



Dashboards Educacionais: Impacto da Complexidade na Carga Cognitiva e Usabilidade

Deivinson Severino da Silva, Universidade Federal de Alagoas,

dsds@ic.ufal.br, <https://orcid.org/0009-0008-6682-1625>

Christian Fabbro, Universidade Federal de Alagoas,

chf@ic.ufal.br, <https://orcid.org/0009-0002-2094-9144>

José Oséas de Oliveira Filho, Universidade Federal de Alagoas,

jose.oseas@educ.al.gov.br, <https://orcid.org/0000-0003-4516-0342>

Resumo: O uso de dashboards educacionais tem crescido como ferramenta de apoio à decisão pedagógica. No entanto, dashboards excessivamente complexos podem aumentar a carga cognitiva dos professores, reduzindo sua utilidade. Este artigo apresenta um estudo avaliativo de três versões de um dashboard com diferentes níveis de complexidade (simples, intermediário, complexo). Participaram da avaliação 13 professores da rede estadual de Alagoas, atuantes no Ensino Médio, que responderam aos instrumentos NASA-TLX reduzido (carga cognitiva), SWAT (carga cognitiva) e SUS reduzido (usabilidade). Os resultados indicam que a carga cognitiva aumenta com a complexidade (V1: 5.0; V2: 5.3; V3: 5.7), com diferença estatisticamente significativa para a dimensão Demanda Mental do NASA-TLX (teste de Friedman, $p = 0,039$). A versão intermediária (V2) apresentou um possível equilíbrio entre utilidade percebida e usabilidade. A versão mais complexa (V3) foi a mais escolhida pelos professores (38%), revelando um possível trade-off entre usabilidade e utilidade percebida. A análise qualitativa das sugestões apontou três categorias principais: clareza visual, personalização e tipos de gráficos. O estudo fornece evidências empíricas sobre a relação entre complexidade e carga cognitiva no contexto da educação básica brasileira, além de contribuir para o design de dashboards educacionais que busquem equilibrar riqueza informacional e esforço cognitivo.

Palavras-chave: dashboards educacionais, carga cognitiva, usabilidade, NASA-TLX, SWAT, SUS.

Educational Dashboards: Impact of Complexity on Cognitive Load and Usability

Abstract: The use of educational dashboards has grown as a tool to support pedagogical decision-making. However, overly complex dashboards can increase teachers' cognitive load, potentially reducing their usefulness. This paper presents an evaluative study of three versions of an educational dashboard with different levels of complexity (simple, intermediate, complex). Thirteen teachers from the state network of Alagoas, working in high school, participated in the evaluation, responding to reduced NASA-TLX (cognitive load), SWAT (cognitive load), and reduced SUS (usability) instruments. The results indicate that cognitive load increases with complexity (V1: 5.0; V2: 5.3; V3: 5.7), with a statistically significant difference for the Mental Demand dimension of NASA-TLX (Friedman test, $p = 0.039$). The intermediate version (V2) presented a possible balance between perceived usefulness and usability. The most complex version (V3) was the most frequently selected by teachers (38%), revealing a possible trade-off between usability and perceived usefulness. Qualitative analysis of suggestions revealed three main categories: visual clarity, customization, and chart types. The study provides empirical evidence on the relationship between complexity and cognitive load in the context of Brazilian basic education and contributes to the design of educational dashboards that seek to balance information richness and cognitive effort.



Keywords: educational dashboards, cognitive load, usability, NASA-TLX, SWAT, SUS.

1. Introdução

O uso de dashboards educacionais tem crescido como ferramenta de apoio à decisão pedagógica, consolidando dados de desempenho, frequência e engajamento para identificar estudantes em risco de evasão ou baixo rendimento (Nguyen *et al.*, 2024). Entretanto, dashboards excessivamente complexos podem aumentar a carga cognitiva dos professores e reduzir sua utilidade (Karademir *et al.*, 2024). Esse fenômeno é particularmente relevante no contexto brasileiro, em que professores da rede pública enfrentam turmas numerosas, jornadas de trabalho extensas e recursos tecnológicos limitados.

No Brasil, iniciativas como a Rede de Inovação para Educação Híbrida (RIEH) (Ministério da Educação (MEC), 2025) buscam ampliar a integração de tecnologias educacionais na rede pública, especialmente no Ensino Médio. Nesse contexto, dashboards podem apoiar a interpretação de dados educacionais e a tomada de decisões pedagógicas, desde que apresentem um equilíbrio adequado entre riqueza informacional e facilidade de uso.

Apesar do avanço em Learning Analytics (LA) e do desenvolvimento de dashboards voltados a professores (Lemes; Dias e Oliveira, 2023), ainda são limitados os estudos que avaliam empiricamente como diferentes níveis de complexidade afetam a carga cognitiva e a usabilidade percebida por professores da educação básica brasileira. A literatura concentra-se, em grande parte, no ensino superior ou em contextos internacionais (Nguyen *et al.*, 2024; Karademir *et al.*, 2024), evidenciando uma lacuna na compreensão desse fenômeno no contexto da educação básica brasileira.

