

UniSenac RS promovendo o uso de Tecnologias Digitais para Avaliação Colaborativa na Formação Interdisciplinar em MBAs

Marcelo Vianna Batista ¹

Jozilda Berenice Candido Fogaça ²

Cláudia Cecília Serafini Mallmann ³

RESUMO

Este trabalho apresenta como o uso de tecnologias digitais permite uma avaliação aprofundada do desempenho de 48 alunos de 4 MBAs do UniSenac RS. Realizada ao final de cada módulo do Projeto Interdisciplinar por meio de formulários do Google Forms, a Avaliação Por Pares Unisenac RS solicita aos alunos que analisem seu desempenho e o dos colegas utilizando uma escala de cinco conceitos (ruim, médio, bom, ótimo e excelente) em 10 questões relativas à projeção para inovação (i.e.: “eficácia dos métodos e instrumentos utilizados nos sprints”), além de 1 pergunta qualitativa sobre sua visão geral de desempenho. Os dados, cruzados entre os membros do mesmo grupo e tratados na ferramenta NotebookLM, possibilita a geração de feedbacks robustos com aproximadamente 500 palavras para cada aluno. Os resultados demonstram que a estratégia avaliativa promove a responsabilidade individual e o reconhecimento do esforço coletivo nos grupos, permitindo uma coleta estruturada de informações e uma avaliação transparente e abrangente. Essa abordagem confere ao docente um olhar endógeno que não seria alcançado por outras ferramentas avaliativas. Ademais, a integração de tecnologias digitais com avaliação colaborativa aprimora a qualidade da avaliação e prepara os alunos para os desafios de um mercado de trabalho mais digital e interconectado. A experiência do UniSenac ilustra como as instituições de ensino podem inovar na formação de profissionais adaptáveis e criativos, explorando a interação entre tecnologia e pedagogia para fomentar inovação na educação superior.

Palavras-chave: ferramentas digitais para educação, avaliação por pares, design estratégico, aprendizagem baseada em projetos, produto educacional.

INTRODUÇÃO

A avaliação em cursos profissionais de pós-graduação carrega uma tensão conhecida: como captar, de modo válido e útil, contribuições individuais que se expressam em tarefas coletivas, típicas de ambientes de inovação e de projetos interdisciplinares? A literatura tem reiterado que avaliação molda aprendizagem de forma

¹ Doutor em Design (área de concentração: Design Estratégico) pela Universidade do Vale do Rio dos Sinos (UNISINOS)-RS, marcelovib@gmail.com

² Mestre em Inclusão e Acessibilidade Social pela Universidade Federação dos Estabelecimentos de Ensino Superior em Novo Hamburgo (FEEVALE)-RS, jbfgaca@senacrs.com.br

³ Doutoranda em Administração pela Universidade do Vale do Rio dos Sinos (UNISINOS)-RS, ccmallmann@senacrs.com.br



mais intensa que o ensino em si, motivo pelo qual deve ser concebida como parte constitutiva do processo formativo, e não como etapa terminal de aferição de “produto” (Boud & Falchikov, 2007). Nesse horizonte, avaliação por pares emerge como estratégia que articula responsabilização, transparência e desenvolvimento do julgamento profissional: quando estudantes apreciam a qualidade do trabalho próprio e alheio, eles internalizam critérios, refinam padrões de comparação e ampliam sua autonomia (Topping, 2009; Falchikov & Goldfinch, 2000; Tai et al., 2018).

Ao mesmo tempo, a efetividade da avaliação formativa depende de feedbacks que sejam tempestivos, acionáveis e alinhados a metas de aprendizagem. Dois marcos conceituais sustentam essa proposição: o modelo de sete princípios de Nicol & Macfarlane-Dick (2006), que enfatiza a autorregulação do estudante, e a síntese de Hattie & Timperley (2007) sobre o impacto do feedback (“para onde vou”, “como estou indo”, “qual o próximo passo”). Em programas voltados a gestão, design e inovação, onde a aprendizagem é orientada a problemas (PBL), faz sentido acoplar métricas de contribuição entre pares, comentários qualitativos ricos e devolutivas personalizadas - incorporando, quando apropriado, affordances digitais para análise de dados e síntese textual.

