

Análise de conteúdo das provas do Enade para os alunos do curso de Bacharelado em Ciência da Computação

Title: Content analysis of Enade exams for Computer Science

Priscila da Silva Neves Lima
Universidade Federal de Goiás
priscila_lima@ufg.br

Ana Paula Laboissière Ambrósio
Universidade Federal de Goiás
apaula@inf.ufg.br

João Lucas dos Santos Oliveira
Universidade Federal de Goiás
joaooliveira@discente.ufg.br

Cedric Luiz de Carvalho
Universidade Federal de Goiás
cedric@ufg.br

Resumo

A partir da análise dos dados educacionais é possível observar o desempenho tanto do estudante quanto das instituições, oportunizando a tomada de decisões para a melhoria do processo de ensino-aprendizagem. O Exame Nacional de Desempenho de Estudantes (Enade) permite avaliar e acompanhar este processo. A educação é uma área fundamental para o crescimento e desenvolvimento de um país. Desta forma, o desafio desta pesquisa é detectar possíveis falhas na formação dos alunos a partir da análise do conteúdo do Enade, além de identificar as áreas de conhecimento que são preditoras do sucesso no resultado do exame. Assim, este texto descreve a implementação de uma metodologia para a análise de conteúdo do Enade, com vistas a indicar as disciplinas que mais impactam no resultado deste exame. Como estudo de caso, foram usados os dados de todos os estudantes do curso de Bacharelado em Ciência da Computação, que realizaram os exames referentes aos anos de 2005, 2008, 2011, 2014 e 2017. As questões dos exames foram classificadas de acordo com as áreas do curso de Bacharelado em Ciência da Computação. Além disso, foram aplicados testes estatísticos nos resultados obtidos pelos estudantes no Enade de modo a reconhecer preditores e estrutura do exame. O resultado foi um relatório baseado em análise de conteúdo que permite analisar o Enade por tema. A partir da generalização dos resultados obtidos, a metodologia poderá ser aplicada a outras áreas de conhecimento, expandindo o escopo de análises possíveis, utilizando os dados do Enade.
Palavras-Chave: Enade, análise de dados educacionais, áreas preditoras de sucesso, análise de conteúdo, metodologia de análise, ciência da computação

Abstract

The analysis of educational data makes it possible to observe the performance of both the student and the institutions, providing opportunities for decision-making to improve the teaching-learning process. The National Student Performance Exam (Enade) assesses and monitors this process. Education is a fundamental area for the growth and development of a country. Therefore, the challenge of this research is to detect possible flaws in training students from the content analysis of Enade, in addition to identifying the areas of knowledge that are predictive of success in the final exam result. Thus, this text describes the implementation of a methodology for the content analysis of Enade, to identify the disciplines that most impact the result of this exam. As a case study, we used data from all students and tests applied to the Computer Science area, for the years 2005, 2008, 2011, 2014, and 2017. The questions of the tests were classified according to the subareas of Computer Science. Besides, statistical tests were applied to the results obtained by students at Enade to identify predictors and structure of the exam. The final result was a report based on content analysis that allows Enade to be analyzed by theme. Based on the generalization of the results obtained, the methodology can be applied to other areas of knowledge, expanding the scope of possible analyzes, using data from Enade.

Keywords: Enade, educational data analysis, predictive areas of success, content analysis, analysis methodology, computer science

1 Introdução

Para avaliar a qualidade das instituições de ensino é fundamental que existam indicadores de desempenho. As avaliações, desenvolvidas por entidades públicas, quando aplicadas em grande escala, podem contribuir na análise de qualidade das instituições. Tais avaliações têm como um de seus objetivos produzir informações sobre a eficiência e qualidade das instituições analisadas que podem ser utilizadas no gerenciamento, de modo a melhorar a qualidade do ensino (Primi, Hutz, & Da Silva, 2011). Na construção de classificações globais, a medição da qualidade do ensino é um desafio, uma vez que é preciso criar indicadores válidos e viáveis para a grande diversidade de instituições e países. Por outro lado, as avaliações nacionais têm vantagem e procuram ampliar a visão sobre a qualidade da formação dos estudantes (Vassalo, 2016).

No Brasil, a entidade responsável por promover estudos, pesquisas e avaliações sobre o sistema educacional é o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), que possui bases de dados provenientes de levantamentos estatísticos e avaliativos em diversos níveis e modalidades de ensino (Fonseca & Namen, 2016). No Ensino Superior, são 3 os indicadores de qualidade fornecidos pelo INEP: Índice Geral de Cursos Avaliados da Instituição (IGC), que avalia a qualidade das instituições de educação superior; Conceito Preliminar de Curso (CPC), que avalia a qualidade dos cursos superiores; e o Conceito Exame Nacional de Desempenho de Estudantes (Enade), que avalia o desempenho dos estudantes a partir dos resultados obtidos na aplicação do exame. Esses indicadores são expressos em escala contínua e em cinco níveis. Os níveis iguais ou superiores a 3 indicam qualidade satisfatória (INEP, 2021c).

O Enade busca contribuir para a melhoria da qualidade dos níveis de escolaridade e é utilizado tanto para o desenvolvimento de políticas públicas da educação, sendo usado no processo de regulamentação pelo Ministério da Educação e Cultura (MEC), quanto para garantir maior transparência dos dados sobre a qualidade da educação superior à sociedade (INEP, 2021b). Os dados coletados podem ser usados em diversos níveis de tomada de decisão: aluno, docente, curso, instituição, governo e sociedade. Sistemas de avaliação educacional como o Enade, coletam e armazenam grande volume de informações referentes às provas aplicadas e aos estudantes e seus resultados. As análises desses dados muitas vezes se limitam a estatísticas descritivas dos dados, sendo os mais populares as notas do exame e o questionário socioeconômico.

Análises de conteúdo das questões nem sempre são exploradas na avaliação oficial, dadas as diferenças existentes entre as áreas de conhecimento e a dificuldade em padronizar uma metodologia aplicável a todas as áreas. Eventualmente, análises de conteúdo pontuais são feitas por pesquisadores interessados no aprimoramento dos seus projetos pedagógicos (Costa & Martins, 2014) ou de formação de professores (Lara, 2007) (Novossate, 2010) (Schwengber, 2013), mas não fazem parte das estatísticas oficiais do INEP.

No sentido de contribuir para uma análise mais abrangente do Enade, que considere o conteúdo programático das questões específicas das áreas de ensino, esta pesquisa visa aplicar uma metodologia que permita classificar as questões nas áreas do conhecimento representadas nas provas e identificar análises que podem ser realizadas a partir desta classificação. Um estudo de caso foi realizado usando os dados do Enade para o curso de Bacharelado em Ciência da Computação (BCC). Diferentemente dos trabalhos encontrados na literatura apresentados na seção de trabalhos relacionados, essa pesquisa é voltada para a análise das provas do Enade com base em classificação de conteúdo.

O trabalho de (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018) propõe uma metodologia que permite classificar as questões do exame nos domínios do conhecimento representados nas provas do Enade e identificar análises que podem ser realizadas usando essa classificação. A metodologia aplica procedimentos gerais utilizando um sistema de *software* desenvolvido pelos

pesquisadores. O algoritmo separa os resultados do aluno por temas¹ que representam tópicos gerais abordados nas provas e gera relatórios que fornecem melhor compreensão da estrutura da prova e do resultado do aluno por domínio de conhecimento. As informações fornecem aos tomadores de decisão novas perspectivas que podem ter grande impacto. Os procedimentos podem ser definidos para gerar relatórios que facilitem o entendimento do desempenho do aluno com base no conhecimento do domínio representado no exame. Neste estudo, aplicamos a metodologia proposta em (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018) ao exame Enade, para verificar sua aplicabilidade aos exames nacionais e o esforço necessário para adaptá-la a um contexto específico.

O impacto desta pesquisa abrange diversos atores do contexto educacional. Para os gestores, ela pode ser importante na tomada de decisões estratégicas no que tange à formação dos futuros profissionais e na implantação de políticas públicas voltadas para a educação superior. Do ponto de vista científico, a pesquisa contribui para uma análise mais abrangente dos dados do Enade. Apesar deste artigo focar o curso de Bacharelado em Ciência da Computação, a metodologia pode ser aplicada em outras áreas² de conhecimento. Além disso, a metodologia também pode ser adaptada para outras provas, como, por exemplo, Exame Nacional do Ensino Médio (Enem), Prova Brasil ou Programa Internacional de Avaliação de Estudantes (Pisa), bastando identificar as áreas de conhecimento para diferentes níveis da educação serem analisados. No contexto acadêmico, métodos podem ser desenvolvidos para melhorar a relação do estudante com as áreas mais críticas de aprendizado e estratégias administrativas e pedagógicas podem ser implementadas de modo a melhorar a relação do professor com o conteúdo ministrado em sala de aula. Essas medidas contribuem diretamente, não apenas na qualidade do ensino, mas principalmente para melhor compreensão da prova do Enade e sua capacidade de avaliação do ensino superior.

Esse trabalho está estruturado em 8 seções. Na seção 2, apresenta-se o Enade, sua estrutura e periodicidade, assim como os conteúdos do curso de Bacharelado em Ciência da Computação cobrados pelo exame. Em seguida, na seção 3, são apresentados trabalhos relacionados às categorias de análises realizadas com os dados do Enade. Nessa seção, são considerados todos os tipos de análises encontradas através de revisão sistemática da literatura sobre os dados do Enade. Logo após, na seção 4, é apresentada a metodologia baseada em análise de conteúdo utilizada na elaboração do trabalho e, na seção 5, é apresentado o estudo de caso dos estudantes do Enade, considerando as provas dos anos de 2005, 2008, 2011, 2014 e 2017. Na seção 6, é apresentada uma descrição da amostra dos estudantes extraída dos dados fornecidos pelo INEP. Na seção 7, são descritos os resultados obtidos sob a perspectiva de contribuir para uma análise mais abrangente do Enade. Por fim, na seção 8, são apresentadas as conclusões do trabalho, suas limitações e direcionamento para continuidade do mesmo.

2 Enade

No Brasil, o Sistema Nacional de Avaliação do Ensino Superior (SINAES), instituído pela Lei n.º 10.861/2004, é responsável por avaliar as Instituições de Ensino Superior (IES). O sistema tem seus processos avaliativos coordenados e supervisionados pela Comissão Nacional de Avaliação Superior (CONAES) e a operacionalização é de responsabilidade do INEP. O SINAES é constituído por 3 partes principais: avaliação das instituições, avaliação dos cursos e avaliação de desempenho dos estudantes através do Enade. O Enade é realizado com base nos conteúdos

¹ Neste texto, usaremos a palavra "tema" para indicar grupos de assuntos relacionados à computação. E a palavra "assunto" para indicar individualmente conteúdos abordados dentro de cada tema.

² Neste texto a palavra "área" indica o campo em que se exerce determinada atividade.

programáticos previstos nas diretrizes curriculares dos respectivos cursos de graduação, bem como suas competências e habilidades oriundas de sua formação (INEP, 2021d).

O Enade é subdividido em três etapas, realizadas em anos subsequentes. Em cada ano, é avaliado um conjunto de áreas de ensino. O ano I abrange as áreas de conhecimento de Ciências Agrárias, Ciências da Saúde e áreas afins. O ano II é formado pelas áreas de Ciências Exatas, Licenciaturas e afins. O ano III é composto pelas áreas de Ciências Sociais Aplicadas, Ciências Humanas e áreas afins. Após todos os anos serem avaliados, o exame volta novamente a avaliar as áreas relacionadas ao ano I, seguindo posteriormente para os demais anos, formando assim um ciclo, onde cada conjunto de áreas é avaliado em um intervalo de três anos (INEP, 2017a).

A sua primeira versão foi realizada em todo o país em 6 de novembro de 2005, com a avaliação de 20 áreas do conhecimento (INEP, 2008b). A área de computação, compreendendo os cursos de Bacharelado em Ciência da Computação, Licenciatura em Computação, Bacharelado em Engenharia de Computação e Bacharelado em Sistemas de Informação, foi avaliada nos anos de 2005, 2008, 2011, 2014 e 2017. Nas versões de 2005 e 2008, o Enade incluiu grupos de estudantes dos cursos de Bacharelado em Sistemas de Informação, Bacharelado em Ciência da Computação e Bacharelado em Engenharia da Computação, selecionados por amostragem, e que se encontravam em momentos distintos de sua graduação: um grupo, considerado ingressante, cursava o final do primeiro ano e outro grupo, considerado concluinte, no final do último ano do curso. Os dois grupos de estudantes foram submetidos à mesma prova.