Diante desse cenário, este estudo avalia três versões de um dashboard educacional com diferentes níveis de complexidade (simples, intermediário e complexo), buscando responder à seguinte questão de pesquisa: *Como diferentes níveis de complexidade em dashboards educacionais impactam a carga cognitiva e a usabilidade percebida por professores do Ensino Médio da rede estadual de Alagoas?*

Para responder a essa questão, foram desenvolvidos protótipos funcionais e submetidos à avaliação de 13 professores, utilizando os instrumentos NASA-TLX reduzido, SWAT e SUS reduzido, além de medidas de utilidade percebida e preferência geral. Como contribuições, o estudo: (1) fornece evidências empíricas sobre a relação entre complexidade e carga cognitiva no contexto da educação básica brasileira; (2) compara diferentes instrumentos para avaliação da carga de trabalho percebida; e (3) evidencia um possível trade-off entre usabilidade e utilidade percebida, com implicações para o projeto de dashboards educacionais.

O artigo está estruturado da seguinte forma: a Seção 2 apresenta os trabalhos relacionados; a Seção 3 apresenta a fundamentação teórica; a Seção 4 descreve a metodologia; a Seção 5 detalha a proposta e o sistema desenvolvido; a Seção 6 apresenta os resultados; a Seção 7 discute os achados; e a Seção 8 apresenta as considerações finais.

2. Trabalhos Relacionados

2.1. Dashboards educacionais para professores

Diferentes abordagens têm sido propostas para desenvolver dashboards voltados a professores. Nguyen *et al.* (2024) apresentaram o TEADASH, desenvolvido com Design Science Research (DSR) em três ciclos, envolvendo co-design com professores e avaliação em cinco cursos reais. Os resultados demonstraram que dashboards podem



reduzir o tempo de análise de dados e apoiar intervenções pedagógicas, embora nem todas as funcionalidades sejam necessariamente utilizadas pelos docentes. Karademir et al. (2024), por sua vez, investigaram requisitos para um Learning Analytics Cockpit voltado à avaliação formativa e, por meio de entrevistas de co-design com 16 professores alemães, identificaram 16 requisitos funcionais, destacando a necessidade de automação com controle docente e de redução da carga de trabalho.

No contexto brasileiro, uma Revisão Sistemática da Literatura (Lemes; Dias e Oliveira, 2023) identificou 20 estudos publicados entre 2010 e 2021, evidenciando a predominância de pesquisas no ensino superior e a carência de investigações voltadas à educação básica brasileira. Lee-Cultura, Sharma e Giannakos (2024) avaliaram um dashboard multimodal com 20 professores, utilizando dados provenientes de sensores e análise de engajamento. Os resultados indicaram que interfaces com elevada densidade informacional podem aumentar a sobrecarga cognitiva, além de evidenciarem desafios relacionados à habilidade técnica, à literacia em dados educacionais e à adoção dessas tecnologias.

2.2. Avaliação da carga cognitiva

O NASA-TLX (Madi; Peng e Rogers, 2022) e o SWAT (Reid e Nygren, 1988) são instrumentos consolidados para avaliação da carga de trabalho percebida, sendo utilizados em diferentes contextos de interação humano-computador. Madi, Peng e Rogers (2022) demonstraram, em um estudo com 85 estudantes de Ciência da Computação, que a carga cognitiva percebida é apenas parcialmente explicada por medidas objetivas de complexidade. Rubio et al. (2004), ao compararem NASA-TLX, SWAT e Workload Profile, observaram diferenças na sensibilidade e no tempo de aplicação dos instrumentos, recomendando seu uso combinado para ampliar a avaliação da carga de trabalho percebida.

2.3. Lacunas e contribuição

A literatura analisada evidencia três lacunas principais: (1) a predominância de estudos sobre dashboards educacionais no ensino superior, com menor atenção à educação básica brasileira (Lemes; Dias e Oliveira, 2023; Nguyen *et al.*, 2024); (2) a escassez de avaliações empíricas que investiguem sistematicamente o impacto de diferentes níveis de complexidade de dashboards sobre a experiência de professores (Karademir *et al.*, 2024); e (3) a necessidade de compreender conjuntamente os efeitos da complexidade sobre carga cognitiva, usabilidade e utilidade percebida.

Diante dessas lacunas, este estudo avalia três versões de um dashboard educacional, classificadas em níveis simples, intermediário e complexo, com professores da educação básica brasileira. A avaliação combina NASA-TLX, SWAT e SUS, permitindo analisar diferentes dimensões da experiência de uso e investigar o possível trade-off entre riqueza informacional, carga cognitiva e usabilidade percebida.

3. Fundamentação Teórica

O estudo fundamenta-se em três conceitos centrais: carga cognitiva, usabilidade e complexidade no design de dashboards educacionais. Esses conceitos estabelecem a base teórica para analisar como diferentes níveis de complexidade podem influenciar a experiência de uso e a percepção dos professores.



