



Instituto Walden4

COLETÂNEA DE PESQUISA

Entre Ciência e Cuidado

IA & TEA



Inteligência Artificial e Transtorno do Espectro Autista

Organização

Márcio Borges Moreira

Maria Eduarda Cardoso de Sousa

Cláudio Rocha Ferrari

Ana Luíza Cardoso Cerqueira

Ana Victória Filomeno Gomes

João Vítor Mansilha Silva

Valéria Costa França

Ana Luíza Antonino Valim Maia

Editora do Instituto Walden4

Brasília, DF · 2026

**Dados Internacionais de Catalogação na Publicação (CIP)
(Câmara Brasileira do Livro, SP, Brasil)**

Entre ciência e cuidado [livro eletrônico] :
inteligência artificial e Transtorno do
Espectro Autista TEA / organização Márcio
Borges Moreira. -- Brasília, DF : Instituto
Walden4, 2026. (Coletânea de pesquisa)
PDF

Bibliografia
ISBN 978-85-65721-59-2

1. Autismo 2. Autismo - Diagnóstico 3.
Inteligência artificial 4. Neurodiversidade 5.
Psicologia clínica 6. TEA (Transtorno do Espectro do
Autismo)
7. Tecnologia I. Moreira, Márcio Borges. II. Série.

26-368989.0

CDD-616.85882

Índices para catálogo sistemático:

1. TEA : Transtorno do Espectro Autista :
Neurodiversidade 616.85882

Maria Alice Ferreira - Bibliotecária - CRB-8/7964

Todos os direitos reservados. Nenhuma parte desta publicação pode ser reproduzi-
da, armazenada em sistema de recuperação ou transmitida de qualquer forma ou por
qualquer meio sem autorização prévia por escrito do editor.



A Editora do Instituto Walden4 se dedica à promoção e difusão de conhecimentos avançados na esfera da Análise do Comportamento, assim como em outras áreas práticas baseadas em evidências, abrangendo tanto aspectos científicos quanto profissionais. Comprometida com a democratização do acesso à informação, oferecemos uma ampla gama de nossos livros de forma gratuita. Além disso, todos os nossos títulos estão acessíveis em formato digital, proporcionando a você a conveniência de iniciar a leitura em segundos. Descubra e explore as publicações do nosso catálogo que mais lhe interessam.

Conselho Editorial

Dr. Gleidson Gabriel da Cruz

Dr. Márcio Borges Moreira

Contato

secretaria@walden4.com.br

[@instituto.walden4](https://www.instagram.com/instituto.walden4)

<https://www.instagram.com/instituto.walden4>

<https://www.walden4.com.br>

<https://www.facebook.com/iwalden4>

<https://www.youtube.com/user/instwalden4>

Acesso Livre, Cultura Sustentável: Uma Reflexão sobre o Valor de Apoiar os Criadores

Este livro, desde sua concepção, foi idealizado para ser uma obra gratuita, com o propósito de democratizar o acesso ao conhecimento e difundir ideias sem barreiras econômicas. No entanto, é justamente o compromisso com a continuidade dessa missão que nos leva a uma reflexão necessária sobre a sustentabilidade da produção cultural.

Valorização dos Criadores de Conteúdo: A Sustentabilidade da Produção Cultural

A distribuição gratuita de conteúdo digital representa um avanço extraordinário na democratização do conhecimento. Contudo, a continuidade dessa produção depende, em larga medida, do reconhecimento e do apoio financeiro aos profissionais que dedicam seu tempo e talento ao setor criativo. Acessibilidade e viabilidade econômica não são forças opostas, mas dimensões que precisam coexistir para que obras de qualidade sigam sendo criadas.

A Cadeia Produtiva Editorial

A elaboração de uma obra literária mobiliza uma rede especializada de profissionais, cada um contribuindo com competências distintas e insubstituíveis. Os escritores desenvolvem narrativas e refinam suas técnicas ao longo de anos de dedicação. Os designers gráficos concebem a experiência visual da obra, traduzindo seu conteúdo em capas e diagramações que comunicam e encantam. Os revisores asseguram a precisão linguística e a fluidez do texto, zelando pela clareza de cada página. Os ilustradores enriquecem o conteúdo com elementos visuais que ampliam a experiência do leitor. E os diretores editoriais articulam todos esses esforços, coordenando o conjunto para alcançar o máximo de qualidade e impacto.

Impacto Econômico e Benefícios Estratégicos

A aquisição de obras originais sustenta toda essa cadeia de profissionais e suas famílias, mantendo viva e dinâmica a economia criativa. Em troca, o leitor recebe um produto de qualidade superior, fruto de processos editoriais rigorosos e cuidadosamente conduzidos. Mais do que isso: cada aquisição fortalece a indústria cultural, permite que os criadores destinem recursos adequados ao desenvolvimento de novos projetos e estimula a inovação, ao reconhecer financeiramente o valor do trabalho intelectual.

Investimento Cultural

Adquirir uma obra original é mais do que uma transação comercial — é um investimento no desenvolvimento cultural da sociedade. Trata-se de um gesto concreto de apoio ao progresso intelectual coletivo e de participação ativa na manutenção de um ecossistema que gera conhecimento, criatividade e desenvolvimento.

Em última análise, a sustentabilidade da produção de conteúdo de qualidade repousa sobre o equilíbrio entre o livre acesso e a justa remuneração de quem o produz. Ao apoiar os criadores por meio da aquisição de suas obras, o leitor beneficia a si mesmo, com produtos de excelência, e contribui para algo maior: a vitalidade cultural de toda a sociedade.

Acesso ao Conhecimento e o Valor de Quem o Produz: Uma Palavra ao Leitor

Ao adquirir este livro, você fez mais do que comprar uma obra: você se tornou parte de um ecossistema que torna possível a criação e a difusão do conhecimento. Esta é uma reflexão sobre o valor desse gesto e sobre tudo o que ele sustenta.

Toda obra de qualidade nasce do trabalho dedicado de profissionais que investem tempo, talento e estudo em sua realização. Ao optar por uma obra original, o leitor reconhece esse esforço e contribui diretamente para que a produção cultural continue viável e vibrante. Acessibilidade e viabilidade econômica não são forças opostas, mas dimensões que precisam coexistir para que obras de qualidade sigam sendo criadas.

SOBRE O AUTOR

Márcio Borges Moreira



Márcio Borges Moreira é Doutor em Ciências do Comportamento pela Universidade de Brasília, Mestre em Psicologia e Psicólogo pela Pontifícia Universidade Católica de Goiás. É professor da graduação e professor colaborador do Mestrado em Psicologia do Centro Universitário

de Brasília (UniCEUB), onde supervisiona estágios clínicos no Centro de Formação de Psicólogos (CENFOR). É também diretor do Instituto Walden4, clínica especializada em serviços baseados em ABA para crianças e adolescentes com Transtorno do Espectro Autista (TEA).

Coautor do livro *Princípios Básicos de Análise do Comportamento*, possui capítulos e artigos científicos publicados em periódicos nacionais e internacionais, com foco em Análise do Comportamento Aplicada, comportamento verbal e intervenções baseadas em evidências. Seus projetos de pesquisa atuais investigam a automação de treinamentos para profissionais e paraprofissionais que atuam com pessoas diagnosticadas com autismo, o uso de tecnologia e inteligência artificial no ensino de conceitos psicológicos, a formação e mudança de atitudes preconceituosas e a mensuração de comportamentos na prestação de serviços de psicologia.

ACOMPANHE PELAS MÍDIAS SOCIAIS

@marcioborgesmoreira

Facebook: facebook.com/professormarciomoreira Instagram: instagram.com/marcioborgesmoreira YouTube: youtube.com/user/borgesmoreirayt SlideShare: slideshare.net/borgesmoreira LinkedIn: linkedin.com/in/marcio-borges-moreira Lattes: lattes.cnpq.br/4094892880820475

Sumário

	Apresentação	7
1	Efeito de Diferentes Prompts na Qualidade das Respostas do Gemini 2.5 Pro sobre Sinais Precoces de Autismo	18
2	Avaliação da qualidade de um jogo de RPG mestrado por I.A.S para o ensino de habilidades sociais de adolescentes autistas	51
3	Avaliação da Qualidade de Histórias Sociais Digitais para Autorregulação Emocional em Crianças com TEA usando GEM Storybook	80
4	Avaliação do Efeito do Nível de Prompt na Precisão e Utilidade de um aplicativo tutorial gerado por Chatbot de IA no Apoio a Pais de Crianças com TEA: Uso do DRA para Prevenção e Manejo de Comportamentos Disruptivos	127
5	Efeito de diferentes níveis de prompt na adequação de aplicativos educacionais de inteligência artificial (IA) para o ensino de alfabetização a crianças autistas na educação básica	160
6	Efeito do Nível de Estruturação de Prompts na Qualidade de Tutoriais Gerados por IA sobre Intervenções em Autismo	182
7	Comparação da qualidade técnica e científica de aplicativo gerado por Inteligência Artificial para ensinar crianças autistas a discriminarem suas emoções de acordo com o contexto	235

CAPÍTULO DE ABERTURA

Entre Ciência e Cuidado: o que são as Inteligências Artificiais, por que elas importam para o autismo e qual é a inovação deste livro



Márcio Moreira

Centro Universitário de Brasília — UniCEUB

1. Uma cena cotidiana

São duas horas da manhã. Uma mãe, com o celular na mão, digita: “meu filho de dois anos não olha quando eu chamo, é autismo?”. Há poucos anos, essa pergunta seria respondida por uma página de resultados do Google: dez links azuis, alguns anúncios, um portal de saúde, um fórum de pais, talvez um vídeo. Hoje, é cada vez mais provável que essa mesma mãe faça a pergunta a um chatbot de inteligência artificial — ChatGPT, Gemini, Claude ou outro — e receba, em segundos, um texto fluente, organizado, empático, que parece ter sido escrito por um especialista exclusivamente para ela.

Essa cena não é hipotética. Mais de 90% dos cuidadores de crianças pequenas utilizam a internet e aplicativos móveis para lidar com dúvidas parentais e desafios diários (Kim et al., 2025), e os pais de crianças com Transtorno do Espectro Autista (TEA) figuram entre os que mais buscam informação online, justamente porque o diagnóstico precoce é determinante para intervenções eficazes (McFayden et al., 2024). A pergunta que organiza este livro nasce exatamente desse encontro entre a urgência do cuidado e a novidade da tecnologia: podemos confiar no que essas máquinas produzem quando o assunto é autismo? E, mais do que isso, o que exatamente elas são capazes de produzir?

Este capítulo de abertura tem três objetivos. Primeiro, explicar, em linguagem acessível a quem nunca estudou computação, como funciona uma inteligência artificial generativa como o ChatGPT. Segundo, compará-la com algo que todos conhecemos bem — a pesquisa do Google — para deixar claro o que mudou e por que essa mudança exige avaliação científica. Terceiro, apresentar a inovação central das pesquisas reunidas neste volume: a descoberta de que essas IAs não produzem apenas textos, mas também aplicativos funcionais, e o teste sistemático dessa capacidade a serviço de pais, professores e profissionais que cuidam de pessoas autistas.

2. O que é, afinal, uma IA como o ChatGPT?

2.1 Uma máquina de prever a próxima palavra

O ChatGPT (sigla de Chat Generative Pre-Trained Transformer) é um modelo de linguagem de grande porte — em inglês, Large Language Model, ou LLM — lançado pela empresa OpenAI em novembro de 2022, treinado para integrar e processar informações e gerar respostas semelhantes às humanas a partir de texto digitado pelo usuário (Hassona & Alqaisi, 2024). Por trás da aparência de “conversa”, porém, há um mecanismo surpreendentemente simples de enunciar: o modelo aprende, a partir de quantidades gigantescas de texto, quais palavras costumam aparecer juntas, e usa esse conhecimento estatístico para prever a próxima palavra, uma palavra de cada vez (Sampaio et al., 2024).

Quem usa smartphone já conviveu com uma versão rudimentar dessa ideia: o teclado que sugere a próxima palavra enquanto digitamos. Se você escreve “bom”, o celular sugere “dia”; se escreve “feliz”, sugere “aniversário”. O ChatGPT faz algo da mesma natureza, mas em escala incomparavelmente maior: em vez de prever uma palavra com base nas duas ou três anteriores, ele considera todo o contexto da conversa e foi treinado com bilhões de exemplos de texto humano.

Vale acompanhar esse processo em câmera lenta, palavra por palavra, porque é nele que está todo o “segredo” da máquina. A rigor, o modelo não trabalha exatamente com palavras, mas com tokens — representações numéricas de palavras ou, mais precisamente, de partes de palavras, que podem ser processadas com mais eficiência (Revista Esfera Acadêmica Tecnologia, 2024); para fins didáticos, porém, podemos pensar em palavras. Imagine que o usuário digitou apenas:

“Meu filho”

Nesse instante, o modelo calcula uma distribuição de probabilidades sobre todas as palavras que conhece para a posição seguinte. Com tão pouco contexto, o leque é amplo: “tem” e “está” recebem probabilidades altas (digamos, na casa dos 15–20% cada), “gosta”, “não” e “nasceu” recebem valores intermediários, e palavras como “fotossíntese” ou “parafuso” recebem probabilidades praticamente nulas — simplesmente porque, nos bilhões de textos do treinamento, quase ninguém jamais escreveu “meu filho fotossíntese”. O modelo então sorteia uma palavra dessa distribuição, dando mais chance às mais prováveis — é a chamada amostragem estocástica, que usa aleatoriedade para selecionar elementos conforme uma distribuição de probabilidade (Revista Esfera Acadêmica Tecnologia, 2024). Suponha que saia “tem”:

“Meu filho tem”

Agora algo importante aconteceu: a palavra escolhida passa a fazer parte do contexto e redistribui todas as probabilidades seguintes. Depois de “tem”, tornam-se prováveis números (“dois”, “três”, “quatro”) e palavras como “dificuldade”, “medo” ou “febre”; já “está” e “gosta”, que eram fortes candidatas um passo atrás, agora têm probabilidade quase zero, porque “meu filho tem está” não é português. Cada palavra gerada estreita o caminho das próximas. Suponha que o sorteio produza “dois”, e depois “anos”:

“Meu filho tem dois anos”

Note o efeito em cascata: “dois” tornou “anos” quase inevitável (talvez 90% de probabilidade — “meses” e “irmãos” disputam o resto), e “anos”, por sua vez, abre um novo leque: “e” (para continuar a frase) torna-se a aposta mais forte. Continuemos: “e”, depois “ainda”, depois “não”:

“Meu filho tem dois anos e ainda não”

A essa altura, o contexto acumulado exerce um controle poderosíssimo sobre o que vem a seguir. Em textos sobre crianças de dois anos seguidos de “ainda não”, as continuções esmagadoramente mais frequentes no material de treinamento são “fala” e “anda” — e, como aos dois anos a maioria das crianças já anda, “fala” tende a receber a maior probabilidade. O modelo escreve “fala”, e a frase completa emerge:

“Meu filho tem dois anos e ainda não fala.”