Na versão de 2011, os estudantes da área de computação que realizaram a prova, pertenciam aos cursos de Licenciatura em Computação, Bacharelado em Ciência da Computação, Bacharelado em Engenharia da Computação e Bacharelado em Sistemas de Informação. Em 2014 e 2017, houve o desmembramento da área de computação e cada curso fez uma prova específica distinta. Portanto, para a versão de 2014, a prova utilizada neste estudo foi exclusivamente aplicada para estudantes do curso de Bacharelado em Ciência da Computação, pois os alunos da modalidade Licenciatura foram submetidos a uma prova específica para essa modalidade (INEP, 2021b). Em 2011, 2014 e 2017, apenas alunos concluintes realizaram as provas. Em 2017, os alunos ingressantes também foram registrados no sistema, apesar de não terem feito a avaliação.

O Enade é estruturado a partir de dois instrumentos: um questionário socioeconômico/questionário do estudante, com a função de compor o perfil dos estudantes, e uma prova, composta de questões objetivas e discursivas, divididas em duas partes. A primeira parte, denominada "formação geral", apresenta-se como "componente comum" às provas das diferentes áreas, investigando competências de raciocínio e conhecimento geral dos alunos. A segunda parte, denominada "componente específico", contempla as especificidades de cada área, tanto no domínio dos conhecimentos quanto nas habilidades esperadas para o perfil profissional. No final do caderno de prova, há um questionário de percepção da prova que investiga a percepção dos estudantes frente à sua trajetória no curso e na instituição de ensino, através de questões objetivas que exploram a função social da profissão e os aspectos fundamentais da formação profissional. O componente específico das provas é definido pelas diretrizes e matrizes elaboradas pela comissão assessora de avaliação da área e pela comissão assessora de avaliação de formação geral do Enade. Essas matrizes compreendem conteúdo similar ao apresentado pelas diretrizes curriculares nacionais para os cursos de graduação (INEP, 2021a).

Conforme disposição do Art. 5.º, § 5.º, da Lei n.º 10.861/2004, o Enade constitui-se componente curricular obrigatório. O estudante habilitado ao Enade que não realizar a prova não poderá receber o seu diploma enquanto não regularizar a sua situação junto ao Enade, haja vista não ter concluído o respectivo curso de graduação. A existência de irregularidade perante o Enade impossibilita a colação de grau do estudante, em decorrência da não conclusão do curso, por ausência de cumprimento de componente curricular obrigatório.

Com relação aos conteúdos do curso de Bacharelado em Ciência da Computação, a Resolução n.º 5 CNE/CES/2016 - Conselho Nacional de Educação e Câmara de Educação Superior institui as diretrizes curriculares nacionais para os cursos de graduação na área da computação. Essa resolução abrange, dentre outros, o curso de Bacharelado em Ciência da Computação, com base no Parecer CNE/CES n.º 136/2012 que trata dos conteúdos básicos e tecnológicos da área. Essas diretrizes apontam as competências e habilidades dos egressos dos cursos de Bacharelado em Ciência da Computação, porém não estipulam o currículo a ser trabalhado, deixando a cargo de cada IES, através de Projeto Pedagógico, a definição e organização do curso (MEC, 2016).

Com base nessas diretrizes curriculares, o Enade define para cada ano, por portaria publicada no Diário Oficial da União (DOU), uma lista de conteúdos que servirão como referencial na elaboração das provas. Essa lista de referência para as provas de Bacharelado em Ciência da Computação tem se mantido constante, apesar de algumas pequenas diferenças, principalmente no que tange à profundidade em que os temas são tratados, nos anos de aplicação do exame. Como exemplo, o conteúdo de referência para o componente específico de 2017 foi definido no Art. 7.º da portaria n.º 473, de 6 de junho de 2017, especificamente para o curso de Bacharelado em Ciência da Computação. Portarias semelhantes foram emitidas para os anos de 2005 (INEP, 2005), 2008 (INEP, 2008a), 2011 (INEP, 2011) e 2014 (INEP, 2014).

3 Trabalhos relacionados

O INEP disponibiliza dados de forma anônima e gera relatórios das provas aplicadas ao Enade. Esses relatórios geralmente se resumem a estatísticas descritivas dos dados, que visam descrever e resumir as informações coletadas. Em (Lima, Ambrósio, Ferreira & Brancher, 2019) é apresentada uma revisão sistemática que identifica o uso que vem sendo feito dos dados coletados pelo INEP, para além das análises feitas por ele próprio.

Conforme (Lima, Ambrósio, Ferreira & Brancher, 2019), há inúmeros trabalhos que desenvolvem análises dos dados do INEP. Essas análises são feitas para diversos fins, dentre eles, podemos citar os trabalhos classificados na categoria ‘Conhecimento/Conteúdo’ que procuram fazer um estudo sobre o conteúdo das provas e/ou o conhecimento dos estudantes em determinados assuntos abordados nos exames. Por exemplo, Lopes e Vendramini (2015) avaliam as propriedades psicométricas da prova de Pedagogia no Enade, e Bianchi (2010) analisa o percentual de questões envolvendo a interpretação de informações apresentada em gráficos e tabelas estatísticas no Enade. As investigações feitas com o objetivo de análise ‘Administrativa’ abrangem aspectos de gerenciamento, como o acesso ao Ensino Superior (Guerra, Nakamura & Hruschka, 2014), o impacto que os resultados do Enade têm sobre as questões acadêmicas, como mudanças no currículo e na cultura institucional (Gomes, 2011). Ainda nessa categoria há estudos que fazem predição de risco de evasão escolar em instituições públicas utilizando dados do INEP (Teodoro & Kappel, 2020).

As análises que objetivam o ‘Desempenho/Rendimento’ avaliam os resultados dos estudantes e/ou das instituições nos exames. Dentre essas análises, destaca-se (Souza, 2008), que busca identificar a relação entre o desempenho dos alunos e a sua situação socioeconômica e (Souza et al. 2017) que realiza análise preditiva de desempenho dos alunos dos cursos no Enade com base no perfil socioeconômico e de desempenho no Enem. Outros trabalhos focam o ‘Desenvolvimento e Teste de ferramentas’ com o propósito de auxiliar na análise dos resultados dos exames, como o estudo de Costa (2016), com a ferramenta *Vis-Scholar*, que utiliza estatística para correlacionar conjuntos de dados e métodos de visualização dessas correlações, de modo a facilitar a visualização dos dados, e o trabalho de Branco Neto, Durigon, Avila & Oliveira Junior (2020), que desenvolveram uma ferramenta para a manipulação de expressões algébricas.

“Avaliação/Estrutura do exame” refere-se a trabalhos que procuram entender de que forma o exame é estruturado e até fazem uma avaliação comparando com outros exames como, por exemplo, (Verhine & Dantas, 2005). ‘Formação/Qualificação Docente’ agrupa trabalhos onde os dados dos exames são analisados de modo a aprimorar a formação do docente e relacioná-la com o desempenho dos estudantes nas provas. Nessa categoria, encontra-se o trabalho de Brito (2015) que objetiva determinar como as características do corpo docente de uma IES influenciam o desempenho dos concluintes no Enade. É possível relatar ainda análises ou estudos do ponto de vista longitudinal, transversal e/ou por coorte. Essa categoria de análise foi aplicada com os mais diferentes propósitos, como investigar a democratização do Ensino Superior brasileiro a partir do perfil socioeconômico dos alunos (Picaço, 2016) e estudos sobre evasão (Vieira, 2012), entre outros.

As análises mais recorrentes nos trabalhos relacionados são feitas com o objetivo de estudar as condições de acesso e permanência no ensino superior e há quantidade expressiva de trabalhos que analisaram os dados do Enade com o objetivo de verificar o desempenho e traçar os perfis dos estudantes. O interesse pelos dados do Enade pode estar associado ao seu uso para o acesso ao Ensino Superior no Brasil (Saviani, 2010). O foco nas análises está principalmente em dados socioeconômicos, tais como, grau de instrução dos pais e dados demográficos. O conteúdo de provas é o segundo maior dado utilizado nos estudos. Esses estudos estão focados em analisar desempenho e/ou rendimento dos estudantes.

Um estudo de Lima, Rosa, Ambrósio & Oliveira (2019) apresenta um modelo de aplicação desta metodologia em exames de larga escala para análise por conhecimento de conteúdo, permitindo uma comparação entre as duas aplicações da metodologia. O estudo mostrou que a metodologia é flexível e adaptável podendo ser aplicada a qualquer exame de avaliação. Entretanto, não foram apresentados resultados, apenas uma sugestão de como a metodologia pode ser aplicada aos exames nacionais.

A pesquisa apresentada neste artigo contribui com uma nova e mais abrangente forma de analisar os dados do Enade. Diferentemente dos trabalhos relacionados nessa seção, esse estudo categoriza conteúdo e realiza análise a partir de temas estabelecidos, mostrando ser possível produzir resultados de análise a partir da adaptação da metodologia de (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018) para o Enade. Analisar o conteúdo das questões de uma avaliação permite uma melhor compreensão da estrutura do exame e do conhecimento do conteúdo consideradas mais importantes pelas organizações de avaliação. Fornece ainda, informação que permite melhorar o exame e permite que o mercado tenha uma visão mais clara do profissional que recebe. Do ponto de vista das instituições de ensino, estas podem analisar seus currículos e identificar domínios em que seus alunos se destacam ou são deficientes. Essa forma de análise de conteúdo possibilita melhorar a relação dos estudantes com as áreas críticas de aprendizado.

4 Metodologia

Para realizar a análise do conteúdo de um exame, é necessário entender a sua estrutura e o conhecimento específico verificado por ele. Frequentemente, testes amplamente aplicados liberam documentos informando o conteúdo que será avaliado, assim como sua estrutura e categoria de perguntas. Também são fornecidas informações sobre como as pontuações são calculadas, incluindo a pontuação máxima para cada pergunta e seu peso na nota final. A metodologia proposta em (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018) considera todos esses fatores e pode ser aplicada a qualquer avaliação em larga escala. Ela está estruturada em três etapas de acordo com a Figura 1.

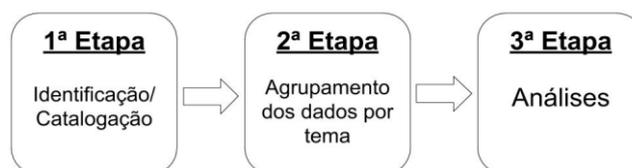


Figura 1: Etapas da metodologia.

1.ª Etapa - Identificação/Catologação

Consiste em identificar as questões do exame que serão analisadas, classificando-as em temas previamente definidos. Isto significa extrair a redação de cada questão do exame (incluindo as alternativas das questões de múltipla escolha) e identificar como ela é avaliada e seu peso na nota final do exame. Cada questão deve ser classificada em um dos temas predefinidos.

2.ª Etapa - Agrupamento dos dados por tema

Para cada aluno, uma nota por tema é calculada. Isso é feito a partir da identificação do resultado do aluno para cada questão, agrupando-as por tema e calculando sua pontuação. Se as questões tiverem pesos diferentes na nota final do teste, pode ser aconselhável manter esses pesos no cálculo das notas por tema.

A forma como essa nota é calculada pode ser definida de acordo com a situação, mas elas devem ser normalizadas de forma a permitir a comparação entre os temas. Uma opção é usar uma média aritmética ponderada, conforme proposto em (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018), ou seja, para cada tema, multiplicar a nota do aluno pelo peso da questão, somar o resultado para todas as questões do tema e dividir pela soma dos pesos das questões relacionadas ao tema.

Outros aspectos da transformação de dados também devem ser considerados, como valores ausentes e *outliers*. Caso o conjunto de dados não tenha sido pré-processado para tratar esses aspectos, isso deve ser feito antes de se proceder ao cálculo da pontuação por tema. A implementação deste procedimento dependerá da abordagem adotada e de como os dados contendo a nota de cada questão são armazenados.

3.ª Etapa - Análises

Técnicas de análises de dados educacionais são definidas e empregadas visando obter informações importantes, melhoria na visualização de dados e a produção de relatórios mais detalhados. Exemplos de técnicas que podem ser usadas incluem: Estatística Descritiva e Estatística Inferencial.

5 Estudo de caso

Conforme visto na seção 4, a metodologia está estruturada em três etapas. Para o contexto do Enade, a 1.ª e 2.ª etapas foram subdivididas em três atividades e a 3.ª etapa foi subdividida em duas atividades, conforme apresentado na Figura 2.