3.1. Teoria da Carga Cognitiva

A Teoria da Carga Cognitiva, proposta por Sweller (1988), considera que a memória de trabalho possui capacidade limitada para processar informações. Quando as demandas de uma tarefa excedem essa capacidade, pode ocorrer sobrecarga cognitiva, comprometendo a compreensão e a tomada de decisão (Sweller, 1988).

A teoria distingue três tipos de carga: intrínseca, relacionada à complexidade inerente da tarefa; extrínseca, decorrente da forma como as informações são apresentadas; e germânica, associada aos processos envolvidos na construção e organização de esquemas mentais. No contexto de dashboards educacionais, a carga extrínseca assume particular relevância, pois pode ser influenciada diretamente por decisões de design, como a quantidade de informações, a complexidade das visualizações e a organização da interface (Karademir *et al.*, 2024).

Assim, o aumento da densidade informacional e da quantidade de funcionalidades pode ampliar o esforço necessário para localizar, interpretar e relacionar informações. Neste estudo, essa relação é investigada por meio da avaliação da carga de trabalho percebida pelos professores diante de dashboards com diferentes níveis de complexidade.

3.2. Usabilidade e Utilidade Percebida

A usabilidade refere-se à facilidade com que os usuários conseguem utilizar uma interface para atingir seus objetivos. Nielsen (1993) relaciona esse conceito a aspectos como facilidade de aprendizado, eficiência, memorização, ocorrência de erros e satisfação (Nielsen, 1993). Em dashboards educacionais, a usabilidade é particularmente relevante, pois os professores precisam interpretar informações de forma rápida e eficiente para apoiar decisões pedagógicas.

A utilidade percebida, por sua vez, representa o grau em que o usuário acredita que uma tecnologia contribui para melhorar seu desempenho (Davis, 1989). No contexto deste estudo, essa percepção está relacionada à capacidade do dashboard de apresentar informações consideradas relevantes para a prática docente, como indicadores de desempenho, frequência e possíveis situações de risco.

Embora relacionadas, usabilidade e utilidade representam dimensões distintas da experiência de uso. Uma interface pode ser fácil de utilizar, mas apresentar poucas informações relevantes, assim como pode oferecer informações consideradas úteis e, simultaneamente, exigir maior esforço para sua interpretação. Essa distinção é fundamental para compreender possíveis relações de compromisso entre simplicidade, riqueza informacional e valor percebido em dashboards educacionais.

3.3. Complexidade no Design de Dashboards Educacionais

O design de dashboards educacionais envolve decisões sobre a seleção, organização e representação visual das informações. Echeverria et al. (2018) destacam a importância de equilibrar uma visão geral dos dados com possibilidades de exploração detalhada, permitindo que diferentes níveis de informação sejam acessados conforme a necessidade do usuário.

Nesse contexto, a complexidade de um dashboard pode ser compreendida a partir de características como densidade informacional, quantidade de funcionalidades interativas, número de filtros e variedade de representações visuais. O aumento desses elementos pode ampliar a riqueza das informações disponíveis, mas também elevar o esforço necessário para sua interpretação.

A literatura sugere, portanto, que dashboards educacionais devem buscar equilíbrio entre riqueza informacional e facilidade de uso, considerando as necessidades



e características de seus usuários (Lee-Cultura; Sharma e Giannakos, 2024). Essa perspectiva fundamenta a comparação realizada neste estudo entre três versões de um dashboard, desenvolvidas com níveis crescentes de complexidade, permitindo investigar como essas diferenças se relacionam com a carga cognitiva, a usabilidade e a utilidade percebida pelos professores.

4. Metodologia

A pesquisa adotou uma abordagem de avaliação de artefato baseada em simulação e teste de conceito (Ghanbari; Tuunanen e Kazan, 2025), seguindo os princípios do FEDS (*Framework for Evaluation in Design Science*) (Venable; Pries-Heje e Baskerville, 2016). Considerando o caráter exploratório e as restrições de tempo, foi realizada uma avaliação artificial com viés formativo, alinhada aos princípios de *short-scoping* (Baskerville; Pries-Heje e Venable, 2026).

4.1. Participantes

Participaram 13 professores da rede estadual de Alagoas, atuantes no Ensino Médio, selecionados por conveniência a partir de contatos com escolas de duas regiões do estado. A maioria (61,5%) possuía mais de 15 anos de experiência docente e utilizava plataformas digitais educacionais com frequência (1–2 vezes por semana).

Consultas exploratórias com professores da rede estadual também contribuíram para identificar necessidades consideradas na definição dos requisitos funcionais do artefato. Essas consultas tiveram caráter exploratório e não foram utilizadas como fonte formal de dados na análise dos resultados.

4.2. Instrumentos de coleta

Para cada versão do dashboard, os participantes responderam:

- **NASA-TLX reduzido** (Madi; Peng e Rogers, 2022): Demanda Mental, Desempenho e Frustração, em escala de 0 a 10;
- **SWAT** (Reid e Nygren, 1988): Demanda de Tempo, Esforço Mental e Estresse, em escala ordinal de três níveis (1 = baixo, 2 = médio, 3 = alto);
- **SUS reduzido** (Brooke, 1996): três itens sobre facilidade de uso, percepção de complexidade e intenção de uso frequente, em escala Likert de cinco pontos;
- **Utilidade percebida**: questão sobre a contribuição do dashboard para decisões pedagógicas, em escala Likert de cinco pontos.

