

## Caracterização de elementos de ensino para retroalimentação do processo de planejamento com o uso de análise de agrupamentos

### *Typification of instruction elements for feedback in the planning process with cluster analysis*

Gustavo Danicki Aureliano Rosa  
Instituto Nacional de Estudos e Pesquisas Educacionais  
Anísio Teixeira  
ORCID: [0000-0003-2531-472X](https://orcid.org/0000-0003-2531-472X)  
[gustavo.danicki@gmail.com](mailto:gustavo.danicki@gmail.com)

Luiz Síveres  
Universidade Católica de Brasília  
ORCID: [0000-0003-4735-6066](https://orcid.org/0000-0003-4735-6066)  
[luiz.siveres@gmail.com](mailto:luiz.siveres@gmail.com)

### Resumo

*Esta pesquisa objetivou caracterizar elementos de organização didático-pedagógica do ensino para feedback no ciclo de planejamento, a partir de dados de percepção de estudantes. Esse objetivo se insere na contribuição para a criação de melhores ambientes de ensino a partir da análise de dados e sua interpretação pedagógica no ambiente educacional. Com o uso de análise de agrupamentos os dados de percepção foram segmentados em função da semelhança entre as observações, considerando as respostas a itens selecionados (relação docente discente, exigência do curso, conteúdos, avaliações e apoio recebido pelo estudante) de questionário do estudante do Exame Nacional de Desempenho dos Estudantes – ENADE, das edições de 2014 e 2017 com dados de licenciaturas do Distrito Federal. O algoritmo de agrupamento utilizado foi o Partitioning Around Medoids (PAM), alimentado por matriz de distâncias de Gower para o tratamento de conjuntos de dados com variáveis mistas. O número ótimo de agrupamentos foi determinado com a medida média de Silhueta. Os resultados possibilitaram caracterizar elementos para o ajuste do ensino frente determinadas características de percepção dos estudantes sobre o curso de graduação. A análise dos agrupamentos conduziu à identificação da interação entre docentes e discentes como elemento de ensino passível de melhoria. Esse elemento de ensino conduziu aos elementos de “definição de procedimentos de ensino” e “especificação de recursos de ensino” como pontos para análise na retroalimentação do processo de planejamento pedagógico.*

**Palavras-Chave:** *Análise de dados educacionais; Planejamento de ensino; Melhoria de ambiente de aprendizagem.*

### Abstract

*This research aims to outline the elements of pedagogical-didactic organization (ODP) in teaching, incorporating them into the feedback for the planning cycle through data from students perception surveys. The referred objectives contributes to advances in the learning environment through data analysis and pedagogical interpretation in the learning environment. The collected data was grouped according to similarities between observations, considering answers to selected items (teacher/student dynamic, content, evaluations and support received by students) in questionnaires from the 2014 and 2017 editions of ENADE (Brazilian National Exam for the Assessment of Students Performance), considering data of undergraduate students from the Federal District. The algorithm applied was the Partitioning Around Medoids (PAM), feeded by a distance matrix calculated with Gower distance for datasets with mixed variables. The optimal number of groups was determined by the application of Average Silhouette Width. The results made it possible to tipify elements so as to adjust instruction planning to certain perceptions from the students about the undergraduation. Assessment of the groups suggests the interaction between students and professors is an educational element suitable for improvement. Considering such educational element, the elements “establishing teaching procedures” and “teaching resources specification” were selected as items for assessment in the fedback for the teaching planning.*

**Keywords:** *Educational data analysis; Instructional planning; Improvements in the learning environment.*

## Introdução

A melhora do processo de ensino por meio da retroalimentação do planejamento pedagógico é atividade inerente ao esforço educacional sistemático, observado o objetivo do ensino que é a facilitação das aprendizagens de estudantes (Haydt, 2006). Nesse contexto, diversos insumos podem ser utilizados para retroalimentar o processo de planejamento, dentre os quais: (a) os resultados de avaliações educacionais e de aprendizagem e; (b) os resultados da análise e interpretação pedagógica de dados sobre a percepção dos estudantes submetidos a determinado conjunto organizado e sistemático de práticas de ensino. Entretanto, cabe apontar que uma análise do ensino apenas em função dos resultados alcançados pelos estudantes em avaliações, ou mesmo de suas percepções imediatas acerca do ensino em determinado programa educacional, pode não alcançar o conjunto de variáveis intervenientes no processo de ensino.

Variáveis externas ao contexto de ensino imediato podem ter influência importante sobre os resultados e percepções de estudantes. Assim, instituições de ensino devem orientar sua prática por meio do uso de dados variados. Tanto dados inerentes à avaliação da aprendizagem dos estudantes, quanto dados de percepções sobre determinados elementos de ensino devem ser integrados em ciclos analíticos para que possam ser alcançadas conclusões significativas do ponto de vista pedagógico para o incremento da atividade de ensino, seu planejamento e replanejamento à luz de retroalimentação pertinente para a criação de melhores ambientes de aprendizagem. Nesse sentido, e considerando melhores práticas da organização do ambiente de ensino – a partir de elementos centrados no aprendiz, no conhecimento, na avaliação ou comunidade, por exemplo -, como apontado por Bransford, Brown e Cocking (2007), elementos de planejamento – como i) previsão de objetivos educacionais; (ii) especificação de itens e subitens do conteúdo; (iii) definição de procedimentos de ensino; (iv) indicação de recursos de ensino; e (v) estabelecimento da forma e conteúdo das avaliações em aderência aos objetivos (Haydt, 2006) -, podem ter insumo importante a partir da análise e interpretação pedagógica de dados associados à percepção de estudantes sobre (i) a relação professor-aluno para o estímulo discente; (ii) a exigência do curso para a adequada mobilização discente, (iii); conteúdos abordados em disciplinas ou componentes curriculares; e (iv) a aderência entre avaliações da aprendizagem e conteúdos.

A análise de dados educacionais tem sido objeto de áreas mais amplas de análise de dados, sendo que no âmbito educacional pode-se citar por exemplo a vertente denominada *Learning Analytics* (Chiappe & Rodriguez, 2017) e a Mineração de Dados Educacionais ou *Educational Data Mining* (EDM) (Baker, Isotani & Carvalho, 2011). Nesse cenário, a disseminação de métodos analíticos atuais e baseados na disponibilidade de dados e capacidade de processamento computacional para docentes, analistas e gestores educacionais, assim como estudantes, pode implicar em impacto significativo para a prática educacional que se quer baseada em evidências (Chiappe & Rodriguez, 2017).

Em vista de um cenário de práticas de gestão baseadas em dados, os planejamentos institucional e pedagógico de Instituições de Educação Superior (IES), necessitam de fluxos constantes e definidos de coleta, análise e interpretação de dados das mais variadas fontes. Essa necessidade advém da variedade de cursos, estudantes e processos envolvidos nesse tipo de organização. Nesse contexto, e considerando a existência de dados coletados e tratados periodicamente em um processo nacional de avaliação como o Sistema Nacional de Avaliação da Educação Superior (SINAES), as instituições de educação superior possuem um rol de dados disponível que podem sustentar o seu processo de planejamento pedagógico e a posterior

retroalimentação para replanejamento. Como apontam Tives, Canedo, Ladeira e Fagundes (2018), a análise de dados do Exame Nacional de Desempenho dos Estudantes (ENADE), em especial do questionário do estudante, com o uso de Mineração de Dados pode fornecer informações importantes para a gestão de IES. Entretanto, a baixa utilização de dados do ENADE em estudos com o uso de Mineração de Dados é ressaltado por Lima, Ambrósio, Ferreira e Brancher (2019), em revisão de estudos que se utilizam desses dados oficiais.

No contexto apresentado, de existência de técnicas analíticas apropriadas e arcabouço conceitual pedagógico com potencial de agregar valor ao processo de planejamento pedagógico das IES, com a concomitante escassez de estudos que se utilizam desses meios, por exemplo, foi identificado o problema do baixo uso de dados do questionário do estudante submetido ao ENADE, em ciclo de análise, interpretação pedagógica e integração de áreas para o planejamento de IES - na revisão do planejamento das condições de organização didático pedagógica do ambiente de ensino, e eventuais sistemas de apoio, para que se possa proporcionar melhores ambientes de ensino aos estudantes. Tal situação motivou a questão de pesquisa subjacente a este estudo que buscou contribuir com discussões sobre um *framework* analítico que unifica técnica de análise de dados (análise de agrupamentos) com arcabouço de interpretação pedagógica para melhoria do ensino por meio do planejamento pedagógico e sua retroalimentação, com o uso de dados de percepção de estudantes. A investigação proposta tem potencial para sustentar processos institucionais internos das IES na implantação de áreas multidisciplinares de análises de dados pedagógicos, assim como integrar diferentes instâncias existentes nessas instituições, como Núcleos Docentes Estruturantes (NDE), coordenações de curso, Comissão Própria de Avaliação (CPA) e áreas de atendimento estudantil para apoio psicopedagógico e psicológico, por exemplo.

De forma detalhada, o objetivo do estudo, dado o cenário abordado, foi *caracterizar elementos de organização didático pedagógica do ambiente de ensino, e eventuais sistemas de apoio, passíveis de atenção em ciclo de retroalimentação do planejamento pedagógico para que a revisão de planejamento possa proporcionar melhores ambientes de ensino aos estudantes*. Os dados analisados vincularam-se a diferentes recortes de microdados do questionário do estudante preenchido por formandos no âmbito do ENADE das edições de 2014 e 2017 de licenciaturas no Distrito Federal. Tais dados trataram da Organização Didático Pedagógica (ODP) dos cursos de graduação e versaram sobre elementos do ensino nos quais os estudantes participaram durante a graduação. Os elementos de ensino de interesse do estudo foram selecionados no questionário segundo sua aderência a arcabouço teórico abordado para fundamentação da pesquisa, de modo que tivessem escopo limitado e suficientemente informativo para o subsídio de análises do ponto de vista pedagógico. A partir da definição das variáveis de investigação, foi utilizada abordagem de análise de agrupamentos para a descoberta de perfis de percepção discente sobre o ensino no âmbito das licenciaturas, realizando-se em seguida a análise dos agrupamentos descobertos para o embasamento de *insights* para a retroalimentação do processo de planejamento pedagógico de cursos.

## **Elementos de organização didático pedagógica para dimensionamento na retroalimentação de planejamento para melhoria do ambiente de ensino**

O processo de ensino e aprendizagem possui uma diversidade de elementos passíveis de escrutínio, a depender do foco analítico e de seus objetivos. Os objetivos podem variar desde àqueles vinculados à análise de insumos que impactam sobre a comunidade escolar ou acadêmica, até àqueles relacionados com contextos mais restritos ao ambiente imediato da sala de aula com elementos vinculados à dinâmica pedagógica no ambiente de interação entre

docentes e estudantes. Nesse contexto, os elementos de interesse para o estudo realizado foram vinculados à organização didático pedagógica do ensino, ou sistemas de apoio, aderentes às funções de planejamento pedagógico e sua retroalimentação, por meio de variáveis *básicas* do ensino e que seriam vinculadas à fatores transversais aplicáveis a diversas etapas da educação formal.

