

PLANO DE ANÁLISE ESTATÍSTICA (PAE)

Pré-registro analítico · Edital Itaú/Lemann 2026 · UEM/PROFCIAMB

Coordenadora científica responsável: Dra. Leilane T. F. Schwind (UEM/PROFCIAMB).

Pré-registro público: Open Science Framework (OSF) · DOI a ser emitido na Fase 1 (M6).

1. Pergunta de pesquisa

As práticas pedagógicas vivenciais (experimentos, feiras de ciências, saídas de campo, projetos integradores e clubes investigativos) influenciam significativamente o engajamento escolar, o pertencimento e a percepção de transformação dos estudantes dos Anos Finais do Ensino Fundamental em escolas estaduais de tempo integral do Paraná, e essa influência atua como fator de equidade educacional para sub-populações em vulnerabilidade?

2. Hipóteses operacionais

- H1 — A frequência de práticas vivenciais (variável independente) está positivamente associada ao escore de engajamento escolar ($\beta > 0$; $p < 0,05$).
- H2 — A frequência de práticas vivenciais está positivamente associada ao escore de pertencimento ($\beta > 0$; $p < 0,05$), controlando por raça/cor, gênero, NSE e PCD.
- H3 — Estudantes dos 8º e 9º anos apresentam menor engajamento médio que estudantes dos 6º e 7º anos (efeito principal do ano), mas a diferença é atenuada quando há frequência elevada de práticas vivenciais (interação ano \times práticas significativa).
- H4 — O efeito das práticas vivenciais sobre engajamento e pertencimento é maior (interação positiva) para estudantes pretos, pardos, indígenas, beneficiários do Bolsa Família e PCD.

3. Desenho amostral

Desenho longitudinal de duas ondas (linha de base no início da Fase 2 e seguimento ao final da Fase 2), em amostra estratificada por NRE (Maringá, Londrina, Apucarana, Paranavaí), com aninhamento natural: estudantes (nível 1) em turmas (nível 2), em escolas (nível 3), em NREs (nível 4).

4. Cálculo amostral

Cálculo via G*Power 3.1 para regressão multinível com efeito médio ($f^2 = 0,02$), poder estatístico de 0,80, $\alpha = 0,05$, e ICC esperado entre 0,05 e 0,10. n mínimo = 1.952 estudantes. Margem de segurança operacional de 28% define meta de coleta em ≥ 2.500 respondentes válidos por onda, distribuídos em 145 escolas.

5. Variáveis

5.1 Variáveis dependentes (escala Likert 1–5, validada em piloto)

- Engajamento escolar (subescalas comportamental, emocional e cognitiva).
- Pertencimento à escola.
- Sustentabilidade (educação ambiental e práticas vivenciais).
- Protagonismo estudantil.
- Transformação percebida.

5.2 Variável independente principal

- Frequência de práticas vivenciais (multiselect: experimentos, feiras de ciências, saídas de campo, visitas técnicas, projetos integradores, clubes investigativos), operacionalizada como contagem (0–6) e como score ponderado.

5.3 Variáveis de controle e moderadoras

- Raça/cor (categorias IBGE: branca, preta, parda, amarela, indígena, não-declarado).
- Gênero (feminino, masculino, não-binário, outro, não-declarado).
- NSE (proxy: beneficiário do Bolsa Família — sim/não).
- Pessoa com deficiência (PCD — sim/não).
- Idade, série, município, anos de tempo integral, rede.

6. Plano de análise

6.1 Análise descritiva

- Médias, desvios-padrão, frequências e proporções por NRE, série, gênero, raça/cor, NSE e PCD.

- Comparação da composição amostral observada × composição censitária esperada (painel de equidade).

6.2 Análise psicométrica (Fase 1 · M6)

- Alfa de Cronbach para cada subescala (critério $\geq 0,70$).
- Análise Fatorial Confirmatória (AFC) com índices CFI $\geq 0,95$, RMSEA $\leq 0,06$, SRMR $\leq 0,08$.
- Invariância de medida entre grupos (gênero, raça/cor, série) via AFC multigrupo.

6.3 Análise inferencial

- Testes t-Student e ANOVA one-way para comparações entre grupos (controle de inflação de erro tipo I via Bonferroni).
- Modelos de regressão multinível (HLM): null model → modelo com VI principal → modelo com controles → modelo com interações (H4).
- Cálculo de ICC, R^2 marginal e R^2 condicional (Nakagawa & Schielzeth).
- Análise de cluster k-means para identificação de perfis de engajamento × resistência.

6.4 Análise longitudinal (onda 1 × onda 2)

- Modelos de mudança intra-individual (random intercepts and slopes).
- Estimativa de efeitos de exposição cumulativa às práticas vivenciais.

7. Tratamento de dados ausentes

Dados ausentes serão tratados via imputação múltipla por equações encadeadas (MICE, $m = 20$), assumindo Missing At Random (MAR). Análises de sensibilidade serão conduzidas sob cenários MNAR (pattern-mixture models).

8. Critérios de inclusão e exclusão

- Inclusão: estudantes regularmente matriculados no 6º ao 9º ano em escolas estaduais de tempo integral das 4 NREs amostradas, com TCLE e TALE assinados.
- Exclusão: respostas com mais de 50% do instrumento em branco; respostas com padrão evidente de invalidez (straight-line responding); duplicatas técnicas detectadas via client_token.

9. Software e reprodutibilidade

- Análise quantitativa: R 4.x (lme4, nlme, lavaan, mice, dplyr, ggplot2) e Stata IC.
- Análise qualitativa: NVivo + Lovable AI Gateway (Gemini 2.5 Pro) para codificação assistida.
- Reprodutibilidade: scripts versionados em repositório público; dados anonimizados depositados sob licença CC BY 4.0.

10. Ciência Aberta

- Pré-registro completo do PAE no OSF (Fase 1 · M6).
- Materiais públicos: instrumento, codebook, manual operacional, scripts de análise.
- Dados anonimizados publicados em repositório institucional ao final do projeto (Fase 3 · M18).
- Auditoria pública via códigos demo na plataforma e log completo de operações.

11. Ética e LGPD

O protocolo será submetido ao CEP/Plataforma Brasil sob a coordenação da Dra. Leilane Schwind / UEM (CEP coordenador 104), com TCLE para responsável legal e TALE para estudante menor. Toda a coleta e armazenamento de dados observam integralmente a LGPD (art. 7º, I e IV), com anonimização irreversível na exportação pública.