

Contribuições da análise estatística implicativa no tratamento e análise de questões no campo da avaliação

Saddo Ag Almouloud

saddoag@gmail.com, saddo@ufpa.br

¹Universidade Federal do Pará, R. Augusto Corrêa, 1 - Guamá, Belém - PA, 66075-

Resumo

O objetivo deste artigo é apresentar um estudo aprofundado sobre a aplicação da Análise Estatística Implicativa (ASI) no tratamento e análise de questões relacionadas à avaliação educacional, especialmente no ensino da matemática. O texto revisita conceitos fundamentais da avaliação, discute desafios práticos e teóricos, e destaca a utilidade da ASI para compreender estruturas cognitivas e hierarquias de competências a partir de grandes bases de dados, como PISA e TIMSS, além de estudos nacionais. O estudo utiliza a metassíntese qualitativa para revisar e interpretar pesquisas que aplicaram a ASI em avaliações educacionais, seguindo etapas rigorosas de seleção, leitura sistemática, análise comparativa e formulação de interpretações integradoras. A busca bibliográfica, baseada no protocolo POT (População, Outcome, Tipo de pesquisa), identificou 11 artigos relevantes, majoritariamente de Antoine Bodin e colaboradores, que abordam a análise de questões de avaliação com o uso da ASI e outras ferramentas estatísticas. O conjunto de estudos revisados evidencia a relevância da ASI como ferramenta estatística avançada para analisar avaliações educacionais, especialmente em matemática. A ASI permite revelar hierarquias cognitivas, dependências entre itens e estruturas do conhecimento, complementando métodos tradicionais e contribuindo para uma avaliação mais significativa e contextualizada. Também evidencia desafios práticos e teóricos da avaliação, a importância de critérios claros, a limitação dos testes padronizados e a necessidade de abordagens multiculturais e didáticas para interpretar resultados internacionais e promover a inclusão educacional. A integração da análise epistemológica e didática com métodos estatísticos como a ASI representa um avanço para a compreensão e melhoria das avaliações educacionais. Como perspectiva, o autor pretende analisar dados de Exame Nacional de Desempenho dos Estudantes do Brasil, no intuito de revelar hierarquias cognitivas, dependências entre itens e estruturas do conhecimento, complementando métodos tradicionais e contribuindo para uma avaliação significativa e contextualizada, aspectos que não são, em geral, evidenciados em outros tipos de análises estatísticas.

Palavras-chave: Avaliação, Análise estatística implicativa, Metassíntese qualitativa.

Contribuciones del análisis estadístico implicativo en el tratamiento y análisis de cuestiones en el campo de la evaluación

Resumen

El objetivo de este artículo es presentar un estudio en profundidad sobre la aplicación del Análisis Estadístico Implicativo (ASI) en el tratamiento y análisis de cuestiones relacionadas con la evaluación educativa, especialmente en la enseñanza de las matemáticas. El texto revisa conceptos fundamentales de la evaluación, discute retos prácticos y teóricos, y destaca la utilidad de la ASI para comprender estructuras cognitivas y jerarquías de competencias a partir de grandes bases de datos, como PISA y TIMSS, además de estudios nacionales. El estudio utiliza la metasíntesis cualitativa para revisar e interpretar investigaciones que han aplicado el ASI en evaluaciones educativas, siguiendo rigurosos pasos de selección, lectura sistemática, análisis comparativo y formulación de interpretaciones integradoras. La búsqueda bibliográfica, basada en el protocolo POT (Población, Resultado, Tipo de investigación), identificó 11 artículos relevantes, en su mayoría de Antoine Bodin y colaboradores, que abordan el análisis de cuestiones de evaluación con el uso de ASI y otras herramientas estadísticas. El conjunto de estudios revisados evidencia la relevancia de ASI como herramienta estadística avanzada para analizar evaluaciones educativas, especialmente en matemáticas. La ASI permite revelar jerarquías cognitivas, dependencias entre ítems y estructuras del conocimiento, complementando métodos tradicionales y contribuyendo para una evaluación más significativa y contextualizada. También evidencia desafíos prácticos y teóricos de la evaluación, la importancia de criterios claros, la limitación de los tests estandarizados y la necesidad de abordajes multiculturales y didácticos para interpretar resultados internacionales y promover la inclusión educativa. La integración de la análisis epistemológico y didáctico con métodos estadísticos como la ASI representa un avance para la comprensión y mejoría de las evaluaciones educativas. Como perspectiva, el autor pretende analizar datos del Examen Nacional de Desempeño de los Estudiantes de Brasil, con el fin de revelar jerarquías cognitivas, dependencias entre ítems y estructuras del conocimiento, complementando métodos tradicionales y contribuyendo para una evaluación significativa y contextualizada, aspectos que no son, en general, evidenciados en otros tipos de análisis estadísticos.

dependencias entre ítems y estructuras de conocimiento, complementando los métodos tradicionales y contribuyendo a una evaluación más significativa y contextualizada. También pone de manifiesto los retos prácticos y teóricos de la evaluación, la importancia de criterios claros, la limitación de las pruebas estandarizadas y la necesidad de enfoques multiculturales y didácticos para interpretar los resultados internacionales y promover la inclusión educativa. La integración del análisis epistemológico y didáctico con métodos estadísticos como el ASI representa un avance para la comprensión y la mejora de las evaluaciones educativas. Desde perspectiva, el autor pretende analizar datos del Examen Nacional de Rendimiento Estudiantil para revelar jerarquías cognitivas, dependencias entre ítems y estructuras de conocimiento, complementando métodos tradicionales y contribuyendo a una evaluación significativa y contextualizada, aspectos que, en general, no se evidencian en otros tipos de análisis estadísticos.

Palabras clave: Evaluación, Análisis estadístico implicativo, Meta síntesis cualitativa.

Contributions of implicative statistical analysis in the treatment and analysis of issues in the field of evaluation

Abstract

The purpose of this article is to present an in-depth study on the application of Implicative Statistical Analysis (ISA) in the treatment and analysis of issues related to educational assessment, especially in mathematics teaching. The text revisits fundamental concepts of assessment, discusses practical and theoretical challenges, and highlights the usefulness of IAA for understanding cognitive structures and skill hierarchies from large databases, such as PISA and TIMSS, as well as national studies. The study uses qualitative meta-synthesis to review and interpret research that has applied IAA in educational assessments, following rigorous steps of selection, systematic reading, comparative analysis, and formulation of integrative interpretations. The literature search, based on the POT (Population, Outcome, Type of research) protocol, identified 11 relevant articles, mostly by Antoine Bodin and collaborators, which address the analysis of assessment questions using ASI and other statistical tools. The set of studies reviewed highlights the relevance of ASI as an advanced statistical tool for analyzing educational assessments, especially in mathematics. ASI reveals cognitive hierarchies, dependencies between items, and knowledge structures, complementing traditional methods and contributing to a more meaningful and contextualized assessment. It also highlights practical and theoretical challenges in assessment, the importance of clear criteria, the limitations of standardized tests, and the need for multicultural and didactic approaches to interpret international results and promote educational inclusion. The integration of epistemological and didactic analysis with statistical methods such as ASI represents a step forward in understanding and improving educational assessments. As part of her research, the author aims to analyze data from the National Student Performance Exam in Brazil to shed light on cognitive hierarchies, the links between items and knowledge structures. This approach aims to complement traditional methods and promote more relevant and context-specific evaluation, aspects that are often overlooked by other statistical methods.

Keywords: Evaluation, Implicative statistical analysis, Qualitative meta-synthesis.

Contributions de l'analyse statistique impliquée dans le traitement et l'analyse des problématiques dans le domaine de l'évaluation

Résumé

L'objectif de cet article est de présenter une étude approfondie sur l'application de l'analyse statistique implicative (ASI) dans le traitement et l'analyse des questions liées à l'évaluation éducative, en particulier dans l'enseignement des mathématiques. Le texte revisite les concepts fondamentaux de l'évaluation, discute des défis pratiques et théoriques, et souligne l'utilité de l'ASI pour comprendre les structures cognitives et les hiérarchies de compétences à partir de grandes bases de données, telles que PISA et TIMSS, ainsi que d'études nationales. L'étude utilise la mété-synthèse qualitative pour examiner et interpréter les recherches qui ont appliqué l'ASI aux évaluations éducatives, en suivant des étapes rigoureuses de sélection, de lecture systématique, d'analyse comparative et de formulation d'interprétations intégratives. La recherche bibliographique, basée sur le protocole POT (Population, Outcome, Type de recherche), a identifié 11 articles pertinents, principalement d'Antoine Bodin et ses collaborateurs, qui traitent de l'analyse des questions d'évaluation à l'aide de l'ASI et d'autres outils statistiques. L'ensemble des études examinées met en évidence la pertinence de l'ASI en tant qu'outil statistique avancé pour analyser les évaluations éducatives, en particulier en mathématiques. L'ASI permet de révéler les hiérarchies cognitives, les dépendances entre les items et les structures de connaissances, complétant ainsi les méthodes traditionnelles et contribuant à une évaluation plus significative et contextualisée. Elle met également en évidence les défis pratiques et théoriques de

l'évaluation, l'importance de critères clairs, les limites des tests standardisés et la nécessité d'approches multiculturelles et didactiques pour interpréter les résultats internationaux et promouvoir l'inclusion éducative. L'intégration de l'analyse épistémologique et didactique avec des méthodes statistiques telles que l'ASI représente une avancée pour la compréhension et l'amélioration des évaluations éducatives. Dans le cadre de ses recherches, l'auteur souhaite analyser les données de l'Examen national de performance des élèves au Brésil pour mettre en lumière les hiérarchies cognitives, les liens entre les items et les structures de connaissances. Cette approche vise à compléter les méthodes traditionnelles et à favoriser une évaluation plus pertinente et adaptée au contexte, des aspects souvent négligés par d'autres méthodes statistiques.

Mots clés : Evaluation, Analyse statistique implicative, Meta-synthèse qualitative.

1. INTRODUÇÃO

O objetivo deste artigo é examinar o uso da análise estatística implicativa (ASI), segundo a concepção de Gras (1979, 1996, 2013), no tratamento e análise de questões relacionadas à avaliação. Uma revisão de trabalhos utilizando a ASI como ferramenta estatística para o estudo das questões de avaliação será apresentada para fornecer uma visão geral dessa pesquisa, destacando o papel das inferências e da generalização na interpretação dos resultados e chamando atenção para os riscos associados à adaptação excessiva do conhecimento aos requisitos da revisão.