O planejamento de ambientes de ensino envolve diferentes aspectos de foco, a depender da centralidade do ambiente no aprendiz, no conhecimento, na avaliação, ou na comunidade; ou ainda na articulação organizada desses aspectos, como recomendado (Bransford, Brown & Cocking, 2007). Para esta pesquisa foram considerados os aspectos aprendiz, conhecimento e avaliação. Conforme estudos organizados pelo Conselho Nacional de Pesquisa dos Estados Unidos (Bransford, Brown & Cocking, 2007), a centralidade no aprendiz envolve a organização do ambiente sobre os conhecimentos, habilidades e atitudes dos estudantes em iniciativas baseadas sobretudo no ensino por diagnóstico da situação dos discentes, para a geração de um ambiente estimulante e adequado. Os ambientes centrados no conhecimento vinculam suas iniciativas no auxílio do ambiente para que os estudantes se tornem instruídos por meio de métodos que possibilitem a compreensão do conhecimento e sua aplicação ou transferência em contextos diferentes daqueles em que ocorreu a aprendizagem. Finalmente, a centralidade na avaliação converge para os fatores do ambiente de ensino vinculados à existência de oportunidades para o *feedback* e revisão do ensino compatíveis com os objetivos de aprendizagem definidos (Bransford, Brown & Cocking, 2007).

Considerando que o ambiente de ensino deva articular coerentemente os aspectos aprendiz, conhecimento e avaliação, de acordo com as necessidades e características dos estudantes, os elementos de organização didático pedagógica enumerados abaixo se configuram como objeto de interesse.

1. **Aprendiz:** (a) a relação professor-aluno para o estímulo discente, e (b) a exigência do curso para a adequada mobilização discente;
2. **Conhecimento:** (a) os conteúdos abordados nas disciplinas;
3. **Avaliação:** (a) compatibilidade das avaliações da aprendizagem com os conteúdos.

Se o aspecto aprendiz do ambiente de ensino pressupõe iniciativas baseadas sobretudo na mobilização de conhecimentos, habilidades e atitudes prévias dos estudantes para que o ambiente de ensino possua fundação estável e de longo termo, a relação entre professor-aluno para o estímulo discente seria processo importante desse tipo de iniciativa. No mesmo contexto, a adequada estruturação do curso (exigência) para a mobilização discente seria fator central para a organização das iniciativas centradas no aprendiz dos ambientes de ensino (Bransford, Brown & Cocking, 2007). O primeiro elemento de interesse se vincularia a elemento interpessoal, configurando-se assim como objeto de características mais informais, enquanto o segundo, observada sua vinculação com a estrutura do curso, se vincularia a característica mais formal do ambiente de ensino. Entretanto, cabe ressaltar que tanto a relação professor-aluno quanto a estrutura do curso, apontadas como mais informais ou formais para os fins expositivos do estudo, vinculam-se ao planejamento do ambiente de ensino e por sua vez à uma organização sistemática e planejada de interações e estruturas pedagógicas

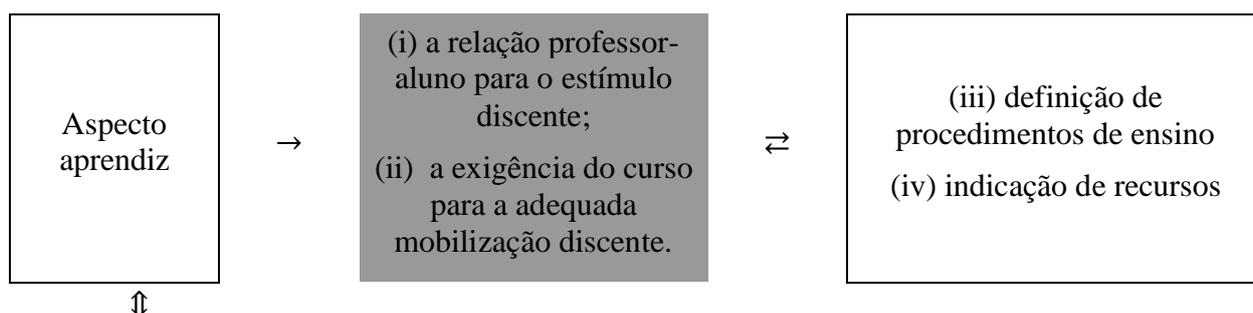
O aspecto conhecimento do ambiente de ensino, por sua vez, se associa à definição articulada e coerente de conteúdos para fins específicos de organização de um itinerário formativo. Especificamente no ambiente de ensino, conhecimentos organizados são normalmente referenciados na forma de conteúdos, cuja articulação em planejamento ocorre por meio da definição de objetivos para a aprendizagem discente a serem alcançados com o transcurso de um programa educacional ou parte dele por quem aprende, por exemplo. Os

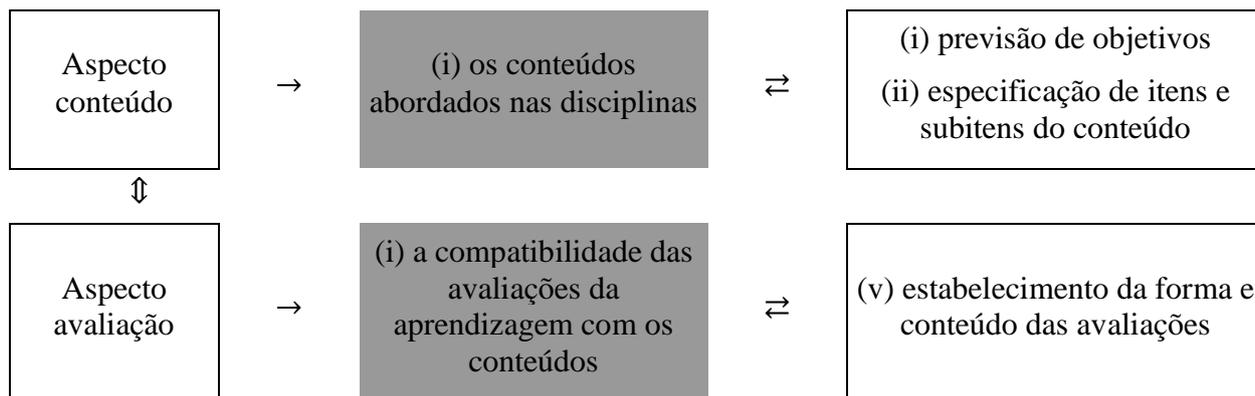
elementos constituintes do aspecto conhecimento também auxiliariam quem aprende à alcançar uma organização facilitada e compreensiva de estruturas disciplinares (Bransford, Brown & Cocking, 2007).

O terceiro aspecto é associado à avaliação que se constitui como fator de *feedback* e revisão para a melhoria da aprendizagem em comparação com práticas pedagógicas ou objetivos estabelecidos. As práticas em avaliação se subdividiriam naquelas de caráter formativo e naquelas de caráter somativo, sendo as formativas estruturadas primariamente para o fornecimento de *feedback* para a melhoria do ensino e da aprendizagem com o ajuste ou revisão de elementos de organização do ensino, como os abordados anteriormente, por exemplo. As avaliações somativas teriam foco na mensuração dos conhecimentos internalizados ou no desempenho, principalmente para fins de promoção ou ranqueamento discente (Bransford, Brown & Cocking, 2007). Assim, o corpo de conhecimentos previstos como conteúdos em um currículo deve guardar compatibilidade com as avaliações, seja na vertente formativa ou somativa abordadas.

Cada um dos elementos de interesse relacionados que se associam a ambientes de ensino, podem ser vinculados a pontos de planejamento que os sustentam em um plano organizado. Segundo Haydt (2006), o planejamento de aula conteria os seguintes pontos chave que passariam por elaboração e reelaboração no contexto do aperfeiçoamento do processo de ensino: (i) previsão de objetivos educacionais a serem alcançados pelos estudantes; (ii) especificação de itens e subitens do conteúdo que serão trabalhados em determinadas unidades curriculares ou conjunto de unidades ou aulas; (iii) definição de procedimentos de ensino com a organização de atividades de aprendizagem; (iv) indicação de recursos a serem utilizados durante atividades para despertar o interesse dos estudantes, facilitar a compreensão e estimular a participação; e (v) estabelecimento da forma e conteúdo das avaliações em aderência aos objetivos determinados, conteúdos abordados e forma das atividades na interação entre professores e estudantes. O planejamento de aula seria a sistematização de requisitos de aprendizagem, atividades e seus recursos, com avaliações, para a interação entre professores e estudantes numa determinada dinâmica de ensino para a viabilização da aprendizagem (Haydt, 2006). O esquema apresentado na Figura 1 apresenta os objetos de interesse da pesquisa – de organização didático pedagógica do ensino -, como elementos de ligação entre os aspectos do ambiente de ensino e pontos de planejamento abordados.

*Figura 1: Objetos de interesse da pesquisa como elementos de ligação entre aspectos do ambiente de ensino e elementos de planejamento*





Fonte: dados da pesquisa.

O aspecto aprendiz da organização do ambiente de ensino, embora tão importante quanto todos os demais aspectos diretamente abordados neste estudo – conhecimento e avaliação –, vincula-se a elementos chave da organização dos demais, observado que descobertas como a do efeito reverso da expertise implicam que a organização do ensino depende de variáveis como o domínio que o aprendiz possui e consolida à medida que cumpre sua trajetória no currículo ou programa educacional dentro da área (Kalyuga, 2010; Plass, Kalyuga, Leutner, 2010). Nesse sentido, o fator conhecimento prévio do estudante repercute diretamente sobre a organização do ambiente de ensino para que este seja eficaz para o estímulo da aprendizagem. Assim, a organização de procedimentos e recursos de ensino, além da própria adequação do conteúdo e conseqüentemente da avaliação da aprendizagem, dependem da adequada consideração do aprendiz e suas necessidades de aprendizagem. Esse necessário ajuste dinâmico pode ser exemplificado por meio do desenvolvimento de soluções que se utilizam de modelos cognitivos do aprendiz e a performance típica de um indivíduo com expertise no domínio de problemas, para o fornecimento de *feedback* pertinente no momento da performance do estudante em problemas dentro de ambientes simulados, por exemplo. Esse tipo de produto que pode ser assistido por computadores, busca associar conhecimento abstrato de um domínio com situações problema reais, com a organização do ensino ocorrendo a partir de um modelo situacional da cognição do aprendiz em função de interações com as situações problema e a modelagem comparativa de pessoa com expertise na resolução daquele tipo de problema. Essa integração procedimento/recurso pode prover ao estudante auxílios localizados e ajuda personalizada quando em interação com o recurso, além de material para interação organizada com um docente para o adequado estímulo do estudante para que estude e aprenda de forma mais efetiva, por exemplo, subsidiando procedimentos de ensino por parte do professor (Lajoie, 2009).

Nesse cenário, em todos os elementos do ambiente de ensino abordados, e os respectivos pontos de planejamento do ensino que se vinculam a eles, ocorre a centralidade de concepção ao redor da aprendizagem do estudante, seja em uma via mais ou menos diretiva, ou uma estrutura mais ou menos centrada na atividade discente, por exemplo. Ou seja, observado que o ensino trata da prática deliberada, sistemática e organizada para que se facilite a aprendizagem de algo por outrem, pode-se dizer que exista um espectro de diretividade e estruturação das atividades e recursos a serem utilizados em determinados tipos de interação entre docentes e discentes para o alcance de objetivos de aprendizagem pelos estudantes. Assim, o ensino como prática sistemática e organizada abrange em seu espectro desde (i) a vertente expositiva de conteúdo pelo docente que em seguida recebe dúvidas e complementa a interação com os estudantes por meio de explicações e exemplificações, até (ii) aquelas situações de aprendizagem planejadas para um trabalho discente mais autônomo, como uma sala de aula invertida por exemplo. Em

qualquer situação planejada, o ensino serve à aprendizagem, sendo que tal argumento é sustentado pela colocação de Simon (2000) de que tanto o ensino quanto o professor possuem papel acessório à aprendizagem do estudante ou de Hiebert, Morris, Berk e Jansen (2007) e Tyler (1949) quanto à importância de objetivos claros para a orientação do exercício do ensino por professores. Nesse sentido, Bach, Water, Frega e Müller (2014) e Beck e Rausch (2014) apontam uma tendência de que fatores de desempenho docente no exercício do ensino afetam a percepção de estudantes em suas oportunidades de aprendizagem. De forma geral, qualquer que seja a forma de estruturação de oportunidades de aprendizagem aos estudantes via ensino, evidências apontam para a tarefa docente fundamental, que é a de realizar uma espécie de transformação do conhecimento, como aponta Bruer (1993). Nesse ponto da transformação do conhecimento que o docente realiza para proporcionar as melhores condições para a aprendizagem dos estudantes, são ressaltadas características dominantes no exercício do ensino como, por exemplo, a clareza por meio de explicações e exemplificações, como aponta Rodrigues Júnior (2002).