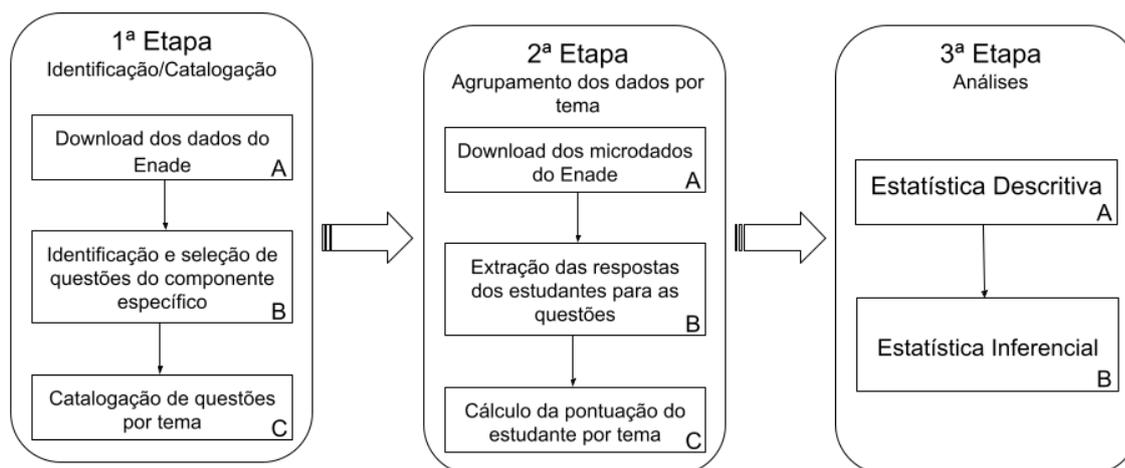


Figura 2: Etapas da metodologia aplicada ao Enade.

1.ª Etapa - Identificação/Catologação

A - Download dos dados do Enade

Para o curso de Bacharelado em Ciência da Computação, os pesquisadores selecionaram, diretamente da página do INEP³, as provas e os relatórios de 2005, 2008, 2011, 2014 e 2017. Ambos, provas e relatórios, foram utilizados pelos pesquisadores na aplicação da metodologia. A redação de cada questão do exame foi extraída, a forma como cada questão é avaliada foi identificada, assim como seu peso na nota final do exame.

Os relatórios contêm os índices de facilidade e discriminação calculados para cada questão. O índice de facilidade é calculado levando-se em conta a taxa de acerto de cada questão. O índice de discriminação é calculado pela correlação ponto-bisserial (pb)⁴. Ambos são fornecidos pelo INEP nos relatórios de síntese de área que podem ser acessados no sítio do órgão. Conforme consta nos relatórios de síntese de área, para o índice de facilidade, questões acertadas por 86% dos estudantes, ou mais, são consideradas muito fáceis. No extremo oposto, questões com percentual de acerto igual ou inferior a 15% são consideradas muito difíceis (Tabela 1) (INEP, 2008b).

Por outro lado, para medir o poder de discriminação da questão, ou seja, para ser considerada apta a avaliar os alunos dos cursos, uma questão deve ser mais acertada por alunos que tiveram bom desempenho do que pelos que tiveram desempenho ruim. Um índice que mede essa capacidade das questões e que foi escolhido para ser utilizado no Enade é o denominado, correlação ponto-bisserial (pb). A Tabela 2 apresenta a classificação das questões de acordo com seu poder de discriminação. Questões com fraco índice de discriminação são anuladas (INEP, 2008b).

³ Disponível em: <http://portal.INEP.gov.br/enade>. Acesso em: 12/01/2020.

⁴ É um coeficiente de correlação utilizado quando uma variável é dicotômica. No Enade, esse índice mede o poder de discriminação da questão, ou seja, para ser considerada apta a avaliar os alunos dos cursos, uma questão deve ser mais acertada por alunos que tiveram bom desempenho do que pelos que tiveram desempenho ruim.

Tabela 1: Índice de Facilidade.

Índice de Acerto	Índice de Facilidade
Maior ou igual que 0,86	Muito Fácil
Entre 0,61 e 0,85	Fácil
Entre 0,41 e 0,60	Médio
Entre 0,16 e 0,40	Difícil
Menor ou igual que 0,15	Muito Difícil

Fonte: INEP

Tabela 2: Índice de Discriminação.

Índice de Discriminação	Classificação
Maior ou igual 0,40	Muito Bom
Entre 0,30 e 0,39	Bom
Entre 0,20 e 0,29	Médio
Menor ou igual que 0,19	Fraco

Fonte: INEP

Os índices de Facilidade e Discriminação para as provas analisadas são apresentados na seção 7 ‘Resultado e Discussão’, subseção 7.2 ‘Catalogação das questões por índice de Facilidade e Índice de Discriminação’.

B - Identificação e seleção de questões do componente específico

A prova do Enade é composta de conhecimentos gerais e específicos, entretanto o foco desta pesquisa é apenas o componente específico. Foi necessário identificar essas questões para o estudo.

Vale ressaltar que, somente nos anos de 2014 e 2017 houve uma prova específica voltada exclusivamente para os alunos do curso de Bacharelado em Ciência da Computação. Nas outras provas, as diferentes áreas de computação compartilharam questões. Assim, foi realizada uma seleção de acordo com os filtros e variáveis fornecidos pelo INEP, nos Dicionários de Variáveis que acompanham os microdados. Essa filtragem teve foco nas questões respondidas pelos estudantes dos cursos de Bacharelado em Ciência da Computação.

Ao todo 150 questões específicas de Bacharelado em Ciência da Computação foram selecionadas, sendo 136 objetivas e 14 discursivas. A Tabela 3 traz a relação de questões selecionadas em cada ano. Importante frisar que os números que aparecem na Tabela 3 na coluna ‘Questões Objetivas’ são os números identificadores das questões nas provas - sendo relevantes para permitir o rastreamento da informação para os leitores interessados.

Tabela 3: Relação de questões selecionadas em cada ano.

Ano	Questões objetivas	Questões discursivas
2005	41*, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 56*, 57*, 58, 59, 60, 61, 62, 63, 64*, 65, 66*, 67, 68, 69*.	55, 70 ⁱ
2008	11, 12, 13, 14, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38.	20, 39, 40
2011	9, 10, 11, 12, 13 ⁺ , 14, 15, 16, 17*, 18, 19, 20, 21, 22, 23*, 24*, 25, 26, 27, 28, 29, 30, 36, 37, 38*, 39, 40*.	D3, D4, D5
2014	9, 10, 11*, 12, 13, 14, 15, 16, 17, 18*, 19, 20, 21*, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31*, 32, 33*, 34*, 35.	D3, D4, D5

2017	9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21*, 22, 23, 24, 25*, 26*, 27, 28, 29, 30, 31, 32, 33, 34, 35.	D3, D4, D5
------	---	------------

Legenda: (*) Questões anuladas com índice de discriminação fraco. (+) Questões desconsideradas. (i) Questões discursivas anuladas.

C - Catalogação de questões por tema

Os relatórios de síntese de área dos anos de 2008, 2014 e 2017 apresentam a relação dos conteúdos do curso de Bacharelado em Ciência da Computação que correspondem a cada uma das questões das provas. Porém, essa relação não é fornecida para os anos de 2005 e 2011. Dessa forma, cada questão foi lida por dois especialistas em computação e relacionada a um tema do curso de Bacharelado em Ciência da Computação. É importante ressaltar que não houve conflito para a classificação entre os especialistas. A partir dos conteúdos de referência estabelecidos pela comissão assessora de área do INEP (Portarias emitidas para os anos de 2005 (INEP, 2005), 2008 (INEP, 2008a), 2011 (INEP, 2011), 2014 (INEP, 2014) e 2017 (INEP, 2017e), definidos com base nas diretrizes curriculares nacionais dos cursos, aprovadas e instituídas pelo Conselho Nacional de Educação (CNE) do Ministério da Educação (MEC). Com base nessas diretrizes, os especialistas identificaram 11 temas para catalogação das questões. Os temas, siglas e os conteúdos relacionados a esses temas estão descritos abaixo:

Algoritmos, Estrutura de Dados e Programação (AEDP): Conceitos Básicos; Técnicas de Projeto: Divisão-e-Conquista, Guloso, Recursão, *Backtracking*; Notação e Análise Assintótica de pior e melhor caso; Análise de Complexidade de Problemas: Classes de Problemas P, NP, NP-Completo e NP-Difícil; Listas; Pilhas, Filas; Árvores: Binárias, AVL, Preto e Vermelho; *Heaps*; Tabelas *Hash*; Tipos Abstratos de Dados; Conjuntos; Mapas; Algoritmos de Pesquisa e Ordenação; Organização de Arquivos.

Arquitetura de Computadores e Circuitos Digitais (ACCD): Organização de Computadores; Conjunto de Instruções, Mecanismos de Interrupção e de Exceção; Barramento, Comunicações; Interfaces e Periféricos, Hierarquia de Memória; Multiprocessadores; Multicomputadores; Arquiteturas Paralelas.

Banco de Dados (BD): Modelagem e projeto de banco de dados; Bancos de dados relacionais e orientados a objetos; Linguagens de consulta e manipulação de dado; Sistemas Gerenciadores de Bancos de Dados: arquitetura, gerenciamento de transações, controle de concorrência, recuperação, processamento e otimização de consultas; Bancos de dados distribuídos.

Computação Gráfica e Processamento de Imagem (CGPI): Transformações geométricas em duas e três dimensões, Recorte e visibilidade, Transformações projetivas, Definição de objetos e cenas tridimensionais, Modelos de iluminação e tonalização (*shading*), Texturas e Mapeamentos, Rasterização e Técnicas de antiserrilhado (*antialiasing*), Percepção visual humana, Amostragem, realce, filtragem e restauração de imagens; Segmentação de imagens; Compressão e comunicação de imagens; Noções de visão computacional e reconhecimento de padrões;

Engenharia de Software (ES): Processos de desenvolvimento de *software*, qualidade de *software*, técnicas de planejamento e gerenciamento de *software*, engenharia de requisitos, métodos de análise e de projeto de *software*, verificação, validação e teste, manutenção, documentação.

Ética, Computador e Sociedade (ECS): Aspectos sociais, econômicos, legais e profissionais de computação; Aspectos Estratégicos do Controle da Tecnologia; Ética e Responsabilidade Profissional.

Inteligência Artificial (IA): Linguagens Simbólicas; Resolução de Problemas por Busca; Esquemas para representação do conhecimento: Lógicos, em rede, estruturados, procedurais;

Formalismos para a representação de conhecimento incerto; Redes Bayesianas, Conjuntos e Lógica *fuzzy*, Aprendizado de máquina; Aprendizado Indutivo; Árvores de decisão; Redes Neurais; Algoritmos Heurísticos; Computação Evolutiva.

Linguagens Formais e Autômatos, Compiladores e Computabilidade (LFACC): Gramáticas; Linguagens Regulares, Sensíveis ao Contexto, Livres de Contexto; Tipos de Reconhecedores; Autômatos de Estado Finito, Determinísticos e não Determinísticos; Máquina de Turing; Autômatos de Pilha; Hierarquia de *Chomsky*; Tese de *Church*, Teorema da Incompletude de Gödel; Conceituação de tradutores: compiladores, interpretadores e montadores; Fases de um compilador; Análise léxica; Análise sintática; Análise semântica; Tratamento e recuperação de erros; Geração e Otimização de Código Intermediário; Tabela de símbolos; Ambientes de *run-time*; Ambiente de interpretação.

Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG): Cálculo Proposicional, Lógica de Primeira Ordem, Conjuntos, Relações, Funções, Ordens Parciais e Totais, Álgebra Booleana, Estruturas Algébricas, Combinatória. Além de Conceitos Básicos sobre Grafos e Subgrafo; Isomorfismo, Matrizes de Adjacência e Incidência, Caminhos e Ciclos. Árvores: Caracterização de Árvores, Cortes de Arestas, Cortes de Vértices. Conectividade: Conectividade de Vértices e Arestas; Ciclos Eulerianos e Hamiltonianos. Emparelhamentos. Coloração de Vértices e de Arestas e Planaridade

Redes de Computadores, Sistemas Distribuídos e Telecomunicações (RCSDT): Topologias, sinalização no meio de transmissão, Protocolos serviços de comunicação, Arquiteturas de protocolos, Interconexão de redes, Planejamento e gerência de redes, Segurança e autenticação, Avaliação de desempenho: teoria das filas, cadeias de *Markov*, monitoração; Conceitos de Sistemas Distribuídos; Comunicação e Sincronização em Sistemas Distribuídos; Modelos e Arquitetura de Sistemas Distribuídos.

Sistemas Operacionais (SO): Gerência de processos/processador, Comunicação, Concorrência e Sincronização de Processos, Gerenciamento de Memória, Alocação de Recursos e *Deadlocks*, Sistemas de Arquivos, Gerenciamento de Dispositivos de Entrada/Saída.

2.^a Etapa - Agrupamento dos dados por tema

A - Download dos microdados do Enade

O *download* dos microdados do Enade foi realizado por meio do site do INEP⁵ para os anos de 2005, 2008, 2011, 2014 e 2017. Esses arquivos contêm dados referentes a todos os alunos que realizaram o Enade naquele ano. Assim, foi necessário extrair os dados referentes aos alunos do curso de Bacharelado em Ciência da Computação, usando informações contidas nos próprios microdados, identificando área, e subárea conforme o caso, do curso do aluno.

