas em termos de sua função. A primeira enfatiza a classificação e julgamento enquanto a última apoia a aprendizagem. Uma questão chave dessa questão na literatura sobre Avaliação para as Aprendizagens é se essas duas funções de avaliação são incompatíveis ou se poderiam ter uma sinergia eficaz. (LAM, 2013, p. 70, tradução nossa)

O objetivo desse estudo de Lam (2013) foi investigar se os testes internos em uma escola poderiam criar sinergias produtivas entre avaliação somativa e formativa na melhoria do desempenho dos alunos nos testes e no aprendizado autorregulado em Hong Kong. Os resultados comprovaram a eficiência da abordagem formativa concomitante com avaliações somativas para melhorar o desempenho dos alunos de nível primário em testes/exames internos da escola, contribuindo para o desenvolvimento da capacidade de autorregulação na aprendizagem de idiomas.

Refletindo sobre a possibilidade de coexistência da Avaliação da Aprendizagem (AdA) e da Avaliação para a Aprendizagem (ApA), associadas às funções somativas e formativas, respectivamente, Fernandes (2009, 2019), Borralho, Cid e Fialho (2019) consideram haver três concepções possíveis entre as formas de coexistência de tais avaliações. A primeira é aquela que concebe tais modelos avaliativos como dicotômicos, opostos, com objetos, finalidades e funções diferentes, tendo “pouco ou nada tendo a ver uma com a outra e, por isso, serem opostas e com inserções pedagógicas e relações com ensino e a aprendizagem muito distintas” (FERNANDES, 2019, p. 151). Na segunda concepção as avaliações somativas e formativas são vistas como duas dimensões

pertencentes a um mesmo constructo, existindo uma espécie de continuum entre os dois polos extremos: a avaliação formativa informal e avaliação somativa formal, havendo práticas avaliativas híbridas.

Mas trata-se de uma formulação interessante que nos remete para uma outra discussão relacionada com a existência de um conjunto de práticas de avaliação dificilmente consideradas como formativas ou somativas, mas que poderão ter características híbridas ou ainda não caracterizadas. O problema reside em investigá-las, em caracterizá-las e em identificar as práticas e o tipo de dados e evidências que é produzido. (FERNANDES, 2019, p. 152)

Essa continuidade pode ser observada no Quadro 12, proposto por Dolin e colaboradores (2018).

Quadro 12 – Dimensões das propostas e práticas avaliativas

Formativa < ----- > Somativa				
	Formativa Informal	Formativa Formal	Somativa Informal	Somativa Formal
Foco Principal	Quais são os próximos passos na aprendizagem?		Qual foi o alcance até o momento?	
Objetivo	Informar as próximas etapas de ensino e de aprendizagem		Monitorar o progresso diante dos planos	Registrar conquistas individuais
Como as evidências são coletadas?	Como uma parte comum das atividades de aula	Como uma parte introduzida nas atividades de aula	Como uma parte introduzida nas atividades de aula	Testes ou tarefas separadas das atividades de aula
Quem faz o julgamento?	Estudantes e Professores	Professores e Estudantes	Professores	Professores ou examinadores externos
Quais medidas são tomadas?	<i>Feedback</i> para estudantes e professores	<i>Feedback</i> para estudantes e sobre o ensino	<i>Feedback</i> para estudantes e sobre o ensino	Informar estudantes, pais, outros professores e etc.
Epíteto	Avaliação para a Aprendizagem	Correspondência ¹⁸	Nivelamento ¹⁹	Avaliação da Aprendizagem
Exemplos de modelos de Feedback	<i>Feedback</i> verbal no decorrer do processo	<i>Feedback</i> escrito no trabalho de classe	Respostas a teste informal ou questionário	Relatório sinóptico sobre o alcance em relação aos objetivos do curso

Fonte: DOLIN et al (2018, tradução nossa)

¹⁸Tradução nossa para Matching.

¹⁹Tradução nossa para Dip Stick.

Pela terceira concepção, as duas funções estão associadas a conceitos distintos, no entanto, são complementares, partilham aspectos comuns, tais como a coleta de informação do que os alunos sabem e são capazes de fazer.

[...] a AdA e ApA são complementares, ou seja, dois conceitos distintos um do outro, mas não dicotômicos que se complementam ou complementam nos seus propósitos, mas, sobretudo, no que se refere às contribuições que podem dar para apoiar/melhorar as aprendizagens dos alunos (FERNANDES, 2019, p. 152).

Segundo Fernandes (2019), do ponto de vista epistemológico, AdA e ApA não podem ser interpretadas como dimensões de um mesmo constructo. Por outro lado, considerá-las como dicotômicas talvez não faça sentido, “uma vez que, por exemplo, ambas contribuem para gerar informação fundamental para distribuir *feedback* de elevada qualidade” (p. 153). Um mesmo instrumento avaliativo pode ter propósitos diferentes, sendo utilizado na AdA ou na ApA. Dolin e colaboradores (2018) corroboram desta ideia, afirmando que

é importante reconhecer que os fatores formativos e somativo referem-se a diferentes *propósitos* de avaliação e não a diferentes tipos ou *formas* de avaliação. Como veremos adiante, dados do mesmo tipo podem, em algumas circunstâncias, ser usados de forma formativa e somativa. É o impacto da avaliação em uma instância específica e o uso feito dos dados que identificam a avaliação formativa ou somativa e não a forma dos dados (DOLIN et al, 2018, p. 53, tradução nossa).

Sobre as possíveis articulações dessas modalidades avaliativas, Santos (2019) identifica na literatura a existência de três linhas de pensamento.

A primeira seria aquela que considera não haver possibilidade de articulação, pois uma mesma prática avaliativa não pode ter, com igual grau de importância, propósitos distintos. Aqui, Santos (2019) destaca os autores Perrenoud e Vial como defensores dessa ideia. Nesse caso, os autores interpretam as avaliações formativa e somativa como oriundas de conceitos dicotômicos.

A segunda considera ser viável a utilização formativa de avaliações somativas, mas não considera a existência de uma avaliação formativa que seja utilizada em uma perspectiva somativa. Destacam-se Bennett, Foster, Noyce e Wiliam como defensores dessa linha de pensamento (SANTOS, 2019).

Taras (2005, 2009), representante da segunda linha, considera que toda avaliação é inicialmente somativa e o que se faz a seguir é o que a torna formativa. Se a avaliação

for interrompida no julgamento, ela será somativa. No entanto, é impossível fornecer um *feedback* que ajude os estudantes a atingir determinados objetivos sem que o professor tenha realizado um julgamento. Por isso, toda Avaliação Formativa é composta por uma Somativa acrescida pelo *feedback*

avaliação formativa = avaliação somativa + *feedback*

(TARAS, 2005, 2009)

A terceira seria aquela que considera a possibilidade de articulação entre duas modalidades de avaliação, uma avaliação formativa pode ter características somativas e vice-versa. Tal concepção é defendida por Fernandes (2009) e Hadji (2001).

toda avaliação, mesmo no centro da ação, tem uma dimensão cumulativa. Sempre faz o balanço das aquisições dos alunos. E toda avaliação tem – ou deveria ter, em um contexto pedagógico – uma dimensão prognóstica, no sentido que conduz – ou deveria conduzir – a um melhor ajuste ensino/aprendizagem. (HADJI, 2001, p. 19-20)

Portanto, considerando a segunda, ou até mesmo, a terceira possibilidade, constata-se que a prova é um instrumento amplamente aceito e culturalmente implementado na escola e na universidade. Na impossibilidade da implementação da *Dupla Diversificação Avaliativa*, pergunta-se: Como o *feedback* poderia ser utilizado em uma prova para atribuir a esta, também, uma função formativa?

Como discutido no capítulo 4, uma possibilidade seria implementar variantes das provas tradicionais. A prova em fases é uma excelente alternativa à prova tradicional que oferece aos professores a possibilidade de regular o ensino e a aprendizagem, oferecendo *feedbacks formativos*. No entanto, a limitação de tempo para cumprir o conteúdo programático pode ser um problema para essa implementação.

Outra possibilidade é a utilização de mais *feedbacks formativos* nas provas tradicionais. Ao invés de habituais notas e pontuações comumente utilizados, a implementação de comentários sobre o que errou e como fazer corretamente podem oferecer às provas tradicionais características mais formativas.

Há dois aspectos importantes que devem ser pontuados. Primeiramente, não pretendemos reduzir a avaliação formativa a uma avaliação somativa acrescida de *feedback*. A avaliação formativa é um conceito amplo e complexo. O foco aqui é discutir meios para atribuir às provas e testes, instrumentos avaliativos comuns em Matemática,

um caráter mais formativo de um modo pragmático. Analogamente, a redução dos *feedbacks* avaliativos aos *feedbacks* escritos em um teste é uma visão limitadora.

No entanto, é possível oferecer uma contribuição à “fórmula” de avaliação de Taras (2005, 2009). Partindo de uma avaliação somativa, quanto mais utilizarmos *feedbacks* formativos, quanto mais utilizamos esse teste para regular ensino e aprendizagem, mais transformamos esse ‘instrumento’ avaliativo, atribuindo-lhe características formativas.

avaliação somativa + feedbacks formativos → avaliação formativa

A visão dicotômica entre as avaliações somativa e formativa representa um obstáculo para a utilização da avaliação voltada para a aprendizagem. Neste sentido, oferecer aos professores uma “formula” capaz de atribuir às provas e testes, tão incorporados na prática docente, características formativas, pode ser positivo.

