

Quem conclui uma capacitação apoiada por IA?

Engajamento, não perfil.

O que 791 bolsistas de uma capacitação linguística apoiada por IA revelam sobre seleção, engajamento e inclusão — e por que, nesta coorte, os atributos de perfil de entrada preveem mal a conclusão.

Alexandrine Brami

— Cofundadora & CEO, Lingopass · ENS Paris-Saclay · MSc em Ciências Políticas (Sciences Po Paris)

Apresentado na Round Table 6 (Inteligência Artificial em Pequenas e Médias Empresas) da 6ª edição do Colóquio Internacional CERALE–FGV EBAPE, Rio de Janeiro, 8–10 de junho de 2026.

Versão 1.0 para discussão · setembro de 2026.

Data · Setembro de 2026

Linha editorial · Capital humano, tecnologias educacionais, inteligência artificial aplicada à aprendizagem

Status · Working paper. Comentários e críticas são bem-vindos: research@lingopass.com.br

Citação sugerida · Brami, A. (2026). *Quem conclui uma capacitação apoiada por IA? Engajamento, não perfil* (Lingopass Research Working Paper Nº 003). Apresentado na 6ª edição do Colóquio Internacional CERALE–FGV EBAPE, Rio de Janeiro. Lingopass.

Sumário

Principais conclusões.....3

Sobre este working paper.....4

Resumo executivo.....4

Abstract & Keywords.....4

PARTE I — Contexto e método.....6

 Introdução: duas dívidas em aberto.....6

 O paper e o Colóquio CERALE 2026.....7

 O programa Empowering Talent.....7

 O que a ciência já sabia.....8

 Dados e método.....10

PARTE II — Os achados.....13

 Resultados, em três atos.....13

 Evidência qualitativa.....18

PARTE III — Implicações e teoria de mudança.....19

 Discussão: micro, meso e macro.....19

 Implicações para política e prática.....20

 Da evidência à ação: a teoria de mudança.....22

 Limitações.....23

 Pesquisa futura.....24

Conclusão.....25

Sobre a autora.....25

Como citar · Referências (APA).....26

Apêndice A — Perfis dos clusters.....27

Apêndice B — Notas sobre não-resposta e robustez.....27

Apêndice C — Depoimentos e análise temática.....28

Apêndice D — Desempenho preditivo do onboarding e resultados completos.....30

Principais conclusões

Os resultados polarizam. 35% nunca acessaram a plataforma; 22% concluíram os três idiomas (36% concluíram ao menos um), e 32% formam o cluster de alta dedicação. Não é uma curva de esforço — a distribuição é fortemente polarizada (dois regimes de comportamento, não necessariamente duas subpopulações naturais).

Quem a pessoa é não prevê quem conclui. Demografia, condição socioeconômica, personalidade (PCM/Kahler), estilos de aprendizagem e disponibilidade declarada predizem a conclusão com desempenho preditivo ausente a fraco ($AUC \approx 0,49-0,57$), insuficiente para justificar uso como critério de seleção nesta coorte. O indicador inicial mais informativo foi a ativação no onboarding, associada a 3,3x mais conclusão (IC95% 2,4–4,6; $AUC 0,64$). Como o estudo é observacional, essa associação não demonstra efeito causal.

Não identificamos disparidades estatisticamente robustas nos desfechos observados. Entre os melhores, 33% são de baixa renda, 61% de escola pública e 64% primeira geração na universidade. No recorte analisado, estudantes de baixa renda e racialmente sub-representados apresentaram conclusão semelhante com maior tempo de plataforma. O padrão é compatível com diferenças de condição, mas não mede capacidade intrínseca nem exclui outras explicações. Nesta coorte, não identificamos disparidades estatisticamente robustas de conclusão por raça, renda ou gênero após a correção para múltiplas comparações; isso não demonstra ausência de desigualdade, nem permite atribuir o padrão à mediação por IA — e não elimina desigualdades de acesso e de condição material.

Tese — Nesta coorte, a ativação inicial foi mais informativa que os atributos de perfil para identificar risco de não conclusão de uma capacitação apoiada por IA. O achado justifica testar intervenções de onboarding, suporte e pertencimento; não demonstra, isoladamente, o efeito causal dessas intervenções. A implicação para PMEs e governos é direta: não usar perfil como critério de seleção, e testar — com desenho e medição adequados — investimentos em ativação inicial e pertencimento.

Recomendações — (1) empresas e PMEs devem validar localmente qualquer regra de pré-seleção por perfil antes de usá-la, priorizando experimentos de onboarding, suporte e pertencimento nas primeiras semanas, com medição de efeito incremental antes de escalar recursos, e tratando comunidade, monitoria e encontros ao vivo como componentes de suporte e hipóteses de retenção a testar, não como custos acessórios a cortar; (2) o setor público não deve usar modelos preditivos de perfil não validados como critério automático de exclusão — critérios sociais e territoriais de elegibilidade com base legal continuam apropriados — e deve especificar resultado por engajamento e conclusão por coorte, não por demografia; (3) o terceiro setor e financiadores devem distribuir bolsas amplamente em vez de usar atributos de perfil não validados para restringir acesso, investir em pertencimento desde o primeiro dia e

reportar impacto como identificação de participantes com persistência e conclusão demonstradas, não como seleção.

Sobre este working paper

Em uma frase — Numa coorte de 791 bolsistas de uma capacitação linguística apoiada por IA, o perfil de entrada não prevê quem conclui; a ativação inicial — o onboarding — mostrou associação com quem conclui.

Método — Clustering k-means, regressão logística e Random Forest com AUC em validação cruzada de 5 partições, testes de Fisher, Mann-Whitney e qui-quadrado, e correção de Benjamini-Hochberg para múltiplas comparações, sobre registros individuais pseudonimizados de 791 alunos, com resultados divulgados de forma agregada (556 respondentes ao assessment de entrada), e entrevistas qualitativas em três grupos.

Para quem — Gestores de PMEs, formuladores de política pública e organizações do terceiro setor que desenham, financiam ou compram capacitação e precisam decidir como (ou se) selecionar beneficiários por perfil.

Ressalva metodológica — Estudo observacional de uma única coorte (2023), um único programa e um único parceiro: os achados são associação, não causa; a validade preditiva do onboarding refere-se à conclusão da capacitação, não ao desempenho profissional subsequente.

Resumo executivo

Este working paper analisa uma coorte longitudinal proprietária de 791 bolsistas em um programa multilíngue de 90 dias apoiado por IA (francês, espanhol e inglês; nível A0), operado pelo Lingopass para alunos do grupo educacional Ânima (Brasil, 2023). Usando clustering não supervisionado, modelos preditivos validados por validação cruzada e correção para múltiplas comparações, mostramos forte polarização comportamental, com concentração entre não ativação/abandono precoce e alta dedicação; que atributos de perfil de entrada — demografia, condição socioeconômica, personalidade e estilos de aprendizagem — preveem a conclusão com desempenho preditivo ausente a fraco ($AUC \approx 0,49-0,57$); que a ativação inicial (onboarding) foi o sinal precoce mais informativo entre os itens testados, com desempenho moderado; e que estudantes de baixa renda e racialmente sub-representados atingem os mesmos desfechos investindo mais tempo de plataforma, sem que nenhuma dimensão de perfil sobreviva à correção de múltiplas comparações. Argumentamos que perfilar aprendizes não deveria ser usado como critério de seleção nesta coorte, e que testar intervenções de ativação inicial e pertencimento é uma agenda prioritária para uma capacitação inclusiva apoiada por IA.

Abstract & Keywords

Who completes AI-supported training? We analyze a proprietary longitudinal cohort of 791 scholarship students in a 90-day, AI-supported multilingual program (French, Spanish, English; level A0) run by

Lingopass for students of the Ânima education group (Brazil, 2023). Using unsupervised clustering, cross-validated predictive models, and multiple-comparison corrections, we find: (1) polarized outcomes — 35% never log in and only 22% complete all three courses (36% complete at least one), with a high-engagement cluster of 32%; (2) demographics, personality (PCM/Kahler), learning styles (Kolb, Felder-Silverman, Dunn & Dunn) and self-declared availability predict completion no better than chance (AUC $\approx$ 0.49–0.57); (3) one early-engagement behavior — onboarding — is associated with 3.3× higher completion (AUC 0.64); (4) low-income and racially under-represented students reach the same outcomes while investing more platform time; (5) no profile dimension — including the Dunn & Dunn sociological preference (group vs. solo) — survives multiple-comparison correction; the role of the social dimension appears only in qualitative interviews, not as a quantitative predictor. Findings extend the long-standing critique of learning-styles-based instruction (Pashler et al., 2008) from theory into real deployment, and complement evidence that AI helps or harms learning depending on design (Bastani et al., 2025). We argue that learner profiling should not be used as a selection criterion in this cohort, and that testing early-activation and belonging interventions is a priority agenda for inclusive, AI-supported capability building.

Keywords: AI in education; learning analytics; learning-styles myth; language economics; digital inclusion; engagement design; cohort analysis.

PARTE I

Contexto e método

Esta parte situa o problema — a persistência do perfilamento de aprendizes apesar da falta de evidência, e a pergunta sobre se a inteligência artificial ajuda ou atrapalha —, apresenta o programa estudado (Empowering Talent) e detalha os dados e o método usados para testar, em deployment real, se atributos de entrada preveem quem conclui a capacitação.