Este método visa estabelecer regras de associação a partir de um conjunto de dados que inclui diferentes sujeitos e variáveis. Seu objetivo inicial é responder à seguinte pergunta: “Se um objeto possui uma propriedade, ele também possui outra?” Quando a resposta é totalmente positiva, a conclusão é fácil. No entanto, caso contrário, é possível observar certas tendências emergentes.

A ASI visa evidenciar certas tendências dentro de um conjunto de propriedades. Ao contrário de outros métodos estatísticos que permitem produzir regras de associação, a ASI se distingue pelo uso de uma medida não linear que responde a critérios essenciais. Essa medida se baseia na intensidade de implicação, que avalia o grau de surpresa associado a uma regra e, assim, elimina as regras triviais que podem já ser conhecidas pelos especialistas.

Além disso, a intensidade de implicação pode ser complementada pelo nível de validade determinado a partir da entropia de Shannon, se esse método de cálculo for escolhido pelo usuário. Nesse contexto, a medida considera a validade da regra, mas também a de sua contraposição. Concretamente, quando uma regra de associação é considerada válida — ou seja, quando o conjunto de itens A se aproxima fortemente do conjunto B —, torna-se lógico e natural esperar que a contraposição também seja válida, ou seja, que o conjunto de itens não B esteja fortemente relacionado ao conjunto de itens não A. Essas duas medidas originais são complementadas por uma medida clássica baseada no tamanho do suporte da regra e em sua confiança (frequência condicional de B sabendo A). Assim, combinando as três medidas, é possível definir uma medida relevante que possua as qualidades das três medidas (se considerarmos o uso da teoria entrópica), ou seja, a

resistência ao ruído, uma vez que a contraposição da regra é levada em consideração, e a rejeição de regras triviais. Para mais informações, o leitor interessado pode consultar Gras et al. (2008) e o capítulo 1, Parte 1 da referida obra.

Graças a essa medição original, o software CHIC¹ (Classificação Hierárquica, Implicativa e Coesitiva) possibilita calcular regras de associação a partir de um conjunto de dados. CHIC e ASI têm sido usados em uma ampla gama de áreas de pesquisa.

O CHIC permite construir dois tipos de hierarquia e um gráfico. A hierarquia mais conhecida é a hierarquia das semelhanças. O índice de semelhança foi definido em Lerman I. C. (1981) e permite construir uma hierarquia ascendente. De forma semelhante, a intensidade de implicação pode ser usada para construir uma hierarquia orientada. Além disso, o CHIC oferece a possibilidade de gerar um gráfico original, chamado gráfico implicativo, que permite ao usuário selecionar as regras de associação e as variáveis que deseja ver aparecer.

Quanto à avaliação, baseamo-nos nos trabalhos do IREM de Besançon, que fornece uma introdução geral ao conceito de avaliação no contexto do ensino da matemática e enfatiza os diferentes tipos de avaliação.

O documento aborda os desafios práticos relacionados à implementação de práticas de avaliação eficazes, incluindo a necessidade de superar os hábitos e as percepções estabelecidas de avaliação dos professores e alunos, em que as notas muitas vezes têm prioridade sobre a aprendizagem. Outro desafio é também lidar com a natureza demorada dos métodos de avaliação diversificados, especialmente durante a transição das avaliações somativas para as avaliações formativas. A implementação de práticas de avaliação requer superar hábitos profundamente enraizados, lidar com pressões externas e enfrentar desafios práticos relacionados ao tempo e aos preconceitos do professor.

No que diz respeito à distinção entre avaliação normativa e avaliação baseada em critérios, este documento apresenta duas abordagens avaliativas. A avaliação normativa compara o desempenho de um aluno com o de um grupo de referência e é o método mais amplamente utilizado. A classificação pode ser ajustada para garantir que os alunos médios obtenham resultados em conformidade com a média, procurando obter uma distribuição de notas próxima de uma curva normal. Os testes que não correspondem a essa lógica

¹ Para ter mais informações sobre CHIC e sua versão RCHIC, acessa o seguinte link: <https://sites.univ-lyon2.fr/asi/13/index.php?page=18>

são geralmente questionados, enquanto aqueles que distribuem adequadamente as notas, são considerados válidos. No entanto, esse modo de avaliação pode desmotivar os alunos com dificuldades e não refletir com precisão seus conhecimentos e competências reais.

Ao contrário, a avaliação baseada em critérios facilita o diálogo com o aluno graças à comunicação explícita dos critérios de avaliação, ao mesmo tempo que promove a harmonização na fase somativa. Esta abordagem valoriza os progressos realizados e as qualidades individuais do aluno, independentemente da sua posição em relação ao grupo. O professor dispõe assim de indicadores precisos sobre os conhecimentos adquiridos pelo aluno, que também sabe orientar seus esforços. Quando os objetivos são definidos coletivamente pela equipe pedagógica ou sob a égide de uma instância superior, o professor atua como guia no processo de aprendizagem.

Os autores da obra destacam os principais fatores que garantem que os métodos de avaliação reflitam fielmente os conhecimentos e as competências dos alunos. Isso implica a definição e a comunicação de objetivos claros. A avaliação deve incidir sobre o trabalho dos alunos em relação a critérios pré-determinados, sem fazer comparações entre indivíduos. É essencial que esses critérios sejam transparentes, a fim de promover o diálogo com os alunos e melhorar a coerência nas avaliações somativas. A utilização de abordagens de avaliação variadas permite representar o conjunto das competências de um aluno, por meio da amostragem de diferentes tipos de produções. Além disso, é necessário reduzir ao máximo o risco de atribuir erroneamente ou recusar incorretamente conhecimentos, adaptando os métodos de questionamento para as avaliações formativas e somativas.

Este documento aborda as possíveis limitações dos testes padronizados na avaliação. Os testes padronizados e os questionários de múltipla escolha (QME) podem alterar a natureza do que é medido, valorizando mais a capacidade de validar uma resposta do que de elaborar um raciocínio, ou avaliando principalmente os conhecimentos memorizados em detrimento da compreensão real.

Os QME favorecem a lembrança de informações aprendidas, mas muitas vezes negligenciam a capacidade de reestruturação e a criatividade diante dos desafios pedagógicos. Essa característica os torna menos adequados para certas disciplinas, como matemática, em que essas competências são muito importantes. Além disso, o formato de múltipla escolha apresenta o risco de os alunos escolherem respostas aleatoriamente, o que pode distorcer a avaliação.

Em resumo, os testes padronizados correm o risco de transformar o objeto da avaliação, restringir a medição das competências gerais, levar a respostas aleatórias e tornar a interpretação demasiado dependente das escores globais.

Este documento enfatiza a importância de encontrar um equilíbrio entre uma avaliação objetiva e a valorização do progresso de cada aluno. Ele sugere destacar as qualidades individuais, definir critérios claros para promover o diálogo e garantir a coerência, bem como utilizar métodos imparciais para reduzir as diferenças de avaliação. Ele também recomenda diversificar os métodos de avaliação

para refletir todas as competências do aluno. Para que uma avaliação seja relevante, ela deve basear-se em objetivos precisos, critérios transparentes, uma variedade de métodos e uma atenção especial para evitar erros de interpretação.

2. METODOLOGIA DE PESQUISA

Neste artigo adotamos a metassíntese qualitativa que, segundo Alencar e Almouloud (2017), é uma forma de revisão sistemática em pesquisas bibliográficas. Ela auxilia na identificação de melhorias nas teorias, metodologias e tendências, além de destacar pontos que exigem maior investigação.

Os autores ainda asseveraram que a metassíntese qualitativa possibilita a análise aprofundada das investigações primárias. As análises são realizadas por meio do estabelecimento de relações entre estudos, considerando comparações, semelhanças e diferenças. A partir da interpretação dos dados, desenvolve-se a metassíntese qualitativa, que apresenta o objeto investigado em um patamar avançado de abstração e compreensão.

Matheus (2009, apud Alencar e Almouloud, 2017) delineia as etapas essenciais da metassíntese qualitativa, apontando dois principais métodos para sua condução: a integração ou agregação dos dados provenientes das investigações; e a interpretação das evidências qualitativas. Conforme o autor, ambos os métodos seguem uma sequência de etapas conforme a Tabela 1.

Tabela 1-Etapa da metassíntese qualitativa. Fonte: Construído apoiando-se em Alencar e Almouloud (2017, p. 2010)

Etapa	Descrição
a) Identificação do interesse do pesquisador	Definição clara do objetivo da pesquisa; objetivo específico orienta busca e seleção; interesse fundamentado na relevância acadêmica
b) Seleção criteriosa dos elementos	Coleta de dados com procedimentos rigorosos para validade da metassíntese qualitativa; análise detalhada dos estudos com base nos objetivos
c) Leitura e releitura sistemática	Identificação de interpretações relevantes; preparação de resumos e fichamentos para organizar resultados
d) Apresentação das relações	Destacar resultados obtidos; comparações iniciais evidenciam semelhanças, diferenças e complementariedades
e) Formulação de afirmações abrangentes	Afirmações fundamentadas sobre resultados, preservando contexto original de cada pesquisa
f) Elaboração de nova interpretação integradora	Contempla todos os estudos analisados e responde ao objetivo inicialmente proposto

Para a realização de uma busca com rigor metodológico, utilizamos o protocolo de estratégia POT (População, *Outcome*, Tipo de pesquisa), para sumarizar e analisar estudos já realizados sobre o tema em questão.

P – População: Artigos em plataformas nacionais e estrangeiras.

O - Out come (desfecho): Fazer o levantamento e a análise da importância da Análise estatística implicativa (ASI) no tratamento e análise de questões da avaliação.

T - Tipo de pesquisa: Pesquisas que analisam as avaliações de grande escala, apoiando-se, entre outros tipos de ferramentas estatísticas, a ASI. Uma busca nas plataformas CAPES, ERIC, PROFMAT, BDTD, DART EUROPE, FUNDAÇÃO DIALNET, Focus Université Paris-saclay, Theses.fr, Repositório Científico de Acesso Aberto de Portugal (RCAAP), não permitiu encontrar nenhuma tese ou dissertação, cujo foco é o uso da ASI na análise de questões da avaliação de grande escala. No entanto, uma busca usando-se o google permitiu identificar onze artigos, cujo foco está na análise de questões de avaliação, usando-se, entre outras ferramentas estatísticas, a ASI. Todos esses artigos, na sua maioria (dez), foram escritos por Antoine Bodin² e colaboradores.