A avaliação por pares (peer assessment) ocupa, há pelo menos três décadas, um lugar de destaque em propostas de reconfiguração da avaliação no ensino superior. Longe de ser mera técnica de distribuição de notas, ela é compreendida como **mecanismo de socialização dos critérios de qualidade** e de desenvolvimento de um **julgamento avaliativo** (evaluative judgement) capaz de sustentar práticas profissionais em contextos complexos e colaborativos (Boud; Falchikov, 2007; Boud; Soler, 2016; Tai et al., 2018) [Não verificado]. Em ambientes orientados à inovação - como MBAs que combinam gestão, design e tecnologia, o trabalho é predominantemente coletivo, distribuído e interdependente. Nesses contextos, **aferir contribuição individual** por instrumentos exclusivamente docentes incorre no risco de invisibilizar o papel de cada membro e de reduzir a avaliação a marcos finais de entrega (produto), em detrimento dos **processos** (interações, decisões, trade-offs).

Este artigo examina um desenho avaliativo implementado no Centro Universitário UniSenac-RS em 2/3 de um ciclo formativo de quatro MBAs, com **seis grupos interdisciplinares** formados por alunos de cada MBA em uma Unidade Curricular intitulada **Projeto Interdisciplinar**, entre os meses de abril e dezembro de 2024. O arranjo do desenho avaliativo por pares em questão combina: (i) **instrumento**



estruturado em escala Likert de 5 pontos com dez itens atrelados a dimensões-chave de trabalho em projeto; (ii) pergunta aberta de apreciação; (iii) coleta via **Google Forms** (com consentimento/LGPD explicitados no próprio formulário); (iv) tratamento dos dados em **planilha** e, por fim, (v) **feedback individual (~500 palavras)** gerado a partir de prompts padronizados em LLM, sempre com revisão docente para garantir adequação pedagógica e guardrails éticos.

Do ponto de vista pedagógico, o estudo ancora-se em três pilares. Primeiro, o alinhamento construtivo (Biggs & Tang, 2011) que orienta a coerência entre objetivos, tarefas autênticas e avaliação — aqui, operacionalizado no nexos entre desafios de inovação, colaboração interdisciplinar e critérios de desempenho. Segundo a aprendizagem experiencial (Kolb, 1984), segundo a qual ciclos de experiência–reflexão–conceituação–experimentação são catalisados por feedbacks significativos e por práticas de coavaliação em equipe. Terceiro, a pedagogia ancorada na Aprendizagem Baseada em Problemas (PBL), imputando a proporá a avaliação para aprendizagem explicitando critérios transparentes, feedbacks acionáveis e oportunidades de re-entrega elevam a autorregulação e a qualidade final (Nicol & Macfarlane-Dick, 2006). Avaliação por pares, quando ancorada em rubricas e com treino de avaliadores, produz ganhos duplos: (i) amplia repertório de julgamento dos estudantes e (ii) distribui o trabalho avaliativo sem colapsar a qualidade (Falchikov, 2005; Topping, 1998). Dois riscos clássicos - *leniency/severity* e *friendship bias* - são mitigados por calibração (exemplares-âncora), dupla leitura e amostragem pelo docente, em uma avaliação formativa contínua e rubricas explícitas durante o desenvolvimento do projeto, e não apenas no momento do “pitch”.