Ao final, os participantes indicaram sua versão preferida (V1, V2, V3 ou nenhuma) e puderam registrar sugestões de melhoria.

4.3. Procedimento

A coleta ocorreu entre 15 e 23 de maio de 2026, por meio de formulário eletrônico. O fluxo compreendeu: (1) leitura das instruções; (2) acesso aos dashboards; (3) exploração livre das três versões, sem ordem obrigatória; (4) preenchimento dos instrumentos para cada versão; e (5) indicação da preferência geral e registro de sugestões.

4.4. Análise dos dados

Foram calculadas médias e desvios padrão para NASA-TLX, SWAT, SUS e utilidade percebida em cada versão. As diferenças entre as versões foram avaliadas pelo teste de Friedman, equivalente não paramétrico à ANOVA de medidas repetidas. A preferência geral foi analisada por frequências e percentuais, e as respostas abertas por análise de conteúdo categorial.



4.5. Aspectos éticos

A participação foi voluntária e anônima, sem coleta de dados pessoais identificáveis. Os dashboards utilizaram exclusivamente dados simulados de estudantes fictícios, não envolvendo informações acadêmicas reais.

A avaliação exploratória e formativa, baseada em dados simulados e realizada em condições artificiais, permitiu analisar a percepção dos participantes sobre as diferentes versões sem exposição de dados educacionais reais.

5. Proposta / Sistema

5.1. Definição dos Requisitos e Níveis de Complexidade

Os requisitos das três versões do dashboard foram definidos a partir de duas fontes: (1) recomendações da literatura sobre dashboards educacionais e (2) necessidades levantadas em consultas exploratórias com professores da rede estadual de Alagoas. A literatura destaca a importância de visualizações claras, filtros para segmentação, alertas e apoio à análise de dados (Karademir *et al.*, 2024; Nguyen *et al.*, 2024). As consultas exploratórias indicaram como necessidades prioritárias a visualização rápida de frequência e notas, a comparação entre alunos e turmas e a identificação de padrões relacionados a desempenho, engajamento e risco de evasão.

A partir desses requisitos, as versões foram definidas por níveis crescentes de complexidade funcional e informacional, considerando três critérios: (i) quantidade e diversidade de informações apresentadas; (ii) número de mecanismos de interação e filtragem; e (iii) complexidade das análises e visualizações disponibilizadas. A **V1 (Simples)** priorizou indicadores essenciais, como frequência, nota média e identificação visual de risco, sem filtros ou análises comparativas. A **V2 (Intermediária)** acrescentou filtro por turma, evolução temporal das notas, informações de engajamento e comparação com a média da turma. A **V3 (Complexa)** incorporou múltiplos filtros, análise da relação entre frequência e nota, *heatmap* de desempenho, alertas de risco e múltiplas métricas.

5.2. Desenvolvimento dos Dashboards

Foram desenvolvidas três versões funcionais de um dashboard educacional utilizando HTML5, Tailwind CSS, JavaScript (Vanilla) e Chart.js, com hospedagem no GitHub Pages. Para garantir condições equivalentes de avaliação, as três versões utilizaram a mesma base de dados simulados, composta por 45 alunos fictícios distribuídos em nove turmas de Ensino Médio, gerada com *seed* fixa para assegurar reprodutibilidade. As versões foram disponibilizadas aos participantes por meio de links públicos.

A Figura 4 apresenta as três versões desenvolvidas, evidenciando o aumento progressivo da densidade informacional, das possibilidades de interação e da complexidade das análises.

5.3. Dados Utilizados

Todos os dados exibidos nos dashboards são simulados e gerados por algoritmo com *seed* fixa. Foram criados 45 alunos fictícios distribuídos em nove turmas de Ensino Médio (1º A, 1º B, 1º C, 2º A, 2º B, 2º C, 3º A, 3º B e 3º C). Os indicadores incluem frequência, nota média, notas de avaliações, engajamento, último acesso e nível de risco.