Nº	Autor(s)	Título	Ano de publicação
01	Aquino, Rafael Santos de, Carneiro-Leão, Ana Maria dos Anjos, Acioly-Régnier, Nadja Maria, Andrade, Vladimir Lira Véras Xavier de	Análise Estatística Implicativa e Análise de Variância: estudo estatístico comparativo sobre o desempenho escolar em sala de aula multicultural	2021
02	Bodin, Antoine	L'évaluation du savoir mathématique,	1989
03	Bodin, Antoine	Problèmes didactiques de l'évaluation du savoir mathématique.	1994
04	Bodin, Antoine	L'évaluation du savoir mathématique : Questions et méthodes.	1997
05	Bodin, Antoine	. Ce qui est vraiment évalué par PISA en mathématiques. Ce qui ne l'est pas. Un point de vue français	2017
06	Bodin, Antoine et Giovannini, Maria Lucia	L'utilisation de l'analyse implicative et cohésitive pour exploiter un test standardisé de mathématique	2017
07	Bodin, Antoine et Grapin, Nadine	Un regard didactique sur les évaluations du PISA et de la TIMSS : mieux les comprendre pour mieux les exploiter	2018

08	Bodin, Antoine e Almouloud, Saddo Ag	Traitement de données et évaluation	1991
09	Bodin, Antoine, Trouche, Lu	Penser PISA dans la perspective du renouvellement de l'enseignement des mathématiques en Algérie	2019
10	Gras, Régis et Bodin. Antoine	L'A.S.I., analyseur et révélateur de la complexité cognitive taxonomique	2017
11	Khaled, Hayette; Bodin, Antoine; Couturier, Raphael	L'analyse statistique implicative (ASI) appliquée aux données de programme international pour le suivi des acquis des élèves (PISA), 2017.	2017

Tabela 2 – Lista das obras encontradas

Três artigos (2, 3 e 5) do mesmo autor foram excluídos dessa análise, pois tratam da avaliação de forma geral não implicando o uso da ASI.

Na próxima seção, iniciamos a metassíntese dos diferentes artigos, tendo por base, dois grupos de questões. O primeiro grupo de questões diz respeito à estrutura de questões propostas pelas avaliações: Quais são os principais fatores a considerar para garantir que os métodos de avaliação refletem com precisão o conhecimento e as habilidades dos alunos? Quais são as possíveis desvantagens do uso de testes padronizados para avaliação? Como os professores podem equilibrar a necessidade de avaliação objetiva com o desejo de valorizar o progresso e as qualidades individuais dos alunos? Quais são os potenciais vantagens e limitações da abordagem de "multi avaliação" em comparação com os métodos tradicionais de avaliação? Etc.

O segundo grupo de questões diz respeito ao uso de ferramentas estatísticas, em especial da ASI, na análise dos achados oriundos das avaliações tratadas nas diferentes pesquisas analisadas: Quais são os potenciais benefícios do uso de métodos modernos de análise de dados para melhorar a qualidade da avaliação em matemática? O que é análise implicativa e como ela pode ser usada na avaliação do ensino da matemática? Quais são os principais benefícios de usar análises implicativa e coesitiva para examinar os resultados de um teste diagnóstico padronizado? Como os estudos validam os resultados da análise ASI em comparação com outros métodos estabelecidos de avaliação educacional?

Essas perguntas são respondidas nas diversas análises dos artigos.

² Pesquisador em didática e avaliação, Antoine Bodin é responsável por um grupo de pesquisa e desenvolvimento no Institut de Recherche sur l'Enseignement des Mathématiques d'Aix-Marseille, focado na avaliação do núcleo comum de conhecimentos e habilidades. Como consultor da Escola Segunda Chance em Marselha para auxiliá-la na renovação de seu currículo e na melhoria de sua plataforma de formação, ele também é

responsável por uma equipe de pesquisadores e professores, organizando e conduzindo cursos de formação para professores, além de produzir documentos para a formação e avaliação dos alunos. Ele é Diretor do CEDEC (Centro para o Desenvolvimento e Avaliação de Habilidades) <http://www.cedec-e2c.fr/>, além de autor e desenvolvedor de um projeto vencedor de seleção nacional de investimentos para o futuro financiados pelo Estado.

3. METASSÍNTESE DOS ARTIGOS SELECIONADOS

Iniciamos nossa análise pelo artigo de Bodin e Almouloud (1991). Este artigo é uma publicação do Instituto de Matemática e Ciência da Computação de Rennes, que trata de um estudo sobre processamento e avaliação de dados. Este texto está relacionado a uma oficina sobre o uso da ASI para analisar os dados coletados durante um experimento de "multi avaliação" realizado no IREM³ de Besançon, envolvendo 300 estudantes do ensino médio. A multi avaliação é um estudo do conhecimento dos alunos baseado em um conjunto de diferentes tipos de testes, mas sujeito ao mesmo tema de ensino ou área conceitual. Métodos estatísticos clássicos (ordenação plana, análises de correlação) já permitem fazer um certo número de observações e propor algumas interpretações didáticas. No entanto, diante de uma ou mais situações organizadas para coletar informações sobre o conhecimento de um grupo de estudantes (avaliação), surgem várias questões: Os dados obtidos podem contribuir para melhorar nosso conhecimento sobre a organização do conhecimento individual, sobre dependências e hierarquias? Seria possível obter as mesmas informações de forma mais econômica, removendo partes do interrogatório? Os dados tornam

possível afirmar algo coerente sobre o conhecimento de um aluno específico (que é o problema da avaliação pedagógica)?

Os métodos de análise estatística implicativa desenvolvidos por Gras (1979) e sua equipe trazem novas perspectivas para esse tipo de pesquisa, especialmente porque o método de análise implicativa entre itens agora é complementado por um método de análise implicativa entre classes.

Os autores dessa publicação oferecem uma visão geral das complexidades e desafios de avaliar o conhecimento dos alunos, e como métodos estatísticos avançados podem contribuir para uma compreensão mais profunda das estruturas e dependências subjacentes.

O gráfico das implicações (Figura 1) entre elementos foi estudado, mas ainda há trabalho a ser feito em relação às implicações interclasses. É certo que o número de itens aqui é muito importante e que, para um primeiro workshop, seria mais sensato se limitar a cerca de quinze itens (o que foi feito em um workshop paralelo). Aqui, o objetivo era trabalhar em um plano de avaliação necessariamente complexo e ver o que a análise implicativa poderia trazer que não fosse já fornecido por outros métodos.

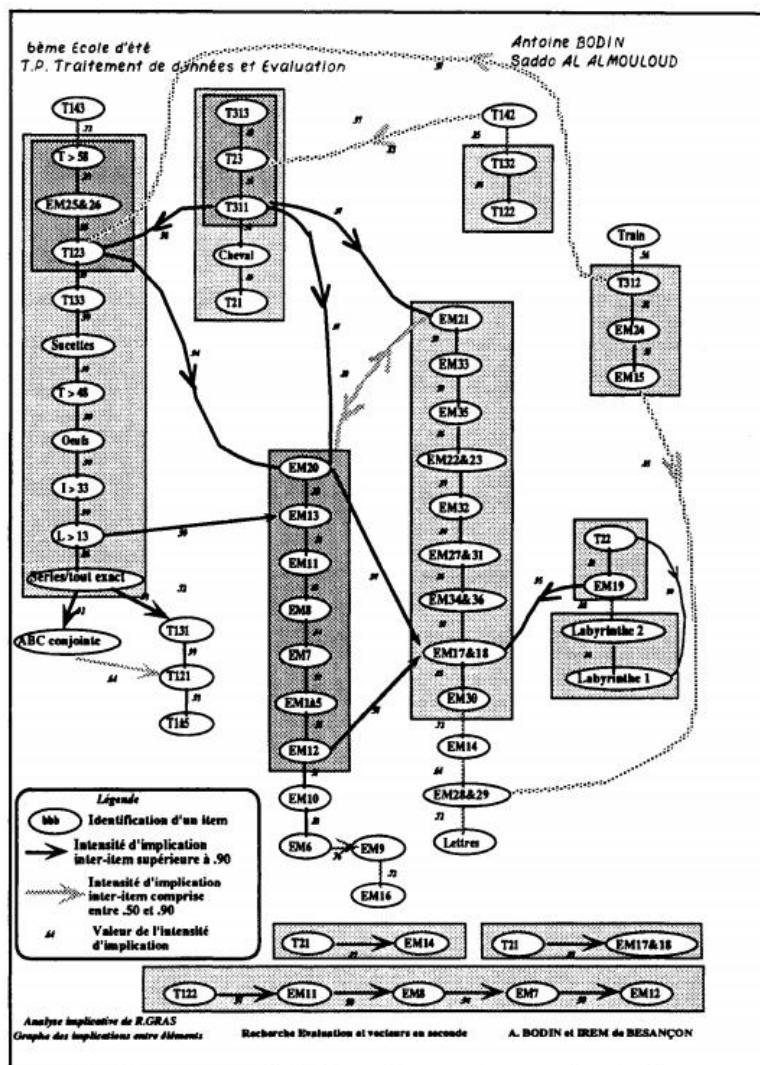


Figura 1: Grafo implicativo. Fonte: Bodin e Almouloud (1991, p.125).

Bodin (1997) explora a avaliação do conhecimento matemático no contexto da didática, examinando as teorias e métodos de avaliação. A pesquisa em didática matemática foca nos fenômenos e sistemas relacionados à avaliação e em como ela interage com o contrato didático. A questão do significado e da medição é levantada, ressaltando a importância de compreender o significado do conhecimento matemático além das capacidades mensuráveis.

As distinções entre avaliação somativa e formativa são abordadas, assim como o impacto do contrato didático nas práticas de avaliação. O autor critica métodos tradicionais de avaliação, incluindo testes padronizados, e propõe alternativas teóricas como análise implicativa e teoria da resposta a itens (TRI). Esses métodos permitem uma exploração detalhada das respostas dos estudantes, ajudando a construir escalas de competências independentes do grupo avaliado.

Exemplos concretos são usados para ilustrar os conceitos abordados, como a análise de questões específicas sobre geometria e sua aplicação em diferentes níveis escolares. A importância do contexto, das inferências e da generalização na avaliação é discutida, com uma advertência para não anular o significado do conhecimento relacionado à adaptação excessiva ao exame.