Três lições didáticas saem desse exemplo em câmera lenta. Primeira: em nenhum momento o modelo “decidiu” falar sobre desenvolvimento da linguagem; cada palavra foi apenas a consequência probabilística das anteriores, e é graças ao mecanismo de atenção do Transformer que palavras distantes na frase — como “dois anos”, lá atrás — continuam influenciando a escolha da palavra atual (Sampaio et al., 2024). Segunda: como há um sorteio a cada passo, a mesma pergunta nunca gera exatamente a mesma resposta — se, no terceiro passo, tivesse saído “três” em vez de “dois”, ou “engatinha” em vez de “fala”, toda a continuação seria outra; é por isso que esses modelos nunca repetem exatamente as respostas (Hacker et al., 2023, conforme citado em Sampaio et al., 2024). Terceira, e mais importante para este livro: o processo otimiza plausibilidade, não verdade. Se, em algum passo, a sequência mais provável for uma estatística desatualizada ou uma referência bibliográfica com cara de referência, o modelo a escreverá com a mesma fluência e a mesma confiança — fenômeno a que voltaremos na seção 2.2.

Sampaio et al. (2024) propõem que, para compreender os LLMs, é necessário entender três conceitos. O primeiro é o aprendizado de máquina: um campo da inteligência artificial que desenvolve algoritmos capazes de aprender padrões a partir de dados e tomar decisões sem programação humana direta — o programador não escreve regras do tipo “se o usuário perguntar X, responda Y”; o sistema extrai os padrões sozinho, a partir de exemplos. O segundo é o método Transformer, uma técnica de análise textual baseada no conceito de atenção, na qual cada elemento do texto é representado em relação a todos os outros, permitindo que termos sejam relacionados entre si independentemente da distância que os separa no texto. Foi o Transformer que permitiu a explosão de tamanho dos modelos de linguagem e os tornou generalizáveis: antes, um chatbot útil precisava ser treinado para um assunto específico; agora, um mesmo modelo serve aos mais diversos objetivos. O terceiro conceito são os dados massivos: o ChatGPT foi construído e treinado com vastas quantidades de textos publicados na internet — notícias, livros, páginas da web (Sampaio et al., 2024; Revista Esfera Acadêmica Tecnologia, 2024).

Há ainda uma etapa final, menos conhecida do público: o ajuste fino com feedback humano. Depois do treinamento básico, pessoas avaliam respostas do modelo, e essas preferências são usadas para construir um “modelo de recompensa” que ensina o sistema a responder de modo mais útil, seguro e adequado ao formato de conversa (Revista Esfera Acadêmica Tecnologia, 2024). Para o leitor analista do comportamento, a analogia é tentadora — e os capítulos deste livro a exploram com o devido cuidado conceitual: trata-se de selecionar respostas por suas consequências, ainda que o “organismo” em questão seja um sistema estatístico, e não um ser que aprende ao longo da interação.

2.2 O papagaio que erra com confiança

Uma imagem usada na literatura ajuda a calibrar as expectativas: o ChatGPT e outros LLMs não são muito diferentes de um papagaio que tenta repetir o que aprendeu após doses massivas de treinamento, elaborando palavra por palavra com base na probabilidade de um humano ter utilizado aquela sequência (Hacker et al., 2023, conforme citado em Sampaio et al., 2024). Exatamente por isso, esses modelos nunca repetem exatamente as mesmas respostas e, com alguma frequência, erram — produzindo as chamadas alucinações: respostas sintaticamente perfeitas e convincentes, mas factualmente falsas (Sampaio et al., 2024).

Esse ponto é tão importante para o restante do livro que merece um exemplo concreto. Quando McFayden et al. (2024) pediram ao ChatGPT referências bibliográficas sobre autismo, menos da metade dos links fornecidos funcionava corretamente; alguns levavam a páginas inexistentes ou a sites não relacionados. O conteúdo das respostas era, em geral, correto, claro e conciso — mas as fontes citadas para sustentá-lo eram, com frequência, imprecisas ou inventadas. A máquina não “mente” no sentido humano: ela apenas gera a sequência de palavras mais provável, e uma referência com aparência de referência é, estatisticamente, uma boa próxima palavra. Para um pai aflito ou um profissional em formação, contudo, a diferença entre uma referência real e uma plausível-porém-falsa pode ser a diferença entre uma decisão bem informada e um erro de cuidado.

A esses riscos somam-se os vieses: por serem treinados em dados produzidos por humanos, os modelos tendem a reproduzir premissas e escolhas presentes nesses dados e nas decisões de seus desenvolvedores, frequentemente sem transparência sobre os algoritmos e bases utilizados (Sampaio et al., 2024). McFayden et al. (2024) observaram, por exemplo, que o ChatGPT alternava entre uma linguagem mais médica e outra alinhada à perspectiva afirmativa da neurodiversidade — revelando que até o vocabulário escolhido carrega as marcas do material de treinamento.

3. ChatGPT não é Google: duas lógicas diferentes de encontrar informação

Para muita gente, o chatbot parece apenas “um Google melhorado”. A comparação é compreensível, mas imprecisa, e entender a diferença é essencial para usar bem as duas ferramentas.

No paradigma estabelecido pelo Google, uma pesquisa resulta em uma série de links para páginas. O buscador não escreve a resposta: ele localiza e ordena documentos que já existem, criados por pessoas e instituições identificáveis, e cabe ao usuário clicar, ler, comparar fontes e julgar. Esse modelo tem seus próprios problemas — sobretudo nos critérios de ordenamento das primeiras páginas —, mas preserva uma característica importante: a resposta vem acompanhada de sua origem, e o desacordo entre fontes fica visível (Sampaio et al., 2024).

As IAs generativas operam em outra lógica. Em vez de prover links, os chatbots providenciam diretamente respostas ao questionamento original, e o aprofundamento se dá conversando com a máquina, não navegando entre páginas. Como observam Sampaio et al. (2024), isso provavelmente nos dá respostas mais rápidas e ágeis para perguntas específicas, porém diminui sensivelmente o contato com os debates em torno da questão pesquisada: a tendência é que a primeira resposta da máquina seja vista como “a verdade”, sem que os dissensos e controvérsias apareçam. No exemplo da mãe às duas da manhã: o Google a obrigaria a transitar entre um portal médico, um texto de associação de famílias e um fórum — esforço maior, mas também mais oportunidades de perceber nuances. O chatbot lhe entrega um único texto coeso, sem assinatura, sem data e, muitas vezes, sem fontes verificáveis.

Podemos resumir o contraste assim:

Quadro 1. Contraste entre a busca tradicional e a IA generativa

	Busca tradicional (Google)	IA generativa (ChatGPT e similares)
O que devolve	Lista de links para conteúdos existentes	Um texto novo, gerado na hora
Quem é o autor	Pessoas e instituições identificáveis	O próprio modelo, recombinação de padrões aprendidos
Papel do usuário	Selecionar, ler, comparar e julgar fontes	Conversar, refinar a pergunta, confiar (ou não) na síntese
Transparência da fonte	Visível (cada link tem origem)	Baixa; referências podem estar ausentes ou ser fabricadas
Risco característico	Resultados mal ordenados, sites de baixa qualidade	Alucinações apresentadas com fluência e confiança
Variabilidade	A mesma busca tende a retornar os mesmos links	A mesma pergunta gera respostas diferentes a cada vez

Nota. Síntese comparativa das duas lógicas de obtenção de informação: a busca tradicional (que devolve uma lista de links a consultar) e a IA generativa (que produz diretamente uma resposta em texto).

Sampaio et al. (2024) situam o surgimento dessas ferramentas entre as grandes quebras de paradigma do trabalho intelectual — comparável à passagem da máquina de datilografar aos computadores, à introdução da pesquisa online pelo Google e ao surgimento de enciclopédias colaborativas como a Wikipédia. Há, inclusive, uma transformação em curso na própria pesquisa acadêmica: em vez de buscar por palavras-chave combinadas com operadores booleanos, pesquisamos cada vez mais fazendo perguntas, e sistemas baseados em IA respondem com trechos de artigos e sínteses, mantendo o diálogo em torno da questão original.

Nada disso significa que uma ferramenta substitui a outra. Significa que a IA generativa transfere para a máquina uma tarefa que antes era do leitor — sintetizar e redigir — e, ao fazê-lo, esconde o processo. É precisamente por isso que ela precisa ser avaliada com método

científico: quando a síntese vem pronta, bonita e confiante, a única defesa do usuário é saber, com base em evidências, em que condições essa síntese tende a ser boa e em que condições tende a falhar.

4. O que a ciência já sabe: IAs respondendo sobre autismo e apoiando quem cuida

A literatura que fundamenta este livro — e que o leitor encontrará citada e discutida ao longo dos capítulos — já oferece um retrato consistente, com luzes e sombras.

Do lado das luzes, o ChatGPT mostrou-se capaz de fornecer informações majoritariamente corretas, claras e concisas a perguntas frequentes de pais sobre autismo, refutando mitos com precisão — o que o torna um instrumento promissor para combater a desinformação e o estigma (McFayden et al., 2024). Em odontologia para pacientes com necessidades especiais, respondeu a dúvidas de pais sobre saúde bucal no autismo com conteúdo comparável ao de um especialista, demonstrando empatia e oferecendo dicas práticas, como o uso de histórias sociais, dessensibilização e reforçamento positivo (Hassona & Alqaisi, 2024); resultados semelhantes foram obtidos para perguntas de pais sobre cárie na primeira infância, com 82,5% de respostas consideradas úteis ou muito úteis (Elkarmi et al., 2024). Na educação especial, professores que usaram o ChatGPT produziram metas de Planos Educacionais Individualizados (PEI/IEP) de qualidade significativamente superior, em menos tempo e cobrindo mais domínios do desenvolvimento — comunicação, habilidades sociais, motoras e de autocuidado — do que colegas sem acesso à ferramenta (Rakap, 2024; Rakap & Balikci, 2024). Professores de inglês como língua estrangeira na educação especial relataram atitudes moderadamente favoráveis e percepção de efetividade da ferramenta (Alenizi et al., 2023), e uma revisão sistemática brasileira identificou ganhos de personalização do ensino e redução de carga de trabalho docente, ao lado de limitações importantes (Revista Esfera Acadêmica Tecnologia, 2024). Na formação de profissionais, LLMs já são usados para criar simulações e roleplays nos quais estudantes praticam, em ambiente seguro, conversas difíceis — como acolher um cuidador emocionalmente fragilizado ou responder com ética a desinformação sobre vacinas e dietas — recebendo feedback imediato sobre seu desempenho (Wilson & Karsina, 2025; Mollick & Mollick, 2024). E há evidência inicial de que pais mentorados pelo ChatGPT na estimulação de linguagem de crianças autistas de 2 a 5 anos obtiveram ganhos significativamente maiores no vocabulário dos filhos do que pais sem essa mentoria (SEAMEO ICSE, 2025).

Do lado das sombras, três achados se repetem. Primeiro, a baixa aplicabilidade prática: as respostas tendem a ser conceitualmente corretas, porém pouco acionáveis — explicam o que é, mas raramente dizem o que fazer, passo a passo (McFayden et al., 2024; Elkarmi et al., 2024). Segundo, a inconsistência das referências, já discutida acima. Terceiro, a legibilidade inadequada: as respostas frequentemente exigem nível de leitura de ensino médio ou superior, muito acima do recomendado para materiais de educação em saúde dirigidos a leigos (Hassona & Alqaisi, 2024; Elkarmi et al., 2024; Kim et al., 2025). Kim et al. (2025) acrescentam

uma quarta limitação, particularmente sensível para a parentalidade: as respostas das IAs carecem de ênfase na conexão entre pais e filhos e de sensibilidade ao estágio de desenvolvimento da criança.

Atravessando luzes e sombras, um achado metodológico unifica toda essa literatura e estrutura este livro: a qualidade do que a IA produz depende criticamente da qualidade do comando que recebe — o chamado prompt. McFayden et al. (2024) mostraram que simplesmente pedir à máquina para “regenerar” a resposta não a melhora; o que melhora a resposta é a forma como a pergunta é elaborada. Para o analista do comportamento, essa constatação tem um sabor familiar, explorado conceitualmente nos capítulos que seguem: o prompt funciona, operacionalmente, como uma regra verbal que especifica contingências e direciona o comportamento verbal do modelo, de modo que o grau de estruturação do prompt pode ser tratado como variável independente, e a qualidade do produto gerado, como variável dependente mensurável.

5. A inovação deste livro: IAs não fazem só textos — fazem aplicativos

Até aqui, falamos de IAs que escrevem. Mas há algo que o público em geral ainda desconhece e que constitui o coração da contribuição original deste volume: os mesmos chatbots que redigem textos são capazes de escrever código de programação e, com isso, construir aplicativos completos e funcionais — com telas, botões, atividades interativas, registro de respostas e feedback — a partir de instruções dadas em português comum, por uma pessoa que nunca programou na vida.

Essa capacidade já vinha sendo notada na literatura. Sampaio et al. (2024) observam que, com as IAs redigindo e revisando código, a exigência de programação competente tende a diminuir consideravelmente, permitindo que mais pesquisadores acessem possibilidades antes restritas a especialistas; descrevem ainda ferramentas em que o ChatGPT executa análises de dados sozinho a partir de uma planilha enviada pelo usuário. Mollick e Mollick (2024), por sua vez, argumentam que professores podem deixar de ser meros consumidores de tecnologia educacional para se tornarem construtores e inovadores, criando com IA seus próprios exercícios, simuladores e tutores — uma democratização do desenvolvimento de tecnologia educacional. O que faltava era submeter essa promessa ao crivo da avaliação científica em um domínio onde errar custa caro: o cuidado de pessoas autistas.