Se a avaliação por pares socializa critérios, o **feedback** abre caminho para a **ação regulatória** do estudante. Dois marcos permanecem incontornáveis. Primeiro, os **sete princípios** de bom feedback para aprendizagem autorregulada (Nicol; Macfarlane-Dick, 2006), que incluem esclarecer critérios, cultivar diálogo, promover autovalidação e oferecer informação utilizável. Segundo, a síntese de Hattie e Timperley (2007) sobre as três perguntas estruturantes do feedback eficaz: **para onde vou (metas e critérios)**, **como estou indo (desempenho atual)** e **qual o próximo passo (feed-forward)**. Em cursos profissionais, feedback útil é aquele **acionável**, sensível ao tempo do projeto e capaz de indicar **próximas decisões** (o que refatorar, o que validar, o que abandonar).

O desenho do sistema avaliativo por pares privilegia (i) análise intra-aluno (médias por item e variações entre avaliadores), (ii) análise intra-equipe (dispersão e



padrões de concordância) e (iii) apreciações qualitativas oriundas dos campos abertos e das devolutivas personalizadas. No quadro abaixo apresentamos os instrumentos desenvolvidos.

Quadro 1 – síntese do sistema avaliativo por pares

Instrumentos para coleta O instrumento foi implementado em Google Forms, com uma seção de itens estruturados em escala Likert de 5 pontos e uma questão aberta de apreciação global. A escala operou com âncoras consistentes (1=desempenho inicial/insatisfatório; 5=desempenho excelente/sistêmico). O conjunto de 10 itens cobre dimensões críticas do trabalho em projeto interdisciplinar, com foco tanto em produto (entregas, evidências) quanto em processo (colaboração, comunicação, gestão de tempo, utilização de métodos). A questão aberta solicita comentários justificados sobre contribuições, pontos fortes e <i>feed-forward</i> (o que priorizar no próximo ciclo).
Evidências de validade de conteúdo (procedimento) A validade de conteúdo foi endereçada por alinhamento construtivo entre objetivos da UC, rubricas de projeto e itens do instrumento; complementa-se por perícia docente (ajustes semânticos dos itens) e convergência com referenciais da literatura (avaliação formativa, PBL, design/gestão da inovação). Como limitação, não foram realizados <i>cognitive interviews</i> ou <i>pilot testing</i> formal com amostra externa; esse aspecto é indicado na Agenda de pesquisa.
Procedimentos de aplicação • Janelas: duas aplicações ao final de módulo em 2024 (junho e dezembro), imediatamente após as entregas e apresentações, para maximizar recall e validade consequential. • Logística: os estudantes preencheram o formulário online em ambiente supervisionado/remoto, com orientação prévia sobre critérios e propósito formativo. O tempo de resposta típico variou entre 6 e 15 minutos, dependendo do volume de justificativas textuais. • Anonimato e clima: o desenho privilegiou anonimização do respondente perante o avaliado, preservando, no entanto, a rastreabilidade pedagógica no nível do docente/coordenador para fins de auditoria e suporte (mitigação de reciprocidade/halo é discutida em “Qualidade e vieses”). • Encerramento: cada rodada foi encerrada após atingir taxa de resposta suficiente por equipe (reportada em Resultados), precedendo o processamento e a geração de devolutivas.
Pipeline de processamento e geração de feedbacks 1. Exportação e preparação ○ Exportação do Google Forms para planilha Excel. ○ Checagens: inconsistências (entradas duplicadas), faltas (itens em branco), outliers (valores fora da escala), conformidade de <i>timestamps</i>. ○ Padronização de identificadores (pseudônimos/IDs por aluno), mapeamento de pares por equipe e normalização dos rótulos de itens. 2. Estatística descritiva (por aluno e por equipe) ○ Médias e medianas por item; dispersão (DP, amplitude). ○ Índice de contribuição relativa (ICR): média individual – média da equipe (por item e geral), para sinalizar aportes acima/abaixo da linha de base do time. ○ Distribuição de concordância entre avaliadores (boxplots internos, <i>heatmaps</i> por item/equipe — relatados em texto/tablas na seção Resultados). 3. Confiabilidade e estrutura ○ α de Cronbach (por aplicação; por conjunto de itens), com cautelas interpretativas (dependência do nº de itens e tau-equivalência). ○ ω de McDonald (quando amostra permitir) como estimativa alternativa.



