ANEXO II - PLANO DE AÇÃO PARA ALCANCE DAS METAS DE APRENDIZAGEM
INSTRUÇÃO NORMATIVA SME Nº XX de XX de maio de 2026

1. Introdução e Propósito

A presente Nota Técnica tem por finalidade consolidar a proposta metodológica do modelo de metas baseada em quartis de proficiência, estabelecendo uma alternativa estrutural em relação aos índices sintéticos tradicionais, especificamente o Índice de Desenvolvimento da Educação Básica (IDEB) e o Índice de Desenvolvimento da Educação Paulistana (IDEP). Este documento fornece subsídios técnicos para a gestão da SME, das DREs e das Unidades Educacionais de Ensino Fundamental, fundamentando-se nos princípios de equidade e do foco na aprendizagem. O desenho normativo aqui proposto prioriza a ação pedagógica direta, contrastando com a natureza predominantemente diagnóstica dos índices compostos.

2. Síntese do método de Metas por Quartis (Prova São Paulo)

As metas de proficiência para o ano de 2026 foram construídas a partir de um modelo estatístico baseado na distribuição observada das proficiências das Unidades Educacionais (UEs) na Prova São Paulo de 2025 (N = 621 escolas) e de testes de robustez a partir de 2023. O modelo combina três elementos: uma estrutura de classificação por quartis dentro de cada combinação de série e componente curricular; metas de crescimento diferenciadas por nível de proficiência; e um modelo probabilístico Normal-ogiva para estimar a proporção de estudantes que alcançará o nível Adequado ou Avançado ao final do ciclo. O modelo de quartis para a Prova São Paulo (PSP) foca exclusivamente na proficiência, individualizando as unidades escolares para garantir que o esforço de

melhoria seja calibrado conforme o ponto de partida de cada agrupamento de componente curricular no ano do ciclo.

2.1 Estratificação e equidade horizontal

A estratégia de estratificação separa os componentes e anos do ciclo da rede em quatro grupos (Q1 a Q4) utilizando os percentis P25, P50 e P75 como pontos de corte. A inovação técnica reside na aplicação desta divisão de forma estrita dentro de cada combinação de componente curricular (Língua Portuguesa-LP, Matemática-MT, Ciências Naturais-CN) x ano escolar. Este procedimento mitiga vieses de comparação entre anos dos ciclos com bases curriculares distintas e assegura que uma escola seja comparada apenas a seus "pares curriculares", elevando a validade da comparação e mensuração interna.

2.2 Atribuição de ganhos diferenciados

As metas de crescimento são normativas e crescentes conforme o nível de proficiência inicial: escolas com a maior parte dos anos nos quartis superiores recebem, no conjunto, metas absolutas maiores, a partir de um princípio de equidade.

Diferentemente do pressuposto justificado do IDEB – princípio logístico – o qual prevê que à medida que o indicador aumenta após determinado esforço inicial, a melhora vai se tornando cada vez menor próximo ao topo da proficiência (ou seja, aumentar as proficiências é mais difícil nos níveis mais elevados, e o fluxo tem um ‘limite natural’); nos ancoramos, aqui, em outros pressupostos:

1 – as escolas que, no conjunto, possuem as menores proficiências, historicamente, têm maiores dificuldades e complexidades cotidianas e, na média, estão entre aquelas de menor nível socioeconômico;

2 – as escolas com maior dificuldade precisam de um desafio factível em seu horizonte, para que as altas expectativas não se tornem metas irreais e prejudiciais ao esforço destas unidades educacionais;

3 – a literatura mostra que um ano de baixa aprendizagem significa 10 pontos na escala de proficiência[1]: no caso da Prova São Paulo, para os estudantes que apresentam padrão de desempenho no "Abaixo do Básico" e "Básico", isto é o suficiente apenas para mantê-los nesses níveis, ano após ano. Até 2018, na Rede Municipal de Ensino, o crescimento médio por ano em Língua Portuguesa foi de 4,8 pontos de proficiência e 3,8 pontos em Matemática, nos anos iniciais. Nos anos finais, houve crescimento médio de 3,0 pontos em Língua Portuguesa e 1,6 pontos em Matemática.