CAPÍTULO 7 - CONSIDERAÇÕES FINAIS



*É preciso ter esperança,
mas ter esperança do verbo esperançar;
porque tem gente que tem esperança do verbo esperar.
E esperança do verbo esperar não é esperança,
é espera.
Esperançar é se levantar,
esperançar é ir atrás,
esperançar é construir,
esperançar é não desistir!
Esperançar é levar adiante,
esperançar é juntar-se com outros para fazer de outro modo.*
Paulo Freire

7.1 Reflexões sobre a coleta

Depois das sugestões da banca na qualificação, a tese foi reconstruída. Essa reconstrução partiu das novas leituras às novas interpretações. A Análise Narrativa foi, nesse sentido, um elemento importante nessa reconstrução do texto e para o meu reposicionamento, enquanto pesquisador. Na pesquisa pragmática, em geral, busca-se encontrar a verdade. Uma descoberta! Comprovar algo, validando uma teoria. Nos estudos realizados antes da qualificação, ainda influenciado por essa visão, busquei identificar/comprovar divergências entre corretores de questões discursivas. Em seguida, consegui identificar a existência de uma relação entre características do *feedback* oferecido pelos respondentes aos instrumentos de coleta e as pontuações por eles atribuídas a essas mesmas questões. Os resultados indicaram que os professores que atribuíam melhores pontuações também ofereciam melhores *feedbacks* (faço uma breve discussão sobre esse resultado no capítulo 5).

Com a ressignificação da tese realizada após a qualificação, essa linha de tratamento perdeu o sentido. Uma tese que defenda o rompimento de classificações aos estudantes não pode ser construída a partir da classificação dos professores. Em substituição à comprovação da conexão entre os professores mais aptos a desenvolver bons *feedbacks* e aqueles que atribuem melhores pontuações, foi desenvolvida uma discussão sobre a construção de *feedbacks* mais pedagógicos, mais voltados à aprendizagem, que denominei de *formativos*. Este é um ponto em que vale a pena investirmos enquanto comunidade acadêmica. Um bom *feedback* é um elemento do ensino com grande potencial para o desenvolvimento das aprendizagens (BLACK; WILIAM, 1998, WILIAM, 2011; DOLIN et al 2018). O poder do *feedback* para o ensino e a aprendizagem e as características de um *feedback formativo* deveriam ser inseridos nos cursos de formação, inicial e continuada, de professores, deveriam ser desenvolvidos em novas e amplas pesquisa sobre o tema.

Retomando a discussão sobre a reconstrução da tese. A Análise Narrativa me ajudou na construção de um novo olhar, mais amplo, mais aberto às diferenças. Uma forma de atuar politicamente contra discursos hegemônicos, como discutido na tese de Silva (2020). O foco se deslocou da identificação das limitações docentes, formados e em formação, para a compreensão dessas limitações. Não se trata de comprovar, e sim, prioritariamente, de compreender características relevantes apresentadas pelos corretores, valorizando os saberes e não saberes ali presentes. Sejam eles acadêmicos, experenciais, culturais, sociais e históricos.

A Análise Narrativa oferece fundamentos teórico-metodológicos para reconhecer, também, que os dados obtidos em uma pesquisa como esta são analisados sob o ponto de vista do pesquisador. Esta é a minha análise, a minha leitura. O meu olhar sobre o que foi escrito pelo outro. Dessa forma, os resultados constituem uma interpretação da verdade, e não a verdade em termos absolutos. Uma interpretação científica, pois dialoga com a literatura e se utiliza, neste caso, de elementos da Análise Pragmática para organizar e tratar os dados.

‘A prova é uma foto e a avaliação é um filme’ é uma analogia que acho uma boa representação que diferencia o instrumento prova do processo avaliativo. Algo similar pode ser utilizado em relação à *folha de correções*, nosso instrumento de coleta. A *folha de correções* forneceu imagens que refletem a compreensão dos respondentes sobre a

correção de questões, sobre a avaliação, sobre o erro e o feedback. Não a compreensão em sua totalidade, mas parte dela. A parte que foi escrita. Muitas ideias não foram transcritas ao papel, outras foram modificadas para atender ao que eles acreditavam ser a resposta esperada pelo pesquisador. Algo similar ao famoso problema da idade do capitão²⁰. Por isso, acredito que tratar esses dados para descobrir a verdade em uma perspectiva pragmática seria limitado e inadequado.

Às vezes, as respostas e as falas de estudantes e professores se transformam em dados que categorizamos e analisamos. Como se essas respostas representassem ‘a verdade absoluta’ desses respondentes, independente das escolhas do pesquisador. Nessa visão científica (pragmática), a pesquisa é isenta e neutra. A Análise Narrativa permite interpretar essas respostas como uma ‘verdade relativa’, temporal, histórica, social e política (GALVÃO, 2005). Interpretada sob uma perspectiva científica, amparada por estudos da área, mas também situada em um contexto de não neutralidade.

A minha escolha inicial foi utilizar a Análise Pragmática, à qual já estava mais acostumado. No entanto, a Análise Narrativa tornou-se cada vez mais presente à minha interpretação, à minha análise. De fato, utilizei duas formas de pensar, talvez, esta utilização seja uma contribuição metodológica relevante ao campo, reforçando a necessidade do desenvolvimento consciente do caráter dialógico da pesquisa em Educação Matemática, principalmente aquela que utiliza a produção de estudantes e professores, seja de forma oral ou escrita.

7.2 Reflexões sobre os resultados

A investigação realizada com professores atuantes e em formação sobre correção de questões abertas de Matemática, naturalmente, pode ser classificada como uma experiência sobre a realidade e não vivenciada na realidade. Existem fatores presentes na sala de aula que poderiam interferir nas respostas desses profissionais. Outros aspectos poderiam ser identificados, alguns poderiam não ser. No entanto, acreditamos que esses

²⁰ Em um estudo desenvolvido por uma equipe de professores do I.R.E.M. de Grenoble (1980) propôs a alunos do ensino fundamental 1 que respondessem ao seguinte problema: Num barco existem 26 carneiros e 10 cabras. Qual a idade do Capitão? Dos 97 alunos participantes, 75 deram respostas a partir da combinação dos números presentes no enunciado.

resultados representem a realidade sob um ponto de vista. Uma realidade situada historicamente e analisada a partir das minhas escolhas enquanto pesquisador.

De modo análogo, os saberes e não saberes dos corretores que participaram deste estudo influenciaram suas pontuações e seus comentários. Os aspectos relacionados à avaliação que emergiram da interpretação dos erros, apresentados no capítulo 5, não devem ser considerados como uma generalização do modo com que os professores corrigem suas provas. Esses aspectos devem ser interpretados como elementos presentes na prática avaliativa dos professores participantes, e que, por analogia, podem ser considerados elementos úteis à compreensão das práticas avaliativas, mais especificamente, da correção de itens abertos de Matemática.

Retornando à pergunta de tese:

Quais são os aspectos relacionados à avaliação que emergem da interpretação dos erros dos estudantes na correção de itens, e como estes aspectos influenciam a avaliação em Matemática?

A partir da produção dos corretores, participantes dessa pesquisa, foram identificados os seguintes aspectos relacionados à correção de questões abertas de Matemáticas, discutidos anteriormente no capítulo 5:

- (1) Critérios de correção → ausentes ou vagos;
- (2) A posição em que o erro ocorre → influencia a correção e a pontuação;
- (3) A matriz curricular → desconectada com a atividade de correção;
- (4) Ênfase na resposta final;
- (5) Desvalorização da interpretação do problema.

Além destes aspectos, identificamos

- (6) limitações e potencialidades na construção de *feedbacks* formativos.

A falta de critérios bem definidos utilizados para conduzir a correção de itens, aspectos identificados nesta pesquisa, podem representar elementos presentes em nossa cultura escolar, sustentada no, e pelo, exame, características de uma possível falha em nossa formação profissional. Atrevo-me afirmar, que raros cursos de licenciatura em Matemática no Brasil promovem debates relacionados à correção de questões de Matemática, às subjetividades presentes e ao tratamento do erro. Nesse sentido, o estudo

de critérios e do tratamento dado ao erro no ensino de Matemática, como discutido no capítulo 4, pode oferecer significativas contribuições à formação de professores. Vale destacar que alguns dos respondentes identificaram essas possíveis falhas em sua formação, manifestando em seus comentários dificuldades em estabelecer/utilizar critérios, de forma segura, para corrigir as questões.

Essa carência na formação de professores de Matemática pode ser identificada ao analisarmos a desvalorização (em termos de pontuação nos itens) das estratégias de resolução e a despreocupação entre as habilidades apresentadas nas respostas e o ano de escolarização. Essa desvalorização pode denotar uma possível desconexão da atividade docente com aspectos curriculares em um sentido mais amplo. Poucos professores demonstraram incômodo com a solução do estudante D, inadequada à faixa etária e de escolaridade. Mesmo dentre aqueles que demonstraram por escrito, esse incômodo não foi refletido em termos de pontuação. O maior incômodo dos corretores participantes era com o erro! Como alertam Borasi (1998) e Cury(2007), nós professores, geralmente, execramos o erro!

Outro resultado que merece destaque se refere ao currículo. Falta conexão com outras disciplinas! Falta, até mesmo, conexão interna, entre tópicos da mesma disciplina. A Matemática é diluída em tópicos fragmentados, como se estes fossem desconectados e independentes na própria Matemática. Os participantes ignoraram o fato de que o método de solução por contagem utilizado pelo estudante D indicava um atraso no domínio de habilidades necessárias àquele ano de escolarização. Essa despreocupação, refletida através da alta pontuação média no item D, destacou o outro lado dessa moeda.

Há, ainda, outro ponto de vista sobre essa questão. Os corretores valorizaram as respostas alternativas. Algo positivo! As soluções D e E obtiveram, como discutido anteriormente, pontuações médias altas. No entanto, há uma dúvida do real motivo dessa valorização. Foram as estratégias alternativas que, de fato, foram valorizadas ou a resposta correta? Ou seja, será que essa pontuação alta ocorreu pelo fato de que os estudantes D e E chegaram à resposta final esperada? Aparentemente, as soluções alternativas não se configuram um problema para o docente, desde que atinja o seu objetivo: a resposta correta.