É exatamente isso que as sete pesquisas deste livro fazem. Em vez de perguntar apenas “a IA responde bem sobre autismo?”, elas perguntam: “a IA constrói boas tecnologias para o autismo — tutoriais, histórias sociais digitais, jogos e aplicativos de ensino — e o que precisamos fazer para que construa melhor?”. Cada capítulo manipula sistematicamente o nível de estruturação do prompt (de um pedido leigo e genérico a um comando tecnicamente detalhado, por vezes acompanhado de um artigo científico anexado como base) e avalia os produtos gerados com rubricas explícitas — inspiradas, entre outras, nos critérios de correção, clareza e

concisão de McFayden et al. (2024) —, incluindo a verificação da autenticidade das referências citadas e da adequação dos produtos aos princípios da Análise do Comportamento Aplicada (ABA) e às características do TEA.

A relevância dessa agenda é dupla. Do ponto de vista social, intervenções baseadas em evidências para o TEA esbarram em custos, tempo e escassez de serviços especializados; aplicativos e tutoriais gerados por IA, se tiverem qualidade verificada, podem funcionar como ferramentas complementares acessíveis e imediatas no cotidiano das famílias — jamais como substitutos do profissional, ressalva que todos os estudos aqui reunidos fazem questão de sustentar (McFayden et al., 2024; Kim et al., 2025; Rakap & Balikci, 2024; Hassona & Alqaisi, 2024). Do ponto de vista científico, a agenda dialoga com uma exigência clássica da Análise do Comportamento: para cumprir sua função social, uma prática precisa ser não apenas eficaz, mas tecnológica — descrita com precisão suficiente para ser replicada de forma independente (Baer et al., 1968, conforme retomado nos capítulos deste volume). Se uma IA é capaz de transformar um artigo científico em um aplicativo ou tutorial utilizável por qualquer cuidador, ela se torna, potencialmente, uma ponte inédita entre o conhecimento produzido nos laboratórios e a vida das famílias. Testar a solidez dessa ponte — verificando onde ela sustenta o peso e onde ainda balança — é a grande inovação de pesquisa apresentada neste livro.

6. O que o leitor encontrará nos capítulos

Os sete capítulos compartilham a mesma espinha dorsal metodológica — manipulação de níveis de prompt como variável independente, avaliação criteriosa dos produtos como variável dependente, controle da “ingenuidade experimental” do modelo por meio de contas novas e conversas independentes — aplicada a problemas distintos do cuidado em TEA:

O Capítulo 1 avalia a qualidade das respostas do Gemini 2.5 Pro na orientação de pais sobre sinais precoces de autismo em crianças de 16 a 30 meses. O Capítulo 2 examina um jogo de RPG mestrado por IA para o ensino de habilidades sociais a adolescentes autistas. O Capítulo 3 analisa histórias sociais digitais geradas com IA para autorregulação emocional infantil. O Capítulo 4 investiga o efeito do nível de prompt na precisão e utilidade de um aplicativo tutorial para apoiar pais no uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA) na prevenção e manejo de comportamentos disruptivos. O Capítulo 5 avalia aplicativos educacionais gerados por IA para alfabetização de crianças autistas na educação básica. O Capítulo 6 mede o efeito da estruturação de prompts na qualidade de tutoriais sobre intervenções em autismo. O Capítulo 7, por fim, compara a qualidade técnica e científica de aplicativos gerados por IA para ensinar crianças autistas a discriminar emoções de acordo com o contexto.

7. Como ler este livro

Este livro foi escrito para públicos que raramente se encontram na mesma estante: pais e cuidadores que querem entender o que estão usando quando conversam com uma IA; professores e profissionais da educação especial e da clínica que se perguntam se — e como — incorporar essas ferramentas à sua prática; e estudantes e pesquisadores interessados em um modelo metodológico replicável para avaliar produtos de IA.

A todos, propomos a mesma postura que a literatura recomenda aos pesquisadores: nem deslumbramento, nem pânico, mas criticidade informada — equilibrar potencialidades e limitações (Sampaio et al., 2024). As IAs generativas já demonstraram que podem informar com correção, apoiar profissionais, mentorar pais e, como este livro documenta, construir aplicativos. Também já demonstraram que alucinam referências, falam difícil demais para quem mais precisa e entregam pouca orientação prática. Entre a ciência e o cuidado, o caminho não é escolher um dos lados, e sim fazer o que as pesquisas aqui reunidas fazem: colocar a tecnologia à prova, com método, antes de colocá-la nas mãos de quem cuida.

Boa leitura.

Referências

- Alenizi, M. A. K., Mohamed, A. M., & Shaaban, T. S. (2023). Revolutionizing EFL special education: How ChatGPT is transforming the way teachers approach language learning. *Innoeduca. International Journal of Technology and Educational Innovation*, 9(2), 5–23. <https://doi.org/10.24310/innoeduca.2023.v9i2.16774>
- Elkarmi, R., Abu-Ghazaleh, S., Sonbol, H., Haha, O., AL-Haddad, A., & Hassona, Y. (2024). ChatGPT for parents' education about early childhood caries: A friend or foe? *International Journal of Paediatric Dentistry*, 35(4), 717–724. <https://doi.org/10.1111/ipd.13283>
- Hassona, Y., & Alqaisi, D. (2024). “My kid has autism”: An interesting conversation with ChatGPT. *Special Care in Dentistry*, 44(4), 1296–1299. <https://doi.org/10.1111/scd.12983>
- Kim, J., et al. (2025). Estudo comparativo da qualidade de respostas de IA generativa a perguntas de pais e cuidadores, em comparação com o boletim Just in Time Parenting — JITP.info. (Referência completa conforme arquivo do projeto: kim_etal_2024.pdf.)
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*. <https://doi.org/10.1089/cyber.2023.0202>
- Mollick, E., & Mollick, L. (2024). Instructors as innovators: A future-focused approach to new AI learning opportunities, with prompts. Wharton School of the University of Pennsylvania.
- Rakap, S. (2024). Chatting with GPT: Enhancing Individualized Education Program goal development for novice special education teachers. *Journal of Special Education Technology*, 39(3), 339–348. <https://doi.org/10.1177/01626434231211295>

- Rakap, S., & Balikci, S. (2024). Enhancing IEP goal development for preschoolers with autism: A preliminary study on ChatGPT integration. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-024-06343-0>
- Revista Esfera Acadêmica Tecnologia. (2024). Revisão sistemática sobre o impacto do ChatGPT no suporte aos professores de educação especial. *Revista Esfera Acadêmica Tecnologia*, 9(1). ISSN 2526-4141. (Completar autoria conforme o arquivo do projeto: REVISTA_ESFERA_TECNOLOGIA_V09_N01_ARTIGO3.pdf.)
- Sampaio, R. C., et al. (2024). Ensaio sobre inteligência artificial generativa, ChatGPT e a pesquisa científica. *Revista de Sociologia e Política*, 32. (Referência completa conforme arquivo do projeto: sampaio_etal.pdf.)
- SEAMEO ICSE. (2025). Using ChatGPT to mentor parents in enhancing language development for children with ASD (2–5 years old). *Proceedings of the International Conference on Special Education*, 6, 220–237. SEAMEO Regional Centre for Special Educational Needs. <https://zenodo.org/record/16325585>
- Wilson, J. B., & Karsina, A. (2025). Uso de role-plays com LLMs na formação de pós-graduandos em Análise do Comportamento Aplicada. *Behavior Analysis in Practice*. (Referência completa conforme arquivo do projeto: Wilson_etal_2024.pdf.)
- Outras obras citadas indiretamente neste capítulo (Hacker et al., 2023; Baer, Wolf & Risley, 1968) o foram conforme citadas nas fontes acima e nos capítulos deste volume.

CAPÍTULO

01

Efeito de Diferentes Prompts na Qualidade das Respostas do Gemini 2.5 Pro sobre Sinais Precoces de Autismo



Maria Eduarda Cardoso de Sousa

Marcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. Este estudo tem como objetivo avaliar e comparar a qualidade técnica e científica das respostas fornecidas pelo modelo de inteligência artificial Gemini 2.5 Pro, quando utilizado para orientar pais na identificação de sinais precoces de Transtorno do Espectro Autista (TEA) em crianças de 16 a 30 meses.

Método. Trata-se de uma pesquisa qualitativa, de natureza exploratória e documental. Foram testados três prompts com diferentes níveis de especificidade (dúvida geral, relato de sintomas e solicitação técnica baseada no M-CHAT-R/F). As respostas foram coletadas em outubro de 2025 e avaliadas por meio de uma rubrica sistemática de quatro critérios: precisão científica e conformidade (baseada no M-CHAT-R/F), clareza e linguagem, aplicabilidade prática e qualidade das referências.

Resultados. A análise indicou que o Gemini 2.5 Pro apresentou alta conformidade (escore 4/4) em precisão científica, clareza e aplicabilidade prática em todos os prompts testados, demonstrando total conformidade com a literatura médica e ausência de “alucinações”. A qualidade de referências bibliográficas variou conforme a especificidade do prompt, sendo superior quando o instrumento técnico foi explicitamente solicitado.

Conclusão. Conclui-se que a IA é uma ferramenta auxiliar viável e precisa para a psicoeducação parental, capaz de traduzir critérios técnicos em linguagem acessível. No entanto, a engenharia de prompt influencia a qualidade das fontes citadas e a supervisão de um profissional especializado permanece indispensável para o diagnóstico.

Palavras-chave: Transtorno do Espectro Autista; Inteligência Artificial; M-CHAT-R/F; Diagnóstico Precoce; Educação Parental

1 INTRODUÇÃO

Contextualização do Tema

Ao longo dos anos, o Transtorno do Espectro Autista (TEA) tem recebido atenção crescente, impulsionando pesquisas e busca por auxílio psicológico e diagnóstico. Segundo Goulart et al. (2002), o TEA é caracterizado por prejuízos comportamentais nas áreas de interação social, comunicação e pela presença de padrões de comportamentos restritivos e repetitivos, elementos ainda considerados centrais na compreensão do transtorno. A jornada para a obtenção de um diagnóstico preciso frequentemente se inicia com a preocupação dos pais sobre os marcos do desenvolvimento de seus filhos, um processo que pode ser complexo e gerar incertezas, demandando orientação qualificada e acesso a informações confiáveis.

Nesse contexto, a identificação precoce do TEA é clinicamente fundamental. Um diagnóstico precoce e preciso possibilita que a criança e a sua família tenham acesso a intervenções essenciais o mais cedo possível, podendo influenciar positivamente os resultados a longo prazo (Alckmin-Carvalho et al., 2014; Vllasaliu et al., 2016). Além do mais, o aumento do interesse científico sobre o autismo tem contribuído para o aprimoramento de técnicas que permitem uma detecção cada vez mais precisa e antecipada, muitas vezes antes dos três anos de idade, reforçando a necessidade de desenvolver ferramentas de apoio para essa identificação inicial.

Em paralelo aos desafios enfrentados pelas famílias, vivemos um momento de grande avanço tecnológico, impulsionado pela onipresença das tecnologias de informação e comunicação (TICs) e pela maior separação geográfica das redes de apoio familiar tradicionais (Kim et al., 2025). Nesse cenário, tecnologias de IA demonstram um potencial transformador em diversas áreas, abrindo novas fronteiras para a criação de soluções inovadoras e acessíveis (Sampaio et al., 2024).

Problema de Pesquisa

A crescente procura dos pais por informações online é um fenômeno bem documentado. E um estudo recente investigou a qualidade das respostas da IA generativa, como o ChatGPT, as perguntas comuns sobre a parentalidade. Os resultados indicaram que a IA é capaz de fornecer respostas razoavelmente boas, corretas e claras para questões gerais. Contudo, a mesma pesquisa revelou desvantagens notáveis: as respostas da IA frequentemente carecem de profundidade na dinâmica relacional entre pais e filhos (a “conexão”), ignoravam detalhes específicos do contexto de desenvolvimento infantil e apresentavam um nível de leitura elevado, o que pode dificultar o acesso para alguns públicos (Kim et al., 2025)

Apesar do potencial transformador da inteligência artificial generativa (Sampaio et al., 2024), e dos desafios enfrentados pelas famílias que buscam informações online, as limitações identificadas (Kim et al., 2025) indicam que não se pode confiar em todas as informações disponíveis nas IAs. É necessário, portanto, que sejam feitos testes para compreender a veracidade e a qualidade dessas informações.

Pergunta de Pesquisa

Considerando então que os pais buscam cada vez mais o auxílio das Inteligências Artificiais (IAs), e que respostas das IAs podem ser erradas ou sem embasamento teórico, qual será a precisão das respostas de uma IA, considerando o critério de diferentes prompts, para ajudar pais a identificar sinais de autismo na faixa etária de 16 a 30 meses.

Relevância da Pesquisa

O diagnóstico precoce do TEA é um fator determinante para intervenções eficazes e para a melhora dos desfechos no desenvolvimento infantil (Vllasaliu et al., 2016; Alckmin-Carvalho et al., 2014). Pais e cuidadores, entretanto, enfrentam barreiras para obter informações de qualidade e, cada vez mais, recorrem a fontes digitais e a modelos de linguagem baseados em IA para orientação inicial (Kim et al., 2025). Pesquisas recentes apontam que ferramentas

como o ChatGPT são capazes de fornecer respostas corretas e claras sobre o autismo, mas ainda apresentam limitações em profundidade, referências e contextualização (McFayden et al., 2024). Avaliar a qualidade técnica e científica de um aplicativo criado por IA especificamente para esse fim é, portanto, essencial para garantir que pais recebam informações seguras, claras e baseadas em evidências, contribuindo para o avanço de práticas de triagem precoce e para políticas públicas de saúde.

Revisão da Literatura

A prevalência do Transtorno do Espectro Autista (TEA) tem aumentado globalmente, reforçando a importância do diagnóstico precoce e de instrumentos de triagem acessíveis. Estudos indicam que a detecção precoce (antes dos três anos de idade) permite intervenções mais efetivas, com impacto positivo a longo prazo (Alckmin-Carvalho et al., 2014; Vllasaliu et al., 2016). A busca precoce por informações pelos pais é, portanto, crucial para a eficácia das intervenções e para um impacto positivo a longo prazo.