- **Correlação item-total corrigida** e inspeção de **itens problemáticos** (padrões de piso/teto; baixas correlações).
 - Quando adequado, **análise fatorial exploratória** (AFE) para sondar dimensionalidade (pré-requisitos: KMO e Bartlett; critérios *eigenvalue* e paralelos).
4. **Análise qualitativa assistida**
- **Base:** comentários abertos dos pares e **feedbacks individualizados** gerados com apoio de LLM (documento “aluno a aluno”).
 - **Estratégia: análise temática** (abordagem híbrida *dedutivo-indutiva*), com **códigos-semente** alinhados às dimensões do instrumento (investigação, projeção, entrega, colaboração, comunicação, gestão do tempo, liderança) e abertura a **códigos emergentes** (p.ex., “negociação de escopo”, “padrões de confiabilidade”, “gestão de risco”).
 - **Unidade de análise:** segmento textual com sentido completo (1–3 frases).
 - **Rigor:** auditoria de trilhas (memórias de codificação, exemplos-âncora) e **contrachecagem** pontual pelo docente. *Limitação:* não houve dupla codificação independente; o procedimento é reconhecido e sinalizado na Discussão.
5. **Síntese de feedbacks individualizados**
- **Entradas:** estatísticas por item, ICR, comentários abertos dos pares (filtrados por relevância e recorrência).
 - **Geração: prompt padronizado** (Apêndice A) solicita síntese em ~500 palavras com estrutura triádica (*onde estou / evidências / próximos passos*), linguagem **profissional** e **não punitiva**, proibições explícitas a extrapolações sem dado, e **verificação cruzada** com os números.
 - **Revisão humana:** verificação do docente para checar **fidelidade** aos dados, **clareza**, **ação recomendada** e **coerência** com a rubrica da UC. Ajustes finais são registrados.

Plano de análise (mapeado às questões de pesquisa)

- **RQ1 (visibilidade e acionabilidade):** descrever **ganhos de observabilidade** ao comparar métricas de pares com julgamentos docentes; examinar **ICR** e casos de **dispersão atípica** que suscitam intervenção pedagógica. Evidências: tabelas por aluno/equipe; excertos de devolutivas (anonimizados).
- **RQ2 (padrões descritivos):** reportar **tendências centrais** e **dispersões** por item/equipe; identificar itens **mais/menos discriminativos**; comparar **junho vs. dezembro** em medidas agregadas.
- **RQ3 (consistência interna):** estimar α e ω por rodada; discutir implicações para interpretação e uso; sugerir **revisões de itens** quando necessário.
- **RQ4 (temas qualitativos):** apresentar **temas recorrentes** nos comentários/feedbacks (p.ex., colaboração, comunicação de decisões, uso de evidências, ritmo de sprints) e conectá-los a princípios de **bom feedback** (feed-up / feedback / feed-forward).
- **RQ5 (implicações institucionais):** derivar recomendações para **governança, escalabilidade e políticas de uso de IA** (padrões de prompt, logs, formação docente, diretrizes de anonimato).

Qualidade dos dados e mitigação de vieses

- **Rubrica e treino:** explicitação de critérios e exemplos-âncora na instrução do formulário; orientação para **justificar** avaliações extremas.
- **Vieses previstos:** **generosidade/severidade, reciprocidade, efeito halo**; mitigação por anonimização perante colegas, e leitura pelo docente com foco em **padrões** (não em casos isolados).
- **Dependência intragrupo:** relatórios trazem estatísticas **por equipe** e alertas quando a **dispersão intra-aluno** excede limiares (sinalizando baixa concordância).
- **Dados faltantes:** *listwise deletion* para análises que exijam vetores completos; *pairwise* quando aplicável. Casos com baixa base avaliativa recebem **nota metodológica** na devolutiva.