A ênfase na resposta final e a pouca importância dada, tanto à interpretação, quanto à elaboração de estratégias, foram aspectos identificados em outras questões presentes na folha de repostas, como por exemplo, na maior pontuação média dada à solução B se comparada à solução C. Pesquisas futuras poderiam investigar o comportamento de corretores em soluções alternativas parcialmente corretas. Se a ênfase está, de fato, na resposta final, e não na valorização de soluções alternativas, a pontuação tenderia a ser baixa. Eu acredito que seriam igualmente pequenas. Suspeito que a valorização da resposta final seja sustentada pelo modo como, culturalmente, valorizamos muito, e negativamente, o erro!

Outros corretores demonstraram reconhecer suas limitações para realizar a tarefa. Diante de uma atividade como essa, alguns professores e licenciandos conseguiram reconhecer a complexidade que é analisar a produção de um aluno e construir feedbacks individuais e escritos. A avaliação é ainda pouco estudada nas universidades e pouco estudada pelos professores formadores de professores. Não é reconhecida em sua complexidade. Avaliamos, normalmente, como fomos avaliados, através da perspectiva reduzida e limitada do exame. A avaliação é comumente interpretada como uma medida, uma nota (LUCKESI, 1998; MORGAN, 2000; HADJI, 2001). Isso não é o suficiente. Precisamos formar professores mais aptos a avaliar seus estudantes. O tema da avaliação precisa estar presente tanto nas aulas das disciplinas pedagógicas, quanto nas aulas de Cálculo, Geometria e Álgebra.

O erro em Matemática, talvez, seja ainda pouco estudado na universidade, mesmo sendo um elemento chave na avaliação, no ensino e na aprendizagem. Um erro pode representar um lapso, uma falha processual (CARMO, 2010). Um erro pode representar um saber em construção, em desenvolvimento. Algo que precisa ser melhor compreendido, que deve ser investigado. Tanto no caso de um lapso ou de um conhecimento em construção, o erro representa um saber em potencial (BORASI, 1985; CURY, 2007, 2012). O modo como interpretamos o erro impacta no modo como avaliamos, como pontuamos e como utilizamos essa avaliação como um agente da aprendizagem. O modo como interpretamos o erro influencia a qualidade do *feedback* e o desenvolvimento de avaliações voltadas às aprendizagens. O *feedback* precisa ser estudado nas licenciaturas e na pós-graduação. Precisamos compreender, analisar,

investigar mais e melhorar as respostas dos estudantes. Para, então elaborarmos com maior propriedade *feedbacks* formativos.

Há uma cultura da resolução de exercícios nas aulas de Matemática. Algo identificado por Skovsmose (2000) como um Paradigma do Exercício. As aulas de Matemática têm como objetivo ensinar os estudantes a resolver exercícios. Saber Matemática é saber resolver exercícios. Alguns corretores sugeriram como uma das medidas, outros como a única medida, a aplicação de lista de exercícios como forma de reduzir as dificuldades dos estudantes.

Consequentemente, o Paradigma do Exercício está em conformidade com a cultura do exame. Se aprender matemática é um sinônimo de conseguir resolver exercícios, então, aqueles que sabem matemática têm, geralmente, um bom desempenho nos exames e vice-versa. Um professor, com alguma experiência, sabe que isso nem sempre é verdade. Existe aquele aluno que participa da aula, faz as tarefas propostas, mas não consegue ter um bom desempenho nas provas escolares.

As características histórico-culturais do ensino de Matemática e as carências na formação em avaliação foram perceptíveis neste estudo. Alguns professores, provavelmente, interpretam a correção de questões avaliativas muito mais como uma tarefa classificatória, do que como um processo pedagógico, denotando aspectos positivistas, possivelmente, oriundos de sua formação acadêmica. A maioria dos licenciandos e professores que participaram deste estudo, provavelmente, não tiveram em sua formação orientações de como proceder na correção de questões ou, pelo menos, nunca foram convidados a refletir sobre. Nunca estudaram aspectos relacionados ao *feedback*. Não conhecem a prova em fases ou a prova com cola. Acreditam que autoavaliação é apenas pedir aos seus estudantes que atribuam uma nota ao próprio desempenho. Talvez, não tiveram a oportunidade de refletir sobre os significados dos erros em Matemática. Em toda a sua formação, o erro foi interpretado como um elemento negativo. Algo que deve ser apagado no caderno. Algo que deve ser evitado.

No entanto, esses aspectos não refletem a totalidade dos respondentes. Nas conchas, encontramos pérolas.

7.3 Sobre as pérolas

Como discutido anteriormente, somente oferecer conceitos e notas aos estudantes como única forma de *feedback* não oferece contribuições relevantes para a aprendizagem (DOLIN et al, 2018). Por outro lado, devolutivas que levariam estudantes a refletir sobre sua produção, identificar erros e impulsioná-los a elaborar novas estratégias, como algumas das elaboradas pelos respondentes, têm por finalidade orientar as aprendizagens. Como elucida Pinto (2019), estes *feedbacks* apresentam características da avaliação formativa que, para Black e Wiliam (1988), configuram-se como elementos de grande impacto na aprendizagem dos estudantes, principalmente aqueles com maiores dificuldades e limitações. Alguns docentes (atuantes e em formação) conseguiram elaborar *feedbacks* com grande potencial pedagógico (*feedbacks* *formativos*).

Alguns respondentes conseguiram elaborar critérios interessantes de correção, similares àqueles discutidos por Abrantes (1995), Nunes Neto e Mendonza (2021). Alguns corretores buscaram valorizar a produção dos estudantes, rompendo a visão dicotômica, identificada por Esteban (2002), entre o certo e errado e entre o saber e o não saber. Um corretor, em uma atitude digna de ser classificada como insubordinada e criativa (D'AMBROSIO; LOPES, 2015), não atribuiu nota à **solução B**, entendendo que havia a necessidade de investigar os saberes presentes naquela solução. A postura investigativa em relação à produção dos estudantes defendida por diversos autores (ESTEBAN, 2002; BURIASCO, 2004; VIOLA dos SANTOS, BURIASCO, CIANI, 2008) foi encontrada nas respostas de alguns dos participantes, como mostram os exemplos de *feedbacks* a seguir.

Voltaremos a pensar juntos para entender o que levou a esse resultado (Corretor 48 ao estudante C).

Gostei muito do seu raciocínio, mas não consegui identificar se você consegue efetuar as operações. Tente refazer o problema com 73 carrinhos em cada caixa (Corretor 1 ao estudante D).

Estes exemplos são pérolas, pois nos permitem supor que há, para esses corretores, uma distinção entre o processo avaliativo e o instrumento prova. A avaliação não se encerra na prova. Essa distinção, explicitada nos *feedbacks* acima, torna a avaliação mais pedagógica, pois está mais voltada à aprendizagem. Por outro lado, também torna a

avaliação mais justa em termos de pontuação, pois alguns corretores não se limitaram a estabelecer pontuação sobre o que foi escrito, entendendo que havia a necessidade de investigar mais os saberes presentes nas soluções.

Perguntaria como ele chegou ao resultado, e que ele deveria nas provas seguintes colocar a resolução (feedback). Se eu solicitasse que o aluno fizesse as operações, nesse caso eu tiraria pontos (critério/justificativa). (Corretor 25 ao estudante B).

Chamaria esse aluno na minha mesa para entregar a avaliação e questionaria se compreendeu o problema, leria e perguntaria se seria possível alguma complementação na sua resposta. Caso positivo, se o aluno chegar à resposta desejada, aumentaria a sua nota, caso contrário, manteria a nota (0,5) e explicaria a ele depois de dar o feedback a todos da turma (feedback). (Corretor 31 ao estudante F).

7.4 Últimos passos e reflexões da travessia

A avaliação promove a regulação do ensino e da aprendizagem. É possível fornecer ao aluno uma oportunidade de entender seu erro. Proporcionar mecanismos para corrigi-lo, contribuindo, assim, para a aprendizagem desse estudante. A avaliação não pode ser reduzida à pontuação em um exame. Os exames como instrumentos avaliativos são limitados em relação ao tratamento do erro e à utilização do tempo.

Na perspectiva somativa da avaliação através dos exames, a aprendizagem deve ocorrer no momento anterior ao exame. Muitos alunos têm o hábito de estudar apenas na véspera de uma prova. A maior parte do que aprendemos na tarefa da realização do exame ocorre no estudo para o exame. Durante o exame, podemos ter esse ou aquele insight, então aprendemos um pouco, caso contrário, nada aprendemos. No modelo de prova tradicional, a aprendizagem posterior à prova também é pequena. Escrever o gabarito na lousa, devolver a prova com *feedbacks* em forma de x (errado), c (certo) ou de pontuações pouco contribuem para a aprendizagem. Nesse sentido, a prova tradicional é limitada em relação ao tempo de aprendizagem, pois sua aprendizagem situa-se prioritariamente no passado.

A prova é, geralmente, limitada também no que se refere ao tratamento do erro. Se os *feedbacks* não permitirem o reconhecimento do que o estudante errou e de como seria o correto, a aprendizagem ocorrerá de forma limitada. Se o fim da avaliação é a aprendizagem, é necessário que o estudante possa aprender com os erros também durante

os processos avaliativos. Os exames, as provas tradicionais, geralmente não oferecem essa possibilidade. Há a regulação da aprendizagem, mas menor que na avaliação formativa. Além disso, se a responsabilidade de um bom desempenho nos exames for atribuída somente aos estudantes, não haverá regulação do ensino.

Como discutido no capítulo 6, as práticas avaliativas devem ser redirecionadas a uma cultura de desenvolvimento de *feedbacks* mais direcionados à aprendizagem. Devemos investir em instrumentos avaliativos em que esses *feedbacks* sejam utilizados pragmaticamente, como a prova em fases, por exemplo. Na impossibilidade da utilização de instrumentos alternativos, que pelo menos, possamos desenvolver maiores e melhores habilidades de construção de *feedbacks* *formativos* para as provas tradicionais.