B - Extração das respostas dos estudantes para as questões

Nessa etapa, foi extraído o vetor resposta dos estudantes⁶, utilizando a ferramenta *R Studio*. As planilhas disponibilizadas pelo INEP⁷ apresentam uma coluna com um vetor para cada aluno, que indica a correção da parte objetiva das questões do componente específico no exame, onde cada posição do vetor refere-se a uma questão. Tais posições do vetor equivalem ao número de cada questão. O número da questão no exame identifica sua posição no vetor. O vetor pode ser

⁵ Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>. Acesso em: 16/01/2020.

⁶ Esses dados extraídos estão em arquivos nas máquinas dos pesquisadores, podendo ser distribuídos aos interessados, contatando os pesquisadores via *e-mail*.

⁷ Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados>. Acesso em: 16/01/2020.

preenchido por 4 diferentes valores, expressando as seguintes informações: 0 = errado, 1 = correto, 8 = anulado pela comissão, 9 = anulado pelo índice de discriminação. Quando um aluno não responde a uma determinada questão, sua resposta é considerada errada, portanto, 0 é atribuído à questão. As pontuações obtidas nas questões discursivas são apresentadas em colunas individualmente, tendo como valor uma nota de 0 a 100.

C - Cálculo da pontuação do estudante por tema

Essa atividade foi completamente desenvolvida e executada pelos pesquisadores deste artigo. Para o cálculo da pontuação do estudante em cada tema, as quantidades de respostas corretas foram somadas e o resultado dividido pela quantidade total de perguntas associadas ao tema. Essas informações foram armazenadas em uma nova coluna adicionada à planilha. Questões anuladas foram desconsideradas nesse cálculo. Outra coluna foi usada para armazenar a média dos alunos nas questões discursivas. Se nenhuma questão discursiva foi associada a um tema, o valor 0 foi atribuído a todos os alunos. Por fim, foi adicionada uma coluna contendo a nota final para o tema, calculada pela média ponderada, considerando 85% para as questões objetivas e 15% para as questões dissertativas, adotando-se a mesma proporção utilizada para calcular a pontuação final no exame.

3.^a Etapa - Análises

A - Estatística Descritiva

Inclui um conjunto de medidas (moda, mediana, média, frequência, intervalo interquartil, desvio padrão etc.) e representações gráficas, que permitem sumariamente, a descrição de um conjunto de dados (Howell, 2010). Nesta pesquisa, foram utilizados média, intervalo interquartil e desvio padrão. Essas medidas foram selecionadas para proporcionar uma visão descritiva básica dos dados, entretanto, isso não impede que outros pesquisadores utilizem diferentes medidas de estatística descritiva.

B - Estatística Inferencial

Abarca testes estatísticos (tais como coeficiente de correlação de *Pearson*, coeficiente de correlação de *Spearman*, Teste-T, ANOVA etc), que possibilitam conclusões acerca da população-alvo, com base em resultados obtidos na amostra selecionada (Howell, 2010). Para este estudo, foram utilizados o coeficiente de correlação de *Pearson* para verificar correção entre o índice de facilidade, índice de discriminação e os temas.

6 Descrição da amostra

A Tabela 4 apresenta uma visão da amostra dos estudantes participantes neste estudo, isto é, estudantes do curso de Bacharelado em Ciência da Computação, e é resultante das estatísticas descritivas geradas pela linguagem R utilizando o *RStudio Cloud*⁸ sobre os microdados do Enade.

⁸ <https://rstudio.cloud/>

Tabela 4: Amostra descritiva dos estudantes concluintes. **Med** representa a média, **DP** o desvio padrão, **Min** a nota mínima, **Max** a nota máxima, **Q₂₅**, **Q₅₀**, **Q₇₅**, representam, respectivamente, primeiro, segundo e terceiro quartis.

Ano	Total alunos	Sexo		Med	DP	Min	Q ₂₅	Q ₅₀	Q ₇₅	Max
		F	M							
2005	6.734	1.272	5.463	30,80	15,35	0	18,20	31,80	40,90	81,80
2008	6.434	946	5.488	31,79	13,89	0	22,20	29,60	40,70	88,90
2011	9.834	1.348	8.496	30,79	15,42	0	19,00	28,60	38,10	90,50
2014	8.278	1.074	7.203	42,48	15,82	0	33,30	42,90	52,40	100
2017	8.489	1.023	7.466	41,58	14,56	0	33,30	41,70	50,00	100

Na Tabela 4, pode-se observar a evolução da nota dos estudantes no decorrer da aplicação do exame, sendo que em 2005 a média das notas foi de 30,80 (DP = 15,35), e em 2017 este valor foi levemente maior, sendo de 41,58 (DP = 14,56). É importante lembrar que, nos anos de 2005 e 2008, a seleção de alunos foi por amostragem e incluía ingressantes e concluintes, enquanto, em 2011, 2014 e 2017, a prova foi aplicada a todos os concluintes. Nessa pesquisa, as análises foram realizadas apenas com as amostragens de estudantes concluintes.

Os estudantes são oriundos de instituições de ensino distribuídas em todo o Brasil. A Tabela 5 mostra esta distribuição por região, ano e categoria de instituição (pública (Púb.) ou privada (Priv.)). Observa-se que a quantidade de cursos em instituições privadas é significativamente maior que a de instituições públicas, e que a maioria das instituições, sejam públicas ou privadas, concentram-se na região Sul (S) e Sudeste (SE).

Tabela 5: Total de instituições participantes.

Reg.	2005			2008			2011			2014			2017		
	Total	Púb.	Priv.												
Br	685	118	567	809	156	653	803	176	627	343	139	204	311	125	186
N	32	8	24	46	12	34	47	17	30	31	22	9	19	8	11
NE	78	30	48	115	40	75	126	52	74	71	44	27	55	34	21
SE	359	34	325	401	48	353	385	48	337	146	35	111	146	42	104
S	150	26	124	165	32	133	161	28	133	66	22	44	65	28	37
CO	66	20	46	82	24	58	84	31	53	29	16	13	26	13	13

Com relação ao sexo dos estudantes, feminino e masculino, apresentados na Tabela 4, a quantidade de estudantes do sexo masculino (M) é expressivamente maior de que estudantes do sexo feminino (F). Isto ocorre em todos os anos analisados, e está em conformidade com estatísticas mundiais (Lima, 2013).

7 Resultados e Discussão

O conjunto de todas as provas (2005, 2008, 2011, 2014 e 2017) analisadas nesta pesquisa contém 150 questões específicas do curso de Bacharelado em Ciência da Computação, sendo 136 objetivas e 14 discursivas. Cada prova é composta de 27 questões objetivas e 3 discursivas, exceto para o ano de 2005, com 28 objetivas e 2 discursivas.

Nas provas analisadas, 20 questões objetivas foram anuladas pelo critério do índice de discriminação, uma questão objetiva (2011) foi desconsiderada pela comissão assessora da área de computação, e uma questão discursiva (2005) foi anulada, pois, o enunciado da questão continha imprecisões que prejudicaram a sua resolução. Neste estudo, as questões anuladas e desconsideradas foram utilizadas apenas na análise de quantidade de questões por temas e índices de facilidade e discriminação. Na análise de acertos e erros, estas não foram contabilizadas. Portanto, o total de questões sob análise nesse estudo, corresponde a 116 questões objetivas e 13 questões discursivas.

Os resultados desse trabalho são apresentados de acordo com a catalogação apresentada na seção 5, 1.^a etapa item C. Inicialmente, mostramos as estatísticas das questões catalogadas por tema de uma forma geral. Logo a seguir, estão descritos os resultados a partir da catalogação das questões por índice de facilidade e índice de discriminação, sempre considerando os temas. Em seguida, são apresentadas a catalogação de temas a partir de acertos/erros dos estudantes, e, finalmente, uma visão geral sobre as questões discursivas.

7.1 Catalogação das questões por tema

A catalogação das questões por tema foi realizada com base em 11 temas, como descrito na seção 4. Esses temas foram definidos a partir da classificação das questões usando os componentes das diretrizes curriculares do curso de Bacharelado em Ciência da Computação e a matriz de temas do Enade, tentando agrupá-los na menor quantidade de temas, de modo a otimizar o processo de análise. Nenhuma questão abordou dois ou mais temas simultaneamente. Vale observar que, apesar de o Enade ter como objetivo avaliar também habilidades, essas não são abordadas neste trabalho.

A Figura 3 apresenta, no gráfico superior, a distribuição da quantidade de questões por tema em cada ano de aplicação do exame. No gráfico de barras ao lado esquerdo, o percentual geral de cada tema nas provas do Enade avaliadas por esse estudo. A barra ao lado direito com o título 'Carga Horária Semestral (horas)' apresenta a carga horária da matriz curricular 2008 do curso de Bacharelado em Ciência da Computação do Instituto de Informática da Universidade Federal de Goiás (BCC INF/UFG), que foi usada nessa análise. A escolha do BCC INF/UFG se deu, pois, o curso é totalmente aderente às diretrizes curriculares da Sociedade Brasileira de Computação (SBC), além de ter tido nota máxima no Enade nas avaliações de 2005, 2008, 2011 e 2014 (Instituto de Informática UFG, 2021) (Instituto de Informática UFG, 2021b) e nota quatro em 2017 (INEP, 2021e), outro fator decisivo para a escolha do BCC INF/UFG foi maior familiaridade dos pesquisadores com esse curso específico.

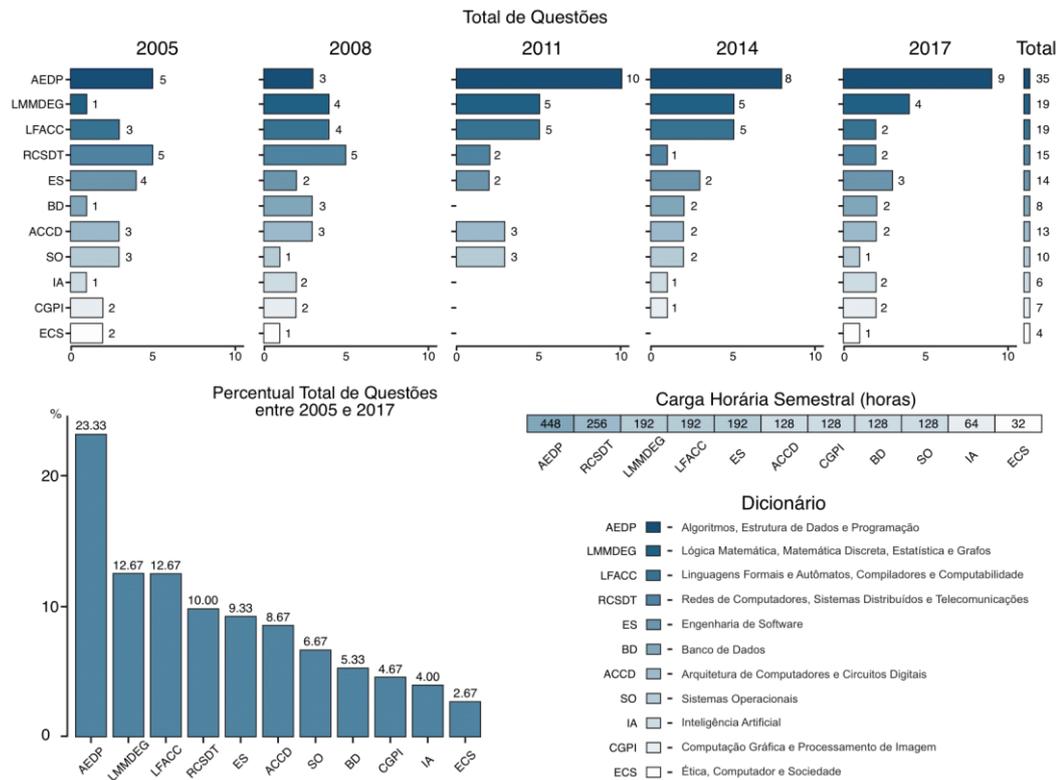


Figura 3: Quantidade de questões por temas em cada ano.

Para proporcionar uma visão comparativa de quantidade de questões e carga horária, a Figura 4, mostra a relação através de gráfico da distribuição dos temas presentes nas provas do Enade e a carga horária da matriz curricular do BCC INF/UFG. Os dados analisados apontam que os temas mais recorrentes nas provas estão congruentes com a matriz curricular do curso. Apesar de não ter sido feito um estudo abrangente sobre o assunto, existem indícios de que as provas aplicadas ao curso de Bacharelado em Ciência da Computação refletem a ênfase dada pela grade curricular na distribuição das disciplinas, ou seja, o número de disciplinas de cada área, considerando o total. Por exemplo, Inteligência Artificial (IA) tem uma disciplina na matriz do BCC INF/UFG, enquanto a área de matemática tem várias. Provavelmente, as questões foram elaboradas considerando a proporção dos conteúdos ministrados aos estudantes.