4 – portanto, a partir do princípio de equidade, as escolas com a maior parte dos anos do ciclo situadas nos quartis superiores (Q3 e Q4) possuem, no conjunto, metas mais elevadas.

5 - uma escola poderá, no mesmo componente e em anos dos ciclos diferentes (nos 5º e 8º anos, por exemplo), ter metas do Q1 e do Q4, respectivamente. Esse princípio permite delinear a desigualdade intra-escolar.

Componente Curricular	**Q1 (Menor Prof.)**	**Q2**	**Q3**	**Q4 (Maior Prof.)**
Língua Portuguesa (LP)	+30 pontos (média de 10 pontos por ano)	+35 pontos (média de 11,67 pontos por ano)	+40 pontos (média de 13,3 pontos por ano)	+45 pontos (média de 15 pontos por ano)
Matemática e Ciências (MT/CN)	+40 pontos (média de 13,3 pontos por ano)	+45 pontos (média de 15 pontos por ano)	+50 pontos (média de 16,6 pontos por ano)	+55 pontos (média de 18,3 pontos por ano)
Nota: Ganhos totais acumulados para o ciclo 2026-2028.				

[1] ALVES, Maria Teresa Gonzaga ; SOARES, José Francisco; XAVIER, Flavia Pereira. Desigualdades Educacionais no Ensino Fundamental de 2005 a 2013: hiato entre grupos sociais. *Revista Brasileira de Sociologia*, v. 4, p. 49, 2016; SOARES, José Francisco; DELGADO, Victor Maia Senna. Medida das desigualdades de aprendizado entre estudantes de ensino fundamental. *Estudos em Avaliação Educacional* (Impresso), v. 27, p. 754-780, 2016; SOARES, José Francisco e ALVES, Maria Teresa Gonzaga. Efeitos de escolas e municípios na qualidade do ensino fundamental. *Cadernos de Pesquisa*, 2013, Maio; 43(149): p. 492–517.

2.3 Escalonamento Temporal (esforço necessário)

A implementação não é linear, adotando uma crescente que respeita o tempo de implementação das intervenções pedagógicas e da formação docente:

2026: 20% do ganho projetado para o triênio.

2027: 35% adicionais do ganho projetado para o triênio, totalizando 55% acumulados.

2028: 45% do ganho projetado para o triênio, totalizando 100% acumulados.

3. Visão Geral dos Modelos de Índices Compostos (IDEB e IDEP)

3.1 Estrutura de dois fatores e o Indicador de Rendimento

IDEB = P x N, onde **"P"** representa o "Indicador de rendimento" (fluxo) e **"N"** a "nota média padronizada" de Língua Portuguesa e Matemática. Este desenho impõe uma troca explícita: o sistema deve maximizar o aprendizado sem comprometer as taxas de aprovação escolar.

3.2 Contextualização e "Faixas" no IDEP

Diferente do IDEB nacional, o **IDEP** municipal agrupa escolas em 6 "Faixas" de metas, geradas pelo cruzamento do Índice de Nível Socioeconômico (INSE, níveis 3 a 5) com o Indicador de Complexidade de Gestão (ICG, agrupado em dois níveis de porte e turnos), e considera as médias das proficiências do 3º, 5º, 7º e 9º anos de Língua Portuguesa, Matemática e Ciências Naturais. Esta matriz (Tabela 1 da Nota Técnica IDEP) visa garantir que as trajetórias de avanço sejam condizentes com a realidade socioeconômica e organizacional de cada unidade.

3.3 Projeção Logística de Longo Prazo

Ambos os modelos utilizam funções logísticas para traçar trajetórias de convergência. No caso do IDEB, as metas foram desenhadas para convergir a um

teto (nota 6,0 ou 9,9) em um horizonte estendido; para os municípios, o tempo de convergência (λ) foi estabelecido em 91 anos (horizonte até 2096), contrastando com a periodicidade trienal das metas da PSP, ou mesmo o período de 5 anos de revisão previsto para o IDEP.