Se por um lado, a ressignificação do erro pode oferecer contribuições à regulação das aprendizagens, por outro, pode oferecer contribuições tão relevantes à regulação do ensino. A análise das soluções dos alunos pode fornecer ao professor orientação sobre o que os alunos não aprenderam e sobre como(re)direcionar estratégias de ensino para suprir carências na aprendizagem. Quais estratégias foram mal sucedidas? Por quê? São reflexões que fazem parte de qualquer profissão.

A visão de que possa existir um bom ensino sem que ocorra uma boa aprendizagem é falaciosa. Uma aula só é boa quando os alunos aprendem. Um exame só é bom se os alunos aprenderam com, para e por causa dele. Uma avaliação só é boa se avaliar o ensino, além da aprendizagem. O modo como o professor avalia está intrinsicamente conectado ao modo como esse docente ensina, e provavelmente, como seu aluno aprende. A avaliação como elemento estruturante do ensino pode, conseqüentemente, ser um rico momento de aprendizagem, devendo ser compreendida como uma ação dialógica.

“A avaliação não é o ato pelo qual A avalia B. É o ato por meio do qual A e B avaliam juntos uma prática, seu desenvolvimento, os obstáculos encontrados ou os erros e equívocos porventura cometidos” (FREIRE, 1981, p. 21).

REFERÊNCIAS

- ALRØ, H; SKOVSMOSE, O. **Diálogo e aprendizagem em Educação Matemática**. 2ed. Tradução: Orlando Figueiredo. Belo Horizonte, Autêntica Editora, 2010.
- ANTUNES, T. P.; MENDES, M. T. Desenvolvimento profissional de um professor ao (re) elaborar uma prova escrita de Matemática. *Amazônia: Revista de Educação em Ciências e Matemáticas*, v. 14, n. 31, p. 22-38, 2018.
- BARLOW, M. **Avaliação escolar**: mitos e realidades. Porto Alegre: Artmed, 2006.
- BARRIGA, A. D. Uma polêmica em relação ao exame. *In*: ESTEBAN, M. T. (org.). **Avaliação**: uma prática em busca de novos sentidos. 4 ed. Rio de Janeiro: DP&A, 2002.
- BARREIRA, C. M. F. Concepções e práticas de avaliação formativa e sua relação com os processos de ensino de aprendizagem. *In*: ORTIGÃO, M. I. R. et al (org.). **Avaliar para aprender no Brasil e em Portugal**: perspectivas teóricas, práticas e de desenvolvimento. Curitiba: CRV, 2019. v. 1, p. 191-218.
- BARRY, V. J. Using Descriptive Feedback in a Sixth Grade Mathematics Classroom. **Action Research Projects**. 9. Nebraska-Lincoln. 2008. Disponível em: <https://digitalcommons.unl.edu/mathmidactionresearch/9>. Acesso em 10 mar 2020.
- BLACK, P., HARRISON, C., LEE, C., MARSHALL, B., WILLIAM, D. **Assessment for learning**: Putting it into practice. Maidenhead, UK: Open University Press. 2003.
- BLACK, P.; WILLIAM, D. Inside the Black Box: Raising Standards through Classroom Assessment. **The Phi Delta Kappan**, v. 80, n. 2, p. 139-148. 1998.
- BOLÍVAR BOTÍA, A. "¿De nobis ipsis silemus?": Epistemología de la investigación biográfico-narrativa en educación. **Revista electrónica de investigación educativa**, v. 4, n. 1, p. 01-26, 2002.
- BORASI, R. Using Errors as Springboards for the Learning of Mathematics: An Introduction. **Focus on learning problems in Mathematics**, v. 7, p. 1-14, 1985.
- BURIASCO, R. L. C. **Avaliação em Matemática**: um estudo das respostas de alunos e professores. 1999. 238 f. Tese (Doutorado em Educação) –Universidade Estadual Paulista, Marília, 1999.
- _____. Algumas considerações sobre avaliação educacional. **Estudos em avaliação educacional**, n. 22, p. 155-178, 2000.
- BURIASCO, R. L. C.; FERREIRA, P. E. A.; CIANI, A. B. Avaliação como prática de investigação: alguns apontamentos. **Bolema**, Rio Claro, v. 22, n. 33, p.69-96. 2009.
- BRASIL. **Parâmetros Curriculares Nacionais (PCN)**. Introdução. Ensino Fundamental. Brasília: MEC/SEF, 1998.
- _____. **Base Nacional Comum Curricular (BNCC)**. Educação é a Base. Brasília: MEC/CONSED/UNDIME, 2017.
- BRUNER, J. Life as narrative. **Social research**, v. 54, n.1, p.11-32, 1987.

BRUNER, J. The Narrative Construction of Reality. **Critical Inquiry**, v. 18, n.1, p.1-21. 1991. Disponível em <http://www.jstor.org/stable/1343711>. Acesso em 3 mar 2021.

CARMO, J. S. Produção de erros no ensino e na aprendizagem: implicações para a interação professor-aluno. *In*: MIZUKAMI, M. G. N.; REALI, A. M. M. R. (org.) **Aprendizagem profissional da docência: saberes, contextos e práticas**. São Carlos, SP: EDUFSCar/INEP/COMPED, 2010.p. 211-227.

CURY, H. N. **Análise de Erros: O que podemos aprender com as respostas dos alunos**. Tendências em Educação Matemática. Belo Horizonte, MG: Autêntica. 2007.

CURY, H. N. Pesquisas em ensino de ciências e matemática, relacionadas com erros: uma investigação sobre seus objetivos. **Educação Matemática Pesquisa: Revista do Programa de Estudos Pós-Graduados em Educação Matemática**, v. 14, n. 2, p. 237-256, 2012.

D'AMBROSIO, B. S. A subversão responsável na constituição do educador matemático. *In*: ENCUENTRO COLOMBIANO DE MATEMÁTICA EDUCATIVA, 16., 2015, Bogotá. **Anais [...]**. Bogotá: Asociación Colombiana de Matemática Educativa, 2015.

D'AMBROSIO, B. S.; LOPES, C. E. Insubordinação Criativa: um convite à reinvenção do educador matemático. **Bolema**, Rio Claro, v. 29, n. 51, p. 01-17. 2015.

DOCIMOLOGIA. *In*: WIKIPÉDIA: a enciclopédia livre. Wikimedia, 2020. Disponível em: <https://pt.wikipedia.org/wiki/Docimologia>. Acesso em: 17 ago. 2020

DOLIN, J.; BLACK, P.; HARLEN, W.; TIBERGHEN, A. Exploring relations between formative and summative assessment. *In*: DOLIN, J.; EVANS, R. (org.). **Transforming assessment**. Cham: Springer, 2018. v. 4, p. 53 – 80.

FAZIO, L.; SIEGLER, R. S. Teaching fractions. **Educational Practices Series**. Geneva. International Academy of Education -International Bureau of Education. v. 22, 2011.

FERNANDES, D. **Avaliar para aprender: fundamentos, práticas e políticas**. São Paulo: Editora Unesp, 2009.

_____. Para um enquadramento teórico da avaliação formativa e da avaliação sumativa das aprendizagens escolares. *In*: ORTIGÃO, M. I. R.et al (org.). **Avaliar para aprender no Brasil e em Portugal: perspectivas teóricas, práticas e de desenvolvimento**. Curitiba: CRV, 2019. v. 1, p. 139 – 164

FIORENTINI, D.; LORENZATO, S. **Investigação em Educação Matemática: percursos teóricos e metodológicos**. 3ed. Campinas: Autores Associados, 2012.

FISCHER, M. C. B. Os formadores de professores de Matemática e suas práticas avaliativas. *In*: VALENTE, W. R. (org.). **Avaliação em Matemática: história e perspectivas atuais**. Campinas: Papyrus, 2008. p. 75 -100.

FORSTER, C. **A utilização da prova-escrita-com-cola como recurso à aprendizagem**. Dissertação de Mestrado do Programa de Pós-graduação em Ensino de Ciências e Educação Matemática. UEL, Londrina, 122f. 2016.

FREIRE, P. **Ação cultural para a liberdade**. 5.ed. Rio de Janeiro: Paz e Terra, 1981.

- GUBA, E. G.; LINCOLN, Y. S. **Avaliação de quarta geração**. Campinas: Editora da Unicamp, 2011.
- GRIS, G.; PALOMBARINI, L. S.; CARMO, J. S. Uma Revisão Sistemática de Variáveis Relevantes na produção de erros em Matemática. **Bolema**, v. 33, n. 64, p. 649-671, 2019.
- HENRY, M. **Didactique des Mathématiques**: une présentation de la didactique en vue de la formation des enseignants. Les Publications de l'IREM de Besançon, 1991. p. 01-10.
- HADJI, C. **Avaliação desmistificada**. Porto Alegre: Artmed Editora, 2001. 136 p.
- HATTIE, J. Influences on student learning. *In*: INAUGURAL LECTURE. University of Auckland, New Zealand. August, 1999. Disponível em: <http://geoffpetty.com/wp-content/uploads/2012/12/Influencesonstudent2C683.pdf>. Acesso em 8 jun 2020.
- HATTIE, J.; TIMPERLEY, H. The power of feedback. **Review of Educational Research**, v. 77, n. 1, p. 81-112, 2007.
- LAM, R. Formative use of summative tests: Using test preparation to promote performance and self-regulation. **The Asia-Pacific Education Researcher**, v. 22, n. 1, p. 69-78, 2013.
- LIMA, M. E. C. C.; GERALDI, C. M. G.; GERALDI, J. W. O trabalho com narrativas na investigação em Educação. **Educação em revista**, v. 31, p. 17-44, 2015.
- LUCKESI, C. C. **Avaliação da aprendizagem escolar**: sendas percorridas. 1992. Tese (Doutorado em Filosofia da Educação). Pontifícia Universidade Católica de São Paulo, São Paulo, 1992.
- _____ **Verificação ou avaliação**: o que pratica a escola. Série Ideias, n. 8, p. 71-80, 1998.
- ELBAZ-LUWISCH, F. O ensino e a identidade narrativa. **Revista de Educação**, Lisboa, v. 11, n. 2, p. 21-33, 2002.
- MACEDO, M. F.; LIMA, A. M. Revolvendo o passado da avaliação educacional e algumas repercussões na escola. **Revista Teias**, v. 14, n. 32, p. 155-171, 2013.
- MARZANO, R.; PICKERING, D.; POLLOCK, J. **Classroom instruction that works**: Research based strategies for increasing student achievement. Alexandria: Association for Supervision and Curriculum Development, 2001.
- MERLE, P. **Les pratiques d'évaluation scolaire**: historique, difficultés, perspectives. Paris: Presses Universitaires de France, 2018.
- MORAES, R. Análise de conteúdo. **Revista Educação**, Porto Alegre, v. 22, n. 37, p. 7-32, 1999.
- MORAES, C. A. P. **Avaliação em Matemática**: pontos de vista dos sujeitos envolvidos na Educação básica. Jundiaí: Paco Editorial, 2012.
- MORGAN, C. Better assessment in mathematics education? A social perspective. In BOALER, J.(org.). **Multiple Perspectives on Mathematics Teaching and Learning**. Westport, CT: Ablex Publishing, p. 225-242, 2000.