Da mesma forma, serviços organizados para o apoio ao discente, como o apoio pedagógico ou psicológico, além da atuação imediata do docente em sala de aula, constituem-se como elementos essenciais no planejamento da oferta sistemática e estruturada de um ambiente de ensino adequado para a viabilização ou facilitação de aprendizagens. Concomitantemente ao acompanhamento de elementos relacionados às funções de planejamento, o dimensionamento do apoio recebido pelo estudante, quando necessário, é importante para o planejamento pedagógico mais amplo.

Assim, para uma adequada retroalimentação de planejamento para melhoria do ambiente de ensino deve existir por parte da instituição educacional o acompanhamento de um conjunto de elementos de ODP do ensino. No caso da pesquisa realizada, os objetos de interesse se vincularam a fatores *básicos* dos ambientes de ensino e possíveis sistemas de apoio, em um *framework* analítico que unifica técnica de análise de dados com arcabouço de interpretação pedagógica.

## Trabalhos relacionados

A utilização de análise de agrupamentos nas áreas de *Educational Data Mining* (EDM) e *Learning Analytics* (LA) é uma tendência. Em revisão realizada por Moissa, Gasparini e Karczinski (2015), a atividade de clusterização respondia por 23% e 10% das técnicas identificadas para tratamento de problemas em EDM e LA, respectivamente. Na investigação da percepção discente relacionada a cursos de graduação ou Instituições de Educação Superior, a análise de agrupamentos é técnica analítica de aplicação corrente, observada a necessidade de descoberta de padrões não óbvios nos dados que agregam a percepção discente. Como revisado por Faria (2014), a aplicação da clusterização em tarefas de EDM para melhoria do ensino *online* se circunscreveria ao agrupamento de materiais semelhantes de cursos ou no agrupamento de estudantes com base em seu aproveitamento e padrões de interação, assim como no agrupamento de estudantes para encaminhamento de ações diferenciadas de acordo com outras características do conjunto dos estudantes. Filatro (2021) aponta a utilidade da análise de agrupamentos para se analisar pessoas com interesses parecidos ou estudantes com estilos de aprendizagem similar, observado o agrupamento de observações mais semelhantes entre si em um conjunto de dados a partir de determinadas variáveis ou atributos de interesse das unidades de observação.

A técnica de análise de agrupamentos objetiva agrupar unidades de observação em função de determinadas características semelhantes que elas possuam. A identificação de tais

grupos *a priori* é potencialmente inviável, observada a quantidade de dados e características envolvidas, e o padrão não óbvio de relacionamento existente que é descoberto via mensuração de semelhança e a aplicação de algoritmo de agrupamento. Ao final de um procedimento de agrupamento é obtida uma estrutura de dados agrupados de modo que tenham sido maximizadas: (i) a homogeneidade entre os elementos dentro dos grupos, e; (ii) a heterogeneidade entre os grupos formados. A estrutura de grupos descoberta com a atividade de agrupamento fornece insumos para a compreensão do fenômeno em estudo por meio da análise das características dos grupos formados (Silva, Peres & Boscaroli, 2016).

No cenário de uso da análise de agrupamentos para a análise da percepção discente no subsídio de ações no âmbito da gestão de IES ou cursos de graduação, Zanella e Lopes (2009) avaliaram a satisfação de estudantes do Departamento de Estatística da Universidade Federal de Santa Maria, quanto à atuação do professor e à infraestrutura oferecida em disciplina de Estatística, com o uso da análise de agrupamento para reunião de cursos de graduação que continham a disciplina em grupos, com base na semelhança apresentada nas respostas às variáveis de interesse do estudo. Sordan, Marinho e Angeloni (2021) aplicaram questionário de modelo para avaliação de qualidade de serviços para análise da satisfação de estudantes, com o uso de análise de agrupamentos para complementação de análise inicial por meio da clusterização dos itens do questionário utilizado, resultando na descoberta de questão não priorizada preliminarmente. Cardoso (2021), por sua vez, utiliza os dados do questionário do estudante, no âmbito do ENADE, para propor sistema de autoavaliação de curso de graduação, com utilização de análise de agrupamentos com a finalidade de identificar possíveis agrupamentos de questões. Santos e Pedroso (2021) analisaram a percepção de egressos e evadidos sobre cursos de graduação com o uso de análise de agrupamentos associada a testes de hipóteses para a verificação de existência de diferença entre as percepções de egressos e evadidos que tenham influenciado a evasão. Os estudos em tela demonstram a aplicação recorrente da análise de agrupamentos para o tratamento de dados sobre a percepção de estudantes no âmbito da gestão da educação superior, sendo que tais dados possuem potencial de uso no ciclo de planejamento pedagógico, seja para mitigar a evasão ou para a melhoria do ensino ou serviços de apoio associados.

Os dados do ENADE constituem-se como fonte importante de insumos para que as IES possam compreender as necessidades e percepções dos estudantes. Como mencionado, Cardoso (2021) fez uso de tais dados para a proposição de sistema de autoavaliação de curso de graduação. Tives, Canedo, Ladeira e Fagundes (2018), apontam a importância dos dados do ENADE, em especial do questionário do estudante, para a orientação da gestão de IES com a aplicação de EDM. Entretanto, embora tais dados integrem fluxos consolidados de coleta, tratamento e disponibilização em um sistema nacional de avaliação da educação superior, seu uso em análises envolvendo EDM, por exemplo, parecem ocorrer em quantidades aquém do esperado. Nesse sentido, Lima, Ambrósio, Ferreira e Brancher (2019) apontam o baixo uso dos dados oficiais de avaliação da educação superior para análises no âmbito de EDM, sendo que a maioria das análises nas produções revisadas se concentram em análises estatísticas descritivas ou de outros tipos. A Tabela 1 apresenta os trabalhos relacionados e suas temáticas associadas ao estudo.

Tabela 1: *Trabalhos relacionados*

Temática	Referências
----------	-------------

Diagnóstico do ensino-aprendizagem e satisfação de alunos em curso de graduação	Zanella e Lopes (2009)
Análise de dados advindos de exame nacional de desempenho discente	Tives, Canedo, Ladeira e Fagundes (2018)
Percepção discente sobre cursos de graduação	Santos e Pedroso (2021)
Análise de dados de exames nacionais de desempenho e aprendizagem discente	Lima, Ambrósio, Ferreira e Brancher (2019)
Avaliação de Cursos de Graduação na perspectiva de discentes	Cardoso (2021)
Avaliação da qualidade percebida em uma Instituição de Ensino Superior	Sordan, Marinho e Angeloni (2021)

Fonte: dados da pesquisa

Em perspectiva, foi identificada lacuna na literatura recente em relação a vinculação dos achados de análises de agrupamentos com dados de percepção de estudantes com arcabouço conceitual de planejamento e interpretação pedagógica, concomitante ao baixo uso de dados do ENADE - em especial o questionário do estudante -, em estudos de EDM. Nesse sentido, este estudo busca vincular descobertas dos perfis de agrupamentos de percepção discente sobre o ensino com interpretação pedagógica dos dados obtidos. Tal vinculação poderia contribuir para a investigação de um *framework* analítico que unifica técnica de análise de dados (análise de agrupamentos) com arcabouço de interpretação pedagógica para melhoria do ensino por meio do planejamento pedagógico e sua retroalimentação. O desenvolvimento de soluções de dados para a melhoria do ensino pode ser beneficiado com essa perspectiva integrada.

## Metodologia

A pesquisa possuiu natureza exploratória, fazendo uso de dados secundários em um ciclo de análise para retroalimentação do planejamento pedagógico, com a aplicação de técnica de análise de agrupamentos sobre dados do questionário do estudante submetido ao ENADE. Assim, para o alcance do objetivo deste estudo - *caracterizar elementos de organização didático pedagógica do ambiente de ensino, e eventuais sistemas de apoio, passíveis de atenção em ciclo de retroalimentação do planejamento pedagógico para que a revisão de planejamento possa proporcionar melhores ambientes aos estudantes-*, foram investigados perfis de percepção do ensino a partir de microdados do questionário do estudante, vinculados ao exame mencionado, nos anos de 2014 e 2017. Os dados analisados foram restritos aos anos de 2014 e 2017 por conta de serem os anos mais recentes com dados disponibilizados quando da realização do estudo e sua relação com edições do exame que tiveram foco em estudantes de cursos de licenciatura.

Foram realizadas duas implementações de técnica específica para descoberta dos agrupamentos em conjuntos de dados, uma para cada recorte de dados das edições de interesse. Os conjuntos de dados utilizados em cada implementação foram os seguintes:

1. Com formandos em todas as licenciaturas respondentes do questionário do estudante da edição 2014 do ENADE, do Distrito Federal;
2. Com formandos em todas as licenciaturas respondentes do questionário do estudante da edição 2017 do ENADE, do Distrito Federal.

Os dados foram restritos ao Distrito Federal por sua peculiaridade em relação aos demais estados do Brasil, como maior renda per capita do país (Brasil, 2018) e elevado Índice de Desenvolvimento Humano (IDH) (Atlas do Desenvolvimento Humano no Brasil, 2018), concomitantemente a elevado índice de Gini (Atlas do Desenvolvimento Humano no Brasil 2018).

Os conjuntos de dados utilizados tiveram origem no instrumento "questionário do estudante", aplicado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) como um dos componentes do ENADE. Dentre os itens do questionário, consta um conjunto de itens a serem respondidos de acordo com a concordância do estudante com afirmativas sobre a Organização Didático Pedagógica (ODP) de seu curso de graduação, dentro de seis pontos de uma escala tipo Likert (de 1 [discordo totalmente] a 6 [concordo totalmente]). Existiram ainda pontos de escolha para opções "não sei responder" e "não se aplica", opções 7 e 8, respectivamente, segundo o dicionário de variáveis dos microdados (Brasil, 2018a, 2018b).

No escopo do estudo foram selecionadas variáveis (itens do questionário) para caracterização dos respondentes (estudantes) em grupos de modo a se obter um perfil de agrupamento desses dados com base em semelhança entre as respostas para as questões selecionadas. Foram selecionadas questões caracterizadas como ordinais e nominais, considerando a aderência aos elementos de pesquisa no estudo.

### **Seleção de variáveis**

A seleção das variáveis utilizadas para o processo de agrupamento foi teoricamente orientada a partir dos elementos de interesse da pesquisa. Assim, a seleção pautou-se primariamente pelo conteúdo de cada assertiva da seção de Organização Didático-Pedagógica do questionário de percepção do estudante e seu melhor ajuste semântico. Demais variáveis do questionário possuíam conteúdo semântico diverso dos elementos de interesse, seja por referenciar pontos que ampliavam o escopo do julgamento de concordância ou informação a ser realizado pelo estudante, seja por não estarem diretamente relacionadas aos pontos chave objeto do estudo. No mesmo sentido, concomitante ao critério semântico, o melhor ajuste do conjunto das variáveis selecionadas ante a medida de validação para a descoberta da estrutura de agrupamentos foi critério utilizado para a seleção. Quatro variáveis selecionadas possuíam natureza ordinal, ocorrendo ainda a seleção de uma variável de natureza nominal, totalizando a seleção de cinco variáveis para o agrupamento de estudantes em função da maior ou menor semelhança de percepção sobre essas variáveis.