4. Comparação detalhada: diferenças estruturais

Dimensão	**Método de Quartis (PSP)**	**Índices compostos (IDEB/ IDEP)**
Objeto da Meta	Proficiência pura (Aprendizagem).	Índice sintético (Aprendizado + Fluxo).
Referencial de Comparação	Pares curriculares (mesmo ano do ciclo/componente).	Municípios e UFs – IDEB; Faixas socioeconômicas (INSE x ICG) - IDEP.
Mecânica de Projeção	Metas normativas, em combinações, por quartil.	Regressão logística de convergência ao teto.
Horizonte Temporal	Ciclo de curto/médio prazo (3 anos).	Convergência secular (91 anos / 2096 para o IDEB, 50 anos para o IDEP).
Comunicação	Distribuição por níveis (Abaixo do Básico ao Avançado).	Nota única sintetizada (0 a 10).
Fator de Ajuste	Normal-ogiva.	Compatibilização PISA/SAEB; ICG e NSE/IDEP.

Nota: a célula "Fator de Ajuste" deve ser lida como "Conversão Normal-ogiva como aproximação" (calculado sobre os dados sintéticos da própria Normal-ogiva) — os QQ-plots por série×componente mostram não-normalidade; resultados são reportados com sensibilidade a σ (45/50/55).

5. Análise de Robustez e checagem de consistência interna

A presente seção tem por objetivo avaliar a consistência interna, a robustez paramétrica e as limitações inferenciais do modelo de metas por quartis, com especial atenção à conversão entre proficiência média projetada e proporção de estudantes nos níveis de desempenho, em particular o nível **Adequado+ (Adequado + Avançado)**.

5.1 Modelo de Conversão: Normal-ogiva e estrutura de agregação

A conversão da proficiência média projetada em proporções de estudantes acima de um ponto de corte é realizada por meio da função de distribuição acumulada (CDF) da Normal padrão. Seja X_{ij} a proficiência do estudante na escola i, assume-se:

$$X_{ij} \sim N(\mu_i, \sigma^2), \mu_i = E[X_{ij}]$$

onde:

μ_i é a proficiência média projetada da escola i após aplicação da meta;

σ é o desvio-padrão intraescolar da proficiência;

c é o ponto de corte associado ao nível **Adequado** para cada combinação série × componente curricular.

A proporção esperada de estudantes em nível Adequado ou superior na escola i é dada por:

$$P_i = Pr\ (X_{ij} \geq c \mid i) = 1 - \Phi\left(\frac{c - \mu_i}{\sigma}\right)$$

onde $\Phi(\cdot)$ denota a função de distribuição acumulada da Normal padrão.

Agregação em nível de rede

Dada a não linearidade da função Φ, a proporção agregada da rede não pode ser obtida pela aplicação direta da função à média das proficiências. A forma correta de agregação é:

$$P_{rede} = \frac{1}{N}\sum_{i=1}^{N} \quad \left[1 - \Phi\left(\frac{c - \mu_i}{\sigma}\right)\right]$$

onde Né o número de escolas.

Essa formulação evita viés de agregação decorrente da desigualdade de Jensen, isto é:

$$\Phi\left(\frac{c - E[\mu_i]}{\sigma}\right) \neq E\left[\Phi\left(\frac{c - \mu_i}{\sigma}\right)\right]$$

Na prática, verificou-se que para Língua Portuguesa, Matemática e Ciências Naturais, a variância entre escolas é relativamente baixa, pois a diferença entre os métodos é próxima a 1 ponto percentual.

5.2 Especificação do desvio-padrão (σ) e análise de sensibilidade

A conversão Normal-ogiva depende criticamente da especificação do parâmetro σ, que representa a dispersão das proficiências.
Duas interpretações são possíveis:

(i) Dispersão intraescolar (modelo operacional)

$$P_i = 1 - \Phi\left(\frac{c - \mu_i}{\sigma_{intra}}\right)$$

Para operacionalização do modelo, adota-se um valor de referência para o desvio-padrão intraescolar (σ), calibrado a partir de evidências empíricas da escala da PSP, tipicamente situado entre 45 e 55 pontos. Essa escolha não decorre de uma hipótese estrutural, mas de uma aproximação operacional utilizada para análise de sensibilidade da conversão Normal-ogiva.

(ii) Dispersão total (modelo estrutural)

$$\sigma total = \sqrt{(\sigma^2 intra + \sigma^2 entre + \sigma^2 erro)}$$

onde:

σ_{entre}: variabilidade entre escolas;

σ_{erro}: erro de medida (SEM).