NASSER, L.; LIMA, D. O.; VAZ; R. F. N.; MENEZES, F. Insubordinate Practices in Mathematics Evaluation. *RIPEM*, v. 9, n. 3, 2019, p. 114-128.

NICOL, D. J.; MACFARLANE-DICK, D. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, v. 31, n. 2, p. 199-218, 2006.

NOIZET, G.; CAVERNI, J-P. **Psicologia da avaliação escolar**. Coimbra: Coimbra Editora, 1985.

NUNES NETO, R. **A atividade de situações problema na aprendizagem do conteúdo de fração fundamentada na teoria de formação por etapas das ações mentais de Galperin com estudantes do 5º ano da Escola Municipal Laucides Inácio de Oliveira**. Dissertação (Mestrado em Ensino de Ciências) –UERR, Boa Vista, 182f. 2015.

NUNES NETO, R. MENDOZA, H. J. G. A atividade de situações-problema na aprendizagem do conteúdo de fração fundamentada na teoria de Galperin *In*: BATISTA, A. J.; VAZ; R; F; N.; SANTOS; S. A. (org.) **Aplicações e reflexões da resolução de problemas para o ensino e aprendizagem de matemática**. Boa Vista: Educitec, 2021. p. 106-135

PERRENOUD, P. **Avaliação: da excelência à regulação das aprendizagens - entre duas lógicas**. Porto Alegre: Artes Médicas Sul, 1999.

PESQUISA. *In*: ETIMOLOGIA: origem e conceito. 2020. Disponível em: <<https://etimologia.com.br/pesquisa/>>. Acesso em 22 ago. 2020

PESSOA, C. Contrato didático: sua influência na interação social e na resolução de problemas. *In*: ENCONTRO NACIONAL DE EDUCAÇÃO MATEMÁTICA, 8. Recife: Universidade Federal de Pernambuco, 2004.

PINTO, J. A avaliação em Educação: da linearidade dos usos à complexidade das práticas. *In*: AMANTE, L; OLIVEIRA, I. (org.) **Avaliação das Aprendizagens: Perspectivas, Contextos e Práticas**. Lisboa: Universidade Aberta, 2016

_____. Avaliação Formativa: uma prática para a aprendizagem. *In*: ORTIGÃO, M. I. R. et al (org.). **Avaliar para aprender no Brasil e em Portugal**: perspectivas teóricas, práticas e de desenvolvimento. Curitiba: CRV, 2019. v. 1, p. 19 – 44

PINTO, N. B. **O erro como estratégia didática**: Estudo do erro no ensino da Matemática elementar. 2. ed. Campinas: Papirus, 2000

PIRES, M. N. M.; BURIASCO; R. L. C. Prova em fases: instrumento para aprender. *In*: SEMINÁRIO INTERNACIONAL DE PESQUISA EM EDUCAÇÃO MATEMÁTICA, 5., 2012. Petrópolis. **Anais[...]**. Petrópolis, 2012.

PIRES, M. N. M.; BURIASCO, R. L.C. Professores dos anos iniciais, a prova em fases e a possibilidade de aprender. **Zetetike**, n. 25. v. 3., p 474-495. 2017.

POLYA, G. **A arte de resolver problemas**: um novo aspecto do método matemático. Ed. Interciência, 1978.

POLKINGHORNE, D. E. Narrative configuration in qualitative analysis. **International Journal of Qualitative Studies in Education**, n. 8. v. 1, p. 5-23.1995

POLKINGHORNE, D. E. Validity issues in narrative research. **Qualitative inquiry**, v. 13, n. 4, p. 471-486, 2007.

RAPHAEL, H. S. **Avaliação escolar**: em busca de sua compreensão. São Paulo: Brasiliense, 1998.

RIBEIRO, V. P.; GODOY, E. V.; ROLKOUSKI, E. Análise de erros: um estudo com ingressantes de cursos de graduação. **Revista BOEM**, v. 8, n. 16, p. 112-133, 2020.

SALES, E. R.; CARMO, J. S.; HENKLAIN, M. H. O. Produção de erros em prova de Matemática: Critérios de correção e atribuição de notas por professores do Ensino Fundamental. *In*: SANTOS, E. M.; LAUTERT, S. L. (org.). **Diálogos sobre o ensino, aprendizagem e a formação de professores**: contribuições da Psicologia da Educação Matemática. Rio de Janeiro, RJ: Autografia, 2020.

SANTIAGO, M. C.; SANTOS, M. P.; MELO, S. C. Inclusão em Educação: processos de avaliação em questão. **Ensaio**: avaliação e políticas públicas em Educação, Rio de Janeiro, v. 25, n.96, p. 632-651, set. 2017.

SANTOS, L. Que critérios de qualidade para uma avaliação formativa? *In*: FERNANES, D. (org). **Avaliação em Educação**: olhares sobre uma prática social incontornável. Pinhais: Editora Melo, 2011. 155-165.

_____. Reflexões em torno da avaliação pedagógica. *In*: ORTIGÃO, M. I. R. et al (org). **Avaliar para aprender no Brasil e em Portugal**: perspectivas teóricas, práticas e de desenvolvimento. Curitiba: CRV, 2019. v. 1, p. 165-190.

SANTOS, L., PINTO, J. Ensino de conteúdos escolares: A avaliação como Fator estruturante. *In*: VEIGA, F. H. (org.). **O Ensino como fator de envolvimento numa escola para todos**. Lisboa: Climepsi Editores, 2018. p. 503-539.

SANTOS, J. R. V.; PASSOS, A. Q.; MARQUES, A. F. Análise de Erros e maneiras de lidar: um olhar a respeito da produção do GEPEMA. *In*: BURIASCO, R. L. C. **GEPEMA**: espaço e contexto de aprendizagem. 1ed. Curitiba, PR: CRV, 2014.

SCRIVEN, M. The methodology of evaluation. *In*: TYLER, R. W. (org) **Perspective on Curriculum Evaluation**. Chicago: Rand Mc Nally, 1967.

SILVA, M. S. **O que podem as narrativas na Educação Matemática brasileira**. 2020. Tese (Doutorado em Educação Matemática). Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista, Rio Claro, 2020.

STOBART, G.; HOPFENBECK, T. N. Assessment for learning and formative assessment. *In*: BAIRD, J; HOPFENBECK, T. N.; NEWTON, P; STOBART, G;STEEN-UTHEIM,A. T.(org.), **State of the Field Review of Assessment and Learning**, Oxford: Oxford University Centre for Educational Assessment, p. 30-50, 2014.

SKOVSMOSE, O. Cenários para investigação. **Bolema**, n.13, v.14, p. 66-91, 2000

- SUCHAUT, B. La loterie des notes au bac: un réexamen de l'arbitraire de la notation des élèves. **Les Documents de Travail de l'IREDU**. 2008. Disponível em: <https://halshs.archives-ouvertes.fr/halshs-00260958v2>. Acesso em 07 jun 2019.
- SUMPTER, D. **Dominado pelos números**. 1 ed. Rio de Janeiro: Bertrand Brasil, 2019.
- TARAS, M. Assessment–summative and formative–some theoretical reflections. **British Journal of Educational Studies**, v. 53, n. 4, p.466-478. 2005
- _____. Summative assessment: The missing link for formative assessment. **Journal of Further and Higher Education**, v. 33, n. 1, p. 57-69. 2009
- TEIXEIRA JR., V. P.; SILVA, F. H. S. Avaliação formativa e a prática de provas: uma análise das concepções de professores de Matemática. **Educação Matemática em Revista**, p. 40-59, 2017.
- TREVISAN, A. L. **Prova em fases e um repensar da prática avaliativa em Matemática**. 2013. 168f. Tese de doutorado (Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática) – Universidade Estadual de Londrina, Londrina, 2013.
- VAZ, R. F. N. **Metodologia didática de análise de soluções aplicada no ensino de frações**. 2013. Dissertação (Mestrado em Ensino de Matemática) – Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2013.
- VAZ, R. F. N.; NASSER, L.; BELFORT, E. Alunos analisando suas próprias soluções: Adição de frações. **Boletim Gepem**, n.65, 2014.
- VAZ, R. F. N.; NASSER, L. Avaliação em Matemática: um estudo sobre multcorreção. In: SEMINÁRIO INTERNACIONAL DE PESQUISA EM EDUCAÇÃO MATEMÁTICA, 7. 2018, Foz do Iguaçu. **Anais [...]**. Foz do Iguaçu, 2018.
- VAZ, R. F. N., NASSER, L. Em busca de uma avaliação mais “justa”. **Com a Palavra o Professor**, v. 4, n. 10, p. 311-329, 2019.
- VAZ, R. F. N., NASSER, L. Um Estudo sobre o Feedback Formativo na Avaliação em Matemática e sua Conexão com a Atribuição de Notas. **Bolema**, Rio Claro (SP), v. 35, n. 69, p. 1-21, abr. 2021
- VAZ, R. F. N.; NASSER, L.; LIMA, D. O. Avaliar para aprender: um ato de insubordinação criativa. **Revista @mbienteeducação**, v. 14, n. 1, p. 214-243, jan. 2021
- WILIAM, D. **Embedded formative assessment**. Bloomington: Solution Tree Press, 2011.
- WITTER, G. P. Reciprocidade e compreensão da leitura. **Psicologia Escolar e Educacional**. Campinas, v. 8, n. 1, p. 91-92, junho de 2004
- ZACCARELLI, L. M.; GODOY, A. S. "Deixa eu te contar uma coisa...": Possibilidades do uso de narrativas e sua análise nas pesquisas em organizações. **Revista Gestão Organizacional**, v. 6, n. 3, 2013.