- **Transparência:** registros internos de versões do instrumento, datas de aplicação, prompts e revisões

Fonte: os autores

O restante do artigo organiza-se da seguinte forma: detalhamos o **método** (contexto, amostra, instrumento, procedimentos, análise e considerações éticas/LGPD); depois, **resultados** quantitativos e qualitativos; em seguida, **discussão** (conexões com a literatura, implicações pedagógicas e gerenciais); por fim, **conclusões, limitações e agenda** para pesquisas futuras.

METODOLOGIA

Estudo de caso explanatório, com métodos mistos (descritivo-quantitativo e análise qualitativa de conteúdo), voltado a examinar um arranjo de avaliação por pares acoplado a feedbacks individualizados em projetos interdisciplinares de quatro MBAs do UniSenac-RS. A avaliação por pares ocorre ao final dos módulos (janela de junho e dezembro de 2024), compondo o ecossistema avaliativo da UC e informando decisões pedagógicas (realocação de papéis, reforço de métodos, ajuste de escopo). Participaram os **discentes regularmente matriculados** nas quatro turmas de MBA que integraram as seis equipes do Projeto Interdisciplinar (*amostra por conveniência*, correspondente ao universo das equipes em atividade no período). Os **pares avaliadores** de cada estudante são os **colegas de equipe** do respectivo ciclo. A participação e o tratamento de dados seguiram as condições explicitadas no instrumento (Ética e LGPD).

RESULTADOS E DISCUSSÃO

Os resultados referem-se à aplicação correspondente ao **Módulo 1 – 2024/1**. A base contém **53 estudantes** distribuídos em **6 equipes interdisciplinares** (identificadas como Grupos 1, 2, 4, 5, 6 e 7). Para cada estudante, foram agregadas as **médias por item** (10 itens em escala Likert de 1 a 5) provenientes das avaliações de seus pares na respectiva equipe. Os resultados são apresentados na granularidade **por estudante, por equipe e por item**.

A **média geral** (média dos 10 itens para cada estudante, depois agregada sobre a amostra) foi 4,52 (DP=0,51; mediana=4,66; min=2,30; máx=5,00; N=53). Em termos práticos, há desempenho elevado na maior parte da amostra e cauda inferior curta (poucos casos com escores substancialmente abaixo da linha de base do time). Para qualificar



“contribuição relativa”, calculou-se o ICR (*Índice de Contribuição Relativa* = média do estudante – média da sua equipe). A distribuição do ICR é, por construção, centrada em zero, com $P5 \approx -0,87$, $P50 \approx +0,12$, $P95 \approx +0,46$ e mínimo = $-2,24$ / máximo = $+0,51$. Em suma: poucos outliers negativos puxam a cauda inferior; a parte superior apresenta ganhos relativos moderados ($+0,3$ a $+0,5$). Estimou-se a consistência interna (sobre as médias por item de cada estudante) α de Cronbach (global, 10 itens) = 0,97 e α por equipe: G1=0,995; G2=0,869; G4=0,953; G5=0,991; G6=0,984; G7=0,833.

A **magnitude elevada** do α é compatível com o número de itens e a **alta redundância conceitual** esperada em dimensões de participação/colaboração. Duas ressalvas metodológicas: (i) o α foi calculado sobre médias agregadas por estudante (reduzindo ruído intra-avaliador); (ii) α elevado não garante unidimensionalidade. Ainda assim, os valores sugerem coerência interna adequada para uso formativo; recomendamos, na sequência do projeto, inspeção de correlações item-total e, quando a razão casos/itens permitir, AFE como análise de robustez.