Este artigo confronta os desafios da avaliação com os da didática da matemática. Define um problema específico para avaliação e propõe caminhos para pesquisa. Os pontos-chave abordados são: os atos e fenômenos relacionados à avaliação, as diferentes concepções do campo, a tensão entre julgar e acompanhar o aprendizado, a ligação com o contrato didático, o papel na comunicação e negociação, bem como o dilema entre medição e significado.

O autor aponta cinco grandes desafios para integrar a avaliação à didática da matemática: primeiro, as avaliações funcionam quase isoladas do restante do sistema didático; segundo, ao focar apenas nos fenômenos, corre-se o risco de separar atos, fenômenos e fatos; terceiro, as práticas avaliativas variam entre julgar, compreender ou apoiar o estudante, alternando comunicação e definição de objetivos; quarto, existe uma tensão entre tentar medir tudo e buscar dar significado às observações; por fim, é fundamental incluir a avaliação na análise didática, mas sem deixar de lado aspectos essenciais.

A integração da avaliação à didática matemática enfrenta desafios relacionados à independência das práticas de avaliação, à necessidade de vincular atos, fatos e fenômenos, e à tensão entre diferentes objetivos e abordagens de avaliação. É necessária uma abordagem cuidadosa e abrangente para garantir uma integração bem-sucedida.

O autor argumenta que a influência da avaliação sobre o contrato didático, mesmo sob uma perspectiva sumativa, frequentemente percebida como distante do aprendizado instantâneo, ajuda a moldar o contrato didático e pode antecipar certos aspectos dele. Quando a avaliação é integrada às situações de aprendizagem e promove o desenvolvimento dos alunos, ela está totalmente alinhada com o contrato didático, tornando-se inseparável dos princípios da disciplina.

Quanto à inseparabilidade entre ensino e avaliação, Bodin (1997) nos assegura que, na matemática, especialmente no contexto escolar, a fronteira entre ensino e avaliação permanece tênue. A resolução de problemas desempenha simultaneamente o papel de fonte de conhecimento e de critério de apreciação. Práticas de verificação de pontuação e procedimentos de avaliação estão incorporados ao núcleo da atividade matemática, alinhando-se com competências específicas de cada domínio e com as exigências dos alunos. O processo de validação dos resultados e avaliação das abordagens utilizadas faz parte da atividade matemática e faz parte da prática atual dos profissionais do setor.

Podemos, portanto, dizer que o impacto da avaliação no contrato didático varia desde sua influência preventiva por meio de avaliações somativas até sua plena integração nas atividades de aprendizagem. Ensino e avaliação estão intimamente ligados, especialmente em matemática, onde a resolução de problemas e a verificação de resultados são componentes inerentes.

No que diz respeito às vantagens potenciais do uso de métodos modernos de análise de dados para aprimorar a qualidade da avaliação em matemática, Bodin (1997) afirma que esses métodos modernos de análise de dados, como a ASI, oferecem perspectivas promissoras para melhorar a avaliação matemática. Essas técnicas podem ajudar a construir hierarquias suficientemente robustas em diversas situações, permitindo a criação de escalas diagnósticas, por exemplo, em geometria espacial, cálculo vetorial e proporcionalidade. Mesmo os testes que não foram inicialmente concebidos para esse fim, como exames e concursos, podem ser analisados com essas metodologias para derivar escalas e avaliar sua eficácia.

Este artigo analisa a tensão entre avaliar os resultados de aprendizagem e compreender sua importância no ensino da matemática, destacando as limitações dos métodos tradicionais de avaliação baseados em medição. O autor recomenda que a dimensão didática e o significado do conhecimento matemático tenham prioridade durante a avaliação. Defende uma tendência para sistemas de avaliação que enfatizam a compreensão e o significado, em vez da simples quantificação do desempenho, com o objetivo de melhorar a qualidade pedagógica. Essa abordagem envolve uma reflexão aprofundada sobre todo o processo de avaliação para garantir coerência entre didática, medição e significado, evitando qualquer compromisso que possa reduzir a relevância do dispositivo. Além disso, é importante utilizar os caminhos identificados com discernimento, mantendo-se atento às questões ideológicas latentes no campo da avaliação.

O artigo argumenta que comunicar objetivos de aprendizagem às vezes leva os alunos a focarem no sucesso nas avaliações em vez da compreensão real, o que pode tornar o conhecimento sem sentido. Essa adaptação transforma o conhecimento, inicialmente gerenciado pelos professores, em conhecimento centrado no aluno, mas às custas do significado mais profundo do conteúdo.

Concluímos, portanto, que os métodos alternativos à medição tradicional ajudam a colocar a medida a serviço do significado na avaliação. Nesse ponto de vista, a ASI é um

método de análise de dados que pode ser usado na avaliação do ensino da matemática para construir hierarquias suficientemente fortes em diversas situações, o que permite a construção de escalas diagnósticas em áreas como geometria espacial, cálculo vetorial, proporcionalidade etc. O método ajuda a mostrar como as diferentes perguntas de avaliação se relacionam entre si, fornecendo uma visão sobre a estrutura do conhecimento e das habilidades avaliadas.

Bodin e Giovannini (2017) apresentam uma análise de dados de um teste diagnóstico padronizado de matemática administrado no início do 6º ano na região de Bolonha, Itália. A análise implicativa e coesitiva possibilitou examinar os resultados desse teste de forma original e complementar a outras análises, identificando os seguintes aspectos: ASI permitiu destacar a estrutura íntima do teste e a estrutura cognitiva dos alunos que fizeram o teste; explorar a complexidade do item, a diferença entre complexidade e dificuldade, e o uso da análise implicativa para todo o teste, bem como para algumas análises parciais, especialmente em relação às capacidades definidas pelos desenvolvedores do teste; analisar as relações entre itens do teste e percepções dos alunos sobre dificuldade e desempenho no teste.

Em resumo, a ASI promoveu uma compreensão mais profunda dos resultados, dos processos cognitivos dos alunos e dos princípios do desenho dos testes, agregando assim valor ao desenvolvimento profissional dos professores e à melhoria contínua das ferramentas de avaliação.

Khaled, Bodin, Couturier (2017), em seu artigo "Análise Estatística Implicativa (ASI) aplicada aos dados do Programa Internacional para Avaliação de Estudantes (PISA)", exploram a aplicação da ASI aos dados PISA, em particular os da literacia Matemática do PISA 2012, para comparar as estruturas cognitivas de estudantes em quatro países (França, Alemanha, Japão, Itália). O ASI é usado para analisar as dependências entre questões e comparar as organizações do conhecimento dos alunos em diferentes países. Os resultados revelam semelhanças e diferenças nos comportamentos cognitivos por meio de uma análise detalhada das cadeias explicativas entre as categorias "Espaço e Formas", "Incerteza e Dados", "Quantidade" e "Variações e Relações".

O estudo mostra que, apesar de objetivos semelhantes a avaliações como a PISA, há variação significativa na ordem das implicações e na complexidade cognitiva das perguntas entre países. Ela lança luz sobre a importância das diferenças curriculares e metodológicas, como revelado pela densidade de envolvimento. Embora a hierarquia geral da complexidade seja respeitada, algumas perguntas não seguem essa hierarquia, sugerindo possíveis diferenças nas estruturas cognitivas dos alunos.

Os autores destacam os desafios relacionados ao acesso a dados e a necessidade de aprimorar ferramentas (como o RCHIC-PISA) para lidar com matrizes de implicação. Em conclusão, a ASI oferece um método inovador para a análise dos resultados internacionais de pesquisas internacionais como a PISA, ao mesmo tempo em que destaca a necessidade de pesquisas didáticas mais detalhadas e coletivas para melhor interpretar as diferenças observadas.

As oportunidades incluem o uso de dados internacionais para benchmarking contínuo, inclusive com dados TIMSS.

Os autores argumentam que, embora o PISA seja amplamente discutido internacionalmente, há poucos estudos que examinem a dimensão cognitiva das questões feitas. Portanto, buscam aprofundar a compreensão da estruturação do conhecimento dos estudantes em diferentes países por meio da ASI.

Uma descrição detalhada do banco de dados PISA e dos obstáculos relacionados à sua análise levou os pesquisadores a projetarem um software específico sob R (RCHIC-PISA) para analisar os dados e aplicar o ASI.

O estudo destaca as ligações implicativas entre os itens do teste PISA em quatro países (França, Alemanha, Itália, Japão). Examina a dificuldade e complexidade das perguntas e compara as estruturas cognitivas usando gráficos explicativos, destacando semelhanças e divergências. Os autores apontam que o acesso limitado aos itens do PISA dificulta análises aprofundadas, mas sugerem estender sua abordagem para outras avaliações internacionais, como o TIMSS.

As principais diferenças observadas nas estruturas cognitivas dos estudantes nos países estudados, baseadas na Análise Estatística Implicativa, são destacadas pela densidade de gráficos explicativos e cadeias explicativas específicas. França e Alemanha têm densidades semelhantes, indicando uma semelhança nos comportamentos cognitivos. A Itália difere consideravelmente da França, sugerindo comportamentos cognitivos mais homogêneos nos estudantes italianos. O Japão difere um pouco dos outros três países.

A análise das cadeias implicativas revela diferenças na estrutura cognitiva. Por exemplo, uma cadeia mais longa foi identificada no caso japonês na categoria "Variações e Relações", mas apenas partes dessa cadeia foram encontradas em outros países. Essas diferenças estão relacionadas à alfabetização matemática entre alunos na França, Alemanha, Itália e Japão, dependendo da densidade dos grafos implicativos e da presença de cadeias implicativas específicas.

O estudo destaca o problema das disparidades linguísticas nos itens do PISA entre diferentes países. Apesar da sofisticação dos procedimentos implementados pelo PISA, os autores observam a dificuldade de garantir a equivalência total entre as várias versões linguísticas dos testes. O estudo sugere que uma análise aprofundada poderia examinar o impacto das diferenças curriculares, abordagens pedagógicas e de avaliação, bem como a formação de professores, baseando-se em dados de questionários contextuais.