As variáveis de natureza ordinal selecionadas, assim como suas associações com o itens de planejamento, foram as seguintes:

- QE\_I28 Os conteúdos abordados nas disciplinas do curso favoreceram sua atuação em estágios ou em atividades de iniciação profissional;
  - Considerando que o favorecimento dos conteúdos para a atuação em estágios ou outras atividades de inserção profissional do estudante ocorre apenas se ajustados aos objetivos de aprendizagem estabelecidos, conjecturou-se que uma relação conteúdos-objetivos poderia ser dimensionada com essa variável para a obtenção de

informações sobre os seguintes pontos de planejamento: (i) Objetivos educacionais a serem alcançados pelos estudantes; e (ii) Especificação dos conteúdos que serão trabalhados;

- QE\_I37 As relações professor-aluno ao longo do curso estimularam você a estudar e aprender. | QE\_I42 O curso exigiu de você organização e dedicação frequente aos estudos.
  - Considerando que as relações professor-estudante estimulam este último apenas na medida que as atividades de aprendizagem são adequadamente organizadas, resultando em curso com exigência adequada, essas variáveis poderiam se relacionar com: (iii) definição de procedimentos de ensino com a organização de atividades de aprendizagem; e (iv) indicação de recursos a serem utilizados durante atividades para despertar o interesse dos estudantes, facilitar a compreensão e estimular a participação;
- QE\_I55 As avaliações da aprendizagem realizadas durante o curso foram compatíveis com os conteúdos ou temas trabalhados pelos professores.
  - Considerando a necessária compatibilidade entre as avaliações e os conteúdos do curso, foi estimado que se poderia obter informações pertinentes ao item de planejamento: (v) estabelecimento da forma e conteúdo das avaliações em aderência aos objetivos determinados, conteúdos abordados e forma das atividades na interação entre professores e estudantes.

Dentro do conjunto de variáveis selecionadas para o estudo, optou-se também pelo uso da variável QE\_I20 que versava sobre o grupo que forneceu apoio decisivo para a superação de dificuldades durante o curso, no caso da existência desse tipo de dificuldade. O uso dessa variável decorreu da conjectura de que serviços planejados para o apoio ao discente, como o apoio psicopedagógico ou psicológico, constituem-se como elementos essenciais em um planejamento integrado para a oferta sistemática e estruturada de um ambiente de ensino adequado para a viabilização ou facilitação de aprendizagens. No caso dessa última variável nominal, as respostas possíveis para a questão (11 opções iniciais) foram recodificadas para 4 opções, tendo-se em vista o melhor ajuste do conjunto das variáveis ante a medida de validação para a descoberta da estrutura de agrupamentos. As recodificações se deram da seguinte forma (original/nova): (A/A), (B/B), (C/C), (D/C), (E/C), (F/C), (G/C), (H/D), (I/D), (J/C), (K/C), sendo que as novas opções (A, B, C e D) foram: A - Não tive dificuldade; B - Não recebi apoio para enfrentar dificuldades; C - Apoio de fora da IES e; D - Apoio da IES.

As medidas associadas a cada variável selecionada por conjuntos de dados a serem segmentados foram as seguintes – Tabela 2:

Tabela 2: *Medidas associadas às variáveis selecionadas por recortes utilizados de conjuntos de dados*

Conjunto de dados	Questões	Freq.	Mínimo	1º quartil	Mediana	Média	3º quartil	Máximo
Formandos em licenciaturas 2014 (DF).	QE_I28	-	1	5	6	5,32	6	6
	QE_I37	-	1	5	6	5,18	6	6
	QE_I42	-	1	5	6	5,41	6	6
	QE_I55	-	1	5	6	5,31	6	6
	QE_I20	A: 1.140; B: 251;	-	-	-	-	-	-

		C: 2.120; D: 195.						
Formandos em licenciaturas 2017 (DF).	QE_I28	-	1	5	6	5,24	6	6
	QE_I37	-	1	5	6	5,15	6	6
	QE_I42	-	1	5	6	5,44	6	6
	QE_I55	-	1	5	6	5,27	6	6
	QE_I20	A: 938; B: 215; C: 1.836; D: 234.	-	-	-	-	-	-

Fonte: dados da pesquisa

### Preparação dos dados

Quanto aos dados ausentes adotou-se a abordagem de caso completo (Hair e colegas, 2009). Foram retirados da base de dados os casos em que associadas às variáveis selecionadas QE\_I28, QE\_I37, QE\_I42, QE\_I55 constavam as informações 7 ou 8 ("não ter elementos para avaliar a assertiva" ou julgá-la não pertinente ao seu curso), pois conjecturou-se que tais seleções podem ter decorrido do desejo de manter tais elementos não respondidos por outros motivos além dos explicitados por tais opções. Professores recém formados dificilmente não teriam elementos para avaliar as assertivas das variáveis selecionadas que tratam todas do exercício da docência, e nenhuma das variáveis selecionadas podem ser julgadas não pertinentes a cursos de licenciatura, observado serem elas fundamentais a qualquer processo de ensino.

### Operacionalização da análise

A seguir serão abordados os procedimentos aplicados sobre cada um dos conjuntos de dados nas implementações propostas. A exposição dos procedimentos estão organizadas da seguinte forma: (i) abordagem da medida de distância utilizada e por meio da qual é obtida uma métrica de comparação entre casos; (ii) exposição do procedimento de agrupamento que tem as medidas de distância como insumo e por meio do qual foram agrupados os casos em uma quantidade  $k$  de agrupamentos; (iii) apresentação de uma medida de validação interna de quantidades de agrupamentos a partir de diferentes execuções do procedimento de agrupamento exposto, variando-se a quantidade de agrupamentos em cada execução para a descoberta da quantidade  $k$  ótima de agrupamentos para cada conjunto de dados.

### Medida de distância utilizada

A noção de distância entre objetos ou observações objetiva medir a semelhança entre os comparados a partir de um conjunto finito de características mensuráveis. Quanto menor a distância entre os casos, mais similares eles são, e vice versa. Distâncias matemáticas dependem da escala de mensuração das variáveis utilizadas para se comparar as unidades de observação. Como aponta Frei (2006), bases de dados com diferentes escalas de mensuração (mistas) devem receber o adequado tratamento para o cálculo das distâncias entre casos para a análise de agrupamentos. O autor aponta a adequação da distância de Gower (1971) quando o conjunto de dados a ser analisado possui natureza mista.

A implementação da distância de Gower (1971) utilizada neste trabalho ocorreu por meio do pacote *cluster* (Maechler, Rousseeuw, Struyf, Hubert & Hornik, 2017) para a linguagem R (R Core Team, 2016). A obtenção da matriz de distâncias foi realizada com a implementação mencionada conforme eq 1, entre cada par de casos  $i$  e  $j$  (Frei, 2006).

$$d(i,j) = \frac{\sum_{k=1}^n W_{ij} \cdot S_{ij}}{\sum_{k=1}^n W_{ij}} (1)$$

A distância entre os casos  $i$  e  $j$  é obtida para esses dois casos para cada variável  $k$ . O peso ( $W$ ) para os casos  $i$  e  $j$  na variável  $k$  ( $W_{ij}$ ) é igual a 1 quando as mensurações em  $i$  e  $j$  da variável  $k$  são conhecidas e igual a 0 na situação de uma ou mais medidas da variável  $k$  serem ausentes. O valor de  $S_{ij}$  representa a contribuição da variável  $k$  para a distância entre  $i$  e  $j$ . O valor de  $S_{ij}$  é alcançado de diferentes formas, a depender do nível de mensuração da variável  $k$ . No caso de variáveis em mensuração nominal, ocorre a aplicação do coeficiente *simple matching*. No caso de variáveis intervalares o valor absoluto da diferença entre os valores na variável  $k$  é dividido pelo intervalo da variável. No caso de variáveis ordinais, as entradas são substituídas por seus postos e então é aplicado o procedimento para variáveis intervalares. Para variáveis com escala de mensuração de razão utiliza-se o mesmo processo utilizado para aquelas de natureza intervalar (Frei, 2006).

### Procedimento para agrupamento

O método de clusterização foi o implementado no algoritmo PAM (*Partitioning Around Medoids*) (Rousseeuw & Kaufman, 1990). No algoritmo utilizado busca-se por  $p$  casos representativos, também chamados de medoides. Esses casos representativos (unidade de observação existente nos dados) representam cada agrupamento formado, devendo assim resumir a estrutura subjacente ao conjunto de dados que foi particionado. Esses medoides são aqueles casos nos quais a distância entre ele e todos os demais casos mais próximos que não são medoides é mínima. Nesse contexto, os casos representativos são aqueles mais centralmente localizados em cada agrupamento (Kassambara, 2018; Maechler, Rousseeuw, Struyf, Hubert & Hornik, 2017).

Com a localização inicial dos medoides nas  $k$  quantidades de agrupamentos a serem formados, cada exemplar que não é um medoide é designado ao seu caso representativo mais próximo (fase de construção). Em seguida, cada medoide selecionado ‘ $m$ ’ e cada exemplar não medoide ‘ $o$ ’ são intercambiados com a busca da otimização da função objetivo que é calculada a cada fase de iteração nas sucessivas trocas (fase de intercâmbio). A fase de intercâmbio objetiva incrementar a qualidade dos agrupamentos com a troca de objetos selecionados como medoides e exemplares não selecionados (Kassambara, 2018). Tal algoritmo é mais robusto do que o *k-means* quanto a *outliers* e ruídos nos dados (Kassambara, 2018; Maechler, Rousseeuw, Struyf, Hubert & Hornik, 2017).

### Medida para determinação do número de agrupamentos

Observado que o algoritmo PAM depende do fornecimento do número  $k$  de agrupamentos nos quais o conjunto de dados será segmentado, uma medida útil para o auxílio nessa determinação de  $k$  é a média da medida de Silhueta (*Silhouette*). Essa medida é um índice baseado em critério interno para validação da quantidade de agrupamentos a ser determinada para o algoritmo de partição (Kassambara, 2018).

Para a determinação de um número ótimo de agrupamentos a ser utilizado, foi analisada a Silhueta média dos agrupamentos gerados a partir de diferentes valores experimentados de  $k$

(quantidade de agrupamentos) - dois a vinte agrupamentos por implementação. As quantidades  $k$  com maiores valores da média da medida de Silhueta e superiores a 0,51 foram escolhidos para a atividade de clusterização, conforme a razoabilidade da estrutura de agrupamentos identificada, como apontam Rousseeuw e Kaufman (1990) – Quadro 1:

Quadro 1: Interpretação para os possíveis valores da medida de silhueta média

Medida da Silhueta	Interpretação
0,71 – 1,00	A estrutura localizada é forte
0,51 – 0,70	A estrutura localizada é razoável
0,26 – 0,50	A estrutura localizada é fraca e pode ser artificial. Outros métodos devem ser aplicados sobre o conjunto de dados
Menor ou igual a 0,26	Nenhuma estrutura substancial foi localizada

Fonte: Rousseeuw & Kaufman (1990) (com adaptações).