Os Padrões por item (toda a amostra), revela enquanto leituras principais: (i) todas as médias situam-se acima de 4,4, com tendência a “teto” compatível com coortes adultas e com observabilidade predominantemente processual; (ii) itens com maior discriminação ($DP \geq 0,57$) foram “respeito” (0,66), “colaboração equilibrada” (0,65) e “ideias/provocações” (0,57), úteis para diferenciar perfis em equipes heterogêneas; (iii) os menores escores relativos concentram-se nos artefatos finais (“documentações/apresentações”) e produção de evidências escritas, sinalizando pontos de cuidado para o próximo ciclo (ver §7 – Feed-forward).

Já **padrões por equipe, com Média (DP)** da média geral por estudante em cada equipe, revelaram como leituras principais: (i) G4 apresenta o maior patamar médio e baixa dispersão — coerência de desempenho intra-equipe; (ii) G5 exibe média mais baixa e dispersão maior, com pelo menos um outlier negativo (média individual $\sim 2,9$), o que puxa o grupo para baixo e sugere desequilíbrio de contribuição; (iii) G1 também mostra dispersão elevada ($DP \approx 0,86$), explicada por um caso extremo ($\approx 2,3$), ao lado de desempenhos muito altos.

A **curva do ICR** corrobora a leitura de **núcleos coesos** com **poucas caudas negativas**:

- **Outliers inferiores:** 1 caso com $ICR \approx -2,24$ (equipe G1) e 1 caso com $ICR \approx -1,45$ (G5).



- **Segmento superior:** 5 casos entre +0,43 e +0,51 (vários em G5), sugerindo **lideranças de contribuição** que sustentam o desempenho da equipe.

Para fins de relato público, os nomes foram **anonimizados**. Na devolutiva pedagógica (documento “aluno a aluno”), a combinação de **ICR, itens críticos e comentários** já subsidia intervenções caso a caso (realocação de papéis, redistribuição de tarefas, *coaching* pontual).

A **análise temática** (sobre os comentários abertos e as devolutivas individualizadas) aponta regularidades coerentes com os padrões numéricos. Em termos positivos, emergem com alta frequência: qualidade/robustez das entregas, participação ativa, respeito às pessoas e às perspectivas, uso de dados/evidências, liderança situacional e dinâmica de *sprints*. Como pontos de atenção recorrentes: documentação final e materiais de apresentação (lacunas de padronização e acabamento), comunicação assíncrona e gestão de prazos (atritos em momentos de *handover*), além de assimetrias pontuais de colaboração equilibrada.

Em termos de *feed-forward* (próximos passos) propostos nas devolutivas:

1. instituir checklists de entrega para as peças finais (documentações e *pitches*), de modo a atacar as variabilidades de padrão;
2. explicitar papéis de comunicação e ritmos de sincronização (e.g., *cadences* semanais e *handoff templates*), reduzindo atritos de prazo;
3. reforçar o uso de exemplares-âncora (*exemplars*) e rubricas para equalizar expectativas entre equipes, sobretudo em “ideias/provocações” e “colaboração equilibrada”.

Por fim, a síntese interpretativa (i) A elevação geral dos escores coaduna-se com o perfil de pós-graduação profissional e com a cultura de projetos implementada na UC; (ii) a discriminação útil em itens de colaboração/ideias sugere que o instrumento capta variação significativa onde interessa para gestão de equipes; (iii) a consistência interna é alta (com ressalvas usuais), sustentando o uso formativo; (iv) equipes com caudas negativas (G1, G5) demandam intervenção diretiva de coordenação (redistribuição de tarefas, *coaching* focal e acompanhamento de *cadence*); (v) a camada de feedback individualizado está alinhada aos achados e já oferece planos de ação concretos no nível micro (estudante) e meso (equipe).

CONSIDERAÇÕES FINAIS



A contribuição do estudo é dupla. Teoricamente, propõe um quadro integrador que conecta avaliação por pares, feedback formativo e desenvolvimento de julgamento avaliativo em contextos de gestão/design/inação, com atenção a requisitos de proteção de dados e a riscos/benefícios do apoio por GenAI.