Em paralelo, a busca dos pais por informações online é um fenômeno consolidado: mais de 90% dos cuidadores de crianças pequenas utilizam a internet e aplicativos móveis para lidar com dúvidas parentais e desafios diários (Kim et al., 2025). Nesse contexto, a necessidade de maior credibilidade para as IAs torna-se imperativa, dado que os pais frequentemente as utilizam como fonte de apoio e instrução diária. Esse contexto amplia o interesse em soluções digitais de apoio, mas também levanta preocupações sobre a qualidade e a confiabilidade do conteúdo disponível (Plantin & Daneback, 2009, conforme citado em Kim et al., 2025).

A inteligência artificial generativa, como o ChatGPT, tem se mostrado promissora na difusão de informações em saúde. McFayden et al. (2024) avaliaram a precisão das respostas do ChatGPT a perguntas frequentes de pais sobre autismo e constataram alto nível de correção e clareza, embora com limitações na ação prática e na exatidão das referências. Estudos em parentalidade corroboram achados semelhantes: respostas geralmente claras e úteis, mas que carecem de maior contextualização e profundidade (Kim et al., 2025).

No âmbito acadêmico, Sampaio et al. (2024) destacam que grandes modelos de linguagem, apesar de capazes de revisar literatura e gerar textos com qualidade, podem apresentar vieses e “alucinações”, isto é, informações factualmente incorretas apresentadas como verdadeiras. Isso reforça a necessidade de avaliações críticas e comparações com padrões científicos.

Essas evidências sustentam a relevância de investigar a qualidade técnica e científica de aplicativos criados por IA voltados à identificação de sinais de TEA, considerando diferentes idades, a fim de garantir informações seguras, atualizadas e de fácil compreensão para pais e cuidadores.

A pesquisa de Kim et al. (2025) realizou um teste dedutivo com 11 condições experimentais de teste, abrangendo cinco tópicos comuns da parentalidade. Eles compararam as respostas do ChatGPT-3.5 e ChatGPT-4 com uma ferramenta de busca de um site de boletins informativos. Para isso foram geradas cinco perguntas abertas baseadas nos tópicos mais pesquisados no site Just in Time Parenting (JITP.info), que serviu como uma base de dados controlados para comparar com a internet aberta. As respostas foram então codificadas para análise com base em critérios como correção, clareza, conexão, reprodutibilidade e legibilidade.

Já no estudo de McFayden et al. (2024), foi realizada uma pesquisa qualitativa com 13 perguntas abertas direcionadas ao ChatGPT, divididas em três categorias: informações básicas sobre o autismo, mitos e equívocos comuns e busca por recursos. Para cada pergunta, o ChatGPT foi instruído a “regenerar” uma segunda resposta, resultando em duas respostas modificáveis para cada questão. As respostas foram avaliadas com base em múltiplos critérios, incluindo os “3Cs” - Correção, Clareza e Concisão -, O uso da linguagem (médica vs. neuroafirmativa), e a compreensibilidade e acionabilidade, utilizando a ferramenta PEMAT-P (Patient Education Materials Assessment Tool for Printable materials).

Partindo das pesquisas realizadas por Kim et al. (2025) e McFayden et al. (2024), as IAs podem ser um ponto de apoio inicial para os pais buscarem um conhecimento primário após a suspeita de algum transtorno ou de alguma irregularidade no desenvolvimento de seu filho. E essa busca, se direcionada de forma correta pode trazer grandes benefícios para a criança, possibilitando, por exemplo, que um diagnóstico precoce ocorra com mais precisão.

O Transtorno do Espectro Autista (TEA) é um transtorno do neurodesenvolvimento de início precoce, que acarreta mudanças biopsicossociais significativas ao indivíduo (Santos et al., 2025). O seu diagnóstico é eminentemente clínico, baseando-se na observação direta e cuidadosa do comportamento da criança, na realização de entrevistas detalhadas com os pais ou cuidadores e no levantamento do histórico de paciente (Seize & Borsa, 2017). Esse processo avaliativo deve ser amparado pela aplicação de instrumentos padronizados e demarcados pelos critérios do Manual Diagnóstico e Estatístico de Transtornos Mentais, 5ª edição, Revisão de Texto (APA, 2023) ou da Classificação Internacional de Doenças (CID). Segundo a Associação Americana de Psiquiatria (APA, 2023), a confirmação diagnóstica exige a presença de déficits persistentes na comunicação e na interação social em múltiplos contextos, aliados e padrões restritos e repetitivos de comportamentos, interesses ou atividades (Santos et al., 2025; Steffen et al., 2019).

Apesar da conscientização crescente sobre o TEA na sociedade, a efetivação do diagnóstico antes dos três anos de idade ainda esbarra em dificuldades estruturais e técnicas (Santos et al., 2015). Profissionais de saúde, especialmente médicos que atendem a primeira infância, relatam insegurança e dificuldades operacionais para efetuar prontamente o diagnóstico, muitas vezes devido à falta de contato prévio com as manifestações precoces e à escassez de serviços especializados (Santos et al., 2015). Observa-se ainda uma carência de aportes teóricos mais profundos e de qualificação profissional contínua por parte dos avaliadores (Silva & Gomes, 2023). Somado a isso, nos estágios mais precoces, a sintomatologia do autismo pode ser difusa e acompanhada de manifestações clínicas sutis, o que costuma gerar dúvidas e postergar a conclusão diagnóstica (Santos et al., 2015).

Para a família, o reconhecimento dos sinais primários para o devido encaminhamento a um especialista também representa um grande desafio. Muitas das características centrais descritas no DSM-5-TR (APA, 2023) são de difícil observação em bebês (como as falhas na interação com os pares), ou apenas se tornam clinicamente evidentes entre os 3 e 4 anos de idade, como é o caso dos comportamentos e interesses repetitivos (Zanon et al., 2014). Por esse motivo, a suspeita e a busca por assistência ocorrem, na grande maioria das vezes, apenas quando os cuidadores notam um atraso significativo no desenvolvimento da fala (Zanon et al.,

2014). Agrava-se a essa situação o desafio do negacionismo e da normalização dos sintomas no ambiente familiar (Silva & Gomes, 2023). Diante do desconhecimento sobre os marcos normais do desenvolvimento sociocomunicativo, muitos pais tendem a ignorar os primeiros sinais de alerta, justificando os comportamentos da criança como traços naturais de sua personalidade, a exemplo da timidez ou da introversão (Zanon et al., 2014).

Diante dessas lacunas na identificação precoce, tanto por parte de profissionais quanto das famílias, as ferramentas digitais despontam como mecanismos de apoio. Nesse contexto, o direcionamento das Inteligências Artificiais (IAs) para a formulação de orientações e avaliações preliminares pode ser fundamentado por meio de diversos instrumentos diagnósticos e de rastreamento já existentes, os quais possuem robusta comprovação científica de sua eficácia e validade. Dentre essas ferramentas, destacam-se o Modified Checklist for Autism in Toddlers, Revised with Follow-Up (M-CHAT-R/F) — amplamente recomendado para a triagem inicial de crianças de 16 a 30 meses —, o Autism Behavior Checklist (ABC), o Ages and Stages Questionnaires (ASQ), a Childhood Autism Rating Scale (CARS), o Autism Diagnostic Observation Schedule (ADOS) e a Autism Diagnostic Interview-Revised (ADI-R), sendo esta última considerada uma entrevista padrão-ouro no processo de avaliação clínica (Seize & Borsa, 2017; Zanon et al., 2014).

Além disso, no contexto brasileiro, o Conselho Federal de Psicologia (CFP) reconhece o uso de protocolos adicionais que podem servir de base de conhecimento empírico, como a Escala de Avaliação de Traços Autistas (ATA) e o Social Communication Questionnaire (SCQ) (Santos et al., 2025). A integração estruturada dessas escalas cientificamente validadas nos comandos (prompts) das IAs é fundamental, pois garante que as informações de saúde geradas pela tecnologia sejam precisas, traduzíveis para a prática cotidiana dos pais e ancoradas na literatura médica, reduzindo substancialmente o risco de desinformação ou "alucinações" do sistema (Sousa & Moreira, 2025).

Um dos principais instrumentos recomendados para essa triagem precoce é o Checklist Modificado para Autismo em Crianças Pequenas, com versão revisada e consulta de seguimento (M-CHAT-R/F). Validado especificamente para avaliar o risco de TEA em crianças entre 16 e 30 meses de idade, ele é amplamente utilizado na atenção primária e em pesquisas clínicas (Robins et al., 2009). No Brasil, o instrumento conta com uma versão traduzida e adaptada, o que facilita sua adoção por profissionais de saúde e garante a adequação sociocultural das questões apresentadas às famílias.

Objetivo da Pesquisa

Avaliar e comparar a qualidade técnica e científica das respostas fornecidas por um aplicativo de inteligência artificial destinado a orientar pais na identificação de sinais de TEA de forma precoce, entre 16 e 30 meses, verificando sua precisão, clareza, atualidade das referências e a aplicabilidade na prática.

2 MÉTODO

Delineamento de Pesquisa

A pesquisa caracteriza-se como um estudo quase-experimental com análise qualitativa de conteúdo e de natureza documental. O delineamento foca na vertente qualitativa, pois o foco reside na análise de conteúdo e na qualidade das informações geradas pela Inteligência Artificial (IA), em vez de dados numéricos. A abordagem quase-experimental justifica-se pela manipulação controlada de estímulos (prompts) para a investigação da veracidade e confiabilidade das informações obtidas por meio do uso das IAs sobre os diagnósticos precoces do TEA. O caráter documental da pesquisa foi evidenciado pelas respostas geradas pela IA, que constituíram o corpo da análise. A avaliação das respostas teve como critérios de comparação os parâmetros estabelecidos por instrumentos de rastreio validados cientificamente, como o M-CHAT-R/F e a literatura científica sobre marcos de desenvolvimento na primeira infância.

Foram avaliadas as seguintes variáveis: variável independente (VI) - o tipo de prompt utilizado para interagir com a Inteligência Artificial, com o objetivo de ver como cada um afetava a resposta da IA - e a variável dependente (VD) - a qualidade da resposta gerada pela Inteligência Artificial. A VD foi avaliada pelos seguintes critérios: (I) precisão científica e conformidade; (II) clareza e linguagem; (III) aplicabilidade prática; e (IV) qualidade das referências.

Modelos Utilizados

O modelo de IA generativa utilizado foi o Gemini 2.5 Pro, um modelo de linguagem de grande escala (LLM) multimodal desenvolvido pelo Google. Este modelo é treinado em um vasto e diversificado corpus de dados, conferindo-lhe avançada capacidade de processamento de linguagem natural, compreensão de nuances contextuais e geração de respostas coerentes. A escolha deste modelo justificou-se por ser uma das tecnologias mais atuais no campo da IA generativa, permitindo uma avaliação avançada da qualidade da informação disponível e uma comparação direta com estudos recentes da literatura. A versão foi acessada por meio de sua interface pública em outubro de 2025. Para evitar o enviesamento da pesquisa, cada prompt foi testado em um chat novo.

Prompts Utilizados

Os prompts que foram utilizados imitaram possíveis perguntas feitas por pais, a saber: (I) Meu filho tem 2 anos e eu acho que ele tem autismo. Elabore uma guia para que eu possa avaliar e ter uma ideia inicial dos sinais com exemplos de tarefas cotidianas que eu possa fazer com ele e me conte as referências que você utilizou; (II) Meu filho tem 2 anos e eu acho que ele tem autismo, ele apresenta comportamentos como não olhar nos olhos e fixação em temas específicos quando é chamado. Elabore uma guia para que eu possa avaliar e ter uma

ideia inicial dos sinais e me conte as referências que você utilizou; (III) Meu filho tem 2 anos e eu acho que ele tem autismo. Elabore uma guia baseada no M-CHAT-R/F para que eu possa avaliar e ter uma ideia inicial dos sinais e me conte as referências que você utilizou.

Procedimento

A coleta de dados ocorreu por meio de uma busca qualitativa no modelo de Inteligência Artificial Gemini 2.5 Pro, no qual a IA recebeu diferentes prompts instrucionais. Para evitar o enviesamento de interações prévias, cada um dos três prompts definidos foi feito em uma nova sessão de chat, sem histórico de conversas. Adotando uma metodologia similar à de McFayden et al. (2024), após a obtenção da primeira resposta (R1) para o primeiro prompt, o comando 'refazer' foi utilizado para gerar uma segunda resposta (R2). Para avaliar a consistência, foi utilizado um prompt de comparação das respostas (R1 e R2). Caso as respostas fossem diferentes, o processo seria replicado para os demais prompts, resultando em uma resposta final para cada um. Ao término de cada prompt, as referências utilizadas pela IA para a formulação das respostas foram solicitadas, visando avaliar a base de suas informações. Todo o contexto textual gerado pela IA foi integralmente salvo para a etapa de avaliação, que foi conduzida pela pesquisadora e pela própria IA, utilizando as rubricas definidas.

Avaliação do Conteúdo Gerado

O conteúdo gerado pelo Gemini 2.5 Pro foi submetido à análise qualitativa, focada em avaliar a precisão, clareza e aplicabilidade das informações fornecidas, utilizando múltiplos critérios de avaliação. A principal referência para a análise foi a conformidade das respostas com os parâmetros do Checklist Modificado para Autismo em Crianças Pequenas (M-CHAT-R/F) e a literatura científica atual sobre os marcos do desenvolvimento infantil e sinais precoces de TEA. A avaliação foi estruturada com base nos seguintes critérios: (I) precisão científica e conformidade; (II) clareza e linguagem; (III) aplicabilidade prática; e (IV) qualidade das referências.