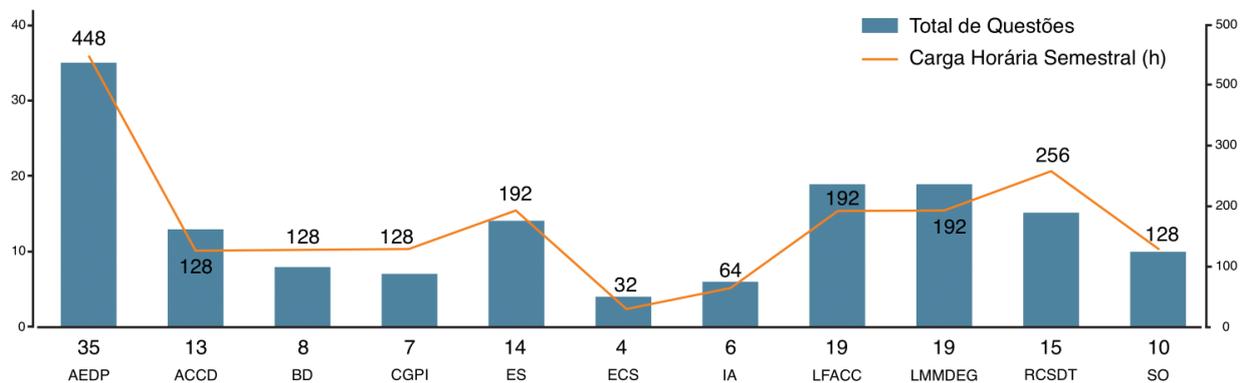


Figura 4: Comparativo entre total de questões e carga horária semestral.

Na análise dos resultados, constatou-se a predominância do tema Algoritmos, Estrutura de Dados e Programação (AEDP), com 23,33% do total de questões. Este tema aborda assuntos como Notação e Análise Assintótica, Análise de Complexidade de Problemas, Listas, Pilhas e Filas. Em segundo lugar ocorre o tema Linguagens Formais e Autômatos, Compiladores (LFACC); e

Computabilidade, e Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG) com 12,67%, com assuntos como Gramáticas, Linguagens Regulares, Autômatos, Máquina de Turing e Fases de um Compilador. Por outro lado, os temas que menos estiveram presentes dentre as questões analisadas são: Ética, Computador e Sociedade (ECS), e Inteligência Artificial (IA), ambos, consecutivamente, com 2,67% e 4%, seguidos de Computação Gráfica e Processamento de Imagem (CGPI), com 4,67%.

Ainda, percebe-se, conforme Figura 3, que a ocorrência do tema Algoritmos, Estrutura de Dados e Programação (AEDP) também é predominante como o mais abordado nos anos 2011, 2014 e 2017. No ano de 2005, há um empate quanto ao maior número de questões entre os temas: Algoritmos, Estrutura de Dados e Programação (AEDP); e Redes de Computadores, Sistemas Distribuídos e Telecomunicações (RCSDT), com 5 questões cada. Em 2008, ocorreram mais questões sobre Redes de Computadores, Sistemas Distribuídos e Telecomunicações (RCSDT), 5 questões.

Quanto à abrangência dos temas nos anos analisados, levando-se em consideração as questões que foram anuladas, é possível observar que, nos anos de 2005, 2008 e 2017, ocorreram questões dentro de todos os temas abordados. Por outro lado, 2011 e 2014 não apresentaram questões sobre os temas Banco de Dados (BD) (2011), Inteligência Artificial (IA) (2011), Computação Gráfica e Processamento de Imagem (CGPI) (2011), Ética Computador e Sociedade (ECS) (2011 e 2014). As questões anuladas não interferiram nesta análise, pois no ano de 2011, essas não estavam relacionadas aos temas Banco de Dados (BD) e Inteligência Artificial (IA). O mesmo ocorreu em 2014.

A Tabela 6 apresenta o percentual das questões anuladas por tema. Os temas Algoritmos, Estrutura de Dados e Programação (AEDP); e Linguagens Formais e Autômatos, Compiladores e Computabilidade (LFACC) concentram os maiores números de questões anuladas, com 4 questões cada tema. O ano de 2005 possui o maior número de questões anuladas, com 7 questões, seguido dos anos de 2011 e 2014, ambos com 6 questões e 2017 com apenas 3 questões anuladas. No ano de 2008, não houve questões anuladas.

Tabela 6: Questões anuladas/desconsideradas.

Tema	N.º Total de questões	Percentual	Ano	Índice de Facilidade
Algoritmos, Estrutura de Dados e Programação (AEDP)	4	18,18%	2011/2014/2017	Difícil/Muito Difícil
Arquitetura de Computadores e Circuitos Digitais (ACCD)	3	13,63%	2005/2011/2014	Médio/Difícil
Computação Gráfica e Processamento de Imagem (CGPI)	1	4,54%	2017	Difícil
Engenharia de Software (ES)	1	4,54%	2005	Difícil
Inteligência Artificial (IA)	2	9%	2005/2014	Difícil/Muito Difícil
Linguagens Formais e Autômatos, Compiladores e Computabilidade (LFACC)	4	18,18%	2005/2011/2014	Difícil/Muito Difícil
Lógica Matemática, Matemática Discreta, Estatística e Grafos(LMMDEG)	1	4,54%	2017	Difícil

Redes de Computadores, Sistemas Distribuídos e Telecomunicações (RCSDT)	3	13,63%	2017	Difícil
Sistemas Operacionais (SO)	3	13,63%	2005/2014	Difícil/Muito Difícil

7.2 Catalogação das questões por Índice de Facilidade e Índice de Discriminação

A Figura 5 apresenta o total de questões por índice de facilidade e índice de discriminação (gráfico comparativo), distribuídas também por ano de realização do exame para o curso de Bacharelado em Ciência da Computação (gráfico flutuação). Essa relação entre índice de facilidade e discriminação permite analisar se as questões com bons índices de discriminação, ou seja, questões mais aptas a avaliar os alunos, são também as questões com menor índice de facilidade, ou seja, acertadas por menor quantidade de estudantes.

Através da Figura 5, também é possível verificar que 77 das questões analisadas foram consideradas difíceis, e 15 muito difíceis. Isso implica que 2/3 das questões específicas foram respondidas corretamente por menos que 40% dos alunos, e nenhuma teve um índice de acerto superior a 85% (muito fácil). Vale ressaltar que esse índice de facilidade reflete a porcentagem de acertos de cada questão, ou seja, quanto maior o percentual de acerto de determinada questão, essa é considerada muito fácil, por outro lado, quanto menor o percentual de acerto, essa é considerada muito difícil. O índice de facilidade não representa uma avaliação da facilidade de resposta da questão, feita por especialistas da área, de forma independente do contexto de aplicação.

O poder de discriminação das questões é apresentado na Figura 5. Nela é possível perceber que, para 63%, ou seja, mais da metade das questões (68 de 108), os índices de discriminação foram bons ou muito bons. Isto pode indicar que as questões conseguiram cumprir seu papel de identificar aqueles alunos que dominam ou não o conteúdo. No entanto, 15% ainda possuem índice de discriminação fraco.

Analisando os gráficos de flutuação da Figura 5, é perceptível que, com o passar dos anos, as questões do Enade ficaram mais difíceis. A questão desconsiderada (2011) pelo INEP e as 20 questões discursivas, anuladas pelo fraco índice de discriminação, não foram classificadas pelo índice de facilidade, contabilizando 115 questões analisadas nesse item.

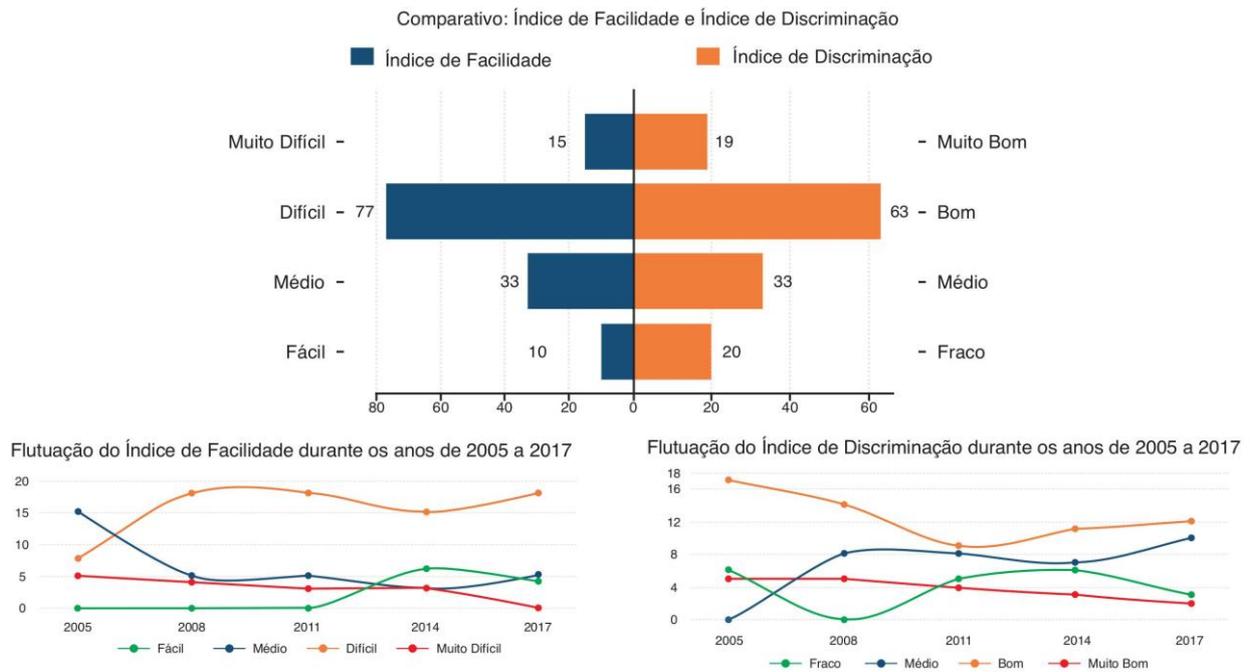


Figura 5: Quantidade de questões por Índice de Facilidade e Índice de Discriminação

Pela perspectiva dos temas analisados, as Figuras 6 e 7 a seguir apresentam, respectivamente, o percentual de questões por tema, de acordo com o índice de facilidade e índice de discriminação. Elas mostram qual o percentual de cada tema com relação aos índices utilizados. Dessa forma, é possível observar que o tema *Ética, Computador e Sociedade (ECS)*, apresenta 75% das questões classificadas com índice de facilidade médio, ou seja, quase 40% dos estudantes acertaram as questões de *Ética, Computador e Sociedade (ECS)*. Os temas *Computação Gráfica e Processamento de Imagens, e Inteligência Artificial (IA)* apresentaram questões com índices de facilidade difícil e muito difícil, isto é, no máximo 40% dos alunos conseguiram acertar essas questões, sendo que as questões muito difíceis foram acertadas por no máximo 15% dos alunos. Outros temas, como *Banco de Dados (BD)* e *Engenharia de Software (ES)* tiveram uma distribuição mais homogênea de índices de facilidade.

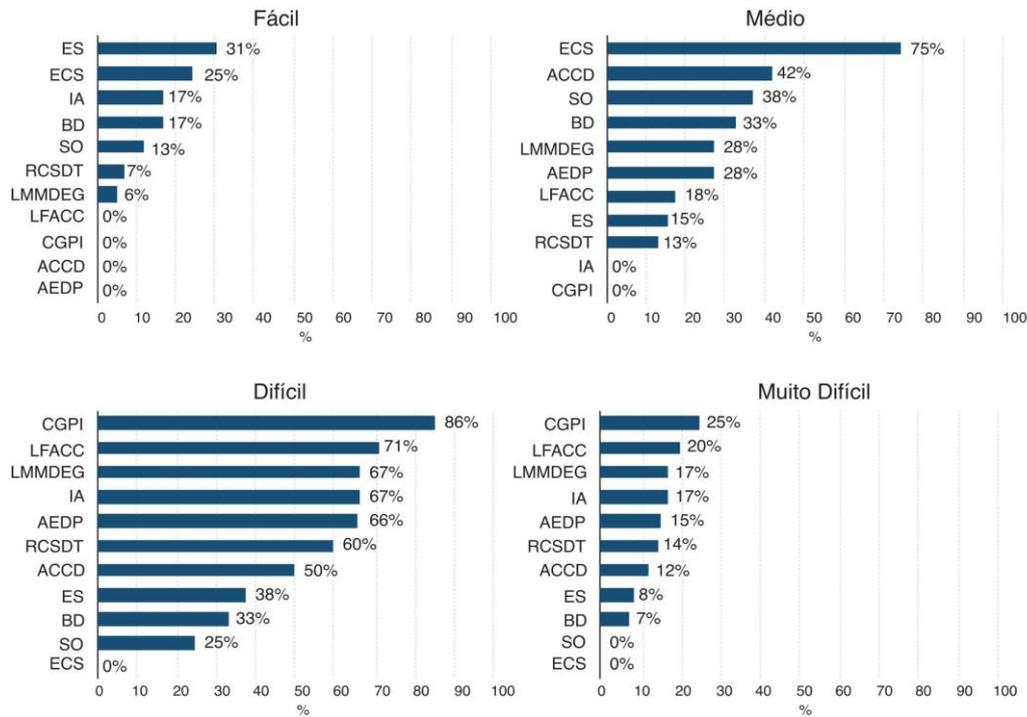


Figura 6: Percentual dos índices de Facilidade por tema.