A literatura identifica efeitos pedagógicos relativamente estáveis da avaliação por pares: (i) maior transparência sobre critérios e padrões; (ii) engajamento cognitivo de ordem superior (comparação, calibração e justificação de juízos); (iii) desenvolvimento de metacognição e autorregulação; e (iv) ganho de accountability e responsabilização nas equipes (Topping, 2009; Falchikov; Goldfinch, 2000) [Não verificado]. Em meta-análises e revisões, observa-se convergência moderada entre julgamentos de pares e de docentes quando há rubricas explícitas, treino e feedback estruturado, o que sustenta sua validade como componente formativo e somativo, desde que o desenho controle vieses previsíveis (reciprocidade, conivência, halo etc.) [Não verificado].

Em síntese, a avaliação por pares deve ser tratada como prática epistêmica: ela torna público o repertório de critérios, promove negociação de significados e distribui a competência de avaliar — condição para a aprendizagem autodirigida ao longo da vida. Em MBAs voltados a problemas reais, essa prática espelha o modo como organizações avaliam projetos: por pares, comitês, clientes e stakeholders.

Praticamente, descreve um pipeline replicável (instrumento + coleta + análise + devolutivas), oferece artefatos (itens, rubricas, prompts, fluxos), reporta resultados descritivos/qualitativos e indica decisões operacionais (anonimato, devolutivas, governança) que podem orientar outras IES.

No plano do desenvolvimento profissional, defendemos que a prática recorrente de avaliação por pares exerce uma função decisiva: cultivar o “evaluative judgement” — a capacidade de decidir sobre qualidade do próprio trabalho e do trabalho alheio, competência-chave para atuação em ambientes complexos e de alto acoplamento entre áreas (Tai et al., 2018; Boud & Soler, 2016). Essa competência dialoga com perfis em T (profissionais com profundidade numa área e abertura transversal para colaborar), virtude associada a práticas de design estratégico e inovação (Brown/IDEO).

Há, contudo, desafios metodológicos e institucionais a endereçar. Em termos psicométricos, é desejável estimar consistência interna do instrumento (e.g., α de Cronbach) e explicitar limites dessa métrica, frequentemente super-interpretada na pesquisa educacional; a literatura recomenda cautela na leitura do α e, quando possível, análises complementares (p.ex., estrutura fatorial exploratória).



Neste sentido a adoção institucional robusta exige tratamento explícito de confiabilidade e validade. Em instrumentos de múltiplos itens (escala Likert), costuma-se estimar consistência interna (e.g., α de Cronbach), com a ressalva — amplamente discutida — de que α depende do número de itens e de pressupostos de tau-equivalência; valores altos não garantem unidimensionalidade, e valores moderados podem ser aceitáveis em medidas formativas ou construtos amplos [Não verificado]. Alternativas e complementos incluem ω de McDonald, análise fatorial exploratória/confirmatória e inspeção de correlações item-total.

No plano da validade, é importante evidenciar: (i) validade de conteúdo (alinhamento dos itens aos objetivos do Projeto Interdisciplinar); (ii) validade de processo (transparência de critérios; treino dos avaliadores); (iii) validade consequential (efeitos do uso da avaliação nas decisões pedagógicas e na aprendizagem).