O software RCHIC-PISA otimiza a aplicação do método ASI a grandes conjuntos de dados como o PISA, especialmente por meio do processamento direto das matrizes de implicação e confiança. Ele foi especificamente projetado para explorar diretamente essas matrizes, obtidas a partir dos dados PISA, eliminando a necessidade de dados brutos e, assim, simplificando o processo analítico. Adaptado à estrutura particular dos bancos de dados PISA — caracterizados por configurações complexas como

livretos rotativos — o RCHIC-PISA garante uma análise de acordo com as especificidades desses dados. Desenvolvido sob R, ele integra módulos permitindo o gerenciamento eficiente das matrizes de envolvimento e confiança, tornando possível examinar grandes volumes de dados sem sobrecarga de informações.

Os autores enfatizam que o RCHIC-PISA aumenta significativamente a eficácia da abordagem ASI aplicada aos dados do PISA ao trabalhar em matrizes de implicação, levando em conta a complexidade estrutural inerente ao estudo e facilitando o processamento ótimo das informações. O PISA-CHCR determina a força das relações implicativas entre as perguntas usando um índice de engajamento de 0,99 (para atingir vícios altamente confiáveis) e um índice de confiança de 0,80 (para garantir que elas abordem uma proporção significativa dos estudantes).

O RCHIC-PISA aborda o problema de dados ausentes e não resposta em conjunto das questões de PISA focando nas informações que realmente estão disponíveis. Diferentemente da abordagem psicométrica da PISA, que atribui valores a estudantes que não responderam a certas perguntas, a RCHIC-PISA depende exclusivamente de respostas autênticas fornecidas pelos alunos. Ele elabora pistas explicativas examinando os pares de perguntas {A; B} ao qual o mesmo grupo de alunos respondeu. Graças a esse método, apenas as respostas dos alunos que lidaram com as duas perguntas em um par são analisadas. A partir dessas comparações, o RCHIC-PISA cria matrizes de envolvimento e confiança, que ele explora diretamente sem retornar aos dados originais de cada estudante.

O estudo confirma a validade das análises ASI ao compará-las com outros métodos reconhecidos na avaliação educacional. Os autores destacam a complementaridade entre ASI e TRI, um método amplamente utilizado para criar e calibrar testes educacionais. Essa comparação possibilita verificar a consistência dos resultados produzidos pela ASI com os do TRI. O TRI se refere à metodologia psicométrica tradicional PISA, incluindo a imputação de valores ausentes. Embora essa técnica seja criticada pelos autores por seu contexto específico, ela serve como referência para situar a ASI em relação aos métodos padrão usados em pesquisas de avaliação de grande escala. O processo de validação baseia-se na paralelização da ASI e do TRI, bem como na discussão dos métodos psicométricos do PISA, a fim de mostrar o interesse e a contribuição complementar da ASI no campo da avaliação educacional.

Os autores destacam **os possíveis vieses introduzidos** ao focar apenas em "dados reais" no RCHIC-PISA, e como esses vieses podem afetar as relações implicativas descobertas. Por exemplo, ao focar apenas em alunos que responderam às duas perguntas em par, o RCHIC-PISA pode introduzir viés de seleitividade amostral. Isso significa que os estudantes incluídos na análise podem não ser totalmente representativos de toda a amostra PISA, pois são estudantes que tinham maior probabilidade de tentar responder ambas as perguntas. Alunos que pulam questões por falta de conhecimento, tempo ou interesse são excluídos, o que pode distorcer relações implicativas. Os autores também argumentam que descartar as não-respostas significa que a informação sobre o porquê os alunos podem

ter deixado certas perguntas é perdida. A falta de resposta pode indicar a dificuldade das perguntas, falta de familiaridade com o tema ou outros fatores que influenciam o comportamento dos alunos. Ao analisar apenas os "dados reais", a análise negligencia essa informação potencial. As relações implicativas descobertas podem ser tendenciosas a favor das relações mais evidentes nos alunos de melhor desempenho ou mais engajados (aqueles que têm mais probabilidade de responder todas as perguntas). Isso pode levar a uma superestimação da força ou relevância de certas conexões cognitivas, já que a análise não leva em conta os padrões cognitivos de alunos que têm dificuldades ou que se desengajam da avaliação.

A alteração dos limiares dos índices de implicação e confiança no RCHIC-PISA teria um impacto significativo nos gráficos implicativos obtidos e nas conclusões tiradas. Por exemplo, a redução do limiar do índice de implicação resultaria em um maior número de conexões nos gráficos implicativos, uma vez que dependências mais fracas seriam incluídas. Isso poderia revelar relações sutis, mas também poderia introduzir ruído ou associações enganosas. Por outro lado, o aumento do limiar levaria a gráficos mais esparsos, destacando apenas as dependências mais fortes e confiáveis. Um limiar mais baixo incluiria relações baseadas em subconjuntos menores de alunos, o que permitiria descobrir padrões específicos de certos subgrupos, mas também aumentaria o risco de incluir relações menos sólidas. Um limiar mais alto garantiria que apenas as relações apoiadas por uma parte mais significativa e consistente da população estudantil fossem consideradas, o que levaria a resultados mais generalizáveis, mas potencialmente menos matizados. Por fim, a escolha dos limites pode influenciar as conclusões sobre as estruturas cognitivas dos alunos. Limites mais elevados oferecem uma visão mais rigorosa, enfatizando apenas as dependências altamente confiáveis. Limites mais baixos oferecem uma visão mais exploratória, revelando potencialmente uma gama mais ampla de relações cognitivas, com o risco de incluir padrões menos estáveis.

Os autores reconhecem o potencial de viés cultural na interpretação dos resultados do PISA em diferentes países. Eles mostram que é difícil obter uma equivalência completa entre as diferentes versões linguísticas das perguntas, apesar dos procedimentos complexos utilizados pelo PISA para garantir a comparabilidade. Eles sugerem que pesquisas futuras poderiam aprofundar as diferenças entre os programas e a pedagogia, a fim de fornecer uma compreensão mais completa. Isso inclui examinar os efeitos das diferenças curriculares, dos métodos de ensino e avaliação e da formação dos professores, usando dados de questionários contextuais.

De acordo com este estudo, as variações nos métodos de ensino entre os diferentes países podem explicar as diferenças observadas nas estruturas cognitivas. Sugere-se que pesquisas futuras se concentrem na influência do conteúdo dos programas, nas abordagens pedagógicas e na formação dos professores, através da análise de dados provenientes de questionários contextuais. Essa abordagem implica reconhecer que as práticas educativas variadas entre as nações podem afetar os resultados obtidos no PISA. O estudo também recomenda a continuação das investigações

sobre as diferenças curriculares e pedagógicas para compreender melhor o impacto das metodologias de ensino no desempenho do PISA.

Este estudo destaca algumas limitações inerentes à abordagem do teste padronizado do PISA para compreender a diversidade cognitiva entre as culturas. É complexo garantir uma equivalência perfeita entre as diferentes versões linguísticas dos itens, o que implica que as especificidades culturais próprias de cada idioma podem influenciar a compreensão e as respostas dos alunos de acordo com seu contexto nacional.

Gras e Bodin (2017) exploram o uso da Análise Estatística Implicativa (ASI) para validar uma hierarquia de complexidade em exercícios de matemática para jovens estudantes, baseada em uma taxonomia de aquisições cognitivas. A taxonomia em análise, apoiada pelos resultados de pesquisas nacionais e internacionais, é estruturada em cinco níveis (de A a E), cada um correspondendo a habilidades matemáticas específicas que vão desde o conhecimento de ferramentas de captação até avaliação crítica. Os autores descrevem esses níveis da seguinte forma:

Nível A: Conhecimento das ferramentas de compreensão do objeto e do fato matemático - *Associar, montar, simular, observar, decifrar, descrever, organizar, calcular*

Nível B: Análise dos fatos e transposição - *Abstrair, prolongar, induzir, analisar, comparar, esquematizar, traduzir, transpor*

Nível C: Compreensão das relações e estruturas - *Reconhecer, construir, justificar, deduzir, analisar, abstrair, aplicar, interpolar*

Nível D: Síntese e criatividade - Organizar, calcular, otimizar, ilustrar, demonstrar, validar, criar, inventar, generalizar, induzir, prever, extrapolar

Nível E: Crítica e avaliação - Controlar, otimizar, prever, criticar, questionar, verificar, tolerar.

O artigo examina como a complexidade cognitiva, avaliada pela ASI, contribui para validar e detectar competências em estudos como o PISA e o TIMSS, destacando assim o papel fundamental dos parâmetros pedagógicos na interpretação dos resultados. Graças às suas representações gráficas, a ASI consegue extrair uma pré-ordem que corresponde quase perfeitamente à taxonomia. Para ajudar os professores, é proposta a taxonomia de Bodin, que adapta a de R. Gras: ela

oferece uma visão detalhada sobre como organizar as competências cognitivas no campo da educação.

Os autores identificam as principais características da taxonomia das competências cognitivas em matemática desenvolvida neste estudo. Trata-se, nomeadamente, das seguintes características:

1. A taxonomia é composta por cinco níveis, identificados de A a E, que representam níveis crescentes de complexidade cognitiva nas competências matemáticas.
2. Os níveis são ordenados de forma que as competências de um nível superior (por exemplo, E) são consideradas mais complexas do que as dos níveis inferiores (por exemplo, A), formando assim uma hierarquia.
3. A taxonomia é especificamente adaptada às competências matemáticas, o que a diferencia de taxonomias gerais como a de Bloom.
4. Ao contrário de algumas taxonomias baseadas em consenso, esta é fundamentada em reflexões didáticas e validações empíricas.
5. A complexidade é definida pelo grau de mobilização e complexidade dos processos de pensamento necessários para resolver um determinado problema.

Cada nível está associado a verbos de ação específicos que ajudam a definir e compreender as competências envolvidas nesse nível.

O nível **A** concentra-se no conhecimento das ferramentas básicas para compreender objetos e fatos matemáticos, envolvendo associações e operações simples. O nível **B** envolve a análise de fatos e a transposição, exigindo a capacidade de coordenação e tradução entre diferentes representações e linguagens. O nível **C** concentra-se na compreensão de relações e estruturas, permitindo a identificação e reestruturação de conceitos. O nível **D** enfatiza a síntese e a criatividade, onde as competências envolvem a organização, a otimização e a criação de novas soluções. Quanto ao nível **E**, ele se concentra na crítica e na avaliação, exigindo a capacidade de controlar, otimizar e questionar os resultados matemáticos.