A utilização de dados simulados segue a estratégia de *free simulation experiments* proposta por Ghanbari, Tuunanen e Kazan (2025), adequada à avaliação de artefatos em estágios iniciais de desenvolvimento. A adoção dessa estratégia também evita a exposição

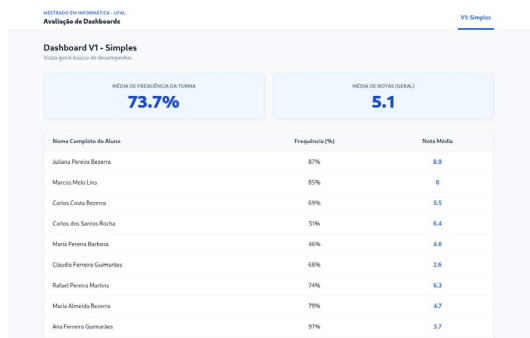


Figura 1. V1 – Simple

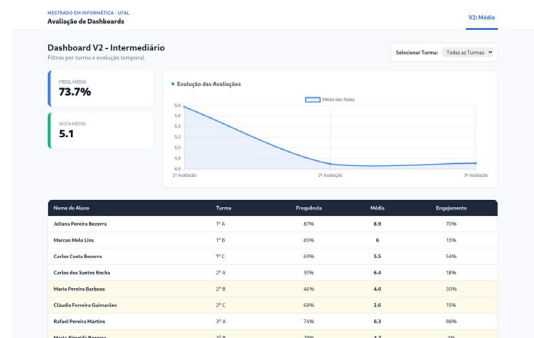


Figura 2. V2 – Intermediário

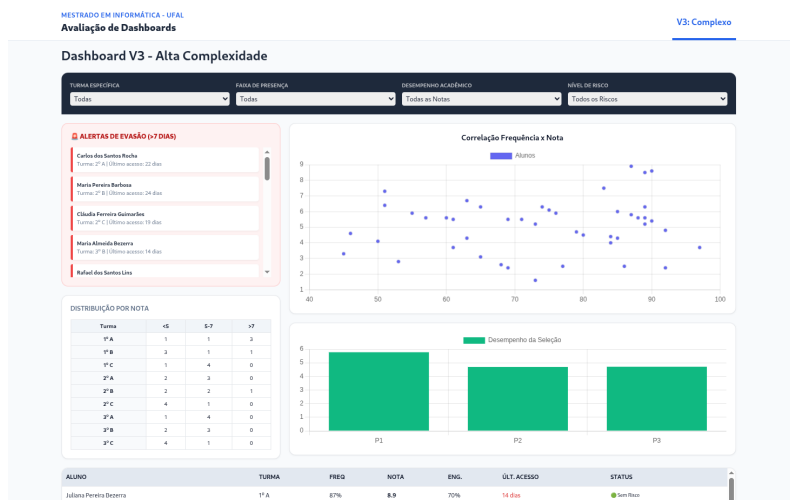


Figura 3. V3 – Complexo

Figura 4. As três versões do dashboard com níveis crescentes de complexidade.

de dados educacionais reais.

6. Resultados

Participaram da avaliação 13 professores da rede estadual de Alagoas, atuantes no Ensino Médio. Os resultados são apresentados quanto à carga cognitiva percebida, avaliada pelo NASA-TLX e SWAT, à usabilidade, medida pelo SUS reduzido, à utilidade percebida, às sugestões dos participantes e à preferência geral pelas três versões.

6.1. Carga Cognitiva: NASA-TLX

A Tabela 1 indica tendência de aumento da carga cognitiva percebida conforme a complexidade do dashboard, com médias gerais de 5,0 (V1), 5,3 (V2) e 5,7 (V3). A maior variação ocorreu na dimensão **Demanda Mental**, que passou de 4,6 na V1 para 6,3 na V3, correspondendo a um aumento de aproximadamente 37%. Em sentido oposto, o **Desempenho** percebido diminuiu de 7,4 (V1) para 6,4 (V3), enquanto a **Frustração** aumentou de 3,0 para 4,4.

O teste de Friedman indicou diferença estatisticamente significativa apenas para a **Demanda Mental** ($\chi^2(2) = 6.48$, $p = 0.039$), sugerindo que o aumento da complexidade esteve associado a maior esforço mental percebido. As demais dimensões não apresentaram diferenças estatisticamente significativas.

Os dashboards estão disponíveis em: V1 - Simple, V2 - Intermediário, e V3 - Complexo.



Tabela 1. Médias e desvios padrão do NASA-TLX por versão

Dimensão	V1	V2	V3
Demanda Mental (0–10)	4.6 (3.1)	4.8 (2.7)	6.3 (2.5)
Desempenho (0–10)	7.4 (2.2)	6.7 (2.2)	6.4 (2.7)
Frustração (0–10)	3.0 (2.0)	4.4 (2.2)	4.4 (3.0)
Média Geral	5.0 (2.4)	5.3 (2.4)	5.7 (2.7)

6.2. Carga Cognitiva: SWAT

Os resultados do SWAT (Tabela 2) apresentaram um padrão menos consistente com o aumento da complexidade. A V2 obteve a menor média geral (1,9), enquanto V1 e V3 apresentaram médias de 2,1. O **Esforço Mental** foi maior na V3 (2,3), em comparação com V1 (2,2) e V2 (2,0). O teste de Friedman não identificou diferenças estatisticamente significativas entre as versões, indicando que o SWAT não detectou variações significativas de carga cognitiva no contexto e tamanho amostral deste estudo.