Resultados de sensibilidade

A análise empírica indica que para Matemática e Ciências Naturais, a variação de σ entre 45 e 55 pontos altera os resultados em menos de 0,5 p.p.; para Língua Portuguesa, a variação de σ entre 45 e 55 pontos é de apenas 2,1 p.p. A formulação assume que as Pi das escolas são tratadas como realizações independentes para fins do estimador pontual de $\mathrm{P}rede$. A correlação entre escolas

da mesma DRE (ICC empírico ≈ 0,03–0,05) é incorporada nos erros-padrão via cluster-escola na validação logística da seção 5.3.

Decisão metodológica

Para fins de política pública, adota-se a especificação onde $\sigma_{intra} \in [45, 55]$ pontos. A componente σ_{entre} não entra no denominador pois a heterogeneidade entre escolas já está capturada pela distribuição dos μ_i individuais. Isso se deve a três razões: coerência com a unidade de intervenção (a escola); evitar dupla contagem da variabilidade entre escolas, já explicitamente modelada via heterogeneidade de μ_i garantir maior estabilidade comunicacional das metas.

Os resultados devem, contudo, ser interpretados como condicionados a essa escolha paramétrica, sendo recomendada a apresentação de análises de sensibilidade.

5.3 Validação interna: consistência da regra de metas

A validação estatística do modelo não é conduzida como teste preditivo, mas como verificação de consistência interna da regra de transformação entre proficiência e probabilidade de atingir o nível Adequado.

Define-se a margem projetada da escola $i$como:

$$Margem_i = \mu_i - c$$

A probabilidade associada é:

$$z_i = \Phi\left(\frac{Margem_i}{\sigma intra}\right)$$

Quando disponível n_i por escola, a simulação apropriada da proporção escolar:

$$Ki \sim Binomial(n_i, P_i)$$

Modelos logísticos do tipo

$$logit\left(E\left[\frac{K_i}{n_i}\right]\right) = \beta_0 + \beta_1 z_i$$

foram estimados com erros-padrão robustos por cluster de escola.

A escolha de uma especificação linear decorre da aproximação conhecida entre a função logística e a normal acumulada, tal que:

$$logit\big(\Phi(z)\big) \approx 1{,}7z \quad para \ \ |\, z \,| < 2$$

para os quartis Q2 a Q4, e para a maioria de Q1, com exceção das escolas de menor proficiência em Matemática.

Interpretação

Os resultados indicam pseudo-R^2 de McFadden elevado, refletindo consistência interna entre margem, probabilidade e coerência matemática da regra de conversão.

Contudo, ressalta-se que esta evidência não constitui validação preditiva externa, uma vez que os dados simulados derivam da própria estrutura do modelo. Assim, essa etapa deve ser interpretada como checagem de coerência, e não como teste empírico de validade causal ou preditiva.

5.4 Premissa de normalidade e limitações de forma

A hipótese de normalidade:

$$X_{ij} \sim N(\mu_i, \sigma^2)$$

é adotada como aproximação operacional.

Análises gráficas (QQ-plots) indicam desvios em relação à normalidade em algumas distribuições empíricas, incluindo: assimetria; caudas mais pesadas; possíveis efeitos de truncamento próximos aos níveis de proficiência.

Esses desvios podem afetar principalmente estimativas próximas aos pontos de corte e componentes com maior heterogeneidade (notadamente Língua Portuguesa).

Apesar disso, a aproximação Normal-ogiva permanece adequada como instrumento de modelagem de primeira ordem, sendo recomendável, em desenvolvimentos futuros, explorar distribuições alternativas (e.g., mistura de normais) e incorporar modelos não paramétricos para validação.

5.5 Regressão à média e ajuste de ganhos

A interpretação dos ganhos observados deve considerar o fenômeno de regressão à média, especialmente entre escolas inicialmente posicionadas nos quartis inferiores de desempenho.