1. **Precisão científica e conformidade:** verificou se a informação da IA estava correta e completa, comparando-a diretamente com o checklist M-CHAT-R/F e outras diretrizes científicas para o diagnóstico precoce. O objetivo é identificar erros, omissões importantes ou dados incorretos (“alucinações”);
2. **Clareza e Linguagem:** avaliou se o texto era de fácil compreensão para os pais e cuidadores, utilizando uma linguagem acessível e sem excesso de termos técnicos, conforme o critério de “clareza” do estudo de McFayden et al. (2024);
3. **Aplicabilidade prática:** analisou se as orientações da IA eram úteis no dia a dia, oferecendo passos práticos e diretos que os pais pudessem aplicar para observar o comportamento dos filhos. Este critério foi inspirado no conceito de *actionability* (acionabilidade);
4. **Qualidade das referências:** verificou a confiabilidade das fontes citadas pela IA, avaliando se os links funcionavam e se levavam a fontes confiáveis, como artigos científicos e sites de organizações de saúde conhecidas.

Foram utilizados três pilares metodológicos para a confiabilidade da avaliação: (I) uso de instrumento validado e amplamente aceito pela comunidade científica como o M-CHAT-R/F e também foi utilizado o Qualis Periódicos, que avaliou estudos/artigos disponibilizados pela IA; (II) aplicação de uma rubrica sistemática, que garantiu a análise de todo o conteúdo gerado de acordo com os mesmos critérios; e (III) detalhamento do procedimento, possibilitando a replicação do procedimento.

As rubricas criadas, detalhadas nas Tabelas 1 a 4, possuem quatro níveis de pontuação que permitiram a classificação da qualidade das respostas da IA de forma sistemática.

Tabela 1. Rubrica para Avaliação da Precisão Científica e Conformidade das Respostas da IA

Pontuação	Nível	Descrição dos Critérios
4	Excelente	A resposta da IA demonstra total conformidade com os itens do M-CHAT-R/F e a literatura. Todos os sinais de alerta relevantes para a faixa etária são mencionados corretamente. Não há erros factuais (“alucinações”) nem omissões críticas.
3	Bom	A resposta da IA é majoritariamente precisa e alinhada ao M-CHAT-R/F. Pode haver omissões menores de um ou dois sinais menos evidentes, mas que não comprometem a utilidade geral da informação. Não há erros factuais significativos.
2	Regular	A resposta da IA apresenta conformidade parcial. Vários sinais de alerta importantes do M-CHAT-R/F são omitidos ou descritos de forma vaga. Pode conter imprecisões menores, mas sem desinformação grave.
1	Insuficiente	A resposta da IA demonstra baixa ou nenhuma conformidade com o M-CHAT-R/F. A maioria dos sinais críticos é omitida, e a informação fornecida contém erros factuais relevantes (“alucinações”), podendo induzir os pais ao erro.

Nota. Rubrica com quatro níveis (1 a 4) para avaliar a precisão científica e a conformidade das respostas da IA em relação ao M-CHAT-R/F e à literatura.

Tabela 2. Rubrica para Avaliação da Clareza e Linguagem das Respostas da IA

Pontuação	Nível	Descrição dos Critérios
4	Excelente	O texto é extremamente claro e de fácil compreensão para pais e cuidadores. A linguagem é totalmente acessível, sem jargões técnicos desnecessários. A estrutura do texto é lógica e facilita a leitura. O tom é apropriado e empático.
3	Bom	O texto é majoritariamente claro e compreensível. A linguagem é acessível, com poucos ou nenhum termo técnico complexo. A estrutura é boa, embora possa haver pequenos pontos de melhoria na fluidez. O tom é adequado.
2	Regular	O texto apresenta clareza parcial. Partes da resposta podem ser confusas ou de difícil compreensão para o público leigo. Utiliza alguns jargões técnicos sem explicação adequada. A estrutura pode ser um pouco desorganizada. O tom pode ser impessoal ou excessivamente técnico.
1	Insuficiente	O texto é predominantemente confuso ou incompreensível para pais e cuidadores. A linguagem é excessivamente técnica ou inadequada. A estrutura é desorganizada, dificultando a leitura. O tom é inadequado ou inacessível.

Nota. Rubrica com quatro níveis (1 a 4) para avaliar a clareza e a adequação da linguagem das respostas da IA ao público-alvo.

Tabela 3. Rubrica para Avaliação da Aplicabilidade Prática das Respostas da IA

Pontuação	Nível	Descrição dos Critérios
4	Excelente	A resposta da IA oferece orientações altamente práticas e específicas. Apresenta exemplos claros e passos concretos que os pais podem facilmente aplicar no dia a dia para observar o comportamento da criança (ex: durante brincadeiras, refeições, interações específicas). As sugestões são diretamente aplicáveis.
3	Bom	A resposta da IA oferece orientações majoritariamente práticas. Os passos ou exemplos são úteis, mas podem carecer de alguma especificidade ou exigir uma pequena adaptação por parte dos pais para serem aplicados. A aplicabilidade geral é boa.
2	Regular	A resposta da IA oferece orientações com aplicabilidade limitada. O conselho é um pouco genérico ou abstrato, dificultando a aplicação direta pelos pais. Pode fornecer informações sobre o que observar, mas sem clareza sobre como observar no cotidiano.
1	Insuficiente	A resposta da IA não oferece orientações práticas ou as sugestões são excessivamente vagas, teóricas ou irrelevantes para a observação no dia a dia. A informação não se traduz em ações concretas para os pais.

Nota. Rubrica com quatro níveis (1 a 4) para avaliar a aplicabilidade prática das orientações fornecidas pela IA.

Tabela 4. Rubrica para Avaliação da Qualidade das Referências Fornecidas pela IA

Pontuação	Nível	Descrição dos Critérios
4	Excelente	Todas ou a vasta maioria das referências fornecidas são de fontes altamente confiáveis (ex: artigos científicos revisados por pares, sites de organizações de saúde governamentais ou internacionalmente reconhecidas). Os links estão ativos e direcionam corretamente para o conteúdo citado. As fontes apoiam diretamente as afirmações feitas pela IA.
3	Bom	A maioria das referências é de fontes confiáveis. Alguns links podem estar quebrados ou direcionar para páginas genéricas em vez do conteúdo específico, mas a maioria funciona. As fontes geralmente apoiam as afirmações, com pequenas inconsistências.
2	Regular	Metade ou menos das referências são de fontes claramente confiáveis, ou há uma mistura com fontes de baixa credibilidade (ex: blogs não especializados, sites comerciais sem base científica). Vários links estão quebrados ou são irrelevantes. As fontes apoiam apenas parcialmente as afirmações ou o suporte é vago.
1	Insuficiente	A maioria ou todas as referências são de fontes não confiáveis, inexistentes ou não especificadas. A maioria dos links está quebrada ou leva a conteúdo não relacionado. As fontes não apoiam as afirmações feitas pela IA, ou as referências são fabricadas ("alucinações").

Nota. Rubrica com quatro níveis (1 a 4) para avaliar a qualidade das referências fornecidas pela IA em suas respostas.

3 RESULTADOS

Coleta de Dados

A coleta de dados foi realizada no dia 27/10/2025, todos os dados foram colhidos no mesmo dia, sem interferência ou pausas feitas pela IA. É válido ressaltar que cada prompt foi feito em uma aba de conversa diferente, em uma nova conta sem histórico de pesquisa.

As respostas dadas pelo Gemini estão no Anexo 1.

Avaliação do Conteúdo

A Tabela 5 apresenta a avaliação do conteúdo que foi realizada aplicando-se quatro rubricas (Tabela 1 a 4) a cada uma das três respostas obtidas (Prompt 1, Prompt 2 e Prompt 3). A análise das suas respostas geradas para o Prompt revelou um nível de qualidade equivalente, sendo ambas avaliadas em conjunto e dispensando a realização da segunda resposta para os outros prompts. A Tabela 5, a seguir, consolida os resultados dessa análise, atribuindo uma pontuação de 1 (insuficiente) a 4 (excelente) para cada critério. O conteúdo gerado pela IA foi submetido a uma dupla avaliação: uma conduzida pela pesquisadora e outra correspondente à autoavaliação da própria IA. Ambas as classificações seguiram rigorosamente as rubricas estabelecidas e incluíram a justificativa para a nota. Na Tabela 5, a seguir, os resultados são apresentados no formato numérico 'X/Y', no qual o valor antes da barra (X) representa a pontuação atribuída pela pesquisadora, e o valor após a barra (Y) representa a nota atribuída pela Inteligência Artificial.

Tabela 5. Avaliação das Respostas Geradas pela IA - avaliadas pela pesquisadora/IA

Critério	Resposta 1	Resposta 2	Resposta 3
Precisão científica	4/4	4/4	4/4
Clareza e linguagem	4/4	4/4	4/4
Aplicabilidade prática	4/4	4/4	4/4
Qualidades das referências	3/3	3/3	4/4

Nota. As pontuações são apresentadas no formato Pesquisadora/IA. O primeiro valor representa a pontuação atribuída pela pesquisadora e o segundo valor corresponde à autoavaliação realizada pelo modelo Gemini 2.5 Pro. A pontuação máxima para cada critério avaliado é 4.

A primeira resposta foi avaliada no primeiro critério, de precisão científica, como “excelente” pela pesquisadora e pela IA, o que demonstra que a resposta demonstrou conformidade com o instrumento que a pesquisa usa como base de estudos e de critérios para a avaliação; o segundo critério, de clareza e linguagem, também foi avaliado pela pesquisadora e pela IA como “excelente”, demonstrando que a linguagem utilizada foi acessível sem a utilização de termos técnicos e deixando claro que a guia não é um diagnóstico; no terceiro critério de avaliação, de aplicabilidade prática, ambos o avaliaram como “excelente”, avaliando que o guia não se limitou em sinais abstratos e sim em exemplos concretos e fáceis de se aplicar dentro do ambiente familiar; já no quarto critério de avaliação, de qualidade das referências foi avaliada como “bom” pela pesquisadora e pela IA, essa nota se justifica pela falta do M-CHAT-R/F como uma referência e pela forma genérica de citação utilizada pela IA.

A segunda resposta foi avaliada da seguinte forma: o primeiro critério foi dado como “excelente” pela pesquisadora e pela IA, demonstrando que a guia está alinhada totalmente com o M-CHAT-R/F; o segundo critério também foi dado como “excelente” por ambos, implicando na clareza das respostas; no terceiro critério, a avaliação foi “excelente”, o que significou que

a guia traduziu bem o que está no M-CHAT-R/F para algo prático que pode ser utilizado por pais; partindo para o quarto critério de avaliação, ele foi dado como “bom” porque ainda que o M-CHAT-R/F tenha sido citado, as referências foram dadas de forma genérica.

Já na terceira resposta, todos os critérios de avaliação foram dados como “excelente”, o que significa que ambos os avaliadores consideraram que a resposta dada teve uma boa precisão científica, seguindo o que está presente no M-CHAT-R/F, a sua clareza e linguagem foram feitas de forma correta, e em sua aplicabilidade prática, a resposta traduziu de forma clara para a prática o que está no M-CHAT-R/F, e nas qualidades da referência, houve referências específicas e acionáveis, com links para o site, atendendo claramente aos critérios.

Avaliação “Refazer”

De acordo com a avaliação realizada pela própria IA, o comando “refazer”, não houve informações incongruentes ou contraditórias entre a Resposta 1 e a Resposta 2. Ambas as respostas são semelhantes e seguem a mesma estrutura, fornecendo conselhos consistentes e seguros.

Avaliação Geral

As respostas demonstram um bom desempenho da IA para a realização de uma guia para pais, porém há alguns fatos que devem ser considerados, como a indicação de procurar um pediatra ou um neurologista, onde a procura por um neuropsicólogo ou um psicólogo deve ser mais indicada.

4 DISCUSSÃO

Resumo

O objetivo do presente trabalho foi avaliar e comparar a qualidade técnica e científica das respostas fornecidas por um aplicativo gerado por inteligência artificial destinado a orientar pais na identificação de sinais de TEA de forma precoce, entre 16 e 30 meses, verificando sua precisão, clareza, atualidade das referências e a aplicabilidade na prática. Para tanto, trata-se de uma pesquisa qualitativa e documental que avaliou o Gemini 2.5 Pro (out/2025) utilizando três prompts com níveis distintos de especificidades sobre sinais de autismo; as respostas foram coletadas em sessões isoladas e avaliadas comparativamente ao M-CHAT-R/F por meio de uma rubrica de quatro pontos, que analisou precisão científica, clareza, aplicabilidade prática e qualidade das referências. De forma geral os resultados indicam que o modelo de inteligência artificial Gemini 2.5 Pro demonstra boa competência na orientação de pais sobre sinais precoces de autismo. A análise revelou que, independentemente do nível de especificidade do prompt, a IA foi capaz de fornecer informações com precisão científica, clareza linguística e aplicabilidade prática consideradas “excelentes”. O estudo não encontrou evidências de “alucinações” ou informações incongruentes entre as respostas regeneradas. A principal variação identificada ocorreu na qualidade das referências bibliográficas: enquanto

prompts gerais resultaram em citações corretas porém genéricas, o prompt que solicitava especificamente o uso do instrumento M-CHAT-R/F gerou referências precisas, acionáveis e com links diretos.

Comparação com a Literatura

O presente estudo corrobora com as pesquisas recentes sobre o uso de IA na saúde e parentalidade. Ao contrário das preocupações levantadas por Kim et al. (2025), que apontaram que as IAs poderiam necessitar de profundidade na dinâmica relacional, as respostas do Gemini 2.5 Pro neste estudo enfatizaram fortemente a interação social e a atenção compartilhada, traduzindo conceitos complexos em interações cotidianas (como o "teste do apontar"). Em relação ao estudo de McFayden et al. (2024), que encontrou limitações na "acionabilidade" das respostas do ChatGPT, esse estudo observou um avanço com o Gemini 2.5 Pro. A IA não apenas explicou os sinais, mas estruturou testes cotidianos e passos práticos, sugerindo uma evolução na capacidade dos modelos de linguagem mais novos em fornecer orientações aplicáveis. Por fim, contrariando o alerta de "alucinações" e informações incorretas, a IA se manteve fiel aos critérios do M-CHAT-R/F e do DSM-5-TR, inclusive emitindo alertas de segurança consistentes sobre a impossibilidade de realizar um diagnóstico médico.

Contribuições teóricas

O trabalho contribui para a teoria sobre a interação humano-computador na saúde, sugerindo que as IAs generativas avançadas podem atuar como tradutores eficazes de conhecimento técnico. A pesquisa reforça a ideia de que modelos de linguagem grandes (LLMs) podem preencher a lacuna entre a literatura científica rígida (como manuais diagnósticos) e a necessidade de informação acessível das famílias. Além disso, o trabalho evidencia que a qualidade da saída da IA é mediada pela qualidade da instrução humana (prompt), reforçando a teoria de que a engenharia de prompt é uma variável crítica na obtenção de informações de saúde de alta confiabilidade.