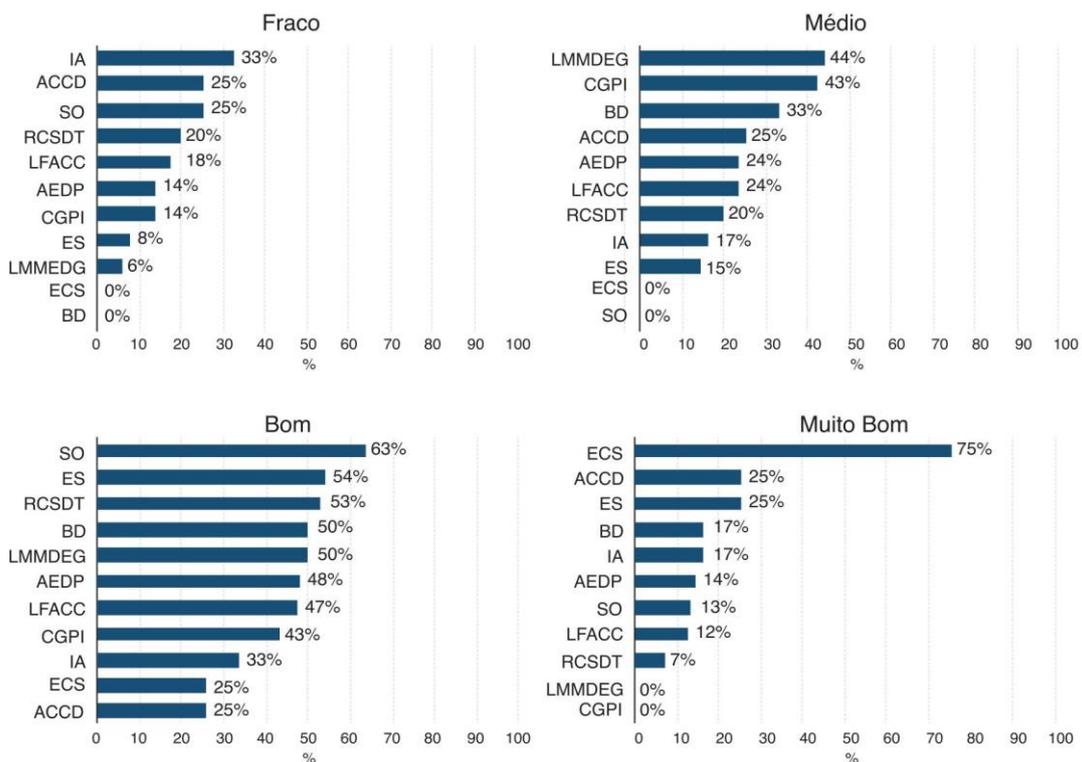


Figura 7: Percentual dos índices de Discriminação por tema.

Conforme Figura 7, o tema Ética, Computador e Sociedade (ECS) possui, dentre os temas estudados, o melhor poder de discriminação, considerando que 100% das questões deste tema foram classificadas com o índice de discriminação bom e muito bom. O tema Inteligência Artificial (IA), obteve o percentual de 33% das questões classificadas com o poder de discriminação fraco.

7.3 Catalogação das questões por Percentual de Acerto/Erro

Para obter o percentual de acerto/erro, foi considerado o vetor de resposta do estudante, checando cada questão (1 acerto e 0 erro) e comparando com o total de questões de determinado tema. A análise, feita do ponto de vista acerto/erro, mostra que, para os estudantes concluintes, nos anos de 2005, 2008, 2011, 2014 e 2017, a distribuição do percentual médio de acerto/erro por tema ocorre conforme Figura 8.

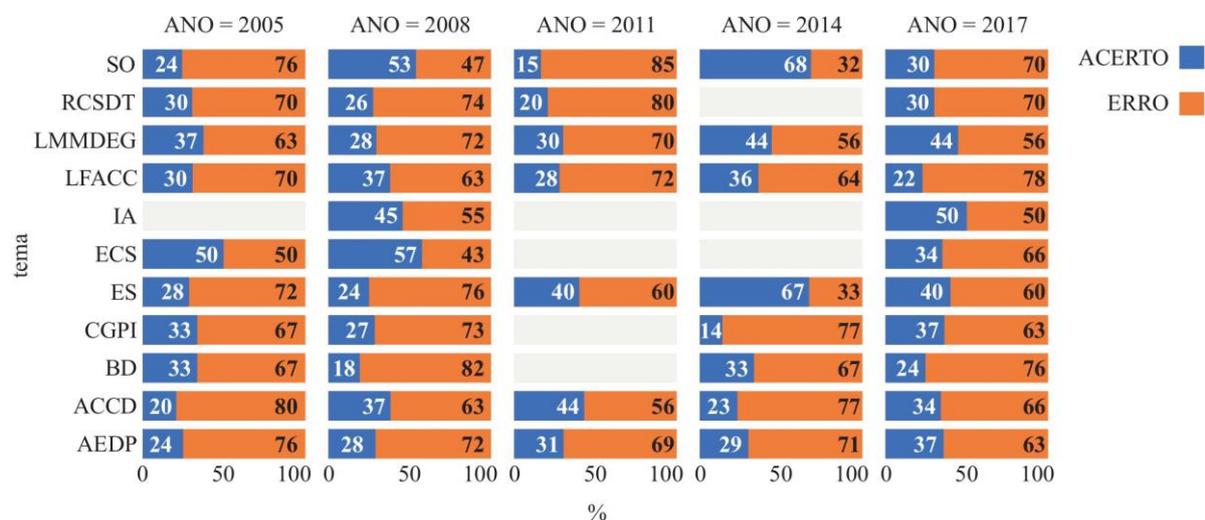


Figura 8: Percentual médio de acerto/erro por tema em cada ano.

Analisando essa distribuição, notou-se que Ética, Computador e Sociedade (ECS) é o tema de maior percentual de acerto entre os estudantes concluintes do curso de Bacharelado em Ciência da Computação nos anos de 2005 e 2008, com 50% e 57%, respectivamente. Essa tendência não pode ser verificada nos anos seguintes visto que não há questões válidas para esse tema nos anos de 2011 e 2014. Porém, em 2017, percebe-se que o percentual de acerto foi menor que em anos anteriores, cerca de 34%. Também junto aos concluintes, Sistemas Operacionais (SO) mostrou grande variabilidade de acerto de um ano para outro, com 15% em 2011 e 68% em 2014. Por outro lado, Engenharia de Software (ES) teve resultados variando de 24% em 2008 a 67% em 2014. Geralmente, os alunos acertaram 34% das questões.

Para uma análise do ponto de vista de nota por tema, a Tabela 7 traz as médias por tema em cada ano. Essa média é calculada de acordo com a fórmula do INEP para a prova do Enade, onde as questões objetivas têm peso de 85% e as questões discursivas de 15%. Nos casos em que não houve questões discursivas catalogadas no tema, foi atribuída nota (0) ao cálculo. Além das médias por tema, são apresentadas as médias referentes à nota do Componente Específico, nota da Formação Geral e nota geral do estudante, obtidas na prova do Enade.

Tabela 7: Média do Componente Específico por tema em cada ano

Tema	Média				
	2005	2008	2011	2014	2017
Algoritmos, Estrutura de Dados e Programação (AEDP)	20,80	27,57	27,43	26,11	36,76
Arquitetura de Computadores e Circuitos Digitais (ACCD)	32,23	31,80	37,63	19,51	44,15
Banco de Dados (BD)	28,16	16,82	-	66,38	33,72
Computação Gráfica e Processamento de Imagens	28,31	23,18	-	12,15	29,62
Engenharia de Software (ES)	23,64	20,27	34,36	57,10	40,04
Ética, Computador e Sociedade (ECS)	49,87	48,74	-	-	29,62
Inteligência Artificial (IA)	-	38,39	-	-	49,88
Linguagens Formais e Autômatos, Compiladores e Computabilidade (LFACC)	25,47	34,75	24,33	30,84	24,80
Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG)	31,47	24,17	26,02	37,68	34,22
Redes de Computadores, Sistemas Distribuídos e Telecomunicações (RCSDT)	25,83	22,55	16,82	-	36,88
Sistemas Operacionais (SO)	20,30	45,84	21,59	57,73	21,98

As médias apresentadas na Tabela 7 mostram uma distribuição desigual para todos os anos. O tema Algoritmos, Estrutura de Dados e Programação (AEDP), por exemplo, mesmo sendo o tema com maior quantidade de questões no ano de 2005, é o tema que possui a segunda menor média, ficando atrás apenas de Sistemas Operacionais (SO). Por outro lado, o tema Ética, Computador e Sociedade (ECS) possui, em todos os anos, média maior que a média do Componente Específico na totalidade.

Em 2014, as médias se mostram ainda mais desiguais, variando de 12,15 em Computação Gráfica e Processamento de Imagens, a 66,38 em Banco de Dados (BD). No entanto, essa desigualdade não se reflete nas médias das notas do Componente específico (39,36) e nota geral (39,36), as maiores médias desde o início da aplicação do exame.

As Figuras 6 e 7 mostram que o tema Ética, Computador e Sociedade (ECS) é o tema com poder de discriminação médio e índice de facilidade muito bom. Essa constatação pode ser verificada através da Figura 8 que mostra que esse foi o tema em que mais os alunos tiveram acertos e também através da Tabela 6 em que as melhores médias se concentram nesse tema. No entanto, é importante ressaltar que, o cálculo do índice de discriminação considera as questões que foram acertadas por alunos que tiveram um bom desempenho.

7.4 Questões Discursivas

No total de 14 questões discursivas, uma foi anulada em 2005, e dentre as outras 13 questões, a classificação compreendeu os temas de Algoritmos, Estrutura de Dados e Programação (AEDP);

Arquitetura de Computadores e Circuitos Digitais (ACCD); Banco de Dados (BD); Engenharia de Software (ES); Linguagens Formais, Compiladores e Computabilidade; Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG); e Sistemas Operacionais (SO). As áreas nas quais as questões discursivas se inserem, mostram que as questões têm sido definidas em áreas básicas do curso de Bacharelado em Ciência da Computação.

A Tabela 8 mostra a média e desvio padrão das questões discursivas dos estudantes concluintes. Os cálculos de média e desvio padrão foram feitos desconsiderando as instâncias de notas zero (0) dos estudantes. Essa tabela mostra que o tema Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG) possui a melhor média dentre as questões discursivas analisadas com valor de 65,47 (desvio padrão de 29,69), enquanto Sistemas operacionais (SO) possui média de 30,00 (desvio padrão de 25,13).

Tabela 8: Tabela com média e desvio padrão das questões discursivas dos estudantes concluintes.

Tema	Mínimo	Máximo	Média	Desvio Padrão
Algoritmos Estrutura de Dados e Programação (AEDP)	5,0	100	52,79	27,65
Arquitetura de Computadores e Circuitos Digitais (ACCD)	10,0	100	49,78	24,62
Banco de Dados (BD)	10,0	100	44,44	20,08
Engenharia de Software (ES)	5,0	100	53,43	23,00
Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG)	10,0	100	65,47	29,69
Sistemas Operacionais (SO)	5,0	100	30,00	24,13

Por outro lado, a Figura 9 apresenta a distribuição do percentual das notas obtidas pelos estudantes concluintes. A análise compreende os cinco níveis de notas e é baseada na Escala *Likert* (Likert, 1932). Na figura, é explicitado que mais que 90% dos estudantes tiveram notas em torno de 0 a 20 nos temas de Algoritmos Estrutura de Dados e Programação (92,25%); Banco de Dados (BD) (92,25%); e Sistemas Operacionais (SO) (99,15%). Enquanto, para essa faixa de notas (0 a 20), mais que 50% dos alunos as obtiveram nos temas de Arquitetura de Computadores e Circuitos Digitais (ACCD) (78,44%); Engenharia de Software (ES) (57%); e Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG) (54,58%).