Introdução: duas dívidas em aberto

A expansão da IA generativa na educação reacendeu duas disputas que este paper conecta com dados de deployment real.

Dívida 1 — o perfil prevê quem conclui? Há quase duas décadas a literatura mostra ausência de evidência de que ensinar conforme o ‘estilo de aprendizagem’ melhora resultados (Pashler, McDaniel, Rohrer & Bjork, 2008; Kirschner & van Merriënboer, 2013; Coffield et al., 2004; Rogowsky et al., 2015). Mesmo assim, programas seguem perfilando aprendizes — e raramente alguém testa, em ambiente real e em escala, se algum atributo de entrada (demográfico, socioeconômico, psicométrico) prevê o desfecho. Nós testamos.

Dívida 2 — a IA ajuda ou atrapalha? Bastani et al. (2025, PNAS) mostram que o mesmo tutor de IA pode prejudicar a aprendizagem quando entrega respostas (alunos usam como ‘muleta’) e ajudá-la quando o desenho impõe guard-rails pedagógicos. Ou seja: o efeito da IA não é uma propriedade da tecnologia — é uma propriedade do desenho. Este paper adiciona o elo que falta a essa conversa: que o desenho que importa é o do engajamento, e que tentar selecionar o aprendiz ‘certo’ é, além de ineficaz, uma questão de inclusão.

Pergunta de pesquisa. Quais atributos observáveis na entrada explicam a variação de conclusão em uma capacitação linguística apoiada por IA — e o que isso implica para o desenho de programas e políticas de inclusão?

Contribuições. O paper oferece quatro contribuições. (i) Empírica: testa, em deployment real e em escala, se atributos de entrada preveem conclusão — algo raro na literatura, que costuma modelar comportamento, não perfil. (ii) Metodológica: combina clustering não-supervisionado, modelos preditivos validados fora da amostra e correção de múltiplas comparações, tratando a não-resposta como informativa. (iii) Conceitual: nomeia a ilusão de taxa-base como o mecanismo gerencial que sustenta o perfilamento sem evidência. (iv) Normativa: deriva, da abordagem das capacidades, por que selecionar por perfil preditivo é, além de ineficaz, regressivo — e propõe o engajamento precoce como sinal alternativo, justo e mensurável.

O paper e o Colóquio CERALE 2026

O trabalho foi apresentado na Round Table 6 — IA em PMEs — do 6º Colóquio CERALE–FGV EBAPE (Rio de Janeiro, 8–10/06/2026), cujo tema foi ‘Oportunidades, riscos e desafios na era da IA’. Ele dialoga com três eixos do colóquio:

Macro–meso–micro. Adotamos a mesma moldura analítica do call para organizar as implicações (Parte III): o aluno (micro), o desenho do programa e a PME (meso), a política de inclusão bi-regional (macro).

IA sob escrutínio crítico. Em linha com a keynote sobre geoeconomia e segurança cognitiva e com os trabalhos sobre ‘ceticismo saudável de IA’ e atrofia cognitiva, nosso achado é não-salvacionista: a IA não seleciona quem aprende, e sem desenho de engajamento a licença vira custo afundado.

Cooperação Europa–América Latina. O programa estudado é um caso latino-americano de capacitação linguística como infraestrutura de inclusão — adjacente às Tracks 1 (políticas públicas, educação) e 6 (IA na educação).

O programa Empowering Talent

O Empowering Talent é o programa social do Lingopass (modelo 4×1: a cada quatro licenças vendidas, uma é doada; alinhado aos ODS 4, 5, 8 e 10). Na edição analisada (1º semestre de 2023, com o grupo Ânima), foram ofertadas até 1.000 bolsas — 800 universitários e 200 professores da rede — para um bootcamp de 90 dias. Esta análise cobre os 791 estudantes matriculados; o braço de professores não entra na modelagem. A jornada soma cerca de 100 horas por aluno: onboarding humanizado (1h), plataforma autoinstrucional EAD alinhada ao CEFR (~60h: videoaulas, podcasts, jogos, quizzes com feedback, gamificação), aulas ao vivo em grupo (~15h, com replays), auxílio contínuo (~24h: tutoria por chat, plantão de dúvidas, suporte e mentoria semanal) e acesso a *microcursos temáticos de empregabilidade e inglês profissional* (por exemplo, uso do LinkedIn e inglês aplicado a áreas específicas). Trata-se de um programa híbrido de capacitação linguística digital — com plataforma autoinstrucional, componentes de IA, aulas ao vivo e suporte humano (tutoria, monitoria, mentoria) —; este estudo não decompõe o efeito de cada componente isoladamente, incluindo o papel específico de IA no mecanismo de aprendizagem. São três cursos de 30 dias, nível A0 (Real Beginner), organizados na sequência estruturada *francês > espanhol > inglês*, que não é arbitrária: como, no Brasil, há exposição prévia desigual ao inglês e ao espanhol mas quase nenhuma ao francês, começar pelo francês equaliza a linha de partida (todos como iniciantes reais), controlando conhecimento prévio. A progressão sequencial era a trilha recomendada, com gating por padrão, mas não estritamente travada: uma minoria concluiu idiomas fora da ordem (35 concluíram espanhol sem francês; 3, inglês sem espanhol) — razão pela qual as conclusões por idioma (Tabela 2) não são perfeitamente monotônicas. A população é majoritariamente feminina (~81% entre concluintes), jovem (mediana 23 anos), com forte presença de baixa renda, escola pública e primeira geração universitária, distribuída por treze estados.

O que a ciência já sabia

Personalização e learning analytics

Sistemas mediados por IA prometem ajustar ritmo e conteúdo ao aprendiz; a learning analytics fornece o instrumental para modelar progressão a partir de traços comportamentais. Meta-análises de sistemas tutores inteligentes (ITS) encontram efeitos positivos, porém heterogêneos (Kulik & Fletcher, 2016; VanLehn, 2011), e a revisão sistemática de Létourneau et al. (2025, npj Science of Learning), com 28 estudos e 4.597 alunos K-12, conclui que os efeitos são ‘geralmente positivos, mas atenuados’ quando o comparador é uma boa tutoria não-inteligente. O horizonte de referência continua sendo o ‘2 sigma problem’ de Bloom (1984), que estimou em dois desvios-padrão o ganho da tutoria individual sobre o ensino em grupo. A learning analytics herdou essa promessa de individualização e a operacionalizou em modelos de risco e recomendação; contudo, a maior parte dessa literatura modela *trajetórias* de engajamento (cliques, tempo, sequência de atividades) e raramente testa se *atributos de entrada* — quem o aluno é antes de começar — preveem o desfecho. É justamente essa lacuna, entre o que se mede (comportamento) e o que se presume (perfil), que este paper ocupa.

A crítica ao perfil e aos estilos de aprendizagem

A tese de que ensinar conforme o ‘estilo’ declarado melhora a aprendizagem carece de evidência experimental (Pashler et al., 2008; Coffield et al., 2004; Rogowsky et al., 2015), e persiste apesar disso (Kirschner & van Merriënboer, 2013; Nancekivell et al., 2020). Nossa contribuição é levar essa crítica do laboratório ao deployment: testamos PCM/Kahler, Kolb, Felder-Silverman, EEPC, DISC e Dunn & Dunn contra um desfecho real (conclusão) em escala. A chamada ‘hipótese de combinação’ (meshing hypothesis) — de que casar instrução ao estilo declarado melhora a aprendizagem — nunca encontrou suporte experimental robusto (Rogowsky et al., 2015), e ainda assim a crença permanece dominante — Newton (2015) encontra que ~89% dos artigos recentes a endossam, e sondagens apontam prevalência acima de 80% entre docentes (Nancekivell et al., 2020). O risco prático não é apenas pedagógico: quando o perfil é usado para *selecionar* quem recebe a oportunidade, a crença sem evidência converte-se em mecanismo de exclusão. Documentamos, ainda, o viés cognitivo que a sustenta — a ilusão de taxa-base (adiante).

A IA ajuda ou atrapalha? Depende do desenho

Bastani et al. (2025, PNAS), em experimento de campo com ~1.000 estudantes de ensino médio, mostram que um tutor de IA sem guard-rails melhora o desempenho durante a prática mas reduz a aprendizagem ao remover o acesso (efeito ‘muleta’), enquanto um tutor com guard-rails pedagógicos mitiga o dano. Trials de tutoria humano+IA reforçam que o ganho vem da combinação e do desenho: o Tutor CoPilot (Wang et al., 2024) elevou o domínio em 4 p.p. no geral e 9 p.p. para alunos de tutores menos experientes; um estudo de ano letivo (Gurung et al., 2025) encontrou +0,36 nível de série para

tutoria humano+IA versus só-IA. A revisão da Stanford SCALE Initiative (2026) — que triou mais de 800 estudos e encontrou apenas ~20 causais de alta qualidade — adiciona o alerta: os ganhos frequentemente valem enquanto o aluno tem acesso à ferramenta, o transfer é difícil, e a IA pode reduzir o raciocínio independente. Em conjunto, a literatura desloca a pergunta de ‘a IA funciona?’ para ‘qual desenho faz a IA funcionar?’ — e sugere que o desenho que mais importa, no nosso caso, é o do engajamento, não o da personalização por perfil.