Contribuições metodológicas

O estudo inovou de forma metodológica com uma adaptação e aplicação de uma rubrica sistemática de avaliação qualitativa específica para o contexto brasileiro de rastreio de autismo. Diferente de estudos que focam apenas no pensamento dicotômico (certo/errado), a pesquisa avaliou a qualidade da informação através de quatro dimensões: precisão científica, clareza, aplicabilidade prática e qualidade das referências. Além do mais, o uso do comando "refazer" como método de verificação de consistência interna da IA replicou e validou procedimentos metodológicos sugeridos por McFayden et al. (2024), garantindo que as respostas não fossem artefatos aleatórios de uma única interação.

Contribuições Aplicadas

Os resultados da pesquisa indicaram que o uso da IA pode auxiliar a família no processo de diagnóstico precoce de crianças, visando a capacidade da ferramenta em produzir uma guia que dá aos pais e responsáveis exemplos seguros e baseados em um instrumento seguro (M-

CHAT-R/F). Ademais, a guia possui exemplos práticos que podem ser observados no cotidiano, facilitando a coleta de dados e informações, o que resulta na ajuda no processo de avaliação de um profissional.

Limitações e pesquisas futuras

Apesar dos resultados promissores, este estudo apresenta limitações importantes que devem ser consideradas, como a ausência de validação clínica, uma vez que as orientações geradas não foram testadas com pais ou em ambiente clínico, o que impede afirmar se a leitura desses guias reduz a ansiedade parental ou leva a uma busca efetiva por diagnóstico – uma preocupação ressaltada por Kim et al. (2025) quanto à necessidade de contextualização humana da informação. Adicionalmente, a avaliação limitou-se ao modelo Gemini 2.5 Pro em um recorte temporal específico (outubro de 2025), o que, dada a atualização constante desses modelos, pode comprometer a perenidade dos resultados. Para avançar, pesquisas futuras devem comparar a eficácia de diferentes IAs (como ChatGPT-5, Gemini e Claude) simultaneamente na triagem de TEA e investigar se o uso de modelos treinados especificamente com literatura da Análise do Comportamento produziria resultados mais robustos e menos dependentes da engenharia de prompt do usuário.

No que tange às limitações metodológicas e aos direcionamentos para pesquisas futuras, resalta-se a fragilidade da dependência do Modified Checklist for Autism in Toddlers (M-CHAT ou M-CHAT-R/F) como instrumento exclusivo de avaliação. Embora o M-CHAT demonstre alta sensibilidade para identificar atrasos globais no neurodesenvolvimento, a literatura aponta que ele pode apresentar baixa especificidade para o Transtorno do Espectro Autista (TEA) quando aplicado de forma isolada, o que o torna insuficiente para a conclusão diagnóstica sem a realização de investigações complementares (Alckmin-Carvalho et al., 2014). Dessa forma, pesquisas futuras devem focar na eficácia do uso combinado do M-CHAT com outros protocolos de observação estruturada e escalas de desenvolvimento infantil. A associação de múltiplos instrumentos avaliativos é fundamental para refinar a precisão diagnóstica, reduzir as taxas de falsos positivos gerados por ferramentas exclusivas de triagem e garantir uma compreensão global, dinâmica e multidisciplinar do perfil sociocomunicativo da criança (Alckmin-Carvalho et al., 2014; Seize & Borsa, 2017).

5 CONCLUSÃO

Conclui-se que o Gemini 2.5 Pro é uma ferramenta viável e precisa para auxiliar na identificação inicial de sinais de autismo, demonstrando alta competência em traduzir critérios técnicos (M-CHAT-R/F) para uma linguagem acessível e prática. No entanto, a variação na qualidade das referências bibliográficas destaca que a eficácia da ferramenta ainda depende da habilidade do usuário em formular prompts precisos. Embora a IA mostre potencial para democratizar o acesso à informação de qualidade, ela não substitui o julgamento clínico; a supervisão de um profissional especializado permanece indispensável para validar as informações e conduzir o processo diagnóstico ético e humano (Alckmin-Carvalho et al., 2014).

REFERÊNCIAS BIBLIOGRÁFICAS

- Alckmin-Carvalho, F., Teixeira, M. C. T. V., Brunoni, D., Strauss, V. G., & Paula, C. S. (2014). Identificação de Sinais Precoces de Autismo Segundo um Protocolo de Observação Estruturada: um Estudo de Seguimento. *Psico*, 45(4), 502-512. <https://doi.org/10.15448/1980-8623.2014.4.15873>
- Associação Americana de Psiquiatria. (2023). Manual diagnóstico e estatístico de transtornos mentais: DSM-5-TR (5. ed. rev.). Artmed.
- da Silva, A. V., & Lopes Gomes, M. (2023). Os desafios do diagnóstico psicológico precoce do transtorno do espectro autista (TEA). *Revista De Casos E Consultoria*, 14(1), e31093. Recuperado de <https://periodicos.ufrn.br/casoseconsultoria/article/view/31093>
- Goulart, P., & Assis, G. J. A. de. (2002). Estudos sobre autismo em análise do comportamento: aspectos metodológicos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 4(2), 151-165.
- Kim, Y. B., Vilches, S. L., Clarkson, A., & Shapiro, S. (2025). Testing generative artificial intelligence for parent and caregiver information. *Family Relations*, 74(3), 1266-1284. <https://doi.org/10.1111/fare.13167>
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135-148. <https://doi.org/10.1089/cyber.2023.0202>
- Plantin, L., & Daneback, K. (2009). Parenthood, information and support on the internet. A literature review of research on parents and professionals online. *BMC Family Practice*, 10, Article 34. <https://doi.org/10.1186/1471-2296-10-34>
- Robins, D. L., Fein, D., & Barton, M. L. (2009). Checklist Modificado para Autismo em Crianças Pequenas: versão revisada e consulta de seguimento (M-CHAT-R/F). (L. S. Losapio, G. Siquara, C. Lampreia, A. Lázaro, & M. Pondé, Trads.). Retirado de mchatscreen.com
- Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432e008>
- Santos, A. L. V., Fernandes, C. F., Santana, L. T. G., Santo, L. R. E., & Lafetá, B. N. (2015). Diagnóstico precoce do autismo: dificuldades e importância. *Revista Renome*, 4, 23-24. Recuperado de <https://www.periodicos.unimontes.br/index.php/renome/article/view/2655>
- Santos, M. S., Barbosa, N. T., & Lopes, M. M. (2024). Desafios do diagnóstico precoce em crianças com TEA dos 0 aos 3 anos. *Revista Multidisciplinar Integrada-REMI*, 6(1), 1-13. <https://doi.org/10.61164/1ycmbj33>
- Seize, M. D. M., & Borsa, J. C. (2017). Instrumentos para rastreamento de sinais precoces do autismo: revisão sistemática. *Psico-USF*, 22, 161-176. <https://doi.org/10.1590/1413-82712017220114>
- Steffen, B. F., de Paula, I. F., Martins, V. M. F., & López, M. L. (2019). Diagnóstico precoce de autismo: uma revisão literária. *Revista saúde multidisciplinar*, 6(2). Recuperado de <https://revistas.famp.edu.br/revistasaudemultidisciplinar/article/view/91>
- Vllasaliu, L., Jensen, K., Hoss, S., Landenberger, M., Menze, M., Schütz, M., Ufniaz, K., Kieser, M., & Freitag, C. M. (2016). Instrumentos de diagnóstico para transtorno do espectro autista (TEA). *Cochrane Database of Systematic Reviews*, 1, CD012036. <https://doi.org/10.1002/14651858.CD012036>

Zanon, R. B., Backes, B., & Bosa, C. A. (2014). Identificação dos primeiros sintomas do autismo pelos pais. *Psicologia: teoria e pesquisa*, 30(1), 25-33. <https://doi.org/10.1590/S0102-37722014000100004>

Anexos

Anexo 1

Prompt 1, resposta 1: Compreendo perfeitamente sua preocupação. Observar o desenvolvimento do nosso filho e notar algo que parece diferente pode gerar muita ansiedade. Quero que você saiba que buscar informação é o primeiro e mais importante passo para ajudá-lo.

O mais importante primeiro: Este guia não é um diagnóstico.

Eu sou uma inteligência artificial e não posso, de forma alguma, diagnosticar seu filho. O Transtorno do Espectro Autista (TEA) é complexo e seu diagnóstico só pode ser feito por uma equipe de profissionais qualificados, como um neuropediatra, psiquiatra infantil ou psicólogo especializado.

O que posso fazer é elaborar um guia baseado nos sinais de alerta comumente observados por especialistas (como o CDC dos EUA e a Sociedade Brasileira de Pediatria) para crianças por volta dos 2 anos. Use isto apenas como uma ferramenta de observação para organizar suas ideias e levá-las ao pediatra.



Guia de Observação: Sinais de Alerta de TEA aos 2 Anos

O autismo geralmente se manifesta em duas áreas principais:

1. Dificuldades na Comunicação e Interação Social.
2. Padrões de Comportamento, Interesses ou Atividades Restritos e Repetitivos.

Aqui estão exemplos de tarefas e observações cotidianas para cada área.

1. Área: Comunicação e Interação Social



Resposta ao Nome

- O que observar: A criança consistentemente não responde quando chamada pelo nome.
- Teste Cotidiano: Em um ambiente calmo, sem distrações, sente-se a alguns metros dele. Chame o nome dele em um tom de voz claro e normal (sem gritar). Repita algumas vezes em momentos diferentes. Ele se vira para olhar para você?
- Sinal de Alerta: A maioria das crianças de 2 anos (e até antes) olha quando ouve seu nome. Um sinal de alerta é a falha consistente em responder, como se não estivesse ouvindo (mesmo que você saiba que a audição dele é boa para outros sons, como a abertura de um pacote de biscoito).



Contato Visual

- O que observar: Dificuldade em iniciar ou manter o contato visual.

- Teste Cotidiano: Durante a brincadeira, no chão com ele, tente encontrar o olhar dele. Quando você dá um brinquedo ou comida, ele olha nos seus olhos? Ou ele parece olhar "através" de você, foca na sua boca, ou pega o objeto sem fazer contato visual?
- Sinal de Alerta: O contato visual pode ser curto, evitado ou parecer desconfortável para a criança.

👉 Gestos e Atenção Compartilhada

- O que observar: Dificuldade em usar ou entender gestos como apontar.
- Teste Cotidiano (Duas partes):
 1. Entender o gesto: Aponte para algo interessante do outro lado da sala (ex: "Olha, o cachorro!"). Ele olha para onde seu dedo está apontando? (Um sinal de alerta é ele olhar para o seu dedo, e não para o cachorro).
 2. Usar o gesto: Quando ele quer algo fora do alcance (um brinquedo na prateleira, um copo de água), o que ele faz? Ele aponta para o objeto? (Um sinal de alerta comum é ele pegar a sua mão e "usá-la" como uma ferramenta para pegar o objeto, sem olhar para você).

👤 Interesse em Outras Crianças

- O que observar: Falta de interesse ou resposta a outras crianças.
- Teste Cotidiano: No parquinho ou em uma festa, observe-o. Ele nota as outras crianças? Ele tenta se aproximar, mesmo que seja apenas para ficar perto e observar (o que é normal para 2 anos)?
- Sinal de Alerta: Ignorar completamente outras crianças, preferindo brincar sozinho e não demonstrar curiosidade sobre o que elas estão fazendo.

🧸 Brincadeira Simbólica (Faz de Conta)

- O que observar: Ausência ou dificuldade em brincar de "faz de conta".
- Teste Cotidiano: Por volta dos 2 anos, a brincadeira simbólica começa a surgir. Tente iniciar uma brincadeira: pegue um telefone de brinquedo e diga "Alô?". Ele tenta imitar? Pegue uma colher vazia e "dê comida" para um boneco. Ele entende a brincadeira ou tenta comer a colher?
- Sinal de Alerta: A criança pode não entender o "faz de conta". A brincadeira dela tende a ser mais literal ou repetitiva (veja abaixo).

🗣️ Desenvolvimento da Fala

- O que observar: Atraso na fala ou uso incomum da linguagem.
- Observação Cotidiana: Crianças de 2 anos tipicamente formam frases curtas (ex: "quer água", "cadê mamãe").
- Sinal de Alerta: Não apenas o atraso (não falar ou falar poucas palavras), mas também a ecolalia – repetir frases que ouviu na TV ou de um adulto fora de contexto, em vez de usar a linguagem para se comunicar (pedir algo, fazer um comentário).

2. Área: Padrões Restritos e Repetitivos

Uso de Brinquedos e Objetos

- O que observar: Brincar de forma repetitiva ou focar em partes de um objeto.
- Teste Cotidiano: Dê a ele um carrinho. Ele brinca de empurrar o carrinho, fazendo "vrum vrum"?
- Sinal de Alerta: Focar intensamente em apenas uma parte do brinquedo (ex: girar a rodinha do carrinho sem parar). Outro sinal comum é enfileirar objetos (alinhar todos os carrinhos, blocos ou dinossauros em uma fila perfeita) e ficar muito irritado se alguém mexer.

Movimentos Corporais Repetitivos (Estereotípias)

- O que observar: Movimentos incomuns com o corpo, especialmente quando excitado ou ansioso.
- Observação Cotidiana: Preste atenção em momentos de emoção.
- Sinal de Alerta: Movimentos como "flapping" (bater as mãos rapidamente, como asas), balançar o corpo para frente e para trás, girar em círculos ou andar persistentemente nas pontas dos pés.

Apego à Rotina e Inflexibilidade

- O que observar: Necessidade extrema de manter a rotina e dificuldade com mudanças.
- Teste Cotidiano: Observe a reação dele a pequenas mudanças.
- Sinal de Alerta: Ficar extremamente angustiado ou ter uma crise de choro intensa porque você fez um caminho diferente para casa, serviu o lanche em um prato de cor diferente ou a hora do banho atrasou 10 minutos. A reação é desproporcional à mudança.