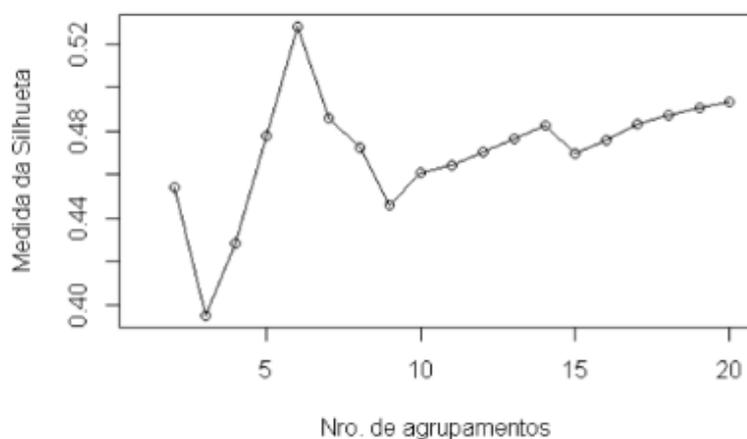
## Resultados e discussão

Os resultados se concentraram na análise de determinação do número mais ajustado de agrupamentos por implementação, com conjecturas sobre a relação entre estruturas identificadas e impactos de cenários sobre o planejamento. Foram analisadas as configurações dos casos representativos de cada agrupamento identificado por implementação e, a partir da configuração desses casos e o comportamento de cada variável de interesse, puderam ser identificados elementos de percepção dos estudantes com repercussão sobre itens de planejamento para retroalimentação da organização de ambientes de ensino.

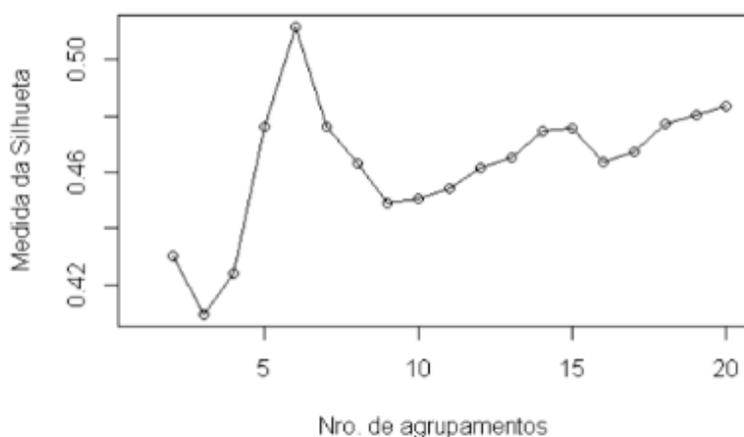
### Determinação da quantidade de agrupamentos

Para a execução do algoritmo com o conjunto de dados para a implementação 1 proposta (com formandos de 2014 de todas as licenciaturas), foram extraídos 6 agrupamentos, observado o maior valor da medida da silhueta (0,5277496), conforme pode ser observado no gráfico 1-a.

Com a análise do gráfico 1-a, embora a medida de silhueta pareça tender a um aumento com o incremento da quantidade de agrupamentos em que se particiona o conjunto de dados, eventual medida superior à alcançada com a quantidade  $k$  definida em 6 representaria nível de granularidade excessivo. Esse alto grau de granularidade (com  $k > 20$ ) representaria baixo impacto para os fins de planejamento, observada a multiplicação de agrupamentos muito pequenos. Em tais situações, a descoberta de agrupamentos poderia não ser útil na alocação de recursos escassos.



(a)



(b)

Gráfico 1: Silhueta média por quantidade de agrupamentos (2-20) para conjunto de dados de 2014 (a) e 2017 (b).

Fonte: dados da pesquisa.

No caso da implementação 2 (com formandos de 2017 provenientes de todos os cursos de licenciatura), a quantidade de agrupamentos com melhor ajuste foi o de 6, observado o maior valor da medida da silhueta (0,5115822), conforme pode ser observado no gráfico 1-b. Como no primeiro recorte de dados, a tendência de aumento da medida de validação com o incremento de agrupamentos também foi constatada, aplicando-se nesse caso as mesmas considerações realizadas para a implementação 1 (gráfico 1-a). A tabela 3 resume a quantidade de agrupamentos descobertos por conjunto de dados particionados e medidas de validação obtidas.

Tabela 3: Quantidade de clusters para cada conjunto de dados

Implementação	Conjunto de dados	Quantidade de agrupamentos	Média de silhueta
1	Formandos em licenciaturas 2014.	6	0,53

2

Formandos em licenciaturas 2017

6

0,51

Fonte: dados da pesquisa

As medidas de silhueta média próximas do limite para a classificação das estruturas de agrupamentos como razoáveis (Tabela 3), parece indicar que poderiam existir melhores aproximações ao problema proposto. Tais aproximações alternativas poderiam se dar por meio: (i) da possível existência de melhor conjunto de questões a serem selecionadas no questionário do estudante, embora com menor aderência das variáveis ao arcabouço conceitual utilizado no estudo; (ii) da elaboração de questões específicas em um instrumento próprio para uso na coleta dos dados de percepção dos estudantes, a partir da sustentação teórica utilizada no estudo ou; (iii) da consideração de aspectos distintos de planejamento em relação aos abordados no estudo, subsidiando a seleção de variáveis distintas.

Potencialmente, algum dos encaminhamentos apontados poderia gerar melhores características de diferenciação entre agrupamentos em cada conjunto de dados, com o respectivo incremento da medida de validação. Entretanto, para o estudo proposto, a razoabilidade das medidas, e as quantidades ótimas descobertas indicam a relevância da análise, considerando as limitações envolvidas no cenário real de planejamento que tem sua efetivação favorecida por quantidades de cenários suficientemente limitados e o pressuposto de métricas no mínimo razoáveis para a análise dos dados.

### Características dos casos representativos entre agrupamentos de cada implementação

Considerando a quantidade de agrupamentos descobertos nas duas implementações realizadas, cada um dos agrupamentos teve uma observação concreta que foi o seu centro (caso representativo), de acordo com o algoritmo de agrupamento. Entretanto, algumas métricas associadas aos agrupamentos foram analisadas, para posterior discussão dos casos representativos de cada agrupamento.

Conforme relacionado na Tabela 4, na implementação com dados de 2014, ocorreu a predominância dos valores mínimos de todas as variáveis no ponto 1 da escala de concordância, para todos os agrupamentos. Os valores máximos se concentraram no topo da escala (6), sendo que o intervalo interquartil das variáveis nos agrupamentos se concentraram em zero (A1 e A5), um (A2, A3 e A6) ou dois pontos (A4). As medianas das variáveis dentro dos agrupamentos tiveram a preponderância de um dos pontos da escala (A1, A4, A5 e A6), existindo agrupamentos destoantes dessa tendência (A2 e A3). Os itens para a variável nominal (I20) tenderam a se distribuir entre os agrupamentos.

Tabela 4: *Medidas associadas às variáveis dentro dos agrupamentos com dados de formandos em licenciaturas de 2014*

Agrupamento	Variáveis	Freq.	Mín.	1º quartil	Mediana	Média	3º quartil	Máx.
A1	QE_I28	-	1	6	6	5,77	6	6
	QE_I37	-	1	6	6	5,75	6	6
	QE_I42	-	1	6	6	5,87	6	6
	QE_I55	-	1	6	6	5,80	6	6
	QE_I20	A: 0;	-	-	-	-	-	-

			B: 0;					
			C: 0;					
			1.293;					
			D: 0.					
A2	QE_I28	-	1	4	5	4,61	5	6
	QE_I37	-	1	4	4	4,18	5	6
	QE_I42	-	1	4	5	4,77	6	6
	QE_I55	-	1	4	5	4,5	5	6
	QE_I20	A: 0;	-	-	-	-	-	-
		B: 0;						
		C: 827;						
		D: 2.						
A3	QE_I28	-	1	5	6	5,37	6	6
	QE_I37	-	1	4	5	5,06	6	6
	QE_I42	-	1	5	6	5,39	6	6
	QE_I55	-	1	5	6	5,35	6	6
	QE_I20	A: 0;	-	-	-	-	-	-
		B: 214;						
		C: 0;						
		D: 0.						
A4	QE_I28	-	1	3	4	3,90	5	6
	QE_I37	-	1	3	4	3,74	4	6
	QE_I42	-	1	3	4	3,94	5	6
	QE_I55	-	1	3	4	4,00	5	6
	QE_I20	A: 229;	-	-	-	-	-	-
		B: 37;						
		C: 0;						
		D: 21.						
A5	QE_I28	-	2	6	6	5,75	6	6
	QE_I37	-	1	6	6	5,71	6	6
	QE_I42	-	1	6	6	5,77	6	6
	QE_I55	-	1	6	6	5,74	6	6
	QE_I20	A: 911;	-	-	-	-	-	-
		B: 0;						
		C: 0;						
		D: 0.						
A6	QE_I28	-	1	5	6	5,50	6	6
	QE_I37	-	1	5	6	5,45	6	6
	QE_I42	-	1	6	6	5,70	6	6
	QE_I55	-	1	5	6	5,45	6	6
	QE_I20	A: 0;	-	-	-	-	-	-
		B: 0;						
		C: 0;						
		D: 172.						

Fonte: dados da pesquisa

Já na implementação com dados dos formandos de 2017 (Tabela 5), como nos agrupamentos gerados a partir dos dados de 2014, ocorreu a preponderância dos valores mínimos no ponto 1 da escala. Em relação aos valores máximos, ocorreu a mesma concentração no topo da escala (6) que foi observada com os dados de 2014. O intervalo interquartil das variáveis nos agrupamentos se concentraram em zero (B4 e B5) e um (B1 e B6) pontos, como no agrupamento com dados de 2014. Entretanto, ocorreram também agrupamentos que possuíam a metade de suas variáveis ordinais com intervalo de um ponto e o restante com dois (B2 e B3). Da mesma forma que nos agrupamentos com dados de 2014, as medianas nos agrupamentos tiveram a preponderância de um dos pontos (B3, B4, B5 e B6), existindo

agrupamentos com medianas destoantes (B1 e B2). Como nos agrupamentos de 2014, os itens para a variável nominal (I20) tenderam a se distribuir entre os agrupamentos.