Evasão, engajamento inicial e pertencimento

Na educação online, a evasão concentra-se no início e o engajamento das primeiras semanas é o preditor mais forte de conclusão (Reich & Ruipérez-Valiente, 2019). A dimensão social é um ingrediente ativo: intervenções de pertencimento têm efeitos duradouros sobre desfechos acadêmicos, sobretudo em grupos sub-representados (Walton & Cohen, 2011, *Science*), e a presença social é central à aprendizagem online (Community of Inquiry; Garrison, Anderson & Archer, 2000). Esses corpos antecipam nossos dois achados centrais: o onboarding como alavanca quantitativa e a dimensão social como o ingrediente que, pela via qualitativa, mais reteve os alunos. A literatura de evasão em MOOCs é particularmente instrutiva: taxas de conclusão de um dígito são a norma — embora o programa analisado não seja um MOOC aberto (tem onboarding humano, monitoria e aulas ao vivo), de modo que não há benchmark diretamente comparável e suas taxas devem ser lidas dentro dessas características, e a janela de decisão concentra-se nas primeiras interações (Reich & Ruipérez-Valiente, 2019). Modelos de previsão precoce de risco em learning analytics confirmam que sinais comportamentais das primeiras semanas superam qualquer variável cadastral na antecipação da evasão — convergente com o nosso achado sobre o onboarding.

Economia da linguagem e abordagem das capacidades

Tratar proficiência como ativo econômico tem base empírica: na Índia, falar inglês fluente associa-se a +34% no salário-hora masculino — retorno comparável a concluir o ensino médio — e mesmo o inglês básico a +13% (Azam, Chin & Prakash, 2013); efeitos semelhantes aparecem na África do Sul (Casale & Posel, 2011) e em imigrantes (Bleakley & Chin, 2004). Seguindo Sen (1999), tratamos a proficiência como uma *capacidade* (capability) que amplia o conjunto de oportunidades efetivamente acessíveis ao indivíduo — trabalho, mobilidade, participação — e, portanto, como infraestrutura de inclusão, não como benefício individual. Essa moldura tem implicação distributiva direta: se o acesso a oportunidades de desenvolver proficiência é desigual, usar atributos de perfil pouco preditivos para restringir esse acesso pode reforçar desvantagens prévias, sem qualquer ganho preditivo demonstrado nesta coorte. O argumento conecta o micro (a progressão do aluno) ao macro (o desenho de políticas de capacitação), que retomamos na Discussão.

Vieses de decisão: a ilusão de taxa-base

Um mecanismo cognitivo ajuda a explicar por que gestores e programas insistem em perfilar mesmo sem evidência. A *negligência da taxa-base* (base-rate neglect) — descrita por Tversky & Kahneman (1974) e formalizada por Bar-Hillel (1980) — é a tendência de superestimar a informação diagnóstica de um traço observado, ignorando sua frequência na população. Em seleção de talento, isso se traduz numa ilusão concreta: observa-se que a maioria dos ‘melhores’ tem certo traço (digamos, perfil analítico) e conclui-se que o traço explica o sucesso — sem notar que ele é igualmente comum entre todos. Nossos dados materializam exatamente esse padrão (68% dos high performers são analítico-dominantes, contra 65% da base inteira; ver Resultados). Nomear o mecanismo importa para a prática: a sensação de que ‘o perfil prevê’ é, em boa parte, um artefato perceptivo, não um sinal estatístico.

Dados e método

Bases. Progresso (2.356 matrículas = 791 alunos × até 3 cursos), teste de nivelamento inicial nos três idiomas e assessment de candidatura (654 candidatos, 240+ campos), além de entrevistas qualitativas em três grupos (concluiu; iniciou e não concluiu; nunca iniciou). O onboarding foi operacionalizado como a conclusão do assessment inicial de perfil (proxy de ativação); a medida de engajamento refere-se ao tempo de auto-estudo (EAD), critério oficial de conclusão.

Unidade e features. Aluno. Features comportamentais (tempo de plataforma, % atividades, dias até concluir, assiduidade, onboarding), socioeconômicas/demográficas e psicométricas reconstruídas (energia Kahler a partir das 18 forças).

Desfecho primário e secundários. O desfecho primário desta análise é a conclusão de ao menos um curso dentro do prazo de 90 dias. Desfechos secundários incluem a conclusão dos três idiomas, o cluster comportamental, o tempo de plataforma e a velocidade de conclusão; resultados para desfechos secundários são reportados de forma explícita como tal ao longo do texto.

Análise. Clustering k-means (k por silhouette); regressão logística e Random Forest com AUC em validação cruzada 5-fold; testes de Fisher e Mann-Whitney; correção de múltiplas comparações de Benjamini & Hochberg (1995); intervalos de Wilson.

Viés tratado. Informative missingness (não responder o assessment é proxy de desengajamento) — testes de perfil restritos aos respondentes; sem claim causal (associação).

Ética. Os dados foram pseudonimizados para análise, com identificadores substituídos por chaves técnicas e divulgação apenas em forma agregada; a publicação foi submetida a avaliação de risco de reidentificação, com supressão ou agregação adicional em subgrupos pequenos. O uso interno e esta apresentação contam com permissão do parceiro institucional.

A amostra

Dos 791 alunos, 556 (70,3%) responderam ao assessment de entrada; a não-resposta foi tratada como informativa (proxy de desengajamento, ver adiante). Variáveis socioeconômicas e demográficas estão disponíveis para esses 556; um subconjunto de 159 — os que completaram o assessment estendido — traz escolaridade, geração universitária, idade e parentalidade. A Tabela 1 resume as características.

Variável / Categoria	n	%
Gênero (n=556)		
Feminino	424	76,3
Masculino	129	23,2
Não declarado	3	0,5
Raça (n=556)		
Branca ou Amarela	286	51,4
Preta, Parda ou Indígena	265	47,7
Não respondeu	5	0,9
Renda — escala 1–6 (n=556)		
Baixa (1–2)	184	33,1
Média (3–4)	205	36,9
Alta (5–6)	167	30,0
Equipamento (n=556)		
Celular + computador	299	53,8
Só computador	131	23,6
Só celular	109	19,6
Tablet	17	3,1
Internet (n=556)		
Em casa	527	94,8
Fora de casa / limitada	29	5,2
Subamostra (n=159)		
Escola pública (ens. médio)	97	61,0

Primeira geração na universidade	101	63,5
Tem filhos	26	16,4
Idade (n=159)		
Mediana 23 anos (IQR 20–29; faixa 17–54)	—	—
Abrangência		
18 instituições · 13 estados	—	—

Tabela 1. Características da amostra. Percentuais sobre a base com dado disponível; a não-resposta é informativa e analisada separadamente.

Especificação e checagens de robustez

Agrupamento. As features comportamentais foram padronizadas (z-score) antes do k-means; o número de grupos foi escolhido por silhueta, com valores comparáveis para $k=2$ e $k=4$ ($\approx 0,72$). Reportamos a solução de quatro grupos por ser mais informativa para a ação, mas a leitura de dois grupos (engaja vs. não engaja) é idêntica em $k=2$ — ou seja, o achado não depende dessa escolha.

Modelos preditivos. Estimamos regressão logística e Random Forest por bloco de preditores (demografia; condição socioeconômica; psicometria; estilos; disponibilidade; engajamento), com AUC em validação cruzada estratificada de cinco partições. Os blocos de perfil ficam na faixa 0,49–0,57 (acaso). O indicador de ativação inicial — conclusão do assessment de onboarding — apresentou AUC de 0,64 para a conclusão de ao menos um curso. As demais variáveis de engajamento são analisadas como descritores da trajetória, pois sua janela de observação pode se sobrepor ao período de definição do desfecho. A concordância entre os dois algoritmos indica que o resultado não é artefato de forma funcional.

Não-resposta informativa. Como não responder ao assessment é, em si, sinal de desengajamento, restringimos os testes de perfil aos respondentes (evitando confundir ausência de dado com desfecho). Análises de sensibilidade tratando a não-resposta como categoria não alteram as conclusões qualitativas.

Inferência. Para proporções, usamos teste exato de Fisher e intervalos de Wilson; para contínuas, Mann-Whitney; para tabelas, qui-quadrado. Como testamos muitas dimensões de perfil, aplicamos correção de Benjamini & Hochberg (1995) dentro da família de testes de perfil — controlando a taxa de falsas descobertas. Nenhuma dimensão sobrevive à correção (Tabela 5), e o resultado é estável entre definições de desfecho (concluir ≥ 1 curso vs. concluir os três).

PARTE II

Os achados

Esta parte apresenta os achados centrais em três atos: a coorte se divide de forma fortemente polarizada entre quem se engaja e quem não se engaja; atributos de perfil não preveem esse desfecho — só a ativação inicial mostrou associação; e esse achado é acompanhado por uma leitura cautelosa de equidade, já que, nesta coorte, grupos sub-representados atingem resultado semelhante investindo mais tempo, e não menos capacidade.