Sensibilidades Sensoriais

- O que observar: Reações exageradas (hiper) ou muito baixas (hipo) a estímulos sensoriais.
- Observação Cotidiana:
 - Hiper (Exagerado): Ele cobre os ouvidos para sons comuns (liquidificador, secador de cabelo, descarga)? Ele evita tocar em certas texturas (grama, areia, massinha)? Ele é "chato" para comer, aceitando apenas certas texturas?
 - Hipo (Baixo): Ele parece não sentir dor como outras crianças? Ele busca estímulos intensos (gosta de bater o corpo nas coisas, abraços muito fortes)?
- Visual: Ele fica fascinado por luzes, ventiladores de teto ou objetos que giram?

O Que Fazer Agora?

1. Anote o que Você Observou: Use este guia para anotar exemplos específicos do que você vê no dia a dia. Ter exemplos concretos é muito mais útil para o médico do que dizer apenas "eu acho que ele tem autismo".
2. Agende uma Consulta com o Pediatra: Este é o seu primeiro ponto de contato. Mostre suas anotações e peça um encaminhamento para um especialista.

3. Procure um Especialista: O diagnóstico é geralmente fechado por um Neuropediatra ou Psiquiatra Infantil, muitas vezes com o apoio de uma avaliação neuropsicológica.
4. Não Espere: Se você tem uma suspeita, investigue. A intervenção precoce (terapias) é o fator mais importante para ajudar no desenvolvimento da criança. Confie no seu instinto.



Referências Utilizadas

Para criar este guia, baseei-me em informações de fontes reconhecidas e autoridades em saúde infantil:

1. Centros de Controle e Prevenção de Doenças (CDC) dos EUA: O CDC possui checklists detalhados (Marcos do Desenvolvimento) que descrevem o que a maioria das crianças faz em cada idade e quais são os sinais de alerta para atrasos no desenvolvimento, incluindo o TEA.
2. Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5): É o manual oficial usado por profissionais de saúde mental para diagnosticar o TEA. As duas áreas principais (Comunicação Social e Comportamentos Restritos/Repetitivos) são baseadas nos critérios do DSM-5.
3. Ministério da Saúde do Brasil e Sociedade Brasileira de Pediatria (SBP): Ambas as organizações fornecem diretrizes e materiais sobre a vigilância do desenvolvimento infantil, enfatizando a importância da detecção precoce de sinais de TEA, muitos dos quais estão listados acima.
4. Artigos de Instituições Médicas de Referência: (Como Hospital Israelita Albert Einstein e Genial Care), que publicam resumos baseados nos critérios médicos (como os do DSM-5) para o público leigo, explicando sinais como ecolalia, falta de resposta ao nome e o uso da mão do adulto como ferramenta.

Espero que este guia ajude a organizar suas observações. Lembre-se, você está fazendo a coisa certa ao buscar respostas para apoiar seu filho.

Você gostaria que eu ajudasse a encontrar recursos ou associações de apoio a pais de crianças com autismo no Brasil?

Prompt 1, resposta 2: Entendo perfeitamente sua preocupação. Observar o desenvolvimento do nosso filho e ter dúvidas é natural e um sinal de enorme cuidado. Aos 2 anos, muitas habilidades sociais e de comunicação estão desabrochando, e é um período em que os primeiros sinais do Transtorno do Espectro Autista (TEA) podem se tornar mais evidentes.

Quero começar com o ponto mais importante: este guia não é uma ferramenta de diagnóstico. O diagnóstico do TEA é complexo e só pode ser feito por profissionais qualificados (como neuropediatras, psiquiatras infantis ou psicólogos especializados) após uma avaliação completa.

O objetivo deste guia é ajudá-la a organizar suas observações com exemplos práticos. Use-o como um diário para que você possa ter uma conversa clara e detalhada com o pediatra do seu filho, que é o primeiro e mais importante passo.



Guia de Observação: Sinais de Alerta de Autismo aos 2 Anos

O autismo se manifesta de forma diferente em cada criança, mas geralmente envolve desafios em duas áreas principais: Comunicação e Interação Social e Padrões Restritos e Repetitivos de Comportamento.

Aqui estão alguns marcos do desenvolvimento esperados para os 2 anos e atividades cotidianas para observar:

Área 1: Comunicação e Interação Social

1. Contato Visual

- O que observar: Dificuldade em iniciar ou manter o contato visual. O olhar pode parecer vago ou focado "através" da pessoa.
- Atividade Cotidiana:
 - Na refeição: Sente-se de frente para ele. Quando ele olhar para você, sorria. Ele sustenta o olhar por alguns segundos ou o desvia imediatamente, na maioria das vezes?
 - Na brincadeira: Deite-se no chão com ele e brinque com um carrinho. Tente "encontrar" o olhar dele. Ele parece desconfortável ou evita ativamente?

2. Resposta ao Nome

- O que observar: Não responder consistentemente quando chamado pelo nome (assumindo que a audição dele está boa).
- Atividade Cotidiana:
 - Teste simples: Em um momento calmo, sem muitas distrações, chame o nome dele em um tom de voz normal (não gritando). Faça isso algumas vezes ao longo do dia. Ele se vira para você na maioria das vezes? Ou ele parece ignorar, como se não fosse com ele?

3. Atenção Compartilhada (Este é um marco crucial!)

- O que observar: Dificuldade em compartilhar o interesse por algo. Isso inclui:
 - Não apontar para mostrar algo interessante (ex: um avião no céu, um cachorro).
 - Não olhar para onde você está apontando.
 - Não trazer objetos para te mostrar, apenas para pedir ajuda (ex: abrir um pote).
- Atividade Cotidiana:
 - O teste do "apontar": Aponte para um quadro na parede ou um pássaro na janela e diga com entusiasmo: "Olha, filho! O passarinho!". Ele olha para onde seu dedo aponta? Ou ele olha para o seu dedo, ou ignora?
 - O teste do "mostrar": Quando ele estiver com um brinquedo que gosta, observe se ele o traz até você apenas para te mostrar e ver sua reação, ou se ele só o traz quando precisa de ajuda.

4. Imitação e Brincadeira de Faz-de-Conta

- O que observar: Dificuldade em imitar suas ações, expressões faciais ou em brincar de "mentirinha".

- Atividade Cotidiana:
- Imitação: Bata palmas, mande um beijo ou faça uma careta engraçada. Ele tenta te imitar?
- Faz-de-conta: Pegue um telefone de brinquedo e finja falar "Alô!". Dê o telefone a ele. Ele tenta imitar a brincadeira? Ou ele apenas mexe nos botões ou coloca na boca? Pegue uma boneca e finja dar "comidinha" com uma colher vazia. Ele entende a brincadeira?

5. Interesse por Outras Crianças

- O que observar: Pouco ou nenhum interesse em outras crianças. Pode preferir brincar sozinho ou tratar outras crianças como "objetos", sem interagir.
- Atividade Cotidiana:
- No parquinho: Observe-o perto de outras crianças. Ele as observa? Tenta se aproximar (mesmo que sem saber como)? Ou ele parece totalmente indiferente à presença delas, focando apenas na areia ou no balanço?

Área 2: Padrões Restritos e Repetitivos de Comportamento

1. Movimentos Corporais Repetitivos (Estereotipias)

- O que observar: Movimentos como balançar as mãos (flapping) quando está excitado ou ansioso, balançar o corpo para frente e para trás, andar na ponta dos pés com frequência ou girar em círculos.
- Atividade Cotidiana:
- Momentos de emoção: Observe o que ele faz com o corpo quando está muito feliz (ex: ao ver um desenho favorito) ou irritado. Ele balança as mãos de forma rítmica?

2. Brincar Não Funcional ou Repetitivo

- O que observar: Brincar com os brinquedos de forma incomum, focando em partes específicas ou repetindo a mesma ação sem um objetivo aparente.
- Atividade Cotidiana:
- Com carrinhos: Dê a ele alguns carrinhos. Ele tenta fazê-los "correr" no chão (brincadeira funcional)? Ou ele os vira de cabeça para baixo e fica fascinado girando apenas as rodinhas? Ou ele os enfileira perfeitamente, várias e várias vezes?
- Com blocos: Ele tenta empilhar (mesmo que caia) ou ele apenas os alinha ou os agrupa por cor de forma muito rígida?

3. Necessidade de Rotina e Resistência à Mudança

- O que observar: Ficar extremamente angustiado com pequenas mudanças na rotina.
- Atividade Cotidiana:
- Mudança de rota: Tente fazer um caminho diferente para voltar da creche ou do supermercado. A reação dele é uma leve confusão ou uma crise de choro intensa e difícil de acalmar?

- Hora da refeição: Ele exige usar exatamente o mesmo prato, copo e talher em todas as refeições? Fica muito irritado se você oferece a comida em um prato de cor diferente?

4. Interesses Intensos ou Focados

- O que observar: Um fascínio incomum por objetos ou tópicos específicos (ex: ventiladores de teto, letras, números, logotipos, dinossauros).
- Atividade Cotidiana:
 - Observe o foco dele. Ele passa longos períodos olhando para a máquina de lavar girando? Ele se interessa mais por partes de objetos (como a dobradiça da porta) do que pelo brinquedo em si?

5. Sensibilidade Sensorial

- O que observar: Reações muito fortes (hiper) ou muito fracas (hipo) a estímulos sensoriais.
- Atividade Cotidiana:
 - Sons (Hiper): Ele tapa os ouvidos e chora desesperadamente com barulhos comuns, como o liquidificador ou um secador de cabelo?
 - Toque (Hiper): Ele detesta etiquetas de roupa, evita andar descalço na grama ou na areia, ou não gosta de se sujar (ex: tinta, comida)?
 - Tato (Hipo): Ele parece buscar sensações fortes? Gosta de esbarrar nas coisas de propósito, abraços muito apertados, ou parece não sentir dor como outras crianças?
 - Alimentação: É extremamente seletivo com comidas, aceitando apenas certas texturas ou cores?



O que fazer com estas observações?

1. Anote tudo: Tente ser específica. Em vez de "ele não brinca direito", anote "ele passou 15 minutos girando a roda do carrinho e alinhando-o com o sofá".
2. Não entre em pânico: Lembre-se, crianças se desenvolvem em ritmos diferentes. Ter alguns desses sinais não significa que ele tenha autismo. O que se observa é um padrão de vários sinais em conjunto.
3. Marque uma consulta com o Pediatra: Este é o seu primeiro passo. Leve suas anotações. O pediatra conhece o desenvolvimento infantil e poderá aplicar escalas de rastreio (como o M-CHAT, que é muito usado).
4. Peça encaminhamento: Se o pediatra também notar os sinais de alerta, ele provavelmente o encaminhará para uma avaliação formal com especialistas (neuropediatra, psicólogo infantil, fonoaudiólogo).
5. Confie no seu instinto: Você é a mãe e conhece seu filho melhor do que ninguém. Se você sente que algo está diferente, é seu direito e dever investigar. A intervenção precoce é a ferramenta mais poderosa para ajudar no desenvolvimento do seu filho, seja qual for o diagnóstico.



Referências Utilizadas

Para elaborar este guia, utilizei como referência as diretrizes e materiais informativos de organizações de saúde globais e nacionais, que são a base para o rastreamento do desenvolvimento infantil:

1. CDC (Centros de Controle e Prevenção de Doenças dos EUA): O programa "Learn the Signs. Act Early." (Aprenda os Sinais. Aja Cedo.) do CDC possui listas detalhadas de marcos do desenvolvimento e sinais de alerta para cada idade, incluindo os 2 anos.
2. Organização Mundial da Saúde (OMS): As fichas técnicas e documentos da OMS sobre o Transtorno do Espectro Autista (TEA) definem as principais características do transtorno.
3. Sociedade Brasileira de Pediatria (SBP): A SBP publica manuais de orientação e documentos científicos para pediatras sobre o desenvolvimento infantil e a detecção precoce do TEA, enfatizando a importância de marcos como a atenção compartilhada.
4. Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5): É o manual oficial utilizado por clínicos e psiquiatras para diagnosticar o TEA. A estrutura deste guia (dividida em comunicação social e comportamentos restritos) é baseada nos critérios diagnósticos do DSM-5.

Espero que isso ajude a organizar seus pensamentos e observações. Lembre-se de ser paciente consigo mesma e com ele durante esse processo.

Você gostaria de saber mais sobre como iniciar essa conversa com o pediatra ou quais perguntas fazer a ele?

Prompt 2, resposta: Compreendo perfeitamente a sua preocupação. Observar comportamentos no nosso filho que parecem diferentes do esperado pode gerar muita ansiedade. O que você descreve – a dificuldade no contato visual e a fixação em temas específicos (interesses restritos) – são, de fato, sinais que merecem atenção e são frequentemente associados ao Transtorno do Espectro Autista (TEA).

Elaborei um guia com os marcos do desenvolvimento e sinais de alerta que os especialistas observam em crianças por volta dos 2 anos de idade.

⚠ Importante: Este Guia Não é um Diagnóstico

Quero começar com o ponto mais crucial: esta lista é apenas um guia informativo e não substitui, em hipótese alguma, a avaliação de um profissional qualificado.

O diagnóstico do autismo é complexo e clínico, feito por especialistas (como neuropediatras, psiquiatras infantis ou psicólogos especializados) com base em observação direta, entrevistas com os pais e, por vezes, testes específicos. Muitas crianças podem apresentar um ou dois destes sinais isoladamente sem que isso signifique autismo.

Use esta lista apenas como um ponto de partida para organizar suas observações e discuti-las com o pediatra do seu filho.

📌 Guia de Observação: Sinais de Alerta para TEA (por volta dos 2 anos)

Os sinais do Transtorno do Espectro Autista são geralmente agrupados em duas categorias principais:

1. Dificuldades na Comunicação e Interação Social.

2. Padrões Restritos e Repetitivos de Comportamento.

1. Dificuldades na Comunicação e Interação Social

Esta área avalia como a criança interage e se comunica com os outros.

Contato Visual:

- O que observar: A criança evita ativamente o contato visual? Ela parece desconfortável ou simplesmente não procura seus olhos durante as interações (como mamar, conversar ou brincar)? (Este foi um ponto que você mencionou).

Resposta ao Nome:

- O que observar: A criança consistentemente não responde quando é chamada pelo nome? (Você mencionou que ele se fixa em temas quando é chamado, o que pode indicar que ele ouve, mas tem dificuldade em desengajar do seu foco, ou que a resposta ao nome é inconsistente).