Tabela 5: *Medidas associadas às variáveis dentro dos agrupamentos com dados de formandos em licenciaturas de 2017*

Agrupamento	Variáveis	Freq.	Mín.	1º quartil	Mediana	Média	3º quartil	Máx.
B1	QE_I28	-	1	4	5	4,37	5	6
	QE_I37	-	1	3	4	3,99	5	6
	QE_I42	-	1	4	5	4,75	5	6
	QE_I55	-	1	4	5	4,41	5	6
	QE_I20	A: 0; B: 0; C: 711; D: 12.	-	-	-	-	-	-
B2	QE_I28	-	1	4	5	4,80	6	6
	QE_I37	-	1	4	5	4,72	6	6
	QE_I42	-	1	5	6	5,27	6	6
	QE_I55	-	1	5	5	5,03	6	6
	QE_I20	A: 0; B: 215; C: 0; D: 0.	-	-	-	-	-	-
B3	QE_I28	-	1	4	5	4,44	6	6
	QE_I37	-	1	4	5	4,35	6	6
	QE_I42	-	1	5	5	4,53	6	6
	QE_I55	-	1	5	5	4,43	6	6
	QE_I20	A: 304; B: 0; C: 0; D: 0.	-	-	-	-	-	-
B4	QE_I28	-	1	6	6	5,76	6	6
	QE_I37	-	2	6	6	5,79	6	6
	QE_I42	-	1	6	6	5,87	6	6
	QE_I55	-	1	6	6	5,75	6	6
	QE_I20	A: 0; B: 0; C: 1.125; D: .	-	-	-	-	-	-
B5	QE_I28	-	1	6	6	5,75	6	6
	QE_I37	-	3	6	6	5,80	6	6
	QE_I42	-	3	6	6	5,89	6	6
	QE_I55	-	1	6	6	5,83	6	6
	QE_I20	A: 634; B: 0; C: 0; D: 0.	-	-	-	-	-	-
B6	QE_I28	-	1	5	6	5,47	6	6
	QE_I37	-	1	5	6	5,33	6	6
	QE_I42	-	1	6	6	5,65	6	6
	QE_I55	-	1	5	6	5,36	6	6
	QE_I20	A: 0; B: 0; C: 0; D: 222.	-	-	-	-	-	-

Fonte: dados da pesquisa

No Quadro 2 são apresentadas as configurações dos casos representativos de cada agrupamento das duas implementações da análise de agrupamentos. Tais configurações tratam do vetor de respostas desses casos representativos para as questões de interesse selecionadas para o estudo, de modo que se pudesse analisar os elementos de ensino essenciais para a retroalimentação do planejamento a partir da percepção dos estudantes. Dentro de cada variável de interesse ordinal (I28, I37, I42, I55), as atribuições dos casos representativos ocorreu dentro dos seis pontos da escala tipo Likert (de 1 [discordo totalmente] a 6 [concordo totalmente]). Os valores foram do mínimo 4 ao máximo 6.

Na implementação com formandos de 2014 os valores mínimo e máximo foram respectivamente, 4 e 6 para todas as variáveis. Já na implementação com formandos de 2017 os valores mínimo e máximo foram 5 e 6, para todas as variáveis, com exceção para a variável I37 que teve tais valores mínimo e máximo em 4 e 6, respectivamente. O vetor de caso representativo com menores valores para os dados de 2014 foi o do agrupamento A4, com todas as variáveis tendo recebido atribuição do ponto 4 da escala de concordância. O vetor de dados com menores valores para os dados de 2017 foi o do agrupamento B3, conforme Quadro 2. Na variável nominal I20, os valores existentes no vetor resposta de cada caso representativo se distribuíram em todo o espectro de possibilidade, em ambas as implementações. Associados aos agrupamentos (A4 e B3) com vetor de resposta com menores valores de concordância, constava associado o item de não recebimento de apoio em caso de dificuldade (item “B – Não recebi apoio” na variável I20).

Nas duas clusterizações foram descobertos agrupamentos (A1, A5 e A6; B4, B5 e B6) que tiveram como casos representativos aqueles que atribuíram pontuação máxima da escala para todas as variáveis de interesse para a caracterização do ensino. Para esses casos, os dados indicam uma percepção favorável generalizada, sendo a variação entre eles referente à variável de apoio para a superação de dificuldades. Ressalta-se que em todos esses casos, nenhum foi caracterizado pelo não recebimento de apoio em caso de dificuldade (item “B – Não recebi apoio” na variável I20). Ou seja, para todos os casos representativos de agrupamentos nos dois conjuntos de dados, as melhores percepções dos elementos de ensino se associaram à inexistência de dificuldade durante o curso, e no caso de dificuldade o estudante recebeu apoio adequado, seja no âmbito da própria IES ou externamente à ela, não ficando assim desassistido.

Quadro 2: Casos representativos de cada agrupamento obtido nas implementações (2014 e 2017) - com elementos de ensino em destaque.

Implementação	Agrupamento	Configuração do caso representativo de cada agrupamento				
		Conteúdos do curso (I28)	Interação docente e discente (I37)	Exigência do curso (I42)	Articulação avaliações/ conteúdos (I55)	Apoio em caso de dificuldade (I20)
Formandos em licenciaturas - 2014	A1	6	6	6	6	C
	A2	5	4	5	5	C
	A3	6	5	6	6	B

	A4	4	4	4	4	A
	A5	6	6	6	6	A
	A6	6	6	6	6	D
	-	M: 5,5	M: 5,1	M: 5,5	M: 5,5	-
Formandos em licenciaturas - 2017	B1	5	4	5	5	C
	B2	5	5	6	5	B
	B3	5	5	5	5	A
	B4	6	6	6	6	C
	B5	6	6	6	6	A
	B6	6	6	6	6	D
	-	M: 5,5	M: 5,3	M: 5,6	M: 5,5	-

Fonte: dados da pesquisa. Legendas: ■: oscilação positiva do elemento de ensino em relação à pontuação dominante (moda) dentro do caso representativo; ■: oscilação negativa do elemento de ensino em relação à pontuação dominante (moda) dentro do caso representativo; ■: configuração de caso representativo com menores pontuações atribuídas ao conjunto dos elementos de ensino dentro da implementação; M: Média da variável; A: Não teve dificuldade; B: Não recebi apoio; C: Recebi apoio de fora da Instituição; D: Recebi apoio da Instituição.

A análise da média de cada variável na percepção dos casos representativos - entre as implementações 2014/2017 -, parece apontar melhora (I37 - Interação docente discente; e I42 - Exigência do curso) ou estabilidade (I28 - Conteúdos do curso e I55 - Articulação avaliações/conteúdos) da percepção dos estudantes formandos. Entretanto, se dentro das variáveis entre implementações ocorreu aumento ou manutenção da média, as variações dentro da configuração do vetor dos casos representativos indica informação útil para a retroalimentação do planejamento, observada a unidade desses elementos dentro do fenômeno 'ensino' e seu planejamento. São destacados no Quadro 2 os elementos de ensino que na percepção de caso(s) representativo(s) de cada agrupamento:

- (i) se destacaram negativamente em relação aos demais elementos, por sua oscilação negativa em relação à moda existente no vetor de resposta do caso representativo – variável I37 Interação docente e discente, nos agrupamentos A2 e A3 da implementação 2014/ no agrupamento B1 da implementação 2014;
- (ii) se destacaram positivamente em relação aos demais elementos, por sua oscilação positiva em relação à moda existente no vetor de resposta do caso representativo - variável I42 Exigência do curso, no agrupamento B2 da implementação 2017.

Ainda no Quadro 2, foram realçados os vetores de resposta dos casos representativos que possuíram pontuações mais baixas para todas as variáveis de interesse (agrupamento A4/implementação 2014; e agrupamento B3/implementação 2017). Para esses casos, embora as pontuações sejam elevadas em termos absolutos da escala (1-6), em termos relativos são as mais baixas, considerando cada vetor de dados dos casos representativos dentro das implementações. Nesses casos representativos, a variável de apoio em caso de dificuldade (I20) com valor "A"

(inexistência de dificuldades) pode indicar a necessidade de melhora no diagnóstico da organização didático-pedagógica com os estudantes. Embora tenha sido apontada a inexistência de dificuldades, a baixa percepção caracterizada pelos representantes desses agrupamentos pode ser objeto de análise detalhada, seja pela coordenação de curso ou pelos NDE, seja pela CPA da IES, por exemplo. Nesse cenário poderiam ser obtidas informações úteis para a retroalimentação do planejamento que de outra forma poderiam não ser identificadas em virtude do apontamento de inexistência de dificuldades para esses estudantes.

Quadro 3: Elementos de ensino e origem do apoio recebido em caso de dificuldade para os casos dos agrupamentos A2 e A3 da implementação 2014, do agrupamento B1 e B2 da implementação 2017.

Elemento de ensino	Percepção negativa	Percepção positiva	Apoio em caso de dificuldade
Interações entre docentes e discentes	▼(2014)	-	B: Não recebi apoio (2014); C: Recebi apoio de fora da Instituição (2014).
	▼(2017)	-	C: Recebi apoio de fora da Instituição (2017).
Exigência do curso	-	▲(2017)	B: Não recebi apoio (2017).

Fonte: dados da pesquisa. Legendas: ▼: oscilação negativa do elemento em relação à moda no vetor do caso representativo; ▲: oscilação positiva do elemento em relação à moda no vetor do caso representativo.

O Quadro 3 apresenta o resumo dos elementos de ensino com percepção positiva/negativa entre implementações, com o dado da variável de apoio recebido. De 2014 a 2017 a percepção negativa sobre a interação entre docentes e discentes associada ao não recebimento de apoio em caso de dificuldade deixou de caracterizar casos representativos. Persistiu entre esses anos a percepção negativa sobre a interação entre docentes e discentes acompanhada do apoio externo à Instituição de Ensino em caso de dificuldades. Em nenhum dos casos de percepção negativa da interação entre docentes e discentes o caso representativo foi caracterizado como tendo recebido apoio da IES. Essa situação pode indicar a existência de um lapso de planejamento ou organização institucional para o tratamento de necessidades estudantis que podem se relacionar com esse elemento de ensino, pois existindo caso de percepção negativa da interação entre docentes e discentes, o apoio recebido pelo estudante não teve origem na estrutura formal da instituição de ensino.

A percepção negativa em relação à interação entre docentes e discentes, para fins de planejamento, não se restringe ao âmbito do relacionamento interpessoal. A interação entre professor e estudante é modulada em sala por meio de determinados (i) procedimentos de ensino para a organização de atividades de aprendizagem; e (ii) recursos para despertar o interesse dos estudantes, facilitar a compreensão e estimular a participação. Assim, no caso do estudante possuir percepção negativa das interações e esta percepção se vincular à inadequação de procedimentos ou recursos de ensino, o fato do apoio para a superação dessa dificuldade não existir ou ser fornecido externamente à instituição pode significar que docentes ou a estrutura de apoio ao estudante (i) não estejam adequadamente instruídos sobre os procedimentos/recursos ou; (ii) não estejam ajustando o uso de procedimentos/recursos em função de diagnósticos com os estudantes, embora possam estar adequadamente instruídos sobre a aplicação/uso dos procedimentos ou recursos. Outra situação possível é que os procedimentos ou recursos não sejam adequados para a organização do ensino e a obtenção de um melhor ambiente de

aprendizagem. Em qualquer situação, o aspecto aprendiz do ambiente de ensino é prejudicado. Considerando que tal aspecto é associado a iniciativas baseadas sobretudo no ensino por diagnóstico da situação dos discentes para a geração de um ambiente estimulante e adequado, conforme apontado por Bransford, Brown e Cocking (2007), pode-se perceber o prejuízo sistêmico para o ensino e seus demais aspectos constituintes: o conteúdo e a avaliação.

O impacto para a retroalimentação do planejamento nas situações levantadas pode se vincular à capacitação de docentes ou de estruturação de apoio psicopedagógico ou psicológico ao estudante, por exemplo. Por outro lado, se o procedimento ou recurso utilizado é ineficaz para a organização das interações no ensino, o ponto de retroalimentação poderia se vincular à revisão de ementas ou planos de ensino, situação que poderia envolver o NDE e a coordenação de curso, com a revisão de objetivos de aprendizagem de componentes curriculares, e a pertinência dos procedimentos ou recursos em relação a conteúdos ministrados para o alcance daqueles objetivos. No caso da interação entre docentes e discentes ser negativamente percebida por fatores de relacionamento interpessoal, poderiam ser revisados ou aperfeiçoados mecanismos para diagnóstico e resolução de conflitos, com a estruturação ou envolvimento de apoio psicológico especializado para estudantes e professores.

Com relação ao recebimento de apoio de fora da instituição quando da existência de percepção negativa sobre a interação entre professores e estudantes, cabe apontar que, embora possa se considerar a existência de apoio externo uma situação desejada (na forma de uma necessária rede de apoio), a ausência de apoio decisivo da instituição na percepção do estudante pode indicar um ambiente de ensino aquém do esperado. Tal ambiente poderia ser aquele em que, embora ocorra aprendizagem adequada, o estudante não é plenamente beneficiado pela organização formal do ensino decorrente do planejamento, investindo esforço maior que o necessário ou em detrimento de outras necessidades de aprendizagem em outros componentes de um currículo ou programa educacional, por exemplo.