Tabela 2. Médias e desvios padrão do SWAT por versão

Dimensão	V1	V2	V3
Demanda de Tempo (1–3)	2.1 (0.9)	1.9 (0.6)	1.9 (0.9)
Esforço Mental (1–3)	2.2 (0.8)	2.0 (0.8)	2.3 (0.8)
Estresse (1–3)	2.1 (0.9)	1.9 (0.7)	2.2 (0.8)
Média Geral	2.1 (0.9)	1.9 (0.7)	2.1 (0.8)

6.3. Usabilidade e Utilidade Percebida

A Tabela 3 mostra redução progressiva da **facilidade de uso** conforme a complexidade aumentou, com médias de 4,4 (V1), 3,8 (V2) e 2,9 (V3). O mesmo padrão foi observado na percepção de que o dashboard **não é confuso**, cujas médias foram 4,2, 3,5 e 3,0, respectivamente.

Em contraste, a **utilidade percebida** foi maior nas versões V2 e V3 (ambas com $M = 3,8$) do que na V1 ($M = 3,4$). O teste de Friedman identificou diferença estatisticamente significativa para a **facilidade de uso** ($\chi^2(2) = 7.85, p = 0.020$), reforçando a vantagem da versão simples nesse aspecto.

Tabela 3. Médias e desvios padrão da usabilidade e utilidade percebida

Dimensão	V1	V2	V3
Fácil de usar (1–5)	4.4 (0.9)	3.8 (1.2)	2.9 (1.2)
Não é confuso (1–5)	4.2 (0.8)	3.5 (1.2)	3.0 (1.4)
Utilidade Percebida (1–5)	3.4 (1.3)	3.8 (1.0)	3.8 (1.3)
Média Geral	4.0 (1.0)	3.7 (1.1)	3.2 (1.3)

6.4. Análise Qualitativa das Sugestões

A análise das respostas abertas revelou três categorias principais. A primeira, mencionada por quatro professores, relacionou-se à **clareza e simplicidade visual**, incluindo sugestões de maior foco e organização das informações. A segunda, apontada



por dois participantes, referiu-se à **personalização**, com destaque para a possibilidade de visualizar o desempenho individual dos alunos. A terceira envolveu a **diversificação das visualizações**, com sugestões de gráficos de barras, pizza e cartões de KPI. Em conjunto, as respostas reforçam a demanda por interfaces claras, mas adaptáveis às diferentes necessidades de análise.

6.5. Preferência Geral dos Professores

A preferência geral é apresentada na Tabela 4. A V3 foi a mais escolhida, com 5 votos (38%), seguida pela V1, com 4 (31%), e pela V2, com 3 (23%). Um participante (8%) não escolheu nenhuma versão.

Tabela 4. Preferência geral dos professores por versão

Versão	Votos	Percentual
V1 – Dashboard Simples	4	31%
V2 – Dashboard Intermediário	3	23%
V3 – Dashboard Complexo	5	38%
Nenhuma	1	8%
Total	13	100%

Em conjunto, os resultados sugerem um possível **trade-off entre usabilidade e utilidade percebida**: a V1 apresentou maior facilidade de uso e menor demanda mental, enquanto V2 e V3 alcançaram maior utilidade percebida, e a V3 foi a versão mais escolhida pelos participantes.

7. Discussão

Os resultados indicam uma relação entre a complexidade dos dashboards e a carga cognitiva percebida. No NASA-TLX, a Demanda Mental aumentou de 4,6 na V1 para 6,3 na V3, com diferença estatisticamente significativa entre as versões ($p = 0.039$). Esse resultado é consistente com a Teoria da Carga Cognitiva, segundo a qual a quantidade e a forma de apresentação das informações podem elevar a demanda sobre a memória de trabalho (Madi; Peng e Rogers, 2022). O achado também converge com estudos que destacam a necessidade de equilibrar riqueza informacional e sobrecarga cognitiva em dashboards educacionais (Nguyen *et al.*, 2024).

7.1. Trade-off entre usabilidade e utilidade percebida

O principal achado foi o possível **trade-off entre usabilidade e utilidade percebida**. A V1 apresentou maior facilidade de uso ($M = 4,4$) e menor percepção de confusão ($M = 4,2$), enquanto V2 e V3 obtiveram maior utilidade percebida ($M = 3,8$). Além disso, a V3 foi a versão mais escolhida (38%), apesar da menor usabilidade e maior carga cognitiva. Esse resultado sugere que, para parte dos participantes, a riqueza informacional pode compensar o aumento da complexidade quando percebida como útil para a tomada de decisão pedagógica.

Esse achado é compatível com estudos que destacam a necessidade de equilibrar simplicidade e quantidade de informações em dashboards destinados a professores (Karademir *et al.*, 2024; Nguyen *et al.*, 2024). No contexto investigado, a maior densidade informacional da V3 pode ter sido percebida como vantajosa para o acompanhamento de múltiplos indicadores, embora tenha exigido maior esforço cognitivo.