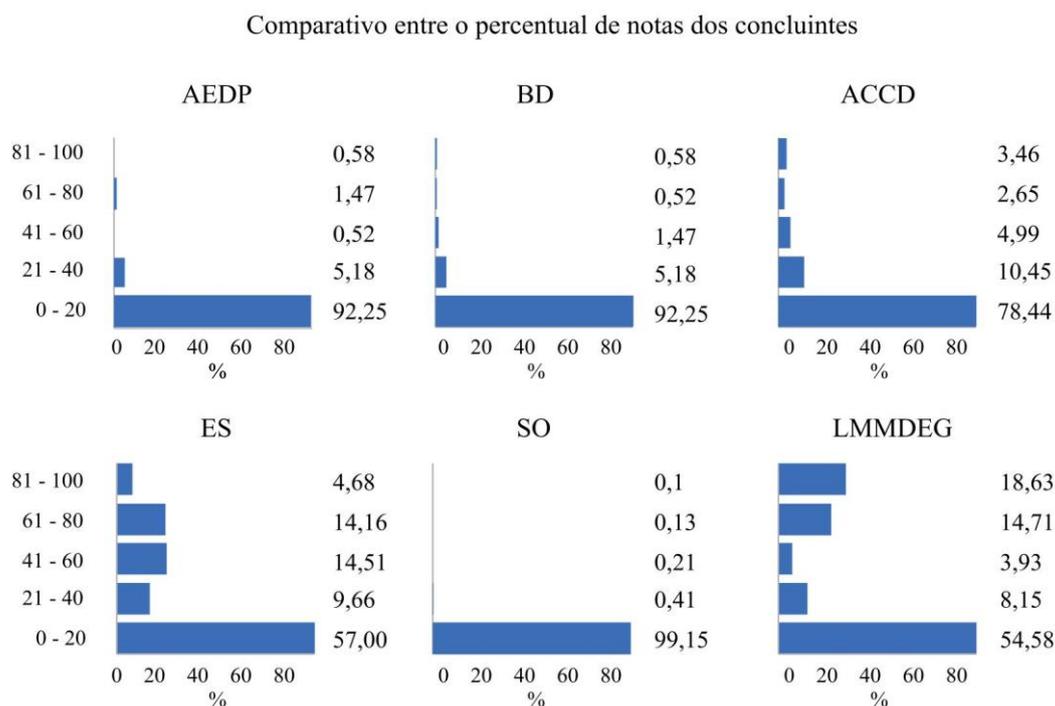


Figura 9: Percentual de Notas – Concluintes.

Os estudantes apresentaram bastante dificuldade em realizar as questões propostas, com índice elevado de notas zero (0). O percentual de notas de 0 a 20 em Sistemas Operacionais (SO), chama atenção, pois esse é um tema abordado na metade do curso de Bacharelado em Ciência da Computação. Era de se esperar que os alunos tivessem conhecimentos mínimos sobre esse tema. A Tabela 9 traz informações detalhadas das questões discursivas tais como o tema e o assunto relacionado em cada questão.

Tabela 9: Detalhes das questões discursivas.

Ano	Questões	Temas	Assunto relacionado
2005	Q55	Arquitetura de Computadores e Circuitos Digitais	Organização de computadores, hierarquia de memória
	Q70	Linguagens Formais e Autômato, Compiladores e Computabilidade	— Questão Anulada —
2008	Q20	Algoritmos Estrutura de Dados e Programação	Tabela <i>hash</i>
	Q39	Linguagens Formais e Autômato, Compiladores e Computabilidade	Gramática
	Q40	Banco de Dados (BD)	Banco de dados relacionais, linguagens de consulta e manipulação de dados
2011	QD3	Algoritmos, Estrutura de Dados e Programação (AEDP)	Conceitos básicos de programação e recursão
	QD4	Algoritmos, Estrutura de Dados e Programação (AEDP)	Conceitos básicos de programação, árvores, recursão
	QD5	Sistemas Operacionais (SO)	Gerenciamento de memória

2014	QD3	Algoritmos, Estrutura de Dados e Programação (AEDP)	Conceitos básicos de programação, <i>backtracking</i> , recursão
	QD4	Banco de Dados (BD)	Modelagem de banco de dados, banco de dados relacionais
	QD5	Algoritmos, Estrutura de Dados e Programação (AEDP)	Técnicas de projeto de algoritmos (divisão-conquista)
2017	QD3	Algoritmos, Estrutura de Dados e Programação (AEDP)	Análise de algoritmo, complexidade e estrutura de dados
	QD4	Engenharia de Software (ES)	Diagrama UML
	QD5	Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG)	Grafos, busca em profundidade

Nas questões discursivas, assim como nas questões objetivas, o tema de maior recorrência foi Algoritmo, Estrutura de Dados e Programação (6 questões discursivas), especificamente abordando conceitos básicos de programação, recursão e tabela *hash*. No extremo oposto, com 1 questão para cada tema, estão Arquitetura de Computadores e Circuitos Digitais (ACCD); Engenharia de Software (ES); e Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG).

8 Conclusões

O principal objetivo do Enade é avaliar a trajetória dos estudantes nos cursos superiores no Brasil. Entretanto, esse exame possibilita avaliar também as instituições e os cursos a elas vinculados. Apesar de fazer parte do cenário da educação superior, existe grande discussão sobre esse modelo avaliativo. Há quem acredite que seja uma forma eficiente de avaliação e outros acreditam que não seja tão eficiente assim (Penin & Messias, 2009) (Matos, Chagas & Menezes, 2016). O fato é que não é possível a criação de processos avaliativos com aplicação em larga escala extremamente precisos. Porém, é papel de toda a sociedade, e não apenas de especialistas, contribuir com melhorias para a qualidade da educação. Isso implica pensar em cobrir todas as dimensões que compõem o conceito “Qualidade da Educação Superior”. Um caminho é realizar pesquisas que venham contribuir para aumentar a eficiência e eficácia de exames como o Enade.

No sentido de contribuir nesse contexto, esse trabalho faz a adaptação de uma metodologia para a análise do conteúdo das provas do Enade, utilizando como estudo de caso o curso de Bacharelado em Ciência da Computação. A análise das questões do componente específico da prova por temas definidos a partir do currículo de referência da área proporciona uma abrangência maior para as análises realizadas a partir da aplicação do exame. Além disso, permite a detecção de falhas na elaboração da prova e na formação dos estudantes, com foco principal em áreas de conhecimento.

Neste estudo, aplicamos a metodologia proposta em (Lima, Ambrósio, Felix, Brancher & Ferreira, 2018) ao exame Enade para verificar sua aplicabilidade aos exames nacionais e o esforço necessário para adaptá-la a um contexto específico. A aplicabilidade do método para um contexto diferente da proposta original é possível desde que, a estrutura e o conhecimento específico verificado pelo exame a ser analisado, seja profundamente dominado pelos pesquisadores. Esse domínio influenciará diretamente no esforço implementado para a adaptação e aplicação. No caso

do Enade, a adaptação e aplicação foram relativamente simples, pois os pesquisadores tinham vasto conhecimento sobre o exame. Entretanto, o trabalho requereu esforço considerável, uma vez que o Enade é um exame complexo.

Um levantamento bibliográfico mostrou que não existem outros estudos sobre o conteúdo das questões do Enade, que buscam categorizar as questões de acordo com a área de conhecimento. Por isso, este estudo mostra-se inovador. Ao explorar o conteúdo das provas, é possível identificar outras categorias de análises com foco na área de conhecimento, inclusive quais áreas de conhecimento são preditoras do sucesso no resultado do exame.

Os dados analisados apontam os temas mais recorrentes, Algoritmos, Estrutura de Dados e Programação (AEDP); Linguagens Formais e Autômatos, Compiladores e Computabilidade (LFACC); e Lógica Matemática, Matemática Discreta, Estatística e Grafos (LMMDEG). Esses temas estão congruentes com a matriz curricular do curso de Bacharelado em Ciência da Computação. Apesar de não ter sido feito um estudo abrangente sobre o assunto, existem indícios de que as provas aplicadas ao curso de Bacharelado em Ciência da Computação refletem a ênfase dada pela grade curricular na distribuição das disciplinas, ou seja, o número de disciplinas de cada área, considerando o total. Por exemplo, IA tem uma disciplina na matriz do BCC INF/UFG, enquanto a área de matemática tem várias. Provavelmente, as questões foram elaboradas considerando a proporção dos conteúdos ministrados aos estudantes. Acredita-se que a adoção do BCC INF/UFG não compromete a análise feita, dado que esse curso tem sua matriz curricular baseada nas diretrizes da SBC, assim como a maioria dos cursos de BCC do Brasil.

Não foi realizada análise das questões para identificar características que determinam a facilidade da questão do ponto de vista do seu conteúdo. Isso deve ser feito utilizando classificação feita pelos avaliadores ou pelos alunos que realizam as provas. O Índice de Facilidade disponibilizado pelo INEP calcula esse índice apenas com base na quantidade de acertos. Trabalhos futuros poderiam tentar classificar as questões por níveis cognitivos, como, por exemplo, usando "Taxonomia de Bloom" (Ferraz & Belhot, 2010). Além disso, vale a pena investigar a relevância dos conteúdos cobrados nas provas para as atividades realizadas pelos egressos no mercado de trabalho. Com relação às análises realizadas aqui, a pretensão consiste em apresentar possibilidades de análises adicionais àquelas fornecidas pelos órgãos competentes e não fazer uma análise exaustiva, considerando todos os aspectos possíveis. Portanto, o escopo adotado para a análise deve ser compreendido como um estímulo para que outras pesquisas sejam conduzidas nesta perspectiva - diante da viabilidade de fazê-la e da riqueza dos dados obtidos.

No curso de Bacharelado em Ciência da Computação, há forte predominância de itens com índice de facilidade difícil e muito difícil, indicando que as provas estão sendo elaboradas acima do nível de conhecimento esperado dos alunos, ou que os alunos não estão atingindo este nível. Menos da metade dos estudantes, 40%, acertaram mais do que 50% das questões. Além disso, índice de facilidade mais alto não significa maior poder de discriminação das questões. Esses aspectos devem ser considerados pelos avaliadores quando elaboram as provas. Estudos mais detalhados devem tentar identificar características que influenciam nestes índices de modo a oferecer diretrizes para a elaboração das questões.

Os estudos apontaram que no ano de 2005, primeiro ano de aplicação da prova, as questões eram mais difíceis se comparadas aos outros anos de aplicação da prova. Por outro lado, no ano de 2014, parece que as questões aplicadas eram mais fáceis. Quanto aos temas, Engenharia de Software (ES) e Sistemas Operacionais (SO) são os que mais tiveram questões fáceis. Por outro lado, Computação Gráfica e Processamento de Imagem (CGPI); e Inteligência Artificial (IA) tiveram mais questões difíceis e o tema Ética, Computador e Sociedade (ECS) teve todas as questões medianamente fáceis. O índice de facilidade das questões mostrou uma distribuição não homogênea, com algumas áreas contendo apenas questões consideradas difíceis ou muito difíceis.

O ideal é que houvesse uma distribuição mais homogênea como apresentado nas áreas de Banco de dados (BD) e Engenharia de Software (ES).

Cerca de 25% das questões foram anuladas devido ao baixo índice de discriminação. Uma análise de outras áreas de conhecimento seria interessante para verificar se esse índice se mantém no Enade todo. De qualquer forma, ele deve ser considerado pela comissão que elabora as provas, em uma tentativa de diminuir esse percentual e elevar a quantidade de questões com alto índice de discriminação, visto que estas são as que apresentam melhores resultados na identificação de alunos com maior rendimento, pois o cálculo do índice de discriminação considera as questões que tiveram o maior número de acertos por alunos de bom desempenho.

A descrição da amostra comprova a desigualdade entre a quantidade de estudantes do sexo masculino e feminino. No curso de Bacharelado em Ciência da Computação, a maioria dos estudantes são do sexo masculino, atingindo em alguns anos 87% dos estudantes. Isso demonstra a necessidade de investigar porque o número de estudantes do sexo feminino é tão baixo no curso de Bacharelado em Ciência da Computação e quais as ações viáveis para reverter este quadro. Uma análise dos temas nos quais as alunas têm melhor desempenho pode influenciar na definição da grade curricular de cursos de Bacharelado em Ciência da Computação de modo a atrair mais mulheres.

Apesar das constatações acima, não é possível afirmar se o Enade é a melhor forma de avaliar a qualidade da educação superior, mesmo que colabore para a melhoria da Educação em posicionamento nas classificações mundiais. Cabe salientar que o Enade não soluciona sozinho a defasagem educacional superior brasileira, porém deve ser usado como instrumento de investigação sobre a formação do estudante durante o caminho acadêmico, sendo necessários estudos e pesquisas constantes de modo a aprimorar esta ferramenta.

Agradecimentos

À Professora Dra. Ana Paula Laboissière Ambrósio (*In memoriam*) pela orientação paciente e sábia desta pesquisa.

À Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) pelo apoio financeiro para o desenvolvimento deste trabalho.