Quanto à percepção positiva em relação à exigência do curso em caso representativo da implementação com dados de 2017, a percepção positiva desse elemento foi acompanhada da informação de que existiu dificuldade e não ocorreu nenhum tipo de suporte. A situação configurada nesse caso representativo de agrupamento pode indicar que ocorreu o adequado ajuste da organização do curso, conforme as necessidades dos estudantes representados, ou que esse conjunto de estudantes pode ser caracterizado como resiliente frente às dificuldades existentes, se ajustando para melhor aproveitamento do ambiente de aprendizagem. No caso da primeira situação, a retroalimentação para o planejamento poderia ocorrer com o levantamento de boas práticas docentes e serviços de apoio ao estudante, para reforço e disseminação em outros componentes curriculares, por exemplo. No caso da segunda situação, esforços para a retroalimentação do planejamento poderiam se vincular à iniciativa institucional de pesquisa com esses estudantes para o levantamento de variáveis que poderiam ter contribuído para uma atitude resiliente, como a predisposição para a atividade autônoma do estudante, por exemplo.

### ***Percepção de elementos de ensino e subsídios para a organização da retroalimentação no ciclo de planejamento***

Considerando a configuração dos casos representativos dos agrupamentos B1 e B2 (Quadro 2), e as variáveis de destaque negativo e positivo apontadas na seção anterior - interações entre docentes e discentes, e exigência do curso, respectivamente -, ressalta-se que ambas possuem relação próxima, considerando sua relação para fins de planejamento, conforme Quadro 4. Ambas estão vinculadas ao fornecimento de informações, para os fins deste estudo, sobre a “definição de procedimentos de ensino” (DPE) e a “especificação de recursos de ensino”

(ERE), embora elas também possam fornecer informações sobre questões de relacionamento interpessoal de professores e estudantes na dinâmica de ensino e aprendizagem.

Quadro 4: Elementos de ensino e aspectos de planejamento para retroalimentação

Elemento de ensino	Ponto de planejamento associado				
	DOE	ECT	DPE	ERE	DAA
1) Os conteúdos abordados nas disciplinas do curso favoreceram sua atuação em estágios ou em atividades de iniciação profissional.	■	■			
2) As relações professor-aluno ao longo do curso estimularam você a estudar e aprender.				■	■
3) O curso exigiu de você organização e dedicação frequente aos estudos.			■	■	
4) As avaliações da aprendizagem realizadas durante o curso foram compatíveis com os conteúdos ou temas trabalhados pelos professores.					■

Fonte: dados da pesquisa. Legendas: DOE: Definição de Objetivos Educacionais; ECT: Especificação de Conteúdos a serem trabalhados; DPE: Definição de Procedimentos de Ensino; ERE: Especificação de Recursos de Ensino; DAA: Definição de forma e conteúdo das Avaliações de Aprendizagem; ■: elemento de ensino/ponto de planejamento neutros no escopo dos casos analisados; ■: elemento de ensino/ponto de planejamento de atenção negativa no escopo dos casos analisados; ■: elemento de ensino/ponto de planejamento de atenção positiva no escopo dos casos analisados.

A obtenção de dados positivos e negativos em variáveis que se refeririam aos mesmos elementos de planejamento poderia aparentar situação contraditória. Entretanto, considerando que diferentes estudantes possuem diferentes necessidades de aprendizagem, seria previsível que o melhor ajuste de uma situação de ensino a um grupo de estudantes poderia representar um ajuste insatisfatório a outro grupo de estudantes, com outras demandas de aprendizagem. Tal situação poderia ensejar diferentes estratégias de organização de estudantes em componentes curriculares com diferentes docentes, procedimentos e recursos de ensino, por exemplo.

Os achados da pesquisa em relação aos elementos de planejamento vinculados ao fator ‘aprendiz’ do ambiente de ensino, reforçam para os formandos analisados (implementação com dados de 2017) a importância da consideração dos conhecimentos, habilidades e atitudes dos estudantes, em linha com pesquisas que evidenciam o impacto do conhecimento prévio do estudante sobre a organização do ensino (Kalyuga, 2010; Plass, Kalyuga, Leutner, 2010).

Ainda em relação aos elementos identificados, a iniciativa centrada no aprendiz para a estruturação do ambiente de ensino pode ter estado aquém do esperado ou requerido pelos formandos com dados analisados. Da mesma forma, os itens de planejamento associados a esses elementos no estudo - “definição de procedimentos de ensino” e a “especificação de recursos de ensino” -, poderiam ter necessitado de organização mais próxima das necessidades previamente diagnosticadas com aqueles estudantes.

Em uma perspectiva unificada de concepção de procedimento de ensino e recursos para o estímulo à aprendizagem, os resultados apresentados reforçam a necessidade de iniciativas como aquelas que se utilizam de modelos cognitivos do aprendiz para o fornecimento de *feedback* na resolução de problemas em ambientes simulados, por exemplo (Lajoie, 2009). Com esse tipo de concepção coesa que objetiva auxiliar a aprendizagem em resolução de problemas complexos, auxiliando ainda na estruturação de melhores interações entre docentes e discentes para o estímulo a se estudar e aprender, docentes e discentes poderiam ter ao seu redor dados e

informações de performance efetiva para o subsídio sistemático e organizado da interação entre eles.

A envergadura desse tipo de melhoria da oferta do ensino pela criação de ambientes variados de aprendizagem, com a articulação de diferentes procedimentos de ensino e especificações de recursos mediadores da relação docente-discente no ambiente de ensino, demandaria esforços integrados para a pesquisa de outras variáveis associadas aos estudantes e ao próprio corpo docente, observado que este último operacionalizaria os diferentes procedimentos de ensino e as diferentes aplicações de recursos. Em situações de ensino a distância com o uso de Ambientes Virtuais de Aprendizagem, a criação de diferentes ambientes e a implementação de variações desejadas em procedimentos e recursos poderia ser mais facilmente realizada, embora esse tipo de iniciativa no ensino presencial não deva ser negligenciada, a despeito de sua maior complexidade de implementação. Nos dois cenários, a distância e presencial, o tipo de ajuste no ensino advindo da retroalimentação no ciclo de planejamento deveria ser precedido pela instituição e atuação de núcleos especializados e multidisciplinares de análise de dados pedagógicos, assim como pela adequada interface desse tipo de estrutura com a CPA, NDE, coordenação de curso, ou áreas de atendimento psicopedagógico ou psicológico, conforme o caso do ajuste projetado a partir do planejamento.

A adoção produtiva de um *framework* analítico como o proposto – e a respectiva obtenção de achados como os realizados –, se basearia sobretudo em sua efetiva inserção no ciclo de planejamento pedagógico das IES, com fluxos de interação definidos com instâncias de gestão ou avaliação interna, tanto no nível da instituição quanto no nível de cada um dos cursos. Da mesma forma, a constituição de núcleo multidisciplinar especializado no tratamento e análise de dados pedagógicos poderia ser beneficiado pela interação capacitada em dados de membros das áreas de sustentação do ensino, como os NDE, coordenações de curso, CPA e setores de apoio ao estudante. Nesse sentido, como apontam Chiappe e Rodriguez (2017), apenas uma efetiva disseminação do conhecimento em análise de dados para docentes, analistas e gestores educacionais pode impactar significativamente a prática educacional, seu planejamento, e retroalimentação que se quer baseada em dados.

No cenário abordado, seria viabilizada ação objetiva e orientada por dados para um ciclo efetivo de planejamento, com destaque para as ações de retroalimentação para a melhoria contínua da oferta educacional. Iniciativas para a orientação do planejamento pedagógico com base em dados, como as exemplificadas, podem contribuir para a consolidação de fluxos produtivos para o enriquecimento da tomada de decisão na gestão da educação superior, visando sobretudo o atendimento e a inclusão de estudantes a partir do adequado diagnóstico de suas necessidades no ambiente educacional.

## Conclusões

O objetivo estabelecido para o estudo, a saber: *caracterizar elementos de organização didático pedagógica do ambiente de ensino, e eventuais sistemas de apoio, passíveis de atenção em ciclo de retroalimentação do planejamento pedagógico para que a revisão de planejamento possa proporcionar melhores ambientes aos estudantes*, foi alcançado, observada a geração de dados e informações com potencial de interpretação pedagógica para o alcance de conjecturas para retroalimentação do planejamento pedagógico.

A aplicação de técnica específica de análise de agrupamentos permitiu a descoberta de padrões em dados, de modo que a própria estrutura de grupos descoberta forneceu informações para a compreensão do fenômeno do ensino e a identificação de pontos de atenção para a

retroalimentação do planejamento pedagógico. Nesse sentido, o estudo em questão contribuiu para a investigação de um fluxo de trabalho que associa técnica de análise de dados com a interpretação pedagógica dos resultados, apontando situações e possíveis instâncias envolvidas no aproveitamento dos dados para a melhoria do ensino e estruturas de apoio ao estudante. O estudo corrobora achados de Tives, Canedo, Ladeira e Fagundes (2018) em relação à utilidade de dados do ENADE, e em particular do questionário do estudante, para o auxílio à gestão de IES com o uso de Mineração de Dados. A importância do maior uso de dados do ENADE em análises como as propostas por Lima, Ambrósio, Ferreira e Brancher (2019) é reforçada, observados os potenciais ganhos decorrentes da interpretação pedagógica de resultados advindos da aplicação de técnicas de EDM, por exemplo.

Com os conjuntos de dados utilizados, e considerando a devida limitação de escopo das conclusões a tais dados, foi identificado que os seguintes elementos de planejamento poderiam passar por escrutínio no processo de replanejamento, segundo a percepção de estudantes sobre elementos de ensino: “definição de procedimentos de ensino” e a “especificação de recursos de ensino”. Com tais dados, um docente teria condições de operar de forma mais eficiente em seu replanejamento, para o alcance do melhor ajuste entre ensino e necessidades dos estudantes para o desenvolvimento das aprendizagens propostas para determinado componente curricular ou curso. Essa possibilidade decorreria do cenário em que tais dados fossem coletados com frequência definida, de modo que a implementação do replanejamento pudesse impactar sobre a continuidade do processo de ensino. Os resultados parecem reforçar a importância de descobertas vinculados à importância do conhecimento prévio do estudante como fator chave para a organização do ambiente de ensino (Kalyuga, 2010; Plass, Kalyuga, Leutner, 2010), observada a associação entre procedimentos e recursos de ensino com o fator ‘aprendiz’ do ambiente de ensino (Bransford, Brown & Cocking, 2007). Iniciativas que buscam a assistência baseada em computador (Lajoie, 2009), de forma a se integrar procedimento e recurso de ensino com a geração de dados e informações úteis para a interação sistemática de docente e discente também é reforçada para a retroalimentação do planejamento pedagógico no cenário de formandos com dados analisados.

Da mesma forma, instâncias de gestão interna de uma IES – como NDE e coordenação de curso -, poderiam obter insumos suficientes para executar ações de sondagem e diagnóstico, de modo que pudessem aperfeiçoar componentes curriculares ou um currículo como um todo.