7.2. Complementaridade entre NASA-TLX e SWAT

A aplicação conjunta do NASA-TLX e do SWAT revelou diferentes padrões de carga cognitiva. O NASA-TLX identificou diferença significativa na Demanda Mental, enquanto o SWAT não apresentou diferenças estatisticamente significativas entre as versões. Descritivamente, a V2 apresentou a menor média geral no SWAT ($M = 1,9$), enquanto a V3 apresentou maior média de Esforço Mental ($M = 2,3$). Esses resultados reforçam a utilidade da triangulação de instrumentos e indicam que diferentes medidas podem responder de maneira distinta às variações de complexidade (Rubio *et al.*, 2004; Reid e Nygren, 1988).

7.3. Implicações para o design de dashboards

Os resultados sugerem que dashboards educacionais não devem ser projetados com um único nível de complexidade. A preferência distribuída entre V1 (31%), V2 (23%) e V3 (38%) indica diferentes necessidades entre os usuários. Uma alternativa seria oferecer **níveis graduais ou customizáveis de visualização**, permitindo acesso inicial a indicadores essenciais e aprofundamento conforme a necessidade.

Esse resultado reforça a importância de envolver professores no processo de design (Karademir *et al.*, 2024; Nguyen *et al.*, 2024). No contexto da educação pública brasileira, iniciativas como a Rede de Inovação para Educação Híbrida (RIEH) (Ministério da Educação (MEC), 2025) podem se beneficiar de soluções que combinem simplicidade de uso e acesso progressivo a informações detalhadas. A implementação desses dashboards deve, portanto, ser acompanhada por ações de formação que desenvolvam a capacidade docente de interpretar e utilizar dados educacionais.

7.4. Ameaças à Validade e Limitações

As principais ameaças à validade foram analisadas considerando as dimensões de validade interna, externa, de construto e de conclusão estatística.

Validade interna. O tamanho reduzido da amostra ($N = 13$) limita o poder estatístico e aumenta a possibilidade de erros Tipo II. Para reduzir esse risco, utilizou-se o teste não paramétrico de Friedman, adequado à comparação de medidas repetidas. A avaliação das três versões pelos mesmos participantes e o uso de múltiplos instrumentos também contribuíram para a comparação dos resultados. Entretanto, como a ordem de exploração foi livre, não foi possível controlar completamente possíveis efeitos de ordem ou aprendizagem.

Validade externa. A amostra foi composta exclusivamente por professores da rede estadual de Alagoas que atuam no Ensino Médio, selecionados por conveniência. Assim, os resultados não podem ser generalizados diretamente para outros níveis de ensino, redes ou regiões. A descrição do contexto e do perfil dos participantes favorece a avaliação da transferibilidade dos achados para contextos semelhantes.

Validade de construto. Foram utilizados instrumentos consolidados na literatura, porém em versões reduzidas do NASA-TLX e da escala de usabilidade, o que pode limitar a abrangência dos construtos avaliados. Para minimizar essa ameaça, os instrumentos foram selecionados com base em estudos sobre avaliações sob restrições de tempo (Ghanbari; Tuunanen e Kazan, 2025; Baskerville; Pries-Heje e Venable, 2026). A coleta anônima e as orientações para que os participantes respondessem com base na experiência imediata também buscaram reduzir possíveis efeitos de desejabilidade social.

Validade de conclusão estatística. O baixo número de participantes limita o poder dos testes estatísticos. Por essa razão, os resultados foram interpretados de forma exploratória, distinguindo diferenças estatisticamente significativas de tendências



descritivas. Assim, a ausência de significância estatística no SWAT não é interpretada como evidência de inexistência de efeito.

Além dessas ameaças, destacam-se como limitações o uso de dados simulados, a avaliação pontual dos dashboards e a ausência de observação em situações reais de ensino. A participação voluntária também pode introduzir viés de autosseleção. Estudos futuros devem ampliar a amostra, incluir diferentes contextos educacionais e avaliar o uso longitudinal dos dashboards com dados reais.

8. Considerações Finais

Este estudo avaliou três versões de um dashboard educacional com diferentes níveis de complexidade, envolvendo 13 professores da rede estadual de Alagoas. Os resultados indicaram tendência de aumento da carga cognitiva com a complexidade, especialmente na dimensão Demanda Mental do NASA-TLX. Em contrapartida, V2 e V3 apresentaram maior utilidade percebida, enquanto V1 obteve melhor facilidade de uso. A V3 foi a versão mais escolhida (38%), sugerindo um possível **trade-off entre usabilidade e utilidade percebida**.

8.1. Contribuições

As principais contribuições deste estudo são: (1) evidências empíricas sobre a percepção de professores da rede estadual de Alagoas, contribuindo para uma área ainda pouco explorada na educação básica brasileira (Lemes; Dias e Oliveira, 2023); (2) triangulação metodológica por meio da combinação de NASA-TLX, SWAT e medidas de usabilidade (Rubio *et al.*, 2004); e (3) evidências de que maior riqueza informacional pode coexistir com maior carga cognitiva e menor facilidade de uso, destacando a necessidade de equilíbrio no design de dashboards educacionais.