Seja ρ a correlação intertemporal das proficiências escolares, o ganho esperado por regressão à média pode ser expresso como:

$$(1 - \rho) \cdot \left(\mu_{rede(s\times c)} - \mu_{i,t_0}\right)$$

Dessa forma, o ganho líquido ajustado é dado por:

$$G_{líq,i} = G_{obs,i} - (1 - \rho) \cdot (\mu_{rede}(s\times c) - \mu_i,t_0)$$

Estimativas empíricas indicam valores de ρ em 0,541 para Matemática e Ciências, implicando que aproximadamente 63% a 65% do ganho observado entre escolas de baixo desempenho pode ser atribuído à regressão à média. A correlação intertemporal ρ em LP foi estimada em 0,494 (2024/2025, N=529 escolas). Com esse valor, $G_{líq}$ (LP Q1) ≈ +1,0 pt/ano: positivo, mas estreito, reforçando a centralidade do $G_{líq}$ como indicador de monitoramento. O indicador central de monitoramento passa a ser o ganho líquido $G_{líq}$, com a regressão à média explicitada.

Esse ajuste é particularmente relevante para o monitoramento de escolas no primeiro quartil (Q1), devendo ser incorporado como indicador central de acompanhamento da política.

5.6 Síntese da robustez

Os resultados indicam que o modelo apresenta alta robustez para Matemática e Ciências Naturais (com distância de 1,6 pontos percentuais de Ciências em relação a Matemática), com baixa sensibilidade a parâmetros, e sensibilidade moderada para Língua Portuguesa, decorrente da maior variabilidade entre escolas e da dependência da especificação de σ.

O modelo deve ser interpretado como um modelo determinístico de deslocamento de distribuição com conversão probabilística aproximada, condicionado a hipóteses paramétricas explícitas. O modelo não deve ser

interpretado como ferramenta de previsão individual ou causal, mas como mecanismo de calibração de metas.

6. Observações

Regressão à Média: escolas com anos em Q1 podem apresentar ganhos expressivos que são, em parte, artefatos estatísticos de flutuações de curto prazo, como vimos na seção 5.5. É imperativo monitorar trajetórias anuais para evitar interpretações equivocadas de sucesso.

Ausência de Indicador de Fluxo: diferente do IDEB e IDEP, o modelo não penaliza a retenção escolar. Por isso o monitoramento constante da frequência escolar é fundamental para o acompanhamento e recomposição das aprendizagens.

Premissa de Normalidade: a adoção da distribuição 'normal' é uma convenção operacional; assimetrias reais na distribuição das notas exigem inspeção contínua (QQ-plots).

Histórico: as metas de Q2, Q3 e Q4 incorporam um efeito esperado da política de formação docente e acompanhamento das aprendizagens estimado em **τ = +10 pontos por ano** (Q2) a **+27 pontos por ano** (Q4) além da trajetória histórica observada até aqui. Para os quartis Q3 e Q4 de Matemática e Ciências Naturais, esse efeito esperado excede o teto documentado na literatura internacional[2] para intervenções intensivas de formação docente **(0,10–0,20 σ por ano = 5 a 10 pontos**). O backtest com os dados reais da PSP 2024/2025 (N=533 escolas) indica que o τ observado já excede esses limiares para todos os grupos Q2–Q4 e todos os componentes, antes mesmo do início do ciclo. Recomenda-se monitoramento bimestral com revisão ao final de 2026, adotando τ_{limiar} = 8 pt/ano (Q2), 12 pt/ano (Q3) e 15 pt/ano (Q4) como gatilhos formais. As metas de Q2–Q4 são normativas

[2] KRAFT, Matthew A.; BLAZAR, David; HOGAN, Dylan. The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, v. 88, n. 4, p. 547–588, 2018 e VISSCHER, Adrie J.; DMOSHINSKAIA, Natasha; PELLEGRINI, Marta; REY-NAIZAQUE, Anny. (When) do teacher professional development interventions improve student achievement? A meta-analysis of 128 high-quality studies. *Teaching and Teacher Education*, 2025

com caráter aspiracional, e a calibração será realizada com base nos primeiros resultados do ciclo.

7. Conclusão e recomendações estratégicas

O método de metas por quartis viabiliza a gestão pedagógica sistemática e contínua, permitindo intervenções diferenciadas por componente curricular e ano do ciclo, ano a ano. Ao mesmo tempo, o IDEB e o IDEP permanecem como ferramentas de monitoramento das políticas públicas e de fluxo em rede.