Os autores afirmam que o método ASI estruturou os elementos em um gráfico dirigido com base na frequência das respostas corretas. Esse gráfico mostrava uma ordem hierárquica das taxas de sucesso, alinhando-se aos níveis de complexidade da taxonomia. Se a ordem da taxonomia fosse respeitada, os elementos passariam do mais complexo (E) para o mais simples (A) ao longo de cada caminho (Figura 2).

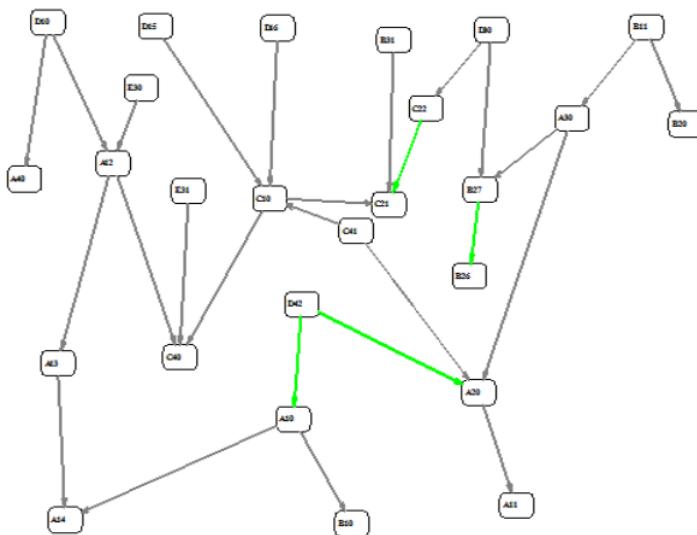


Figura 2- Gráfico implicativo no limiar de 0,79. Fonte: Gras e Bodin (2017, p. 506)

A ASI também permitiu identificar distorções na ordem esperada, revelando casos em que a complexidade real de um elemento diferia de sua classificação a priori. Por exemplo, a análise evidenciou que a questão A12 (Figura 3) exigia uma resolução mais complexa do que se pensava inicialmente.

A ASI gerou uma hierarquia coesitiva que agrupa as respostas bem-sucedidas em classes orientadas, reforçando

assim a ordem presumida da taxonomia. Essa hierarquia respeitou amplamente a ordem taxonômica, com algumas exceções menores relacionadas a elementos de complexidade semelhante. Ao analisar as cadeias e as árvores implicativas, a ASI permitiu validar se um determinado conjunto de elementos poderia ser utilizado para avaliar uma competência específica. Também revelou competências inesperadas que não tinham sido consideradas inicialmente.

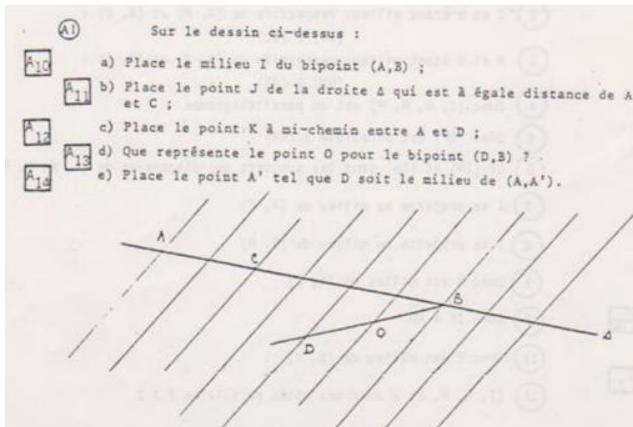


Figura 3: exemplo de prova. Fonte: os autores

Podemos, portanto, dizer que o ASI atuou tanto como analisador quanto como revelador, validando a estrutura hierárquica da taxonomia e revelando nuances na complexidade dos itens e nas relações de competência.

Ao aplicar a taxonomia e a ASI à análise dos dados de avaliação nacionais e internacionais, os autores destacam que a distribuição das questões do PISA e do TIMSS em função dos níveis taxonómicos foi revelada. Isso permitiu comparar a complexidade cognitiva entre diferentes avaliações e programas de estudos. Eles observam [também que o programa de estudos francês se aproximou dos conceitos de complexidade cognitiva do PISA/OCDE desde 2006. As comparações do desempenho de diferentes países no PISA 2012 com base no nível taxonômico das questões mostraram resultados inesperados. Por exemplo, os alunos franceses superaram os alunos japoneses no nível de

conhecimento (nível A), mas tiveram um desempenho inferior no nível de compreensão (nível B). A abordagem permitiu avaliar se um conjunto de elementos poderia avaliar uma competência específica, bem como identificar competências inesperadas.

Bodin e Grapin (2018) apresentam e compararam duas abordagens didáticas distintas para analisar o conteúdo e interpretar os resultados das avaliações internacionais PISA e TIMSS em matemática. A primeira abordagem, conhecida como "cognitiva e didática", está associada à análise estatística implicativa e permite uma nova análise dos resultados do PISA, bem como comparações entre países. A segunda abordagem, entrelaçada com a teoria antropológica do didático, foca principalmente no conteúdo do TIMSS e na distribuição de seus elementos de avaliação em relação a um determinado currículo de matemático.

Os autores destacam a importância da realização de análises didáticas dos conteúdos das avaliações TIMSS e PISA, de modo a possibilitar uma abordagem distinta daquela adotada nos respectivos quadros de referência. Ressaltam também o valor de estudos complementares, como aqueles fundamentados em análise estatística implicativa. Além disso, abordam o impacto potencial dessas avaliações sobre as práticas pedagógicas vigentes.

Esses autores asseveram que a análise estatística implicativa desempenha um papel importante na interpretação dos dados do PISA, pois permite examinar a estrutura global de um conjunto de dados, bem como as dependências orientadas entre variáveis. Este método oferece a possibilidade de estudar o desempenho dos alunos em diferentes itens, bem como as relações entre esses resultados. Permite também identificar as ligações existentes entre os diversos aspectos avaliados pelo PISA. A ASI facilita a identificação de cadeias comparáveis às escalas de Guttman, tornando possível monitorar o progresso dos alunos no desenvolvimento de habilidades específicas. Também ajuda a destacar as semelhanças entre organizações do conhecimento em diferentes países; por exemplo, esse método revela uma proximidade maior entre França e Alemanha do que entre França e Itália ou Japão.

A complexidade cognitiva das questões PISA é analisada usando a taxonomia de Gras-Bodin. Lembramos que essa taxonomia é hierárquica, o que significa que uma pergunta de nível A deve ser mais bem-sucedida do que uma de nível C, já que esta última assume domínio dos níveis inferiores. No entanto, esse princípio deve ser qualificado, pois é ainda mais verdadeiro se as tarefas envolvidas se enquadram no mesmo campo conceitual.

Segundo os autores, a abordagem epistemológica e didática é usada para analisar avaliações externas como PISA e TIMSS, tornando possível definir uma organização matemática de referência a partir do conhecimento em jogo e dos elementos epistemológicos nos quais ela se baseia. Essa organização estrutura o domínio matemático avaliado listando os tipos de tarefas, as técnicas para resolvê-las e os elementos matemáticos (propriedades, teoremas) que justificam essas técnicas.

Em relação à abordagem epistemológica e didática, os autores afirmam que a análise a priori da tarefa é um elemento-chave dessa metodologia. Ela possibilita verificar se a tarefa proposta corresponde ao objetivo de avaliação atribuído. Diferentes parâmetros, como a escolha dos valores das variáveis didáticas, podem ser analisados para evitar enviesar o objetivo da tarefa. Essa abordagem também possibilita elaborar um inventário do conhecimento dos alunos e colocá-los em perspectiva com os currículos escolares e as práticas de ensino. Essa abordagem pode ajudar a identificar possíveis vieses em uma tarefa de avaliação, verificando se ela corresponde ao objetivo de avaliação atribuído. A análise a priori da tarefa permite garantir que a tarefa proposta corresponda ao objetivo pretendido da avaliação. Isso inclui verificar se os parâmetros das tarefas, como a escolha das variáveis didáticas, não alteram esse objetivo.

Para o TIMSS, que visa avaliar o conhecimento dos alunos sobre os currículos escolares, Bodin e Grapin (2018)

afirmam que o uso de uma estrutura epistemológica e antropológica é particularmente apropriado. A definição de uma organização de referência matemática, independente dos currículos nacionais, possibilita estudar o conteúdo da avaliação e interpretar os resultados obtidos em diferentes países em relação aos seus respectivos currículos.

Os autores mostram que a adoção de uma organização de referência matemática independente dos currículos nacionais para a análise do TIMSS possibilita examinar o conteúdo avaliativo de acordo com um quadro imparcial, sem ser afetado pelas especificidades dos currículos escolares de cada país. Essa abordagem também oferece a possibilidade de interpretar os resultados obtidos em escala internacional, colocando-os em perspectiva com os respectivos currículos, facilitando assim a comparação e análise do desempenho dos alunos. A definição de tal organização matemática constitui um arcabouço de referência objetivo voltado para avaliar as conquistas dos alunos independentemente das diferenças curriculares.

Segundo Bodin e Grapin (2018), analisar avaliações como PISA e TIMSS sob a ótica das metodologias didáticas permite compreender melhor o que realmente medem, indo além de rankings ou questões edumétricas. Essa abordagem avalia a validade didática e epistemológica dos testes, contextualiza as aprendizagens em relação ao currículo e práticas pedagógicas, e oferece subsídios para a formação de professores ao elucidar o impacto e relevância dessas avaliações no ensino.

Rafael Santos de Aquino, Ana Maria dos Anjos Carneiro-Leão, Nadja Maria Acioly-Régnier e Vladimir Lira Véras Xavier de Andrade (2021) examinam o impacto do multiculturalismo no desempenho escolar utilizando a análise estatística implicativa (ASI) e a análise de variância (ANOVA) como metodologias analíticas. O estudo realizado no Instituto Federal do Sertão Pernambucano, tem por objetivos são os seguintes: explorar as diferenças de desempenho escolar entre os diferentes grupos culturais de alunos; analisar o impacto dos fatores culturais no desempenho escolar; aplicar simultaneamente os métodos estatísticos ANOVA e A.S.I. para obter uma visão completa dos dados; e propor recomendações para a integração de práticas educativas multiculturais e decoloniais.

Os autores analisam o desempenho escolar diferenciado entre alunos de diferentes culturas: sertanejos, urbanos, indígenas e quilombolas, destacando que os quilombolas apresentam baixo desempenho e taxas de evasão mais elevadas. Além disso, investiga como os elementos culturais afetam o desempenho escolar, desde a formação dos professores até as estratégias pedagógicas escolares.