8.2. Implicações Práticas

Os resultados sugerem o uso de **níveis graduais ou customizáveis de complexidade**, permitindo acesso progressivo às informações conforme a necessidade do professor. Recomenda-se priorizar funcionalidades relacionadas à tomada de decisão pedagógica e envolver os usuários no processo de design. A V2 pode representar um ponto inicial de equilíbrio entre riqueza informacional e carga cognitiva, embora essa hipótese deva ser investigada em amostras maiores.

8.3. Trabalhos Futuros

Como trabalhos futuros, pretende-se: (1) ampliar a amostra para diferentes regiões e níveis de ensino; (2) realizar estudos longitudinais sobre o efeito da familiarização; (3) testar os dashboards em contextos reais, com dados reais de alunos; (4) investigar o uso de Large Language Models (LLMs) para geração de feedbacks personalizados; e (5) avaliar possíveis impactos dos dashboards sobre o acompanhamento e o desempenho dos estudantes (Laet *et al.*, 2020).

Em síntese, os resultados reforçam a necessidade de equilibrar simplicidade, carga cognitiva e riqueza informacional no design de dashboards educacionais. A preferência pela versão mais complexa, apesar da menor usabilidade, sugere que a utilidade percebida pode desempenhar papel relevante na aceitação dessas ferramentas. Para a Informática na Educação, o estudo destaca a importância de envolver professores no design e de utilizar múltiplas perspectivas de avaliação para compreender sua experiência com tecnologias de apoio à decisão pedagógica.



Referências

- Baskerville, R.; Pries-Heje, J.; Venable, J. R. Meds: Methodology for evaluation in design science. **European Journal of Information Systems**, Taylor & Francis, p. 1–18, 2026. Disponível em: <<https://doi.org/10.1080/0960085X.2026.2627280>>.
- Brooke, J. Sus: A quick and dirty usability scale. **Usability Evaluation in Industry**, v. 189, n. 194, p. 4–7, 1996.
- Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. **MIS Quarterly**, v. 13, n. 3, p. 319–340, 1989.
- Ghanbari, H.; Tuunanen, T.; Kazan, E. **Free Simulation Experiments and Concept Testing in Design Science Research**. 2025. SSRN Working Paper.
- Karademir, O. *et al.* I don't have time! but keep me in the loop: Co-designing requirements for a learning analytics cockpit with teachers. **Journal of Computer Assisted Learning**, v. 40, n. 6, p. 2681–2699, 2024. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/jcal.12997>>.
- Laet, T. D. *et al.* Adoption and impact of a learning analytics dashboard supporting the advisor—student dialogue in a higher education institute in latin america. **British Journal of Educational Technology**, v. 51, n. 4, p. 1002–1018, 2020. Disponível em: <<https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1111/bjet.12962>>.
- Lee-Cultura, S.; Sharma, K.; Giannakos, M. N. Multimodal teacher dashboards: Challenges and opportunities of enhancing teacher insights through a case study. **IEEE Trans. Learn. Technol.**, IEEE Computer Society Press, Washington, DC, USA, v. 17, p. 181–201, jan. 2024. ISSN 1939-1382. Disponível em: <<https://doi.org/10.1109/TLT.2023.3276848>>.
- Lemes, T. d. C.; Dias, M. O. d. S.; Oliveira, T. Análise do uso de dashboard como ferramenta de apoio a tomada de decisão em instituições de ensino: uma revisão sistemática da literatura. **RENOTE**, v. 21, n. 1, p. 280–290, 2023.
- Madi, N. A.; Peng, S.; Rogers, T. Assessing workload perception in introductory computer science projects using nasa-tlx. In: **Proceedings of the 53rd ACM Technical Symposium on Computer Science Education (SIGCSE 2022)**. [S.l.: s.n.], 2022. p. 1–7. Ministério da Educação (MEC). **Educação Híbrida: Conceito e Orientações Pedagógicas**. Maceió, 2025. Publicação da Rede de Inovação para Educação Híbrida (RIEH).
- Nguyen, N. B. C.; Lithander, M.; Östlund, C. M.; Karunaratne, T.; Jobe, W. Teadash: Implementing and evaluating a teacher-facing dashboard using design science research. In: MDPI. **Informatics**. [S.l.], 2024. v. 11, n. 3, p. 61.
- Nielsen, J. **Usability Engineering**. [S.l.]: Academic Press, 1993.
- Reid, G. B.; Nygren, T. E. The subjective workload assessment technique: A scaling procedure for measuring mental workload. **Advances in Psychology**, v. 52, p. 185–218, 1988.
- Rubio, S.; Díaz, E.; Martín, J.; Puente, J. M. Evaluation of subjective mental workload: A comparison of swat, nasa-tlx, and workload profile methods. **Applied Psychology: An International Review**, v. 53, n. 1, p. 61–86, 2004.
- Sweller, J. Cognitive load during problem solving: Effects on learning. **Cognitive Science**, v. 12, n. 2, p. 257–285, 1988.
- Venable, J.; Pries-Heje, J.; Baskerville, R. Feds: a framework for evaluation in design science research. **European Journal of Information Systems**, v. 25, n. 1, p. 77–89, 2016.