Resultados, em três atos

Ato 1 — A coorte é polarizada

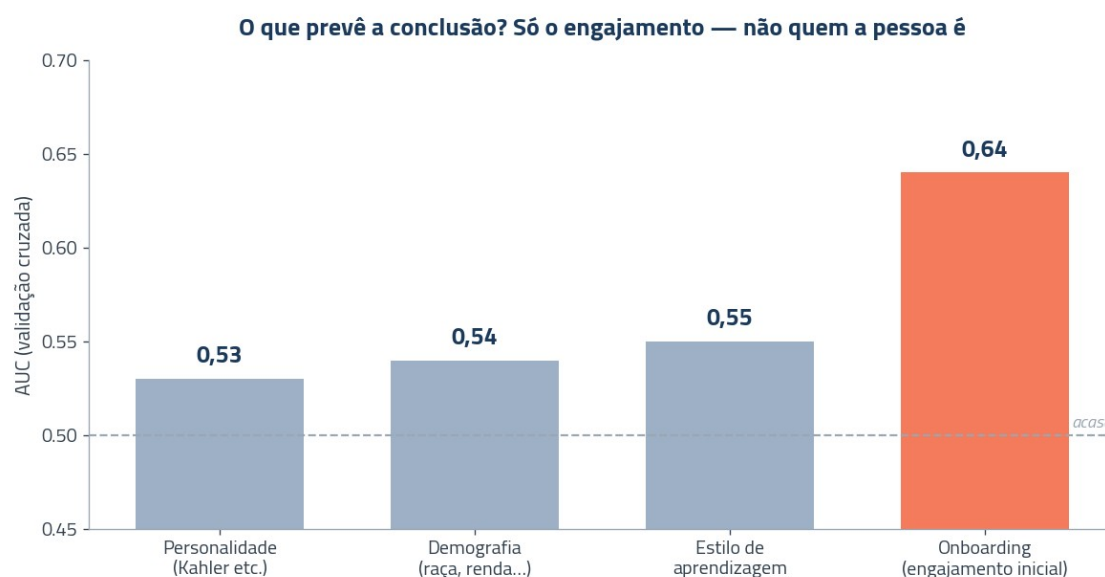
Quatro clusters emergem do comportamento (k-means, k por silhouette): **Fantasma** — não ativados (34,9%, nunca acessaram), **Curiosos** (15,8%, acessam e abandonam — 0% de conclusão), **Constantes** (17,2%, engajam parcialmente — 21% concluem ao menos um curso) e **Maratonistas** (32,1%, alta dedicação — todos concluem ao menos um e 68% concluem os três). A coorte exibe forte polarização comportamental — concentração em trajetórias de não ativação/abandono precoce e em trajetórias de alta dedicação —, e o desfecho se decide cedo. Importa distinguir o cluster de alta dedicação (32,1%) da conclusão dos três idiomas propriamente dita, alcançada por 21,9% (173 alunos); 35,7% concluíram ao menos um curso.

Cluster / desfecho	n	%	Cursos concl. (média)	Tempo EAD mediana (min)
Fantasma (nunca acessou)	276	34,9	0,00	0
Curioso (acessa e larga)	125	15,8	0,00	10
Constante (parcial)	136	17,2	0,21	212
Maratonista (alta dedicação)	254	32,1	2,63	1.078
Concluiu ≥ 1 curso	282	35,7 (IC 32,4–39,1)	—	—
Concluiu os três idiomas	173	21,9 (IC 19,1–24,9)	—	—
Por idioma: FR / ES / EN	245 / 254 / 197	—	—	—

Tabela 2. Desfechos comportamentais e de conclusão ($n=791$).**Figura 1.** Distribuição dos quatro clusters comportamentais ($n=791$).

Ato 2 — Quem a pessoa é não prevê o desfecho

Modelos de demografia, condição socioeconômica, personalidade, estilos de aprendizagem e disponibilidade declarada atingem AUC 0,49–0,57 — desempenho preditivo ausente a fraco (Figura 2), insuficiente para justificar uso como critério de seleção nesta coorte. A taxa de conclusão dos três idiomas é plana entre faixas de renda (28,3% vs 29,3%; Fisher $p=0,92$ para Baixa×Alta; a Tabela 5 reporta o teste na escala ordinal completa, Mann-Whitney $p=0,51$) e não difere significativamente por raça (32,5% vs 24,9%; Fisher $p=0,060$). A energia analítica do Kahler não prediz ($\rho \approx 0,08$, n.s. após correção): 68% dos high performers são analítico-dominantes, mas 65% de toda a base também — ilusão de taxa-base. Nenhuma das dimensões de perfil sobrevive à correção de múltiplas comparações (família de nove testes; Tabela 5), levando a crítica de Pashler et al. (2008) do laboratório ao deployment.

**Figura 2.** Desempenho de classificação por atributos de entrada e ativação inicial (AUC em validação cruzada; 0,50 = acaso). A ativação inicial foi o indicador mais informativo entre os itens testados. Exceto pela ativação no onboarding, indicadores de

engajamento de trajetória (tempo de plataforma, percentual de atividades, dias até concluir) não são tratados como preditores prospectivos nesta versão — ver Limitações.

As AUCs reportadas são estimativas pontuais em validação cruzada de 5 partições; não foram calculados intervalos de confiança por bootstrap nesta versão. Dado o tamanho amostral ($n=791$, com subgrupos de até 159), pequenas variações nessas estimativas são esperadas — a hierarquia entre blocos deve ser lida qualitativamente (perfil $\approx$ acaso a fraco; onboarding, moderado), não como diferença precisa entre valores pontuais.

A Tabela 3 detalha a conclusão dos três idiomas por grupo, com intervalos de Wilson. O padrão é claro: atributos de quem a pessoa é (raça, renda, gênero) não separam o desfecho. O equipamento apresentou o maior contraste bruto entre as variáveis de perfil observadas: 21,1% de conclusão dos três idiomas entre quem estudava apenas por celular, versus 36,6% entre quem dispunha apenas de computador — mas, como o Ato 3 detalha, esse contraste não sobrevive à correção para múltiplas comparações e não demonstra causalidade; é compatível com a hipótese de que restrições materiais merecem monitoramento e teste de desenho específico.

Grupo / Categoria	Conclusão dos 3 (%)	IC95%	n
Raça			
Branca ou Amarela	32,5	27,4–38,1	286
Preta/Parda/Indígena	24,9	20,1–30,4	265
Renda			
Baixa (1–2)	28,3	22,3–35,2	184
Alta (5–6)	29,3	23,0–36,6	167
Gênero			
Feminino	30,2	26,0–34,7	424
Masculino	23,3	16,8–31,3	129
Equipamento			
Só computador	36,6	28,9–45,2	131
Celular + computador	28,1	23,3–33,4	299
Só celular	21,1	14,5–29,7	109

Tabela 3. Conclusão dos três idiomas por grupo (respondentes ao assessment; IC95% de Wilson).

Ato 3 — A ativação inicial é o sinal mais informativo, e a equidade observada exige leitura cautelosa

Concluir o onboarding associa-se a 3,3× mais conclusão de ao menos um curso (45,0% vs 13,6%; razão 3,3×, IC95% 2,4–4,6; AUC 0,64 — acima de todos os conjuntos de perfil somados; Tabela 4). Como o desenho é observacional, o onboarding é o sinal precoce mais informativo, não um efeito causal isolado (ver Limitações e Apêndice D). Entre os high performers: 32,7% de baixa renda, 61,0% de escola pública, 63,5% de primeira geração e 41,5% pretos, pardos ou indígenas (n=159 com dado). No recorte interseccional, estudantes de baixa renda e racialmente sub-representados concluem na mesma proporção, investindo mais tempo de plataforma (mediana 409 vs 358 min; diferença de velocidade pequena mas significativa, $p=0,011$) — resultado semelhante nesta coorte, sob desigualdade de condição; a ausência de diferença detectada não permite atribuir causalmente esse padrão à mediação por IA. Testada também a dimensão sociológica de Dunn & Dunn (preferência declarada por estudar em grupo vs. sozinho), ela não prediz a conclusão (em grupo 27,2% vs sozinho 30,8%; Fisher $p=0,39$).

Nenhuma das dimensões de perfil testadas resiste à correção de múltiplas comparações (Tabela 5): o sinal mais forte é material — o equipamento (só-celular vs computador) — mas mesmo ele não sobrevive ($p=0,010$; $p_{BH}=0,094$). O papel do convívio social aparece apenas pela via qualitativa (as entrevistas apontam comunidade, aulas ao vivo e monitoria como o que reteve), não como preditor quantitativo.

Conjunto de preditores	AUC (CV)	Veredito
Demografia / socioeconômico	0,52–0,57	Ausente a fraco; insuficiente para seleção
Personalidade (Kahler/EEPC/DISC)	0,49–0,53	Próximo ao acaso (ilusão de taxa-base)
Estilos de aprendizagem (Kolb/Felder/Dunn)	0,49–0,57	Ausente a fraco; sem robustez após correção
Disponibilidade declarada	0,51	Próximo ao acaso
Ativação inicial: conclusão do onboarding	0,64	Moderado; útil para priorizar suporte, não excluir

Tabela 4a. Capacidade preditiva por conjunto de características (AUC em validação cruzada).