Os dados analisados são as notas escolares coletadas dos alunos do ensino médio integrado ao técnico agrícola do Instituto Federal de Educação, Ciência e Tecnologia do Sertão Pernambucano (IFSertãoPE).

Por meio do CHIC, os gráficos implicativos mostram que os alunos quilombolas estão associados à classe D e ao maior abandono escolar, enquanto os sertanejos apresentam melhor desempenho relacionado à classe A. A ANOVA revela diferenças estatisticamente significativas entre os grupos étnicos e culturais, confirmado a variabilidade do desempenho escolar.

Os autores asseveram que a ASI traz uma análise dinâmica e probabilística, enquanto a ANOVA garante a confiabilidade por meio de seus parâmetros. O desempenho escolar destaca as insuficiências estruturais da instituição em atender às necessidades de formação. Ao combinar os métodos ANOVA e ASI, é possível obter uma visão mais abrangente sobre como as origens culturais dos alunos afetam seu desempenho escolar, já que ambos proporcionam análises complementares. Enquanto a ANOVA segue uma abordagem mais tradicional, a ASI adota um método probabilístico que evita conclusões deterministas, sendo especialmente adequada para áreas como a educação. O uso conjunto dessas duas técnicas enriquece a análise quantitativa, permitindo uma interpretação mais complexa e multifacetada dos dados relacionados ao rendimento escolar. A ASI permite compreender o objeto de pesquisa como um fenômeno comportamental não estático e sujeito a mudanças baseadas na influência de diversas variáveis, o que é particularmente útil na pesquisa social e educacional.

Os autores identificam os principais desafios e limites da implementação de uma abordagem educacional multicultural que responda às necessidades escolares de diversas populações estudantis. Entre esses desafios, podemos citar a marginalização das culturas tradicionais, como as das comunidades indígenas e afro-brasileiras, que são historicamente marginalizadas em favor da cultura eurocêntrica. Essa marginalização torna difícil para os educadores promoverem uma educação intercultural que atenda às necessidades históricas dessas comunidades tradicionais. Podemos também citar a segregação e o baixo rendimento de certos grupos étnicos e culturais, como os quilombolas, que muitas vezes enfrentam segregação no ambiente escolar, o que resulta em baixos resultados escolares e taxas de evasão mais elevadas. Para remediar essa situação, são necessários uma atenção e um planejamento especiais, a fim de promover a inclusão étnica e cultural.

Os autores destacam que a implementação da educação multicultural requer, além do reconhecimento da diversidade cultural, transformações estruturais nas políticas escolares e práticas pedagógicas, visando assegurar equidade social e educacional. Ressalta-se a importância de considerar aspectos sociais, econômicos e pedagógicos, para além dos dados quantitativos. O estudo aponta diversos fatores determinantes associados ao baixo desempenho escolar de alunos quilombolas e propõe estratégias para que as instituições de ensino implementem intervenções específicas.

Bodin e Trouche (2019) examinam a integração da perspectiva PISA no sistema educacional argelino para renovar o ensino da matemática e preparar o país para a avaliação de 2021. Realizada em cooperação com o Ministério da Educação Nacional da Argélia e a Embaixada da França na Argélia, a formação reuniu inspetores de vários níveis de ensino em várias sessões, com ênfase na colaboração ativa, na análise dos itens PISA, bem como no desenvolvimento e partilha de recursos pedagógicos adaptados ao contexto local.

O artigo apresenta uma análise da melhoria na compreensão das práticas pedagógicas, da relevância da reflexão sobre a avaliação das habilidades dos alunos e da introdução de

novas tecnologias (planilhas, softwares de geometria dinâmica).

Os autores também destacam a parceria internacional, nomeadamente com a ECNU (Xangai), bem como as perspectivas de traduções entre o árabe, o francês e o inglês para os próximos ciclos do PISA na Argélia (2021) nas áreas das ciências, da leitura e da resolução de problemas. Além disso, o relatório destaca e analisa as diferentes vertentes da avaliação PISA, explicando a literacia em várias áreas: compreensão escrita, matemática e científica. Ele destaca os desafios tanto a nível mundial como nacional para os estudantes e cidadãos, no ensino geral e na matemática em particular. Promove o desenvolvimento de competências globais, como o pensamento crítico e a criatividade, por meio de pedagogias baseadas em projetos e aprendizagem por meio do serviço. O quadro de referência matemático para o PISA 2021 é detalhado, incluindo categorias de conteúdo, processos de modelagem e competências-chave do século XXI, necessárias para aplicar a matemática em situações reais.

A análise estatística implicativa revelou sequências de itens em que o sucesso em um item A aumenta a probabilidade de sucesso em um item B, gerando assim uma cadeia de condições suficientes para o sucesso. Essa abordagem permitiu destacar a dependência entre os itens e mostrar que o desempenho em uma seção de uma questão pode afetar o desempenho em outra seção.

Em suma, o uso da análise estatística implicativa neste estudo revelou-se essencial para identificar as relações de dependência entre os itens, compreender as competências necessárias para o sucesso nas várias partes do teste e, potencialmente, destacar competências diferenciadas.

4. CONCLUSÕES E PERSPECTIVAS

Nesta parte, dissertamos sobre os principais resultados desta metassíntese qualitativa e as perspectivas de pesquisas sobre a análise de questões e Enade (Exame Nacional de Desempenho dos Estudantes) usando a ASI.

4.1. Principais conclusões da metassíntese

Os autores dos artigos analisados oferecem investigações detalhadas sobre avaliação do ensino de matemática, abrangendo conceitos fundamentais e metodologias avançadas de análise de dados aplicadas a avaliações nacionais e internacionais. Os trabalhos compilam estudos teóricos, práticos e estatísticos, incluindo o exame de grandes bases de dados como PISA e TIMSS, abordando aspectos didáticos, epistemológicos e culturais relacionados à avaliação em matemática.

Diversos artigos enfatizam a relevância de critérios claros e comunicáveis, privilegiando a avaliação baseada em critérios em detrimento de métodos normativos que efetuam comparações entre alunos. Entre os desafios práticos identificados destacam-se: superação de práticas consolidadas, pressões administrativas e familiares, além da limitação de tempo para implementação de avaliações diversificadas. Aponta-se também o risco de abuso de poder na atuação docente durante o processo avaliativo, reforçando a necessidade de métodos justos e alinhados aos objetivos de aprendizagem.

A distinção entre avaliação normativa e avaliação baseada em critérios é explicitada: a primeira fundamenta-se na comparação entre alunos visando uma distribuição de notas, enquanto a segunda compara o desempenho individual a objetivos previamente estabelecidos, favorecendo o progresso individual e promovendo o diálogo entre professores e estudantes.

Os riscos inerentes aos testes padronizados são ressaltados, como a limitação na aferição de habilidades complexas, possibilidade de respostas aleatórias e dependência excessiva da pontuação total.

No âmbito internacional, a Aplicação da ASI às bases de dados PISA possibilitou a comparação das estruturas cognitivas de estudantes de países como França, Alemanha, Itália e Japão. A análise evidenciou semelhanças e distinções nas correlações entre itens, refletindo possíveis variações curriculares e pedagógicas, ao passo que também reconhece limitações decorrentes de diferenças linguísticas, culturais e de interpretação de dados oriundos de avaliações padronizadas em variados contextos.

A ASI foi utilizada tanto para validar a **taxonomia das competências cognitivas em matemática** — confirmado a coerência hierárquica e identificando nuances na complexidade de elementos específicos — quanto para explorar a organização dos dados PISA, elucidando as dependências entre itens e aproximações nas estruturas cognitivas de diferentes países.

A abordagem epistemológica e didática contribui para identificar possíveis vieses e restrições nas tarefas de avaliação, como adequação dos itens aos objetivos propostos e influência de variáveis didáticas e condições de aplicação, facilitando a interpretação dos resultados relativos aos currículos nacionais e às práticas pedagógicas correspondentes.

O estudo sobre o impacto dos contextos culturais no desempenho acadêmico no Brasil, por meio de ASI e ANOVA, evidencia que estudantes quilombolas enfrentam desafios significativos, manifestados por baixo rendimento e elevada evasão escolar, indicadores de marginalização e segregação cultural. Ressalta-se a importância de práticas educativas multiculturais e decoloniais visando promoção de inclusão e equidade, recomendando intervenções específicas para apoiar esses estudantes e combater desigualdades estruturais no sistema educacional.

4.2. Perspectivas de pesquisa

Em termos de perspectivas, pretendemos analisar o **Exame Nacional de Desempenho dos Estudantes** (Enade). O Enade é um componente obrigatório do currículo do ensino superior brasileiro, essencial para avaliar o desempenho dos estudantes em relação ao programa, bem como as competências e conhecimentos adquiridos. Ele também desempenha um papel fundamental no fortalecimento das políticas públicas destinadas a garantir a qualidade do ensino superior a nível nacional.

O **Exame Nacional de Desempenho dos Estudantes** (Enade) do Ensino Superior, aplicado pelo Instituto Nacional de Estudos e Pesquisas Educacionais (Inep), é realizado em duas etapas:

- **Avaliação teórica**, por meio da Prova Nacional Docente (PND).

- **Avaliação prática:** que permite medir as competências pedagógicas desenvolvidas ao longo da formação.

O Questionário do Aluno visa coletar informações que permitam caracterizar o perfil dos alunos e o contexto de seus processos de formação, relevantes para a compreensão dos resultados dos futuros graduados no Enade e para subsidiar os processos de avaliação dos cursos de licenciatura e das instituições de ensino superior. Instrumento obrigatório, o questionário deve ser preenchido integralmente por todos os alunos matriculados, exclusivamente no sistema Enade.

O preenchimento completo do questionário do instrumento é um dos elementos que caracterizam a participação efetiva do estudante no exame, sendo objeto de verificação no processo de atribuição da sua regularidade antes do Enade.

É da responsabilidade do estudante e do IES monitorizar o andamento do preenchimento deste instrumento. As respostas ao Questionário dos Estudantes são analisadas pelo INEP e agregadas por curso de licenciatura, preservando a confidencialidade da identidade dos respondentes.