Referências

- Almeida Teodoro de e Marco André Kappel, L. (2020). Aplicação de técnicas de aprendizado de máquina para predição de risco de evasão escolar em instituições públicas de ensino superior no Brasil. *Revista Brasileira de Informática na Educação*, 28(0), 838–863. Disponível em: <https://www.br-ie.org/pub/index.php/rbie/article/view/v28p838> doi: 10.5753/rbie.2020.28.0.838 [GS Search]
- Bianchi, J. M. (2010). Leitura e Interpretação de Tabelas e Gráficos pelos Alunos do Curso de Matemática da URI – Campus de Erechim. *Monografia (Graduação) - Universidade Regional do Alto Uruguai e das Missões*.
- Brito, T. F. (2015). Corpo Docente : Fatores Determinantes do Desempenho Discente no ENADE. Dissertação (Mestrado) - Universidade de São Paulo. doi: [10.11606/D.12.2016.tde-21032016-115045](https://doi.org/10.11606/D.12.2016.tde-21032016-115045) [GS Search]
- Costa, J. C. A. (2016). Vis-Scholar: Uma Metodologia de Visualização e Análise de Dados na Educação. Dissertação (Mestrado) - Universidade do Vale do Rio dos S. [GS Search]

- Costa, J. P. D. C., & Martins, M. I. (2014). O ENADE para a licenciatura em física: Uma proposta de Matriz de Referência. *Revista Brasileira de Ensino de Física*, 36(3), 3401 -3401/9. Disponível em: www.sbfisica.org.br [GS Search]
- Ferraz, A. P. d. C. M., & Belhot, R. V. (2010). Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. *Gestão & Produção*, 17(2), 421–431. Disponível em: <http://www.scielo.br/pdf/gp/v17n2/a15v17n2> [GS Search]
- Fonseca, S. O. d., & Namen, A. A. (2016, 3). Mineração em Bases de Dados do INEP: Uma Análise Exploratória para Nortear Melhorias no Sistema Educacional Brasileiro. *Educação em Revista*, 32(1), 133–157. doi: [10.1590/0102-4698140742](https://doi.org/10.1590/0102-4698140742) [GS Search]
- Gomes, P. H. D. F. (2011). Efeitos do Exame Nacional de Desempenho dos Estudantes/ENADE (2008) nas licenciaturas de História e Letras da Universidade de Brasília. *Monografia (Graduação) - Universidade de Brasília*. Disponível em: <http://bdm.unb.br/handle/10483/3195> [GS Search]
- Guerra, P. C., Nakamura, R. Y. M., & Hruschka, E. R. (2014). Estimativa de Demanda Potencial de Matrículas em Ensino Superior usando Dados Públicos e Múltiplos Modelos de Regressão. *Symposium on Knowledge Discovery, Mining and Learning*, 2. Disponível em: <http://www.producao.usp.br/handle/BDPI/48650> [GS Search]
- Howell, D. C. (2010). *Statistical Methods for Psychology* (7ª Ed. ed.). Belmonte: Wadsworth. [GS Search]
- INEP (2005). Relatório de Síntese 2005. Disponível em: <http://portal.inep.gov.br/relatorios>. Acesso em: 26 mar. 2021.
- INEP (2008a). Portaria Inep nº 126 de 07 de agosto de 2008. Disponível em: https://download.inep.gov.br/download/superior/enade/Diretrizes%20Enade/Diretrizes_Computacao_n_126.pdf. Acesso em: 26 mar. 2021.
- INEP (2008b). Relatório de Síntese – Computação. Disponível em: <http://portal.inep.gov.br/relatorios>. Acesso em: 26 mar. 2021.
- INEP (2011). Portaria Inep nº 239 de 04 de agosto de 2011. Disponível em: https://download.inep.gov.br/educacao_superior/enade/legislacao/2011/diretrizes/diretrizes_computacao_n_239.pdf. Acesso em: 26 mar. 2021.
- INEP (2014). Portaria Inep nº 239, de 02 de junho de 2014. Disponível em: https://download.inep.gov.br/educacao_superior/enade/legislacao/2014/diretrizes_cursos_diploma_tecnologo/diretrizes_diploma_tecnologo_analise_desenvolvimento_sistemas.pdf. Acesso em: 26 mar. 2021.
- INEP (2017a). Manual do ENADE 2016. Disponível em: https://download.inep.gov.br/educacao_superior/enade/manuais/manual_do_enade_28092016.pdf. Acesso em: 26 mar. 2021.
- INEP (2017b). Portaria Inep nº 473, de 06 de junho de 2017. Disponível em: https://www.in.gov.br/web/guest/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/19105689/DiarioOficialdaUniao#:~:text=%C3%A1reas%20do%20conhecimento,-,Art.,Geral%20e%20do%20componente%20espec%C3%Adfico. Acesso em: 26 mar. 2021.
- INEP (2021a). Instituto Nacional de Estudos e Pesquisas Educacionais. Avaliação dos Cursos de Graduação. Brasília. Disponível em <http://portal.inep.gov.br/web/guest/avaliacao-dos-cursos-de-graduacao>. Acesso em 12 jan. 2020.

- INEP (2021b). Instituto Nacional de Estudos e Pesquisas Educacionais. Enade. Brasília, 2020b. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enade>. Acesso em: 28 jan. 2020.
- INEP (2021c). Instituto Nacional de Estudos e Pesquisas Educacionais. Indicadores de Qualidade da Educação Superior. Brasília, 2020c. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior>. Acesso em: 28 jan. 2020.
- INEP (2021d). Sistema Nacional de Avaliação da Educação Superior (Sinaes). Disponível em: <http://portal.inep.gov.br/sinaes>. Acesso em: 26 mar. 2021.
- INEP (2021e). Relatórios Públicos. Disponível em: <http://enade.inep.gov.br/enade/#!/relatorioIES>. Acesso em 13 mar. 2021.
- Instituto de Informática UFG (2021). Ciência da Computação. Disponível em: <https://www.inf.ufg.br/p/30138-ciencia-da-computacao>. Acesso em 13 mar. 2021.
- Instituto de Informática UFG (2021b). Enade. Disponível em: <https://inf.ufg.br/p/30125-enade>. Acesso em 13 mar. 2021.
- Lara, I. C. M. de (2007). Exames Nacionais e as "verdades" sobre a Produção do Professor de Matemática. *Tese (Doutorado) - Universidade Federal do Rio Grande do Sul*. [GS Search]
- Likert, R. (1932). A technique for the measurement of attitudes (No. N° 136-165). *Archives of Psychology*. Disponível em: <https://books.google.com.br/books?id=9rotAAAYAAJ> [GS Search]
- Lima, M. P. (2013). As mulheres na Ciência da Computação. *Revista Estudos Feministas*, 21, 793–816. doi: [10.1590/S0104-026X2013000300003](https://doi.org/10.1590/S0104-026X2013000300003) [GS Search]
- Lima, P., Rosa, E., Ambrósio, A., & Oliveira, J. (2019). Applying content analysis to brazilian national exams. *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, 8(1), 139. Disponível em: <https://www.br-ie.org/pub/index.php/wcbie/article/view/8955> doi: [10.5753/cbie.wcbie.2019.139](https://doi.org/10.5753/cbie.wcbie.2019.139) [GS Search]
- Lima, P. d. S. N., Ambrosio, A. P. L., Ferreira, D. J., & Brancher, J. D. (2019, 05). Análise de dados do Enade e Enem: uma revisão sistemática da literatura. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 24, 89 - 107. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1414-40772019000100089&nrm=iso doi:[10.1590/s1414-40772019000100006](https://doi.org/10.1590/s1414-40772019000100006) [GS Search]
- Lima, P. S. N., Ambrósio, A. P. L., Felix, I. M., Brancher, J. D., & Ferreira, D. J. (2018). Content Analysis of Student Assessment Exams. *48th Annual Frontiers in Education (FIE)*. doi: [10.1109/FIE.2018.8659169](https://doi.org/10.1109/FIE.2018.8659169) [GS Search]
- Lopes, F. L., & Vendramini, C. M. M. (2015). Propriedades psicométricas das provas de pedagogia do ENADE via TRI. *Avaliação: Revista da Avaliação da Educação Superior (Campinas)*, 20(1), 27–47. doi: [10.590/S1414-40772015000100004](https://doi.org/10.590/S1414-40772015000100004) [GS Search]
- Matos, K. U. H. S. d., Chagas, S. O., & Menezes, C. R. C. d. (2016). Enade: O Desafio De Uma Avaliação Do Ensino Superior Eficaz Para As Instituições De Ensino. *9º Encontro Internacional de Formação de Professores e 10º Fórum permanente internacional de inovação educacional*, 9. Disponível em: <https://eventos.set.edu.br/inde.php/enfope/article/view/2349/566> [GS Search]
- MEC (2005). Portaria Inep nº 179/2005. Disponível em: http://download.inep.gov.br/download/enade/PORTARIAS_ENADE_2005/Computacao.pdf. Acesso em 28 de jan. 2020.

- MEC (2016). Resolução n.º 5, de 16 de novembro de 2016. *Ministério da Educação, Conselho Nacional de Educação, Câmara de Educação Superior*. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/22073129/do1-2016-11-17-resolucao-n-5-de-16-de-novembro-de-2016-22073052
- Neto, W., Durigon, A., Avila, A., & Junior, E. (2020). Sistema Inteligente para o Ensino-Aprendizagem de Expressões Algébricas. *Revista Brasileira de Informática na Educação*, 28, 359-388. doi: [10.5753/rbie.2020.28.0.359](https://doi.org/10.5753/rbie.2020.28.0.359) [GS Search]
- Novossate, S. (2010). O ENADE e os Documentos Curriculares: Um Estudo sobre a Formação de Professores de Biologia. *Dissertação (Mestrado) - Universidade Federal do Paraná*. [GS Search]
- Penin, S., & Messias, W. (2009). Debate: O Enade é eficiente como forma de avaliação da USP? – *Jornal do Campus*. Disponível em: <http://www.jornaldocampus.usp.br/index.php/2009/09/debate-o-enade-e-eficiente-como-forma-de-avaliacao-da-usp/>. Acesso em: 27 mar. 2021.
- Picanço, F. (2016). Juventude e acesso ao ensino superior no Brasil: Onde está o alvo das políticas de ação afirmativa. *Latin American Research Review*, 51(1), 109–131. doi:[10.1353/lar.2016.0001](https://doi.org/10.1353/lar.2016.0001) [GS Search]
- Primi, R., Hutz, C. S., & Da Silva, M. C. R. (2011). A Prova do ENADE de Psicologia 2006: Concepção, Construção e Análise Psicométrica da Prova. *Revista Avaliação Psicológica*, 10(3), 271–294. [GS Search]
- Saviani, D. (2010). A expansão do ensino superior no Brasil: mudanças e continuidades. *Poiesis Pedagógica*, 8(2), 4–17. doi: [10.5216/rpp.v8i2.14035](https://doi.org/10.5216/rpp.v8i2.14035) [GS Search]
- Schwengber, L. (2013). Exame Nacional de Desempenho de Estudantes: Problematizando Verdades sobre a Formação do Professor de Matemática. *Dissertação (Mestrado) - Universidade de Santa Cruz do Sul*. Disponível em: <http://hdl.handle.net/11624/553> [GS Search]
- Souza, E. S. de (2008). ENADE 2006: Determinantes do Desempenho dos Cursos de Ciências Contábeis. *Dissertação (Mestrado) - Programa Multiinstitucional e Inter-Regional de Pós Graduação - Universidade de Brasília, Universidade Federal do Pernambuco e Universidade Federal do Rio Grande do Norte*. [GS Search]
- Souza, H., Neiva, D., Cavalcanti, R., Rodrigues, R., Gomes, A., & Adeodato, P. (2017). Uma Análise preditiva de desempenho dos alunos dos cursos no ENADE com base no perfil socioeconômico e de desempenho do Enem. *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, 6(1), 684. doi: [10.5753/cbie.wcbie.2017.684](https://doi.org/10.5753/cbie.wcbie.2017.684) [GS Search]
- Vassallo, R. (2016). A difícil tarefa de avaliar a qualidade do ensino superior. Disponível em: <http://jornal.usp.br/universidade/a-dificil-tarefa-de-avaliar-a-qualidade-do-ensino-superior-2>
- Verhine, R. E., & Dantas, L. M. V. (2005). Avaliação da Educação Superior no Brasil: do Provão ao ENADE. *Banco Mundial*, 1 - 48. doi: [10.1590/S0104-40362006000300002](https://doi.org/10.1590/S0104-40362006000300002) [GS Search]
- Vieira, M. d. F. A. (2012). Prevalência de retenção escolar e fatores associados em adolescentes da coorte de nascimentos de 1993 em Pelotas, Brasil. *Revista Panamericana de Salud Publica*, 31(4), 303–310. doi: [10.1590/S1020-49892012000400006](https://doi.org/10.1590/S1020-49892012000400006) [GS Search]