Grupo	Conclusão ≥ 1 curso (%)	IC95%	n
Fez o onboarding	45,0	40,9–49,1	556
Não fez	13,6	9,8–18,6	235
Razão de proporções	≈ 3,3×	—	—

--	--	--	--

Tabela 4. Onboarding (proxy: conclusão do assessment inicial) e conclusão de ≥ 1 curso.

Dimensão (família)	Teste	p	p (BH)	Contraste
Equipamento (só celular vs computador)	Fisher	0,010	0,094	21,1% vs 36,6%
Raça (PPI vs Branca/Amarela)	Fisher	0,060	0,269	24,9% vs 32,5%
Gênero (M vs F)	Fisher	0,148	0,444	23,3% vs 30,2%
Liderança	Qui-quadrado	0,355	0,648	4 níveis
Dunn & Dunn (grupo vs solo)	Fisher	0,392	0,648	27,2% vs 30,8%
Disponibilidade declarada	Mann-Whitney	0,432	0,648	med 2,5 vs 2,5
Renda (escala 1–6)	Mann-Whitney	0,507	0,652	med 3,0 vs 3,0
Kahler (analítico-dominante)	Fisher	0,849	0,876	29,0% vs 28,0%
Felder-Silverman	Qui-quadrado	0,876	0,876	16 níveis

Tabela 5. Perfil × conclusão dos três idiomas: testes por dimensão, com p brutos e corrigidos por Benjamini-Hochberg (família de 9 testes; respondentes ao assessment). Nenhuma dimensão sobrevive à correção (p_BH < 0,05).

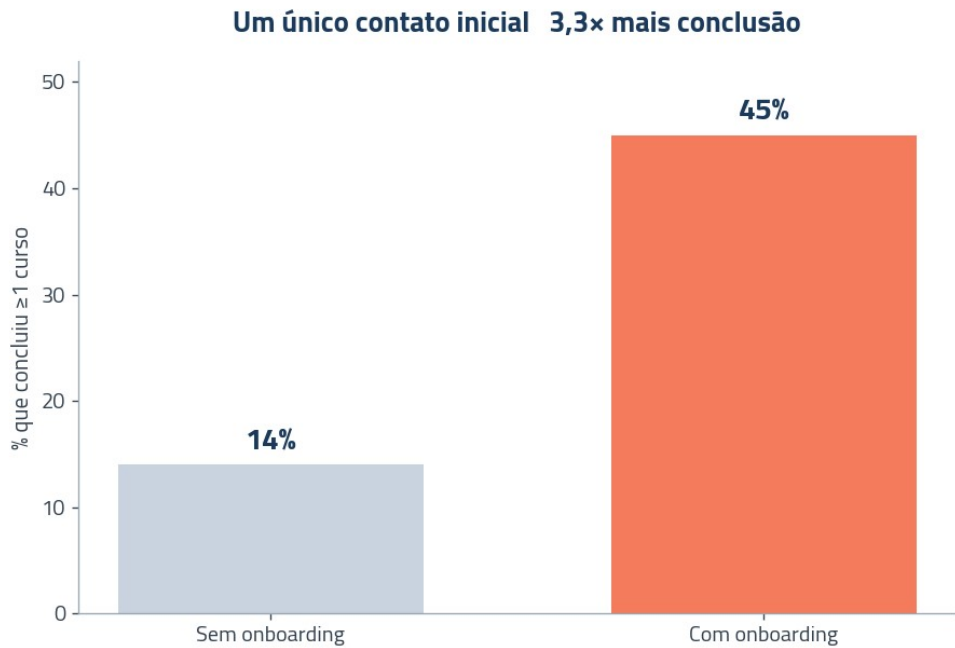


Figura 3. Conclusão de pelo menos um curso conforme realização do onboarding (3,3×).

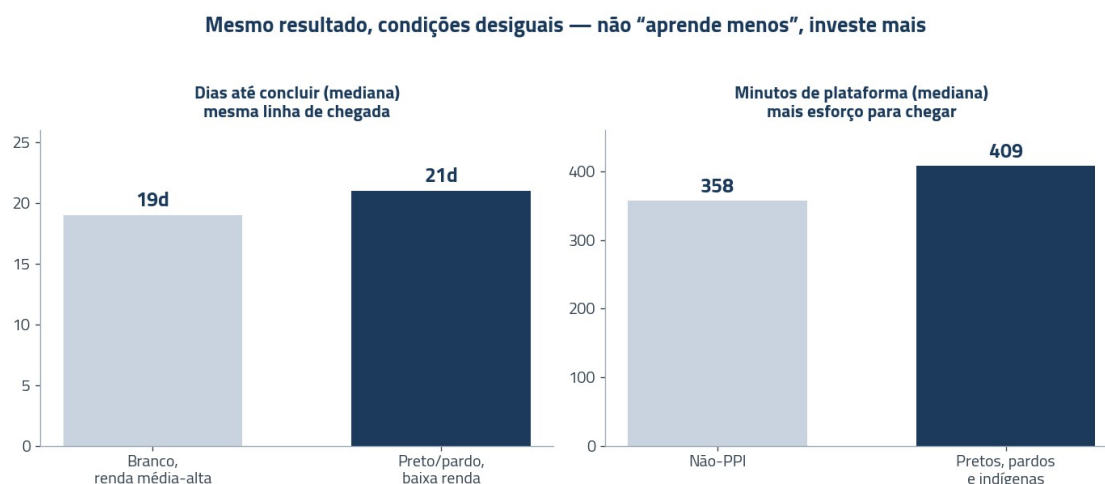


Figura 4. Mesma conclusão, mais esforço: dias até concluir e minutos de plataforma por grupo.

Evidência qualitativa

As entrevistas com os três grupos (concluiu; iniciou e não concluiu; nunca iniciou) convergem com os números e dão-lhes textura. Três temas dominam. Primeiro, *condição, não capacidade*: quem não concluiu cita tempo, trabalho e prioridades concorrentes — nunca incapacidade de aprender. Segundo, *desconhecimento de recursos*: parte dos que largaram não sabia da existência de monitoria, microcursos ou plantão de dúvidas — uma falha de comunicação, não de aluno. Terceiro, *o social como retentor*: quem permaneceu credita à comunidade, às aulas ao vivo e à monitoria a continuidade. As falas abaixo (anonimizadas) ilustram os três casos.

“Sempre quis estudar inglês, mas não pude pagar o curso. Trabalho, pago a faculdade de enfermagem, estudei em escola pública. Abracei a oportunidade — foi apaixonante.”

— Concluiu os três idiomas

“Experiência muito boa, mas tive outras prioridades com a graduação e não consegui finalizar a tempo. Queria ter mais tempo, porque gosto de estudar na plataforma.”

— Concluiu, após o prazo

“A plataforma é fácil, mas não concluí — por falta de tempo e queda de foco. Não sabia que existiam os microcursos.”

— Começou e não concluiu

A leitura qualitativa é coerente com a quantitativa: o que separa trajetórias não é um atributo de perfil, mas a interação entre condição material e desenho do programa (comunicação inicial, andaime social, adequação ao celular).

PARTE III

Implicações e teoria de mudança

Esta parte discute as implicações nos planos micro, meso e macro, traduz os achados em recomendações para empresas, setor público e terceiro setor, propõe uma leitura do programa Empowering Talent como mecanismo de identificação de participantes com persistência e conclusão demonstradas — a teoria de mudança do programa — e reconhece limitações e a agenda de pesquisa futura.

Discussão: micro, meso e macro

Micro (o aluno). Nesta coorte, os atributos de perfil observados na entrada não foram suficientemente informativos para prever a conclusão ou justificar seleção. Perfilar para selecionar não reduz risco — arrisca excluir talento sem ganho preditivo demonstrado: filtros baseados em atributos de perfil podem produzir falsos negativos, pessoas excluídas apesar de poderem concluir, como o caso do bolsista trabalhador de escola pública que concluiu os três idiomas.

Meso (programa e PME). O achado aponta desenho como hipótese prioritária a testar: onboarding ativador nas primeiras semanas e andaime social para quem precisa dele. Dialoga com Bastani et al. (2025): assim como guard-rails podem decidir se a IA ajuda ou atrapalha, o desenho de engajamento é candidato a decidir se a licença vira capacidade ou custo — associação observada nesta coorte, ainda não efeito causal testado. Para uma PME, é a diferença potencial entre treinamento que adere e dinheiro desperdiçado, a confirmar com teste de desenho.

Macro (política e inclusão bi-regional). Nesta coorte, não identificamos disparidades estatisticamente robustas de conclusão por renda, raça ou gênero após correção para múltiplas comparações — o que não elimina desigualdades de acesso e de condição material, nem permite atribuir o padrão à mediação por IA — coerente com a economia da linguagem (Azam et al., 2013) e com a abordagem das capacidades (Sen, 1999). Tratar proficiência como infraestrutura pública de inclusão é viável sem selecionar beneficiários por perfil preditivo — uma agenda comum a Europa e América Latina.