As provas são compostas por uma parte de Formação Geral, comum a todos os cursos em todas as áreas, e uma parte de Componente Específica, própria de cada área de avaliação. A parte de formação geral compreende 10 perguntas, das quais duas são discursivas e oito são de múltipla escolha, envolvendo situações-problema e estudos de caso. A concepção das perguntas é orientada pelos princípios dos direitos humanos. As questões discursivas avaliam aspectos como clareza, coerência, coesão, estratégias argumentativas, uso de vocabulário apropriado e a exatidão gramatical do texto. A componente específica de cada área de avaliação é composta por 30 questões, das quais três são discursivas e 27 são de escolha múltipla, envolvendo situações-problema e estudos de caso.

O teste Enade mede o desempenho dos estudantes em relação ao programa previsto nas diretrizes nacionais dos programas ou no catálogo nacional dos cursos superiores de tecnologia do curso de graduação em questão, a sua capacidade de se adaptar às exigências decorrentes da evolução do conhecimento e as suas competências para compreender temas fora do âmbito específico da sua profissão, relacionados com a realidade brasileira e mundial e com outras áreas do conhecimento.

As análises dos dados do Enade baseiam-se na estatística descritiva e na teoria da resposta ao item (TRI).

Rostirola et al. (2025), por exemplo, realizam uma análise de itens específicos do exame, concentrando-se principalmente nos 27 itens de múltipla escolha relacionados com a **componente específica** do conhecimento do conteúdo matemático. Esses autores utilizaram a teoria clássica dos testes (TCT), a teoria da resposta ao item (TRI) e a análise fatorial exploratória (AFE) em dados de 10.869 participantes extraídos do Enade 2017.

A análise realizada por esses autores revela que o teste possui uma elevada capacidade de diferenciação, com a maioria dos itens apresentando níveis de dificuldade que variam de moderados a elevados. A adequação do modelo bidimensional, integrando conhecimentos pedagógicos e especializados, foi confirmada pela análise fatorial e pelos indicadores de ajuste.

Os autores elaboraram uma escala pedagógica de competências com base em indicadores como resolução de problemas, análise de dados e competências sociocomunicativas. Esta abordagem evidencia a existência de diferentes níveis de competências entre os inquiridos.

A análise multidimensional realizada com base nos dados da Enade revela-se pertinente para compreender o caráter latente dos conhecimentos relativos ao ensino da matemática. Esta abordagem sugere que é possível realizar uma avaliação mais completa das competências profissionais quando várias dimensões são levadas em consideração. Além disso, o estudo recomenda que o INEP integre novas perspectivas computacionais nas suas análises futuras.

Os autores destacam **desafios e limites na avaliação. Por exemplo**, o exame Enade não tem uma distribuição ideal da dificuldade dos itens, com uma forte concentração de itens de dificuldade média e uma falta de itens nas categorias muito fáceis, fáceis e muito difíceis. Essa distribuição assimétrica pode não avaliar plenamente a gama de capacidades dos alunos. Uma percentagem significativa de itens foi classificada como necessitando de revisão com base no seu coeficiente de discriminação, o que indica que eles não fazem uma distinção eficaz entre alunos com diferentes níveis de conhecimento. Uma questão (I10) tinha uma probabilidade muito alta de obter uma resposta correta por acaso, o que sugere que poderia haver uma distração inadequada, comprometendo a validade da questão.

Em relação às implicações para melhorar as políticas e práticas de formação de professores, os autores salientam que uma abordagem multidimensional, em oposição a uma abordagem unidimensional, permite uma medição mais precisa e fiável da característica latente «conhecimentos para o ensino da matemática», uma vez que tem em conta tanto os conhecimentos pedagógicos gerais como os conhecimentos específicos do conteúdo, proporcionando assim uma imagem mais completa da preparação de um professor.

No que diz respeito à identificação das competências predominantes, os autores observam que, ao considerar várias dimensões, torna-se possível identificar as competências predominantes de cada indivíduo ou verificar em que competências cada disciplina precisa de melhorar. Isto permitiria um feedback direcionado e experiências de aprendizagem personalizadas para responder a fraquezas específicas. A formação para o ensino é intrinsecamente complexa, abrangendo conhecimentos pedagógicos gerais e conhecimentos sobre um conteúdo específico. Uma medida bidimensional alinha-se com essa complexidade, levando a uma avaliação mais precisa e relevante.

A nossa perspectiva de análise é uma abordagem que envolva questões didáticas, epistemológicas e culturais relacionadas com a avaliação matemática. Sob essa ótica, a

didática matemática estuda os processos avaliativos e sua relação com o contrato didático, destacando tanto a importância da medição quanto a compreensão do conhecimento matemático para além das capacidades mensuráveis (Bodin, 1989, 1994, 1997).

Além desses aspectos didáticos, a ASI poderá enriquecer a análise da Enade, permitindo compreender não apenas o desempenho, mas também os processos cognitivos subjacentes, a progressão das competências e as especificidades culturais ou contextuais. Assim, poderá melhorar a validade, a pertinência e a equidade das avaliações na educação, mas também

- **Compreender a estrutura cognitiva dos alunos**, revelando a estrutura íntima de um teste e a estrutura cognitiva dos alunos que o realizam. Ela destaca como as diferentes questões (itens) estão relacionadas entre si e como os alunos organizam os seus conhecimentos. Isso vai muito além da simples medição do sucesso ou fracasso em cada questão.

- **Identificar cadeias de competências**: graças à ASI, é possível identificar cadeias de itens em que o sucesso numa questão aumenta a probabilidade de sucesso noutra. Isto permite mapear percursos de aprendizagem ou progressões de competências, o que é valioso para o diagnóstico e a remediação.

Analizar a complexidade e a dificuldade dos testes: A ASI distingue a complexidade de um item da sua dificuldade, o que permite refinar a conceção dos testes e compreender melhor os requisitos cognitivos das diferentes questões. Também ajuda a validar empiricamente taxonomias de competências, como a hierarquia de complexidade cognitiva em matemática.

5. REFERENCIAS

Alencar, Edvonete Souza de Almouloud, Saddo Ag. A metodologia de pesquisa: metassíntese qualitativa. *Revista Reflexão e Ação, Santa Cruz do Sul*, v. 25, n. 3, p. 204-220, Set./Dez. 2017. <https://doi.org/10.1590/S1413-24782022270122>

Aquino, Rafael Santos de, Carneiro-Leão, Ana Maria dos Anjos, Acioly-Régnier, Nadja Maria, Andrade, Vladimir Lira Véras Xavier de. Análise Estatística Implicativa e Análise de Variância: estudo estatístico comparativo sobre o desempenho escolar em sala de aula multicultural. *XI Colloque International A.S.I. Analyse Statistique Implicative XI International Conference Statistical Implicative Analysis Belfort (France) – 3-6 Novembre 2021*. <http://sites.univ-lyon2.fr/ASI/11>

Bodin, A. (1989), L'évaluation du savoir mathématique, *Bulletin de l'Association des Professeurs de Mathématiques*. (APMEP), 368, pp.195- 219, Paris.

Bodin, A. (1994), *Problèmes didactiques de l'évaluation du savoir mathématique*. Cahiers du CIRADE. Université du Québec à Montréal.

Bodin, Antoine (1997). L'évaluation du savoir mathématique : Questions et méthodes. *Recherches en Didactique des Mathématiques*, Vol. 17, n°1, pp. 49-96, 1997.

Bodin, Antoine. Ce qui est vraiment évalué par PISA en mathématiques. Ce qui ne l'est pas. Un point de vue français. In https://antoine-bodin.com/wp-content/uploads/2016/07/ce-que-pisa-evalue_bull-apmep.pdf

Bodin, Antoine et Giovannini, Maria Lucia. L'utilisation de l'analyse implicative et cohésitive pour exploiter un test standardisé de mathématique. Giornale Italiano della Ricerca Educativa, 2017, <https://core.ac.uk/download/322531415.pdf>

Bodin, Antoine et Grapin, Nadine. Un regard didactique sur les évaluations du PISA et de la TIMSS : mieux les comprendre pour mieux les exploiter. Érudit, Volume 41, Number 1, 2018, p. 67–96. <https://www.erudit.org/en/journals/mee/2018-v41-n1-mee04281/1055897ar/>

Bodin, Antoine ; Almouloud, Saddo Ag. Traitement de données et évaluation. *Publications de l'Institut de recherche mathématiques de Rennes*, 1991, fascicule S6 « Vième école d'été de didactique des mathématiques et de l'informatique », p. 123-126. http://www.numdam.org/item?id=PSMIR_1991_S6_123_0

Bodin, Antoine, Trouche, Luc. Penser PISA dans la perspective du renouvellement de l'enseignement des mathématiques en Algérie. [Rapport de recherche] Ministère algérien de l'éducation nationale ; CIEP; Ambassade de France en Algérie. 2019. fffhal-03141731f. <https://hal.science/hal-03141731v1>

BRASIL, Instituto Nacional De Estudos E Pesquisas Educacionais (INEP). Enade.

Brasília: Inep, 2014. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade>.

Gras, Régis, (1979), *Contribution à l'étude expérimentale et à l'analyse de certaines acquisitions cognitives et de certains objectifs didactiques en mathématiques*, Thèse d'Etat, Université de Rennes 1.

GR A S, Régis (Ed.) (1996), *L'implication statistique. Nouvelle méthode exploratoire de données*. La pensée sauvage. Grenoble.

Gras, Régis (2013). Genèse et développement de l'Analyse Statistique Implicative, rétrospective intuitive. <http://sites.univ-lyon2.fr/asi7/10juin2013/10juin2013RG.pdf>

GRAS, Régis et BODIN. Antoine. L'A.S.I., analyseur et révélateur de la complexité cognitive taxonomique, *Educ. Matem. Pesq.*, São Paulo, v.19, n.1, 497-515, 2017. <http://dx.doi.org/10.23925/1983-3156.2017v19i1p497-515>

Khaled, Hayette; Bodin, Antoine; Couturier, Raphael. L'analyse statistique implicative (ASI) appliquée aux données de programme international pour le suivi des acquis des élèves (PISA), 2017.

In: https://www.researchgate.net/publication/324770861_l'analyse_statistique_implicative_asi_appliquee_aux_donnees_d

[e programme international pour le suivi des acquis des élèves_pisa](#)

Rostirola, S. C. M. Henning, E., Siple, I. Z., & Andrade, D. F. de. (2025). Análise da dimensionalidade do Enade 2017 para a Licenciatura em Matemática. *Estudos em Avaliação Educacional*, 36, Artigo e10600. <https://doi.org/10.18222/eae.v36.10600>