APÊNDICE 1 - A FOLHA DE CORREÇÃO



Esta atividade faz parte de uma pesquisa de Doutorado do PEMAT/UFRJ sobre Avaliação das Aprendizagens em Matemática.

Não é necessário se identificar!

Preencha uma das opções abaixo sobre sua formação.

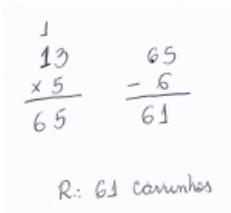
- () Licenciando em Matemática
- () Professor de Matemática da Educação Básica
- () Professor de Matemática do Ensino Superior
- () Estudante ou Professor de outra área: _____

Suponha que você seja o professor de uma turma do sexto ano.

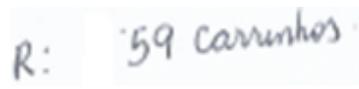
Utilizando o gabarito fornecido, realize a correção das questões dos alunos fictícios A, B, C, D, E e F (todos dessa turma), atribuindo uma pontuação de 0 a 1.

Questão	Gabarito
Pedro possui 5 caixas onde guarda seus carrinhos. Em cada caixa havia 13 carrinhos, até que Pedro retirou 6 carrinhos de uma dessas caixas para brincar. Após essa retirada, quantos carrinhos ficaram nas 5 caixas ao todo?	$5 \times 13 = 65$ $65 - 6 = 59$

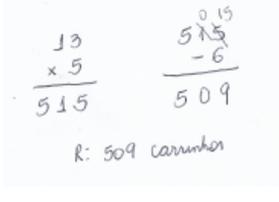
Estudante A

<p>Solução</p> 	<p>Pontuação na questão:</p> <p>Critério/justificativa para a pontuação</p>
<p>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</p>	

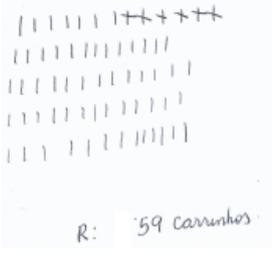
Estudante B

<p>Solução</p> 	<p>Pontuação na questão:</p> <p>Critério/justificativa para a pontuação</p>
<p>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</p>	

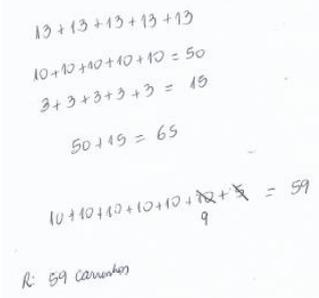
Estudante C

<p>Solução</p> 	<p>Pontuação na questão:</p> <p>Critério/justificativa para a pontuação</p>
<p>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</p>	

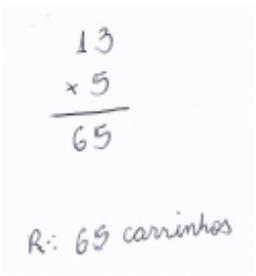
Estudante D

<p>Solução</p> 	Pontuação na questão:
	Critério/justificativa para a pontuação
<i>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</i>	

Estudante E

<p>Solução</p> 	Pontuação na questão:
	Critério/justificativa para a pontuação
<i>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</i>	

Estudante F

<p>Solução</p> 	Pontuação na questão:
	Critério/justificativa para a pontuação
<i>Feedback ao estudante (sugestões, orientações para melhoria da aprendizagem)</i>	

APÊNDICE 2 - AS QUATRO GERAÇÕES DA AVALIAÇÃO AO LONGO DO SÉCULO XX

Este apêndice apresenta uma discussão histórica sobre as transformações da avaliação no século XX que, inicialmente, seria um capítulo da tese. Após a qualificação, tornou-se um apêndice.

A Geração da Medida

A primeira pesquisa de Educação publicada, denominada *A futilidade da labuta ortográfica*, de 1897, relata a criação de um exame de ortografia para verificar se o número de horas de aulas influenciaria no desempenho dos estudantes. Seu autor, Joseph Mayer Rice, identificou que “não havia praticamente nenhuma relação entre o tempo gasto no estudo de ortografia e o êxito posterior no teste” (GUBA; LINCOLN, 2011, p. 29).

A inspiração para a concepção norteadora dessa geração vem dos testes destinados a medir a inteligência e as aptidões, que se desenvolveram na França por Binet e Simon, em 1905 (FERNANDES, 2009). O ministro da Educação francês, atendendo a solicitação dos professores, pediu ao psicólogo Alfred Binet que elaborasse um teste para realizar uma triagem nos sistemas de ensino, a fim de excluir os estudantes com retardamento mental. A partir deste estudo Binet organizou tarefas de acordo com a idade dos sujeitos, cunhando o conceito de “idade mental”.

Seus estudos atravessaram o Atlântico, originando o teste denominado de Stanford-Binet, teste de QI amplamente utilizado nas escolas americanas. Uma importante aplicação para esses testes foi no recrutamento e na seleção de soldados para as forças armadas (GUBA, LINCOLN, 2011).

De acordo com Pinto (2016), a ideia da medição em Pedagogia nos EUA surgiu em 1904 com J. M. Rice e, a partir de 1910,

começam a aparecer as primeiras escalas de medida da escrita, ortografia e cálculo aritmético. Contudo, só em 1923, aparece o Stanford Achievement Test, um teste de rendimento intelectual equivalente ao francês Binet-Simon, embora diferente na sua conceptualização e procedimentos. (PINTO, 2016, p. 14)

Dois fatores principais influenciaram a Geração da Medida. O primeiro se refere à evolução em diversas áreas do conhecimento. O grande sucesso da Matemática e das Ciências Experimentais,

que ocorreram ao longo do séc. XVIII e princípios do séc. XIX, foram também o sucesso de métodos que eram próprios, em particular, o método científico. As ciências sociais foram pressionadas a adotar o método científico para “obter maior credibilidade” (FERNANDES, 2009, p. 45).

O outro fator que impulsionou a implementação dos testes foi a Economia. “Se os seres humanos são o principal elemento na produção de bens e serviços, a responsabilidade do administrador é tornar seu trabalho mais eficiente e eficaz possível” (GUBA; LINCOLN, 2011, p. 31-32). Os testes eram eficazes em verificar e medir “se os sistemas educacionais *produziam* bons *produtos* a partir da matéria-prima disponível” (FERNANDES, 2009, p. 46, grifo do autor)

Na primeira geração, as palavras avaliar e medir eram sinônimas. “A função do avaliador era técnica, ele devia conhecer o arsenal de instrumentos disponíveis” para proporcionar uma avaliação eficiente. Caso não existisse um instrumento adequado para realizar a medida, se esperava que o avaliador tivesse os conhecimentos necessários para criá-lo (GUBA; LINCOLN, 2011, p. 33).

Pinto (2016) destaca que nessa geração a avaliação passa a ser influenciada pela psicometria. A ideia central é que a avaliação era uma medida análoga às medidas das Ciências Físicas, e as divergências encontradas entre os avaliadores eram apenas “erros de medidas”. Na pretensão de que a avaliação fosse um ato puro e perfeito, o objetivo era investigar modos de eliminar a falha humana.

A avaliação, influenciada pela psicometria, procura apenas apreciar os conhecimentos demonstrados por cada aluno. O recurso a medidas tipificadas, a garantia das mesmas condições de aplicação e a comparação dos resultados de cada aluno com o grupo de referência, normalmente o grupo/turma, permite construir uma hierarquia de excelência dentro do grupo. A avaliação utiliza assim um processo de referência normativa, isto é, a comparação dos resultados individuais com um sistema que se institui como norma. (PINTO, 2016, p. 16)

Tais estudos de docimologia²¹ apontaram contradições em relação à ideia que associava a avaliação à medida (como discutiremos nos capítulos seguintes), não sendo “capazes de construir soluções credíveis para combater a falta de rigor evidenciada em termos dos procedimentos de avaliação” (PINTO, 2016, p. 16).

A Geração da Descrição

Ela se inicia pelo fracasso da primeira geração. Um dos fatores limitadores da primeira geração era o fato de que os estudantes eram considerados os únicos objetos da avaliação. “A certa altura acabou por se considerar que seria muito redutor avaliar um sistema apenas com base no desempenho dos seus estudantes” (FERNANDES, 2009, p. 47). Passou a se considerar a possibilidade de o currículo estar impactando negativamente o desempenho dos estudantes na avaliação. Isto porque, logo após a primeira guerra mundial, muitos alunos não conseguiam superar os exames e raramente atingiam o nível elementar, alguns “exibiam necessidades e aspirações que não podiam ser atendidos pelos currículos preparatórios para a faculdade então prevalentes” (GUBA; LINCOLN, 2011, p. 33). Nessa geração a avaliação passou a ser compreendida com uma forte relação com o currículo. Surge a necessidade de avaliar currículos e programas.