Os três níveis conversam entre si. No *micro*, o custo de perfilar não é simétrico: cada falso negativo — um talento descartado na entrada por um sinal que não prediz — é invisível para quem seleciona, mas definitivo para quem foi excluído. No meso, o paralelo com Bastani et al. (2025) é sugestivo: assim como os ‘guard-rails’ pedagógicos podem decidir se a IA ajuda ou atrapalha, o desenho de engajamento (onboarding ativador, andaime social, mobile sem atrito) é candidato a decidir se a licença vira capacidade ou custo afundado — associação a testar, não efeito causal estabelecido nesta coorte; é uma variável de gestão, não de tecnologia, que merece desenho experimental próprio. No *macro*, tratar proficiência como capability (Sen, 1999) e reconhecer que o talento se distribui converte ‘focar recursos nos mais promissores’ — comum sob restrição fiscal — numa política regressiva: ela concentra investimento

onde o retorno marginal de inclusão é menor. Nesta coorte, a ativação inicial foi mais informativa que os atributos de perfil testados. Para programas de inclusão, isso favorece ampliar acesso e usar sinais comportamentais precoces para oferta de suporte, não como atalho para exclusão.

Há também uma leitura de equidade algorítmica. O temor recorrente é que a IA aplicada à educação reproduza desigualdades. Nossos dados sugerem o oposto sob uma condição de desenho: quando o sinal de decisão é comportamental e conquistado em largada equalizada — e não um perfil herdado —, sinais comportamentais produzidos durante o programa podem *revelar* persistência e adesão que credenciais de entrada não capturam; sua validade para desempenho profissional continua a ser hipótese, não resultado deste estudo. O risco de viés migra do dado para o desenho: ele reaparece se o desfecho for definido por velocidade (penalizando quem tem piores condições materiais) em vez de conclusão.

Isso não significa abandonar dados de perfil — significa *separar os usos*. A crítica ao perfilamento vale para *seleção*, não para monitorar equidade ou desenhar suporte (Tabela 6).

Uso de dados de perfil	Recomendação	Justificativa
Selecionar quem recebe acesso	Evitar, salvo critérios sociais explícitos de elegibilidade	Sem poder preditivo suficiente para justificar exclusão
Personalizar suporte e comunicação	Usar com cautela e consentimento	Pode reduzir fricção sem negar acesso
Monitorar equidade de acesso e resultado	Manter	Necessário para detectar disparidades e orientar intervenção
Decidir automaticamente quem “é promissor”	Não usar	Alto risco de falso negativo, opacidade e reprodução de desigualdade

Tabela 6. Usos de dados de perfil: o que a evidência deste estudo recomenda.

Implicações para política e prática

Os achados convertem-se em recomendações concretas para três públicos. O princípio é comum: **não selecionar o beneficiário por perfil preditivo; investir no engajamento precoce e no pertencimento; gerir adesão por sinais de engajamento; monitorar equidade por recortes demográficos; e nunca usar perfil como atalho para excluir**. Abaixo, o que isso significa em cada arena.

Para empresas e PMEs

Para distribuição de bolsas e acesso inicial a programas comparáveis, os atributos de perfil testados não demonstraram poder preditivo suficiente para justificar exclusão. Valide localmente qualquer regra de pré-seleção antes de usá-la, e priorize sinais comportamentais para suporte, não para exclusão.

Concentre orçamento e atenção nas duas primeiras semanas: teste, por experimento ou rollout escalonado, intervenções de onboarding (primeiro acesso guiado e uma primeira pequena vitória em até 48h), medindo o efeito incremental sobre ativação e conclusão antes de escalar como padrão.

Trate comunidade, monitoria e encontros ao vivo como componentes de suporte e hipóteses de retenção a testar — não como custo acessório a cortar.

Garanta funcionamento pleno no celular, onde boa parte da força de trabalho de menor renda estuda.

Mude o KPI de gestão: acompanhe a *taxa de ativação em 14 dias* (preditiva) em vez de recortes de perfil; reporte conclusão por coorte. Em uma frase: o retorno da IA não vem de escolher quem tem perfil para concluir, vem de testar desenhos que ajudem mais gente a concluir.

Para o setor público

Para o Estado, a capacitação linguística e digital é *infraestrutura de inclusão produtiva* (capability, no sentido de Sen): destrava trabalho, renda e mobilidade. As implicações de desenho são diretas — sobretudo sob restrição fiscal, quando cresce a tentação de ‘focar recursos nos mais promissores’.

Órgãos públicos não devem usar modelos preditivos de perfil não validados como critério automático de exclusão. A evidência mostra que isso não melhora resultado e tende a excluir os mais vulneráveis — um risco concreto de discriminação algorítmica. Critérios sociais e territoriais de elegibilidade podem ser apropriados quando têm base legal, finalidade pública clara e justificativa proporcional.

Gerencie adesão por engajamento; monitore equidade por demografia. Indicadores de ativação precoce e de conclusão por coorte permitem agir a tempo e prestar contas sem rotular pessoas; recortes demográficos servem para *monitorar equidade do acesso*, não para selecionar.

Em compras públicas (Lei nº 14.133/2021), traduza isso no termo de referência: requisitos de desenho de engajamento (onboarding, andaime social, mobile), metas de ativação/conclusão por coorte e previsão de *avaliação de impacto independente* — em vez de exigir ‘seleção dos mais aptos’ ou prometer taxas de conclusão sem desenho de retenção.

Trate baixa adesão como falha de comunicação a corrigir (busca ativa, reenvio de convites, canais múltiplos), não como triagem natural: parte de quem ‘não aderiu’ apenas não recebeu ou não entendeu o convite.

Em programas mediados por tecnologia, indicadores de equidade devem acompanhar não apenas conclusão, mas também ativação, acesso a dispositivo, conectividade, tempo investido, permanência e resultados posteriores. A ausência de disparidade detectada nesta coorte não dispensa esse monitoramento — uma agenda comum à Europa e à América Latina, e alinhada aos ODS 4, 5, 8 e 10.

Para o terceiro setor e financiadores

Não use atributos de perfil não validados para antecipar quem concluirá ou para restringir acesso a bolsas. Amplie o acesso quando o objetivo for inclusão, e teste intervenções de ativação, suporte e pertencimento, medindo efeitos sobre adesão, conclusão e equidade antes de escalá-las.

Teste desenhos de pertencimento desde o primeiro dia (cohorts, conversação, monitoria, comunidade ativa): nas entrevistas qualitativas, foi o fator mais citado como retenção, sobretudo por quem precisa de convívio para aprender — ainda não medido quantitativamente como efeito causal.

Reporte impacto como identificação de participantes com persistência e conclusão demonstradas, não como seleção: métricas de engajamento e conclusão por coorte, histórias de trajetória e vínculo explícito com os ODS contam a história certa para o financiador — e a mantêm honesta.

Da evidência à ação: a teoria de mudança

Os achados têm uma implicação que ultrapassa o desenho de programas e toca a forma como o mercado identifica talento. A seleção convencional opera *ex ante*, por perfil estático (currículo, instituição de origem, sinais demográficos). Nesta coorte, os atributos de perfil avaliados apresentaram desempenho preditivo ausente a fraco para a conclusão dentro do prazo ($AUC \approx 0,49-0,57$) e, por excluir candidatos sem acesso prévio, tende a reproduzir desigualdade. Um sprint de aprendizagem, ao contrário, produz um sinal *ex post*, comportamental e conquistado em condições de partida equalizadas (nível A0 declarado): o engajamento e a conclusão observados — não atributos presumidos.

Sob essa leitura, o programa social Empowering Talent pode ser entendido não apenas como intervenção de inclusão, mas como um mecanismo de identificação de participantes com persistência e conclusão demonstradas: oferece a oportunidade a todos e deixa o comportamento demonstrado — o sinal mais informativo encontrado, com desempenho moderado — operar como indicador, conectando participantes com persistência demonstrada a empregadores (matchmaking). É a recomendação derivada dos dados ('medir engajamento precoce, não demografia') convertida em modelo operacional, descrita formalmente na teoria de mudança da organização (Figura 5).

Ressalva sobre validade preditiva

Importa delimitar o escopo da inferência: esta pesquisa estabelece que o engajamento prediz a *conclusão da capacitação* (sinal operacional de adesão e persistência, não medida de ganho linguístico), e **não** o desempenho profissional subsequente. A hipótese de que a performance no sprint se transfere para o trabalho é plausível e teoricamente ancorada (capacidade como sinal), mas permanece *não testada* — constitui o objeto do braço longitudinal do Estudo 2 (acompanhamento de carreira dos participantes com persistência demonstrada no programa). A leitura de matchmaking deve, portanto, ser tratada como hipótese de pesquisa, não como resultado validado.



Figura 5. Teoria de mudança do Empowering Talent: da largada equalizada à conexão com o mercado. As métricas centrais do modelo (conclusão por coorte; acurácia do sinal preditivo) são precisamente as estabelecidas nesta pesquisa.

Registre-se, como evidência preliminar de viabilidade operacional — não como prova de eficácia ou de validade preditiva para desempenho profissional —, que o próprio Lingopass recrutou cinco participantes do programa, residentes em periferias de três estados (SP, BA, MG), alinhado aos ODS 4, 5, 8 e 10.

Conflito de interesse: a autora é cofundadora e CEO do Lingopass, que opera o programa; esta seção aplica os achados, não os produz.

Limitações

Estudo observacional de uma coorte (2023), um programa, um parceiro: associação, não causa; generalização limitada.