Em termos de implementação, anonimato e clima de segurança influenciam a sinceridade das avaliações; revisões e estudos recentes discutem efeitos e trade-offs de anonimização em arranjos formativos, recomendando decisões situadas e transparência de critérios. Em termos regulatórios, o tratamento de dados pessoais é regido no Brasil pela LGPD (Lei 13.709/2018), que exige base legal, finalidade determinada e proteção por desenho, especialmente quando há uso de tecnologias digitais. O tratamento de dados pessoais no contexto educacional é regido, no Brasil, pela Lei nº 13.709/2018 (LGPD), que estabelece princípios de finalidade, adequação, necessidade, transparência e segurança (BRASIL, 2018). Em avaliações por pares, isso implica: (i) base legal explícita (em geral, execução de política educacional/contratual; eventualmente consentimento informado quando adequado); (ii) minimização de dados identificáveis; (iii) comunicação clara a estudantes sobre usos pedagógicos; e (iv) governança para retenção, acesso e descarte.

O caso analisado incorpora uma camada tecnológica adicional: o uso de modelos de linguagem para apoiar a síntese dos feedbacks com base em dados do instrumento e em comentários de pares. A camada adicional de IA generativa (LLMs) na síntese de feedbacks demanda cuidados redobrados: mediação humana (o docente é responsável pelo conteúdo final), registro de propósito e fontes das entradas (itens, médias, comentários), mitigação de vieses e prevenção de riscos de alucinação e inconsistência. Em termos operacionais, boas práticas incluem: (a) prompts padronizados e auditáveis; (b) logs de versões; (c) checagem cruzada com dados originais; (d) linguagem de feedback não punitiva, específica e orientada à ação; (e) revisão docente obrigatória.



À luz de recomendações internacionais (UNESCO, Jisc/QAA), adotamos como princípio a mediação humana e a prestação de contas sobre o uso de GenAI (propósito, fontes, limites), cuidando de privacidade, integridade acadêmica e mitigação de vieses.

AGRADECIMENTOS

Ao Centro Universitário UniSenac RS, à Gerência de Educação Profissional do SENAC-RS e à FECOMERCIO-RS por acreditar na importância da inovação para que, de fato, a educação transforme vidas, expresso em seu Projeto Político Pedagógico e o Plano de Desenvolvimento Institucional estabelecido para o período 2025-2030 – documentos sigilosos, mas tornados públicos parcialmente nos OKRs no Plano Estratégico do Sistema FECOMÉRCIO-RS (2025)

REFERÊNCIAS

BIGGS, J.; TANG, C. *Teaching for Quality Learning at University*. 4. ed. Maidenhead: McGraw-Hill Education/SRHE, 2011.

BOUD, D.; FALCHIKOV, N. (Orgs.). *Rethinking Assessment in Higher Education: Learning for the Longer Term*. London: Routledge, 2007. DOI: 10.4324/9780203964309.

BOUD, D.; SOLER, R. Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, v. 41, n. 3, p. 400-413, 2016. DOI: 10.1080/02602938.2015.1018133.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. *Lei Geral de Proteção de Dados Pessoais (LGPD)*. Brasília, DF: Presidência da República. Acesso em: 21 out. 2025. [Planalto+1](#)

CRONBACH, L. J. Coefficient alpha and the internal structure of tests. *Psychometrika*, v. 16, p. 297-334, 1951. DOI: 10.1007/BF02310555.

DEN OUDEN, E. *Innovation Design: Creating Value for People, Organizations and Society*. London: Springer, 2012. DOI: 10.1007/978-1-4471-2268-5.

FALCHIKOV, N.; GOLDFINCH, J. Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research*, v. 70, n. 3, p. 287-322, 2000. DOI: 10.3102/00346543070003287

FECOMÉRCIO-RS. Plano Estratégico 2025. Porto Alegre, 2025. Disponível em: <https://api.senacrs.com.br/bff/site-fecomercio/v1/file/31e4daaccd7c29618c9327279dc7a522b28479.pdf>. Acesso em: 17 out. 2025.



VAN ZUNDERT, M.; SLUIJSMANS, D.; VAN MERRIËNBOER, J. J. G. Effective peer assessment processes: research findings and future directions. *Learning and Instruction*, v. 20, n. 4, p. 270-279, 2010. DOI: 10.1016/j.learninstruc.2009.08.004