Um programa de destaque dessa geração, de acordo com Guba e Lincoln (2011), foi denominado Eighth-Year Study (Estudo de Oito Anos), que teve a participação de Ralph W. Tyler, membro do departamento de pesquisa educacional da Universidade Estadual de Ohio.

Ralph Tyler é referido como o *pai da avaliação educacional* pelo trabalho que desenvolveu ainda nos anos 30 e 40 do séc. XX e o qual ao longo de 25 anos, veio a ter influência muito significativa na Educação e na avaliação, em particular ao uso de testes. Talvez por isso, Madaus e Stufflebeam (2000) designam o período de avaliação decorrido entre 1930 e 1945 como *Idade Tyleriana*. (FERNANDES, 2009, p. 47, grifo do autor)

O intuito deste programa era demonstrar que estudantes mesmo não capacitados nos currículos tradicionais estariam aptos a cursar o ensino superior. Ralph Tyler coletou informações sobre o nível do aproveitamento dos alunos em relação aos *objetivos*, ou seja, “em relação ao que os alunos haviam ou não aprendido do que seus *professores*

²¹A palavra Docimologia (do grego dokimé, “teste”) foi cunhada por Henri Piéron em 1920. Trata-se do estudo sistemático dos exames, em particular do sistema de atribuição de notas e dos comportamentos dos examinadores e examinados. (DOCIMOLOGIA, 2020)

pretendiam que eles aprendessem”. Os resultados de Tyler foram utilizados para conduzir aprimoramentos e as reformas curriculares. Os instrumentos avaliativos deveriam ser utilizados para avaliar pontos fortes e fracos a respeito de determinados objetivos estabelecidos (GUBA; LINCOLN, 2011, p. 35, *grifo dos autores*).

Em linhas gerais, entendia que os resultados da Educação deveriam provocar mudanças comportamentais nos padrões preestabelecidos pela sociedade. Ou seja, os objetivos propostos precisariam estar em estreita relação com os resultados esperados, pois só assim se concretizaria a mudança de novos valores sociais e culturais. Neste sentido, a avaliação por objetivos também promoveria a revisão do currículo na construção de novas habilidades específicas. A avaliação passaria a ser, portanto, a ligação entre a escola e a sociedade, apontando problemas e interferindo nas possíveis soluções. É um modelo que promove a tomada de decisão (MACEDO, LIMA, 2013, p. 159)

O professor deveria deixar claro seus objetivos para serem alcançados pelos alunos e expor esses objetivos nas práticas curriculares. Nesse sentido, em comparação à geração da medida, há duas importantes mudanças. A primeira que merece destaque diz respeito à referência central da avaliação que

passa ser o conjunto de objetivos pré-estabelecidos, deixando de lado as provas normativas e o grupo/turma. Por outras palavras, há um sistema de referência igual para todos os alunos que frequentem um determinado ciclo de estudos. A avaliação é assim uma comparação entre os objetivos que constituem o sistema de referência e o estado do aluno na consecução desses objetivos (PINTO, 2016, p. 16-17).

A segunda mudança relevante se refere ao papel da medida, a medição torna-se um dos aspectos e não a avaliação em si. Com isso ultrapassa-se a

ideia de simples classificação dos alunos uma vez que a avaliação é perspectivada, tanto com o objetivo de verificar se os alunos atingem os objetivos educativos definidos, como também através da introdução de procedimentos corretivos para melhorar a gestão do programa em curso, através da observação do comportamento dos alunos. A avaliação passa assim a ter um sistema de referência criterial, isto é, o termo de comparação são os critérios previamente definidos que permitem verificar a mestria dos objetivos (PINTO, 2016, p. 17).

As ideias de Tyler influenciaram diversos pesquisadores, dentre eles, podemos destacar, Benjamin Bloom. Na década de 1950, Bloom publicou a teoria da taxonomia dos objetivos, “que prevê uma estrutura de organização hierárquica para os objetivos educacionais e que também classifica os tipos de aprendizagem em cognitivo, afetivo e psicomotor, criando subdomínios para estes campos” (MACEDO, LIMA, 2013, p. 160).

A Geração do Juízo de Valor

Ao reduzir a avaliação a uma operação de medida dos objetivos, a Geração da Descrição apresentava uma falha, pois deixava a avaliação à mercê de uma visão tecnicista. De acordo com Guba e Lincoln (2011), as falhas se tornaram mais evidentes no período pós-Sputnik, pós 1957, com a vitória dos russos na corrida espacial. A geração Tyleriana, essencialmente descritiva, desprezou o que Robert Stake chamou de outra fisionomia ou face da avaliação em seu artigo de 1967: o juízo de valor.

Atribuiu-se ao professor ou corretor o papel de julgador, mantendo igualmente as funções de técnicas relacionadas à medição e à descrição. Michael Scriven evidenciou, em 1967, vários problemas em relação às gerações anteriores: os objetivos eram, em si, problemáticos. De modo similar ao desempenho dos estudantes, os objetivos também deveriam ser avaliados. Além disso, “o juízo de valor depende de critérios para que possa ser feito com base nesses critérios”, entretanto, a inclusão de critérios científicos não era aceita pelos avaliadores (GUBA; LINCOLN, 2011, p. 37).

Temendo ficar para trás na corrida contra os Russos, os países da Europa e os EUA investiram em reformas educacionais, sob o ponto de vista curricular e de ensino-aprendizagem, que privilegiassem o ensino de Ciências e Matemática. Devido à grande expansão e à evolução dos conceitos relacionados à avaliação, Madaus e Stufflebeam denominaram o período entre 1958 e 1972 de *Idade do Desenvolvimento*. (FERNANDES, 2009, grifo do autor).

Algumas ideias poderosas sobre avaliação surgem nessa geração, que também ficou conhecida como geração da avaliação como apreciação do mérito: a ideia de regulação do ensino e da aprendizagem, a coleta de informações deve ir além dos resultados dos testes, a avaliação deve envolver pais, alunos e outros professores, os contextos de ensino e aprendizagem passam a ser considerados e a definição dos critérios passa a ser essencial para se apreciar o valor e o mérito de um dado objeto da avaliação (FERNANDES, 2009).

Havia duas linhas conceituais importantes nessa geração,

uma, aprofundando a perspectiva de Tyler, baseada no desenvolvimento de uma tecnologia ao serviço da definição dos objetivos e dos instrumentos de medida; a outra, desenvolvendo a proposta de Cronbach associa a avaliação ao processo de tomada de decisões (PINTO, 2016, p. 19)

Para Cronbach a avaliação pode ser definida como um processo que visa a coleta de informações que permitam que decisões sejam tomadas sobre um programa educacional. Por ser uma atividade diversificada, a avaliação não pode ser confundida como uma construção de instrumentos de medida e a obtenção de pontuações (MACEDO; VIANA, 2013). Posteriormente, surgem diversas perspectivas que enfatizam novos aspectos a tomar em consideração pela avaliação. Sob influência do paradigma qualitativo que influenciava as Ciências Sociais nos anos 70, destacam-se dois grupos de orientação:

- o interpretativo, do qual Eisner talvez seja “um dos maiores expoentes desta corrente ao sublinhar que o avaliador é um perito em Educação que interpreta a realidade em função do meio cultural saturado de significados” (PINTO, 2016, p. 20);
- o sócio crítico, que procura “revelar, através do diálogo e da negociação com os atores da cena social em estudo, os factos mais relevantes de uma inovação que se pretende introduzir” (PINTO, 2016, p. 20).

Citando uma publicação de Hadji de 1989, Pinto (2016, p. 20) problematiza a concepção de avaliação como um juízo de valor ao considerar que partindo da hipótese que “avaliar é tomar uma decisão entre diferentes possibilidades, através de um julgamento”, na verdade, “o que está no centro da definição, não é a decisão, mas o próprio julgamento”. E, complementa o autor, “estando o julgamento no centro do conceito de avaliação, como é possível assumir que o ato de julgar, em si mesmo, não seja considerado tendo em conta o ponto de vista do avaliador?”

Em 1967, Michael Scriven, em seu artigo intitulado *The Methodology of Evaluation*, propôs a distinção entre a avaliação somativa e formativa, uma diferenciação no que concerne às funções da avaliação. Enquanto a somativa estava relacionada “à prestação de contas, à certificação e à seleção”, a avaliação formativa se voltava “ao desenvolvimento, à melhoria da aprendizagem e à regulação dos processos de ensino e de aprendizagem” (FERNANDES, 2009, p. 49). Na concepção de Scriven (1967), tanto a concepção somativa como a formativa são essenciais.

Em 1971, Bloom, Hastings e Madaus propuseram uma terceira categoria: a diagnóstica. A avaliação diagnóstica “é a que intervém quando se põe o problema a saber se um sujeito possui as capacidades necessárias para empreender uma certa

aprendizagem” (NOIZET; CAVERNI, 1985, p. 15). A avaliação diagnóstica teria como função identificar os pré-requisitos necessários e preparar o aluno para iniciar a nova etapa da aprendizagem (BARREIRA, 2019; PINTO, 2019).

A distinção entre as três modalidades de avaliação – somativa, formativa e diagnóstica – promoveu uma

verdadeira inovação no campo da avaliação escolar com implicações pedagógicas assinaláveis (aposta no aluno, na diversificação de métodos pedagógicos e de avaliação), uma vez que se passou de uma única modalidade, a avaliação final feita através de testes e exames, para três modalidades com objetivos e funções diferenciados. (BARREIRA, 2019, p. 192)

De certa forma, a diagnóstica poderia ser compreendida como um tipo de avaliação formativa. Nessa linha de raciocínio, Hadji (2001) sugere que a avaliação diagnóstica seja classificada como prognóstica quando utilizada em um caráter formativo, pois mais do que traçar um diagnóstico do problema, permite promover ações para a superação de obstáculos no processo ensino-aprendizagem.