Outras instâncias relacionadas ao apoio discente – como seções de Psicologia ou Psicopedagogia da instituição -, poderiam obter informações importantes associadas ao ensino e à suficiência do atendimento existente, para assim dimensionar ajustes de oferta ou a definição de estratégias para focalização de estudantes para atendimento. Como apontaram Santos e Pedrosa (2021), é necessário o incremento de ações para a resolução de questões pedagógicas e psicológicas dos estudantes, assim como questões intrínsecas ao estudante e ao seu bem-estar para o melhor aproveitamento em seu curso, evitando a ocorrência de evasão. Nas situações de aperfeiçoamento apontadas acima a partir de estudo como o proposto, seria alcançada Tomada de Decisão Educacional Orientada por Dados, como revisado por Filatro (2021) – condição que corrobora tais achados associados à prática pedagógica concreta.

Com relação ao incremento da técnica utilizada no estudo, poderia ocorrer ainda a expansão do fluxo de análise de dados. Observada a proximidade das quantidades ótimas de agrupamentos por implementação realizada, no cenário de descoberta de estruturas de agrupamentos com características estáveis ao longo do tempo, poderia se conjecturar o uso de tais casos representativos por agrupamentos como âncoras para classificação de estudantes com base em algoritmo de vizinho mais próximo (*k-Nearest Neighbors – kNN*), por exemplo. Nessa situação a IES poderia passar a atuar de forma prospectiva, antecipando ajustes de planejamento

a partir da atuação orientada por dados das instâncias mencionadas anteriormente, além dos próprios docentes.

Como fator limitante da pesquisa realizada, pode-se elencar o uso de dados secundários, observada a limitação dos dados utilizados ao desenho dos instrumentos de coleta que os geraram. Assim, o desenvolvimento de questionário próprio para a coleta de dados no escopo estrito dos elementos de ensino de interesse poderia ser um incremento para a continuidade do estudo. Da mesma forma, a aplicação de processo analítico análogo ao proposto, em outros contextos de ensino, como na educação básica, profissional ou de pós-graduação *lato e stricto sensu*, poderia contribuir na geração de ajustes cabíveis nos elementos de ensino pesquisados, observados eventuais fatores específicos de ensino relacionados a cada uma dessas etapas de formação. A continuidade das pesquisas poderiam se vincular também ao desenvolvimento de *dashboards* para a apresentação dos dados dos agrupamentos e seus casos representativos, com a aplicação de conceitos de *Data Storytelling*. Da mesma forma, estudos derivados poderiam se vincular à apresentação dos dados e sua interpretação pedagógica (*framework* analítico) a gestores de IES e cursos de graduação para pesquisa sobre a percepção de potenciais impactos que esse tipo de iniciativa traria para a organização educacional como um todo.

A partir do estudo realizado e do potencial de impacto da implementação de práticas de gestão pedagógica baseada em dados, considera-se salutar o desenvolvimento de produtos educacionais para que docentes, gestores, analistas educacionais, psicólogos e pedagogos, por exemplo, possam se apropriar de conceitos, procedimentos e técnicas de análise de dados, EDM, *learning analytics* e ciência de dados. A criação de condições para que esse tipo de arcabouço teórico possa ser aproveitado no contexto imediato de inserção daqueles profissionais contribuiria para a geração de inteligência organizacional de relevo em instituições de ensino. Da mesma forma, a disseminação desse tipo de conhecimento entre aqueles profissionais poderia contribuir para o efetivo surgimento de núcleos especializados e multidisciplinares de análise de dados pedagógicos, local em que o conhecimento finalístico das áreas educacionais, e de seu apoio especializado, faria interface concreta com a análise de dados para o aperfeiçoamento educacional.

## Referências

- Atlas do Desenvolvimento Humano no Brasil. Ranking – todo o Brasil. (2010). Disponível em: <<http://www.atlasbrasil.org.br/>>. Acesso em: 7 outubro. 2018 [GS Search]
- Bach, T. M., Water, S. A., Frega, J. R., & Muller, J. M. S. (2014). Fatores de influência na aprendizagem percebida dos alunos de cursos de administração. Avaliação 19 (1), 13-30. doi: [10.1590/S1414-40772014000100002](https://doi.org/10.1590/S1414-40772014000100002) [GS Search]
- Baker, R. S. J., Isotani, S., & Carvalho, A. M. J. B. (2011). Mineração de dados educacionais: oportunidades para o Brasil. Revista Brasileira de Informática na Educação 19 (2), 3-13. doi: [10.5753/RBIE.2011.19.02.03](https://doi.org/10.5753/RBIE.2011.19.02.03) [GS Search]
- Beck, F., & Rausch, R. B. (2015). Fatores que influenciam o processo de ensino-aprendizagem na percepção de discentes do curso de Ciências Contábeis. Revista Contabilidade Vista & Revista, v. 25, n. 2, p. 38-58. [GS Search]
- Bransford, J. D., Brown, A L., Cocking, R. R. (2007). Como as pessoas aprendem: cérebro, mente, experiência e escola. Conselho Nacional de Pesquisa dos Estados Unidos. São Paulo: Editora Senac São Paulo. [GS Search]

- Brasil. Instituto Brasileiro de Geografia e Estatística. (2018). Renda domiciliar per capita. Disponível em: <[https://ww2.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pnad\\_continua/default\\_renda\\_percapita.shtm](https://ww2.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pnad_continua/default_renda_percapita.shtm)>. Acesso em 7 outubro. 2018.
- Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2018a). Dicionário de variáveis dos microdados do ENADE edição 2014. Disponível em <[http://download.inep.gov.br/microdados/Enade\\_Microdados/microdados\\_enade\\_2014.zip](http://download.inep.gov.br/microdados/Enade_Microdados/microdados_enade_2014.zip)>. Acesso em 7 julho. 2018.
- Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2018b). Dicionário de variáveis dos microdados do ENADE edição 2017. Disponível em <[http://download.inep.gov.br/microdados/Enade\\_Microdados/microdados\\_Enade\\_2017\\_portal\\_2018.10.09.zip](http://download.inep.gov.br/microdados/Enade_Microdados/microdados_Enade_2017_portal_2018.10.09.zip)>. Acesso em 27 outubro. 2018.
- Bruer, J. T. (1993). *Schools for thought: a science of learning in the classroom*. Cambridge: MIT Press. [GS Search]
- Cardoso, A. S. (2021). *Sistemática para Avaliação de Cursos de Graduação: Perspectiva dos Discentes*. (Dissertação). Universidade Federal de Itajubá. Programa de Pós-Graduação em Engenharia de Produção. [GS Search]
- Chiappe, A., & Rodriguez, L. P. (2017). Learning Analytics in 21st century education: a review. *Ensaio: aval.pol.públ.Educ.* vol.25 no.97. Rio de Janeiro. Oct./Dec. Doi: [10.1590/S0104-40362017002501211](https://doi.org/10.1590/S0104-40362017002501211) [GS Search]
- Faria, S. M. S. M. L. (2014). *Educational Data Mining e Learning Analytics na melhoria do ensino online*. (Dissertação). Universidade Aberta. Departamento de Ciências e Tecnologia. Mestrado em Estatística, Matemática e Computação. Especialização - Estatística Computacional. [GS Search]
- Filatro, A. (2021). *Data Science na educação: presencial, a distância e corporativa*. São Paulo: Saraiva Educação. [GS Search]
- Frei, F. (2006). *Introdução a análise de agrupamentos: teoria e prática*. São Paulo: Editora UNESP. [GS Search]
- Gower, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics*. 857-874. [GS Search]
- Hair, J. F., Black, W. C., & Babin, B. J.; Anderson, R. E.; Tatham, R. L. (2009). *Análise multivariada de dados*. 6ª edição. Porto Alegre: Bookman. [GS Search]
- Haydt, R. C. C. (2006). *Curso de didática geral*. 8a edição. São Paulo: Ática. [GS Search]
- Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing teachers to learn from teaching. *Journal of Teacher Education*, Vol. 58, No. 1, p. 47-61. [GS Search]
- Kalyuga, S. (2010). Schema acquisition and sources of cognitive load. In Plass J. L., Moreno, R., Brünken, R. *Cognitive Load Theory*. New York: Cambridge University Press. [GS Search]
- Kassambara, A. (2018). *Multivariate data analysis I: Practical guide to cluster analysis in R: unsupervised machine learning (Edition 1)*. [GS Search]
- Lajoie, S. P. (2009). Developing professional expertise with a cognitive apprenticeship model: examples from avionics and medicine. In Ericsson, K. A. *Development of professional*

- expertise: toward measurement of expert performance and design of optimal learning environments*. New York: Cambridge University Press. [GS Search]
- Lima, P. S. N., Ambrósio, A. P. L., Ferreira, D. J., & Brancher, J. D. (2019). Análise de dados do Enade e Enem: uma revisão sistemática da literatura. *Avaliação*, Campinas; Sorocaba, SP, v. 24, n. 1, p. 89-107. doi: [10.1590/S1414-40772019000100006](https://doi.org/10.1590/S1414-40772019000100006) [GS Search]
- Moissa, B., Gasparini, I., & Kemczinski, A. (2015). Educational Data Mining versus Learning Analytics: estamos reinventando a roda? Um mapeamento sistemático. *Anais do XXVI Simpósio Brasileiro de Informática na Educação (SBIE 2015)*. doi: [10.5753/cbie.sbie.2015.1167](https://doi.org/10.5753/cbie.sbie.2015.1167) [GS Search]
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2017). *cluster: cluster analysis basics and extensions*. R package version 2.0.6. [GS Search]
- Plass, J. L., Kalyuga, S., Leutner, D. (2010). Individual differences and cognitive load theory. In Plass J. L., Moreno, R., Brünken, R. *Cognitive Load Theory*. New York: Cambridge University Press. [GS Search]
- R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Áustria. Disponível em <https://www.R-project.org/>.
- Rodrigues Júnior, J. F. (2002). *Manual para formação do instrutor*. Brasília: Editora Universa. [GS Search]
- Rousseeuw, P., & Kaufman, L. (1990). *Finding groups in data: an introduction to cluster analysis*. New York: Wiley. [GS Search]
- Santos, M. M., & Pedroso, I. G. F. (2021). Percepção discente sobre cursos de graduação em Ciências Agrárias e Humanidades da UNESP. *Educ. Pesqui.*, São Paulo, v. 47, e227954. Doi: [10.1590/S1678-4634202147227954](https://doi.org/10.1590/S1678-4634202147227954) [GS Search]
- Simon, H. A. (2000). Observations on the sciences of science learning. *Journal of Applied Developmental Psychology* 21(1): 115–121. [GS Search]
- Silva, L. A., Peres, S. M., & Boscaroli, C. (2016). *Introdução à mineração de dados: com aplicações em R*. Rio de Janeiro: Elsevier. [GS Search]
- Sordan, J. E., Marinho, C. A., & Angeloni, M. (2021). Avaliação da qualidade percebida em uma Instituição de Ensino Superior: uma pesquisa-ação. *Universidade Federal da Grande Dourados: III Simpósio Nacional de Engenharia de Produção*. [GS Search]
- Tives, H. A., Canedo, E. D., Ladeira, M., & Fagundes, F. (2018). Mining ENADE Data from the Ulbra Network Institution. In *Information Technology – New Generations*. Doi: [10.1007/978-3-319-77028-4\\_39](https://doi.org/10.1007/978-3-319-77028-4_39) [GS Search]
- Tyler, R. W. (1974). *Princípios básicos de currículo e ensino*. Rio de Janeiro: Editora Globo. [GS Search]
- Zanella, A., & Lopes, L. F. D. (2009). Diagnóstico do ensino-aprendizagem e satisfação dos alunos nas disciplinas de estatística da UFMS. *GEPROS, Gestão da Produção, Operações e Sistemas*. Ano 4, nº 3, Jul-Set, 2009, p. 123-140. [GS Search]