O efeito do onboarding pode refletir, em parte, motivação latente — só um experimento separa desenho de disposição.

Subgrupos interseccionais pequenos ($n \approx 89$); informative missingness mitigada, não eliminada.

A medida de engajamento usa o auto-estudo na plataforma (EAD), critério oficial de conclusão; as aulas ao vivo de FR/ES (Zoom) não eram rastreáveis por aluno e entram como evidência de programa ($\approx 90h$), não individual.

A janela temporal das features de engajamento não foi documentada de forma independente do desfecho. Tempo de plataforma, percentual de atividades e dias até concluir podem se sobrepor parcialmente ao período em que o próprio desfecho de conclusão se define; o modelo de engajamento deve ser lido como caracterização da trajetória observada, não como preditor validado prospectivamente com janela de congelamento estanque.

Pesquisa futura: do caso único ao estudo multi-site

Para testar validade externa, **propomos estender** o desenho a outras populações beneficiárias do programa social (refugiados, ensino técnico público e profissionais em transição de carreira), pareando bolsista × desfecho. As hipóteses de pesquisa são: (H1) o perfil apresenta desempenho preditivo ausente a fraco em todas as populações; (H2) o engajamento é o preditor dominante; (H3) controlado o engajamento, raça/renda não predizem; (H4) o efeito sociológico é moderado pelo grau de andaime social de cada coorte. A estratégia prevista combina replicar o pipeline por site, um modelo multinível com efeito aleatório de organização, pré-registro e um braço causal (experimento de onboarding). Trata-se de agenda proposta, ainda não executada.

Conclusão

Este working paper examinou, em uma coorte de capacitação linguística apoiada por IA, se atributos de perfil disponíveis na entrada predizem a conclusão dentro do prazo. Nesta coorte, os atributos avaliados — perfil demográfico, socioeconômico, psicométrico e de estilo de aprendizagem — apresentaram desempenho preditivo ausente a fraco; o único sinal precoce que mostrou associação com quem conclui foi comportamental: a ativação no onboarding. Essa virada — de perfil para ativação inicial — não é apenas um resultado estatístico: aponta uma hipótese concreta a testar, com desenho e medição adequados, em vez de tentar prever e selecionar o aprendiz ‘certo’ — reduzir o atrito das duas primeiras semanas e sustentar o pertencimento de quem já está dentro.

A mesma leitura sustenta a proposta de inclusão: como a desigualdade não apareceu nos desfechos de conclusão observados — apenas no tempo investido para chegar lá —, ampliar o acesso e testar melhorias de desenho pode ajudar a identificar participantes com persistência e conclusão demonstradas, em vez de depender de credenciais de quem já teve acesso prévio. É essa combinação — evidência de que o perfil não prediz a conclusão, e de que a ativação inicial mostrou associação com ela — que orienta a proposta de tratar o desempenho no programa, e não a origem do aprendiz, como sinal a testar mais a fundo.

Sobre a autora

Alexandrine Bami é cofundadora e CEO do Lingopass, empresa de infraestrutura linguística presente em mais de trinta países, que atende grandes organizações privadas e o setor público. Formada pela ENS Paris-Saclay, é mestre em Ciências Políticas pela Sciences Po Paris. Francesa radicada no Brasil desde 2002, atua na articulação entre os ecossistemas de inovação europeu e latino-americano. Sua produção na série Lingopass Research trata de adoção de tecnologia, aprendizagem organizacional e mensuração de capacidade.

Declaração de conflito de interesse & rigor

A autora é cofundadora e CEO do Lingopass, que opera o programa estudado. O interesse é declarado. Justamente por isso, todo o desenho prioriza o rigor: dados pseudonimizados e divulgados de forma agregada, modelos validados fora da amostra, correção de múltiplas comparações e reporte explícito do que é desfavorável (35% nunca acessaram; gap de esforço por condição). O documento tem finalidade de discussão pública e aplicação profissional; os resultados, limitações e métodos são apresentados para escrutínio.

Sobre a série — Lingopass Research é a série de working papers do Lingopass dedicada a quantificar ativos do capital humano em mercados globais. Este é o N° 003.

Aviso legal — Documento para fins de discussão e de prática profissional. Estudo observacional de uma coorte; associação não implica causa.

Como citar

Brami, A. (2026). *Quem conclui uma capacitação apoiada por IA? Engajamento, não perfil*. Lingopass Research, Working Paper N° 003 (versão 1.0, setembro de 2026). Apresentado na Round Table 6 do 6° Colóquio Internacional CERALE–FGV EBAPE, Rio de Janeiro.

Referências (APA)

- Azam, M., Chin, A., & Prakash, N. (2013). The returns to English-language skills in India. *Economic Development and Cultural Change*, 61(2), 335–367.
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *PNAS*, 122, e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, 44(3), 211–233.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate. *Journal of the Royal Statistical Society: Series B*, 57(1), 289–300.
- Bleakley, H., & Chin, A. (2004). Language skills and earnings: Evidence from childhood immigrants. *Review of Economics and Statistics*, 86(2), 481–496.
- Bloom, B. S. (1984). The 2 sigma problem. *Educational Researcher*, 13(6), 4–16.
- Casale, D., & Posel, D. (2011). English language proficiency and earnings in a developing country: The case of South Africa. *The Journal of Socio-Economics*, 40(4), 385–393.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). Learning styles and pedagogy in post-16 learning: A systematic and critical review. LSRC.
- Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment. *The Internet and Higher Education*, 2(2–3), 87–105.
- Gurung, A., Lin, J., Gutterman, J., Thomas, D. R., et al. (2025). Human tutoring improves the impact of AI tutor use on learning outcomes. *AIED* 2025.
- Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist*, 48(3), 169–183.
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, 86(1), 42–78.
- Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P.-M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. *npj Science of Learning*, 10(1), 29.
- Nancekivell, S. E., Shah, P., & Gelman, S. A. (2020). Maybe they're born with it... the learning style myth. *Journal of Educational Psychology*, 112(2), 221–235.
- Newton, P. M. (2015). The learning styles myth is thriving in higher education. *Frontiers in Psychology*, 6, 1908.
- Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9(3), 105–119.
- Reich, J., & Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science*, 363(6423), 130–131.
- Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2015). Matching learning style to instructional method. *Journal of Educational Psychology*, 107(1), 64–78.

Sen, A. (1999). Development as freedom. Oxford University Press.

Stanford SCALE Initiative (2026). The evidence base on AI in K-12: A 2026 review. Stanford University.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221.

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, 331(6023), 1447–1451.

Wang, R. E., Ribeiro, A. T., Robinson, C. D., Loeb, S., & Demszky, D. (2024). Tutor CoPilot: A human-AI approach for scaling real-time expertise. arXiv:2410.03017.

Apêndice A. Perfis dos clusters

A Tabela A1 descreve a composição de cada cluster. O padrão é revelador: a composição *demográfica* é praticamente plana entre os grupos (a fração de pretos, pardos e indígenas e a de baixa renda quase não variam de Fantasmas a Maratonistas), enquanto os indicadores de *engajamento* crescem de forma monotônica (onboarding, tempo de plataforma e percentual de atividades). Em outras palavras: o que separa os grupos não é quem são, é o quanto se engajaram.

Cluster	n	% Fem.	% PPI	% Baixa renda	% Onboarding	Tempo (mediana, min)	% Ativ. (mediana)
Fantasma	276	75,4	45,2	32,5	45,7	0	0,0
Curioso	125	67,8	53,3	34,4	72,0	10	0,7
Constante	136	77,9	50,4	36,3	83,1	212	18,8
Maratonista	254	79,3	45,4	31,3	89,4	1.078	100,0

Tabela A1. Composição dos clusters comportamentais (percentuais sobre a base com dado; tempo em minutos).

Apêndice B. Notas sobre não-resposta e robustez

Dos 791 alunos, 235 (29,7%) não responderam ao assessment de entrada; esse grupo concentra-se entre os Fantasmas e Curiosos, confirmando que a não-resposta é informativa (proxy de desengajamento). Por isso os testes de perfil (Tabelas 3 e 5) restringem-se aos 556 respondentes. As variáveis de escolaridade, geração universitária, idade e parentalidade estão disponíveis apenas para a subamostra de 159 que completou o assessment estendido — daí descreverem o perfil dos high performers (ver Resultados), e não servirem para testar conclusão (sem variação entre concluintes e não-concluintes). A solução de clusters é estável entre k=2 e k=4 (silhueta ≈0,72) e as conclusões preditivas não mudam entre regressão logística e Random Forest nem entre definições de desfecho.

Materiais. Perfis de cluster, tabelas completas de perfil × conclusão com intervalos e p corrigidos, testes de velocidade de conclusão, a reconstrução da energia analítica de Kahler a partir das 18 forças e a síntese das entrevistas qualitativas estão disponíveis nos arquivos do projeto, mediante solicitação e respeitadas as medidas de pseudonimização, minimização e divulgação agregada previstas no protocolo de proteção de dados.