A avaliação prognóstica tem a função de permitir um ajuste recíproco aprendiz/programa de estudos (seja pela modificação do programa, que será adaptado pelos aprendizes, seja pela orientação dos aprendizes para subsistemas de formação mais adaptados a seus conhecimentos e competências atuais). (HADJI, 2001, p. 19)

Problematizando as Três Primeiras Gerações

A coleta de dados sistemática e o desenvolvimento de instrumentos apropriados caracterizaram a primeira geração. Os objetos não humanos, como programas, conteúdos e as estratégias de ensino foram incorporados nos estudos de avaliação na segunda geração. A terceira geração “exigiu que a avaliação se centrasse no juízo de valor, tanto sobre o mérito do objeto da avaliação – seu valor íntimo e intrínseco – quanto sobre sua importância – seu valor extrínseco ou contextual” (GUBA; LINCOLN, 2011, p. 39).

Guba e Lincoln (2011) consideram que há três grandes imperfeições ou defeitos nessas três gerações da avaliação: “a tendência ao gerencialismo, a incapacidade de acomodar o pluralismo de valores e o comprometimento exagerado com o paradigma científico de investigação” (p.39).

Os administradores e avaliadores gerenciam o processo avaliativo e se mantêm em uma posição isenta de responsabilidade. Além disso, há na relação entre o avaliador e o indivíduo a ser avaliado, numa posição injusta e provedora de poder. Os

administradores e avaliadores possuem o poder de determinar o que será avaliado, quando essa avaliação irá ocorrer e como as respostas serão coletadas (GUBA; LINCOLN, 2011). Fernandes (2009) pontua que as responsabilidades por eventuais falhas dos sistemas educacionais normalmente são atribuídas aos estudantes e aos professores, entretanto o baixo desempenho de estudantes nas avaliações normalmente é invariavelmente atribuído aos estudantes.

Se há um juízo de valor, esse julgamento ocorre de acordo com concepções socioculturais dos avaliadores e dos administradores dos estabelecimentos de ensino. De acordo com Guba e Lincoln (2011), “a alegação de que a ciência é isenta de valores pode ser seriamente contestada” (p. 42-43) e, não sendo isenta de valores, os resultados são sujeitos a interpretações diferentes. O “pluralismo de valores de nossa sociedade é uma questão crucial a ser levada em consideração na avaliação” (p. 43), nesse sentido, podemos afirmar que “todo ato de avaliação se torna um ato político” (p. 43).

O comprometimento exagerado com o paradigma científico de investigação, por sua vez, está relacionado à chamada avaliação científica, na qual se utiliza métodos e “instrumentos considerados neutros, normalmente testes, que medem com rigor e objetividade o que os alunos sabem”. Tais instrumentos tendem a ser “pouco ou nada contextualizados, com excessiva dependência da concepção como medida” (FERNANDES, 2009, p. 52).

De modo geral, as três primeiras gerações estavam amparadas pelo positivismo, moldando o desejo dos cientistas sociais de emular os métodos científicos com convicção e entusiasmo, sendo racionais e sistemáticos, ao estilo de Descartes: Penso, logo existo. “Qualquer coisa que esteja sob a avaliação e tenha respaldo da avaliação positivista (científica) é pregada como a coisa certa a fazer” (GUBA, LINCOLN, 2011, p. 45)

A Geração da Avaliação como Negociação e Construção

A denominação que intitula essa seção é adotada por Fernandes (2009). Guba e Lincoln (2011) utilizam o termo avaliação construtivista responsiva (ou construtivista respondente) e Pinto (2016), a avaliação como uma abordagem plural de factos sociais

O termo responsivo foi proposto por Robert Stake, em 1967, e se relaciona dar resposta. “Logo, a coleta da informação tem como objetivo promover a troca, o diálogo entre os sujeitos e sobre a avaliação em todas as fases de seu processo, para a solução de

problemas” (MACEDO; LIMA, 2013, p. 165). Guba e Lincoln (2011) utilizam o termo responsivo para designar uma forma diferente de focalizar a avaliação, onde os parâmetros e limites são negociados.

A avaliação é responsiva não apenas pelo fato de procurar identificar diferentes visões entre os interessados, mas também porque responde às questões pertinentes à coleta de informações subsequente. [...] uma das principais atribuições do avaliador é conduzir a avaliação de modo que cada grupo confronte e lide com as construções de todos os outros, um processo que chamaremos de hermenêutico-dialético (GUBA; LINCOLN, 2011, p. 50).

Já o termo construtivista se refere, segundo Guba e Lincoln (2011), ao paradigma da investigação em substituição ao paradigma norteador das três gerações anteriores, o científico. Ontologicamente, o paradigma construtivista nega a existência de uma realidade objetiva. As realidades são construções sociais. Epistemologicamente, nega a possibilidade do dualismo sujeito-objeto, propondo uma interação entre observador e observado. Metodologicamente,

rejeita a abordagem dominante e manipulatória (experimental) que caracteriza a ciência e substitui por um processo hermenêutico-dialético que aproveita ao máximo – e leva em conta – a interação observador-observado para criar uma realidade construída que seja, tanto quanto possível, fundamentada e esclarecida em um determinado momento. (GUBA; LINCOLN, 2011, p. 53)

As realidades são construções sociais e mentais, e dentro dessas construções se procura compreender e considerar a pluralidade de perspectivas. Ganha importância o trabalho em redor da explicitação, das intenções, dos objetivos, dos meios a utilizar na recolha das informações. A avaliação é compreendida como um processo aberto e negociado entre os vários atores, onde as regras do jogo sejam conhecidas, clarificando e respeitando uma ética do agir avaliativo, que permita uma reflexão crítica sobre a própria ação avaliativa e trace novas perspectivas de futuro enquadrando os novos problemas emergentes (PINTO, 2016).

Em havendo uma pluralidade de abordagens sobre uma mesma realidade, devemos considerar uma diversidade de respostas. Os diversos olhares existentes no processo avaliativo não se excluem, se complementam. A avaliação passa a ser compreendida como

um processo de construção social e político, que envolve uma colaboração entre vários atores, que toma a realidade como dinâmica, que admite a divergência, que lida com resultados imprevisíveis e que nesta sua ação vai gerando a própria realidade (PINTO, 2016, p. 23)

Fernandes (2009) destaca que na avaliação de quarta geração está plenamente integrado o processo de ensino e aprendizagem. A avaliação formativa deve ser a modalidade privilegiada de avaliação devido à sua função de regulação e melhoramento das aprendizagens. “O *feedback*, nas suas mais variadas formas, frequências e distribuições, é um processo indispensável para que a avaliação se integre plenamente no processo de ensino-aprendizagem” (p. 55). A avaliação deve ser reconhecida como uma construção social e atuar para promover aprendizagens e não para julgar ou classificar os estudantes.

Corrigir um teste e atribuir uma nota definitivamente não pode ser encarado como sinônimo de medir o conhecimento de alguém sobre algo, como se achava durante o início do século passado. A ideia de associar a nota de um teste a uma medição é, definitivamente, um mito. O teste está muito mais próximo de uma leitura do que de uma medida. Uma leitura não somente da aprendizagem, mas também do ensino. A variação de notas atribuídas por licenciandos e professores ao corrigir uma mesma prova indica diferentes leituras, estruturadas na subjetividade do corretor.

Se o conhecimento é multifacetado, complexo, construído individualmente e inextricavelmente ligado ao contexto no qual o aprendizado ocorre, conclui-se que nenhum instrumento único é capaz de “medir” esse conhecimento de maneira consistente e significativa (ROMAGNANO, 2011). A partir daí, é necessário que os testes devam ser ressignificados no meio escolar e acadêmico, abandonando o papel de único instrumento avaliativo (comumente utilizado na disciplina de Matemática). As avaliações escolares devem ser mais diversificadas, pois mais instrumentos podem oferecer uma melhor leitura do ensino-aprendizagem, respeitando mais as peculiaridades de cada estudante.

A avaliação como uma negociação e construção constitui uma ruptura epistemológica difícil, mas necessária para modificar a filosofia positivista que acompanha as concepções sobre a avaliação escolar.

Referências do apêndice 2

BARREIRA, C. M. F. Concepções e práticas de avaliação formativa e sua relação com os processos de ensino de aprendizagem. In: ORTIGÃO, M. I. R. et al (org.). **Avaliar para aprender no Brasil e em Portugal**: perspectivas teóricas, práticas e de desenvolvimento. Curitiba: CRV, 2019. v. 1, p. 191-218

FERNANDES, D. **Avaliar para aprender**: fundamentos, práticas e políticas. São Paulo: Editora Unesp, 2009.

GUBA, E. G.; LINCOLN, Y. S. **Avaliação de quarta geração**. Campinas: Editora da Unicamp, 2011.

HADJI, C. Avaliação desmistificada. Porto Alegre: Artmed Editora, 2001. 136 p.

MACEDO, M. F.; LIMA, A. M. Revolvendo o passado da avaliação educacional e algumas repercussões na escola. **Revista Teias**, v. 14, n. 32, p. 17, 2013.

NOIZET, G.; CAVERNI, J-P. **Psicologia da avaliação escolar**. Coimbra: Coimbra Editora, 1985.

PINTO, J. A avaliação em educação: da linearidade dos usos à complexidade das práticas. In: AMANTE, L; OLIVEIRA, I. (org.) **Avaliação das Aprendizagens**: Perspectivas, Contextos e Práticas. Lisboa: Universidade Aberta, 2016.

_____. Avaliação Formativa: uma prática para a aprendizagem. In: ORTIGÃO, M. I. R. et al (org.). **Avaliar para aprender no Brasil e em Portugal**: perspectivas teóricas, práticas e de desenvolvimento. Curitiba: CRV, 2019. v. 1, p. 19 – 44

ROMAGNANO, L. The myth of objectivity in mathematics assessment. **Mathematics Teacher**, v. 94, n. 1, p. 31-37. 2001.