Apêndice C. Depoimentos (anonimizados) e análise temática

Esta seção reúne os depoimentos coletados nas entrevistas. Os depoimentos foram pseudonimizados e minimizados para publicação: foram removidos identificadores diretos (nome, e-mail, telefone, instituição) e reduzidas combinações de atributos que pudessem elevar o risco de reidentificação; a localização é dada em nível de *macrorregião* e a ocupação em nível de *área de formação*, com granularidade reduzida para mitigar o risco de reidentificação — o risco residual não pode ser eliminado por completo em amostras qualitativas. A amostra entrevistada foi estratificada pelos três desfechos (concluiu; iniciou e não concluiu; nunca iniciou) mais um grupo que fez os microcursos.

Análise temática

A codificação dos depoimentos faz emergir cinco temas, todos convergentes com os resultados quantitativos.

1. Condição, não capacidade. Sem exceção, quem não concluiu atribui o desfecho a tempo, trabalho e prioridades acadêmicas concorrentes (provas, graduação) — nunca a dificuldade de aprender. É a tradução qualitativa do achado central: o desfecho não está em quem a pessoa é, mas nas condições e no desenho.

2. O social e o ao vivo como âncora. As aulas ao vivo, a monitoria semanal e a comunidade no WhatsApp são, de longe, o ‘mais gostei’ recorrente — sustentando, pela via qualitativa, o papel do convívio que a variável quantitativa de perfil não captou (ver Evidência qualitativa).

3. Lacuna de descoberta de recursos. Parte dos que não concluíram simplesmente não sabia que certos recursos existiam (‘não sabia que existiam os microcursos’) ou não os usou por falta de tempo — uma falha de comunicação e ativação, não de aluno.

4. Barreira material do celular. Um relato aponta que estudar predominantemente pelo celular esbarra na escassez de atividades adequadas ao formato — eco qualitativo do gradiente de equipamento observado na Tabela 5 (só-celular conclui menos).

5. Idioma como aspiração de mobilidade. A motivação é explicitamente econômica e de carreira — trabalhar fora do país, atuar com idiomas, estudar no exterior —, coerente com a leitura do idioma como capability (ver O que a ciência já sabia).

Uma ressalva honesta tempera o tema 2: o convívio só retém quando bem conduzido — um depoimento registra desconforto com a condução de um encontro ao vivo. O social é alavanca, não garantia; depende de execução.

Depoimentos

Bolsista · Sul · Engenharia Civil · trilhas FR/ES/EN · alta dedicação

“Apaixonada por idiomas; achei o projeto desafiador para conciliar. Adorei a parte de pronúncia com feedback e as aulas ao vivo, que deram dimensão além da plataforma. Usei as dicas dos microcursos no LinkedIn. Queria que o projeto tivesse continuação.”

— Entrevista, 2023

Bolsista · Sudeste · Enfermagem · trilhas FR/ES/EN · alta dedicação

“Sempre quis estudar inglês, mas não pude pagar o curso — trabalho e pago a faculdade de enfermagem. Estudei em escola pública. Abracei a oportunidade, foi apaixonante. Gostei muito dos encontros semanais. Não consegui fazer os microcursos por conta do trabalho.”

— Entrevista, 2023

Bolsista · Nordeste · graduação · foco em francês · concluiu após o prazo

“Experiência muito boa, mas tive outras prioridades com a graduação e não consegui finalizar no prazo. Queria ter mais tempo, porque gosto de estudar na plataforma. As aulas ao vivo foram muito importantes.”

— Entrevista, 2023

Bolsista · Nordeste · foco em espanhol · alta dedicação

“Gostei muito dos cursos; a plataforma é fácil de usar. O melhor foram as aulas ao vivo, o material dinâmico e a monitoria toda quarta-feira. Gostaria de mais conteúdo e exercícios nos microcursos voltados a LinkedIn e empregos.”

— Entrevista, 2023

Bolsista · Sudeste · Tradução/Interpretação · trilhas FR/ES/EN · iniciou

“O projeto passou rápido demais. A monitoria de quarta tirava as dúvidas e gostei da personalização. Mas uso muito o celular e faltam atividades adequadas ao celular. Não concluí os microcursos: prioridade foi o trabalho.”

— Entrevista, 2023

Bolsista · Sudeste · foco em espanhol · iniciou e não concluiu

“A plataforma é objetiva e fácil, mas não concluí — por falta de tempo, planejamento e queda de foco. Não sabia que existiam os microcursos.”

— Entrevista, 2023

Apêndice D. Desempenho preditivo do onboarding e resultados completos

Um sinal só deve informar gestão se seus erros forem conhecidos. Detalhamos o desempenho do onboarding como classificador de "concluir ≥ 1 curso" e a tabela completa de testes de perfil.

Matriz de confusão e métricas

(tratando "fez onboarding" como previsão de conclusão; $n=791$): sensibilidade 88,7%, especificidade 39,9%, precisão (VPP) 45,0%, valor preditivo negativo 86,4%, acurácia 57,3% e AUC 0,64 — melhor que o acaso, insuficiente para decisão automatizada de alto impacto.

	Fez onboarding	Não fez
Concluiu ≥ 1	250 Verdadeiro positivo	306 Falso positivo
Não concluiu	32 FALSO NEGATIVO	203 Verdadeiro negativo

Figura D1. Matriz de confusão: onboarding $\times$ conclusão de ≥ 1 curso ($n=791$).

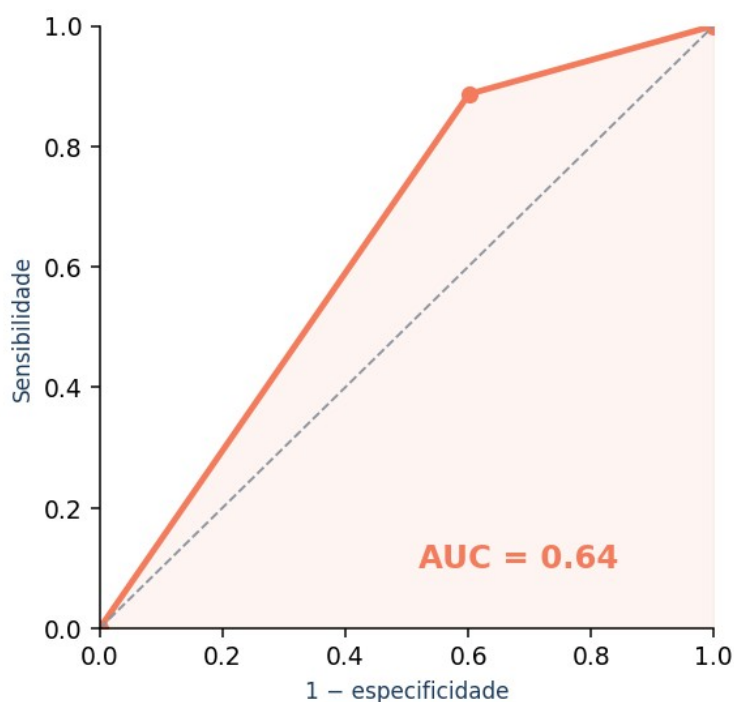


Figura D2. Curva ROC do onboarding (AUC 0,64).

O custo de usar onboarding como filtro (falsos negativos). 32 alunos concluíram ao menos um curso sem ter feito o onboarding — 11,3% de todos os concluintes. Um corte que barrasse quem não fez onboarding excluiria esse talento real. É a tradução operacional do argumento de inclusão: o onboarding informa risco de evasão; não deve virar critério de exclusão.

Razão de conclusão. Fez onboarding 45,0% vs não fez 13,6% > razão de proporções 3,3× (IC95% 2,4–4,6); odds ratio 5,2 (IC95% 3,4–7,8).

Perfil × conclusão dos três idiomas

Tabela completa (respondentes; correção de Benjamini-Hochberg na família de 9 testes). Nenhuma dimensão sobrevive à correção.

Dimensão (família)	Contraste	Teste	p	p (BH)	Sobrevive?
Equipamento (só-celular vs computador)	21,1% vs 36,6%	Fisher	0,010	0,094	Não
Raça (PPI vs Branca/Amarela)	24,9% vs 32,5%	Fisher	0,060	0,269	Não
Gênero (M vs F)	23,3% vs 30,2%	Fisher	0,148	0,444	Não
Liderança	4 níveis	Qui-quadrado	0,355	0,648	Não
Dunn & Dunn (grupo vs solo)	27,2% vs 30,8%	Fisher	0,392	0,648	Não
Disponibilidade declarada	med 2,5 vs 2,5	Mann-Whitney	0,432	0,648	Não
Renda (escala 1–6)	med 3,0 vs 3,0	Mann-Whitney	0,507	0,652	Não
Kahler (analítico-dominante)	29,0% vs 28,0%	Fisher	0,849	0,876	Não
Felder-Silverman	16 níveis	Qui-quadrado	0,876	0,876	Não

Tabela D1. Perfil × conclusão dos três idiomas: p brutos e corrigidos (Benjamini-Hochberg, família de 9). Nenhuma dimensão sobrevive.

Governança de dados e ética. Os dados foram pseudonimizados para análise (identificadores substituídos por chaves técnicas), com divulgação apenas em forma agregada e avaliação de risco de reidentificação; uso restrito a análise interna e a esta apresentação, com permissão do parceiro institucional. A documentação completa de governança — base legal específica, papéis de controlador e operador, canal de oposição/retirada e medidas de segurança além da pseudonimização — está em consolidação para as submissões acadêmicas e a órgãos públicos.