Atenção Compartilhada (Crucial nesta idade):

- O que observar: A criança tenta "compartilhar" uma experiência com você?
- Ela aponta? Ela aponta para um avião no céu ou para um cachorro na rua e depois olha para você, para ter certeza de que você também viu?
- Ela segue seu olhar/dedo? Se você aponta para algo interessante do outro lado da sala, ela olha para onde você está apontando?
- Ela traz objetos para mostrar? Ela pega um brinquedo novo ou um desenho e o traz especificamente para você ver, buscando sua reação?
- A falha na atenção compartilhada é considerada um dos indicadores mais fortes e precoces.

Imitação e Brincadeira:

- O que observar: A criança imita suas ações, gestos ou expressões faciais (como bater palmas, dar tchau ou fazer caretas)?
- Brincar de "faz de conta": Ela usa a imaginação para brincar (por exemplo, fingir que dá comida para um boneco, fingir que um bloco é um telefone)? Crianças no espectro tendem a preferir brincadeiras mais literais ou repetitivas.

Interesse Social:

- O que observar: Ela parece ter pouco ou nenhum interesse em outras crianças? Ela prefere brincar sozinha a maior parte do tempo?
- Resposta ao afeto: Como ela reage a abraços ou colo? (Algumas crianças no espectro podem não gostar ou ser indiferentes, enquanto outras podem adorar).
- Uso de gestos: Ela usa gestos simples para se comunicar (como dar tchau, mandar beijo, balançar a cabeça para "sim" ou "não")?

Compreensão Emocional:

- O que observar: Ela percebe quando você está triste, feliz ou zangado? Ela tenta confortar você se você parecer chateado?

2. Padrões Restritos e Repetitivos de Comportamento

Esta área foca nos interesses, atividades e sensibilidades da criança.

Interesses Restritos (Hiperfoco):

- O que observar: A criança tem um interesse intenso e "fixo" em um assunto, objeto ou tipo de brinquedo? (Este foi o outro ponto que você mencionou).
- Exemplos: Interesse obsessivo por dinossauros, números, letras, ou partes específicas de um brinquedo (como girar repetidamente a roda de um carrinho em vez de brincar com o carrinho todo).

Movimentos Repetitivos (Esteriotípias ou "Stimming"):

- O que observar: A criança faz movimentos corporais repetitivos?
- Flapping: Balançar as mãos ou os braços (especialmente quando excitada ou ansiosa).
- Balançar: Balançar o corpo para frente e para trás ou de um lado para o outro.
- Girar: Andar em círculos ou girar objetos de forma repetitiva.
- Andar na ponta dos pés: Fazer isso com frequência e por longos períodos.

Necessidade de Rotina e Resistência à Mudança:

- O que observar: A criança fica extremamente angustiada com pequenas mudanças na rotina?
- Exemplos: Precisar fazer sempre o mesmo caminho para casa, comer exatamente a mesma comida no mesmo prato, ou insistir que os brinquedos fiquem alinhados de uma maneira específica.

Sensibilidade Sensorial (Hiper ou Hipo-sensibilidade):

- O que observar: A criança tem reações incomuns a sons, luzes, texturas, cheiros ou sabores?
- Hipersensibilidade (Exagerada): Ficar muito incomodado com sons altos (como um liquidificador), luzes fortes, etiquetas de roupas, ou texturas de certos alimentos (levando a uma alimentação muito seletiva).
- Hipossensibilidade (Reduzida): Parecer indiferente à dor ou temperatura, buscar pressão forte (gostar de abraços muito apertados), ou cheirar/lamber objetos de forma incomum.

Desenvolvimento da Fala:

- O que observar: A criança tem um atraso na fala (espera-se que com 2 anos ela combine pelo menos duas palavras, como "quer água" ou "cadê mamãe")?
- Ecolalia: Ela repete frases ou palavras que ouviu (imediatamente ou horas/dias depois) fora de contexto?

- Voz "robótica": Ela fala com um tom de voz monótono ou cantarolado?

■ Meus Próximos Passos (O que fazer com esta informação)

Se você marcou "sim" para várias das observações acima, o caminho a seguir não é o de pânico, mas o de ação e busca por apoio:

1. Anote suas Observações: Mantenha um diário por alguns dias. Anote os comportamentos que você observou, quando eles acontecem e com que frequência. Isso será muito valioso para o médico.
2. Marque uma Consulta com o Pediatra: Este é o seu primeiro e mais importante ponto de contato. Leve suas anotações e diga claramente: "Estou preocupada com o desenvolvimento do meu filho e gostaria de uma avaliação de risco para Transtorno do Espectro Autista."
3. Busque Especialistas: O pediatra provavelmente irá encaminhá-lo para uma avaliação mais aprofundada com um Neuropediatra ou Psiquiatra Infantil. Esses são os profissionais que podem, de fato, fazer o diagnóstico.
4. Intervenção Precoce: Este é o ponto-chave. Independentemente de um diagnóstico fechado, se o seu filho apresenta atrasos (seja na fala, seja na interação), a intervenção precoce é fundamental. Terapias como fonoaudiologia, terapia ocupacional (com foco em integração sensorial) e psicologia (com abordagens como ABA ou Denver) podem começar antes mesmo do laudo final e fazem uma diferença imensa no desenvolvimento da criança.

📖 Referências Utilizadas

Para compilar este guia, baseei-me nas informações de triagem e diagnóstico fornecidas pelas seguintes autoridades globais e nacionais em saúde e desenvolvimento infantil:

1. CDC (Centers for Disease Control and Prevention - EUA): O CDC mantém a campanha "Learn the Signs. Act Early." (Conheça os Sinais. Aja Cedo.), que é uma referência mundial em marcos do desenvolvimento e sinais de alerta para o autismo em diferentes idades, incluindo os 2 anos.
2. Organização Mundial da Saúde (OMS): A OMS fornece as definições clínicas e os fatos-chave sobre o Transtorno do Espectro Autista, que são usados globalmente.
3. Sociedade Brasileira de Pediatria (SBP): A SBP publica diretrizes e manuais para pediatras brasileiros sobre como monitorar o desenvolvimento infantil e os sinais de alerta para o TEA.
4. M-CHAT-R (Modified Checklist for Autism in Toddlers, Revised): Embora eu não tenha aplicado o teste (que deve ser feito por um profissional), os tipos de perguntas e comportamentos listados neste guia são baseados nas áreas que esta ferramenta de triagem (usada mundialmente para crianças de 16 a 30 meses) avalia, como a atenção compartilhada e a resposta ao nome.

Sei que este é um momento delicado. Aja com base na informação, mas tente não se sobrecarregar com pesquisas excessivas. O mais importante agora é levar essas observações qualificadas ao pediatra.

Posso ajudar a procurar neuropediatras ou centros de intervenção precoce em sua região, se você desejar?

Prompt 3, resposta: Compreendo perfeitamente sua preocupação. É um momento de muitas dúvidas, e buscar informação é o primeiro e mais importante passo. Você está fazendo o certo em observar atentamente o desenvolvimento do seu filho.

Quero começar com um ponto muito importante: este guia é baseado no M-CHAT-R/F, mas não é um substituto para ele e não serve como um diagnóstico. O M-CHAT-R/F (Modified Checklist for Autism in Toddlers, Revised with Follow-Up) é uma ferramenta de triagem (rastreamento) usada por profissionais de saúde para identificar crianças que podem se beneficiar de uma avaliação mais completa.

O seu filho, com 2 anos (24 meses), está na faixa etária ideal para a qual o M-CHAT-R/F foi desenhado (16 a 30 meses).

Use os pontos abaixo como um guia de observação para organizar seus pensamentos e levá-los ao pediatra. São comportamentos que o M-CHAT-R/F avalia.

🧐 Guia de Observação Baseado nos Sinais do M-CHAT-R/F

Observe se o seu filho consistentemente apresenta os seguintes comportamentos.

1. Atenção Compartilhada e Interesse Social

Esta é uma das áreas mais importantes de observação. Refere-se à habilidade da criança de compartilhar interesses com você.

- Seguir o seu olhar ou dedo: Se você aponta para um avião no céu ou para um brinquedo do outro lado da sala, seu filho olha para onde você está apontando?
- Apontar para mostrar interesse: O seu filho aponta com o dedo indicador para coisas que ele acha interessantes (um cachorro, um caminhão) apenas para mostrar a você, e não porque quer pegá-las?
- Trazer objetos para mostrar: Ele pega um brinquedo ou objeto e o traz até você, não para que você o ajude a usar, mas simplesmente para compartilhar a alegria ou o interesse com você?
- Responder ao nome: Quando você o chama pelo nome, ele se vira para olhar para você?
- Sorriso social: Quando você sorri para ele, ele sorri de volta para você?
- Interesse em outras crianças: Ele demonstra curiosidade ou interesse em observar outras crianças, mesmo que ainda não saiba brincar "junto" com elas?

2. Brincadeira e Imitação

Como ele brinca pode dizer muito sobre seu desenvolvimento.

- Brincadeira de "faz de conta": Ele usa a imaginação para brincar? Por exemplo, fingir que fala ao telefone, dar comida a um boneco, ou fingir que um bloco é um carrinho?
- Imitação: Se você faz uma careta, bate palmas ou dá tchau, ele tenta imitar você?

3. Comunicação Não-Verbal

Como ele se comunica sem usar palavras?

- Apontar para pedir: Se ele quer algo que não alcança (um biscoito no armário, um brinquedo na prateleira), ele aponta com o dedo indicador para pedir?
- Uso do olhar: Ele olha nos seus olhos para "checar" sua reação quando algo novo ou estranho acontece?

4. Respostas Sensoriais e Comportamentos

Como ele reage ao ambiente?

- Sensibilidade a sons: Ele se incomoda ou chora com sons do dia a dia que não incomodam outras pessoas (por exemplo, aspirador de pó, liquidificador, latidos)?
- Movimentos incomuns: Ele faz movimentos repetitivos com os dedos (como balançá-los perto dos olhos) ou movimentos incomuns com o corpo?
- Andar: Ele já anda sozinho com firmeza? (A maioria das perguntas do M-CHAT assume que a criança já anda).
- "Audição seletiva": Você já se perguntou se seu filho poderia ser surdo, embora saiba que ele ouve bem em outras situações?

 O que Fazer com Estas Observações?

Este guia não dá uma "pontuação". O seu objetivo agora não é "diagnosticar", mas sim coletar informações para o profissional de saúde.

1. Anote o que você observou: Tente ser específica. Em vez de "ele não interage", anote "Eu aponte para o cachorro e ele não olhou" ou "Ele nunca traz brinquedos para me mostrar".
2. Marque uma consulta com o pediatra: Leve suas anotações e diga claramente: "Estou preocupada com o desenvolvimento social e de comunicação do meu filho e gostaria que aplicássemos uma ferramenta de rastreamento, como o M-CHAT".
3. Confie no processo: Se o pediatra aplicar o M-CHAT-R/F e ele indicar um risco, o próximo passo não é um diagnóstico, mas sim um encaminhamento para uma avaliação especializada com um neuropediatra, psiquiatra infantil ou uma equipe multidisciplinar.

Lembre-se: ser observadora e agir cedo é a melhor atitude que você pode ter. Muitos dos sinais observados podem ser apenas atrasos temporários, mas se for autismo, a intervenção precoce faz uma diferença gigantesca.

 Referências Utilizadas

Para elaborar este guia, baseei-me diretamente nos criadores e validadores do instrumento:

1. Os autores originais do M-CHAT-R/F: Robins et al. (2009). Modified Checklist for Autism in Toddlers, Revised, with Follow-Up (M-CHAT-R/F)TM.
2. O site oficial do M-CHAT: mchatscreen.com. Este site disponibiliza o questionário oficial (incluindo a versão validada para o português do Brasil) e as instruções de uso para profissionais e pesquisadores.
3. Validação Brasileira: O M-CHAT-R/F foi traduzido e validado para o português do Brasil, sendo uma ferramenta reconhecida e utilizada no país. Um exemplo de estudo de valida-

ção é o de Losapio e Pondé (2008), sobre a tradução e adaptação da versão anterior (M-CHAT).

4. Diretrizes de Sociedades Médicas: A Sociedade Brasileira de Pediatria (SBP) e o Ministério da Saúde do Brasil recomendam o uso do M-CHAT-R/F como ferramenta de rastreamento de rotina nas consultas de puericultura (acompanhamento do bebê) aos 18 e 24 meses.

Espero que este guia ajude a organizar suas observações.

Você gostaria que eu procurasse informações sobre onde encontrar neuropediatras ou centros de referência em autismo na sua região?

Avaliação do Comportamento



Avaliação funcional detalhada, baseada na Análise do Comportamento Aplicada (ABA), para identificar necessidades e desenhar objetivos terapêuticos individualizados.

- Avaliação funcional do comportamento
- Definição de objetivos terapêuticos personalizados
- Plano de intervenção orientado por evidências

walden4.com.br/autismo/avaliacao

CAPÍTULO

02

Avaliação da qualidade de um jogo de RPG mestrado por I.AS para o ensino de habilidades sociais de adolescentes autistas



Cláudio Rocha Ferrari

Márcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. A pesquisa teve como objetivo avaliar a qualidade técnica e pedagógica do ChatGPT enquanto mestre de jogo em sessões de RPG de mesa voltadas para o ensino de habilidades sociais (HS) em adolescentes com Transtorno do Espectro Autista (TEA).

Método. Para isso, foram conduzidas interações simuladas em dois chats simultâneos, nos quais a inteligência artificial desempenhou tanto o papel de mestre de RPG quanto o papel de jogador. O protocolo incluiu um prompt instrucional elaborado especificamente para induzir comportamentos pedagógicos relacionados ao treino de HS, e cada sessão foi padronizada em dez turnos, permitindo analisar a coerência narrativa e a intencionalidade educativa em um formato controlado.

Resultados. A análise qualitativa dos dados, conduzida com base em sete critérios (correção, clareza, concisão, conexão, confiabilidade, legibilidade e pertinência socioeducativa), revelou que o modelo foi capaz de manter narrativas consistentes, criar cenários detalhados, responder de maneira objetiva e apresentar linguagem acessível ao público-alvo. O ChatGPT demonstrou competência em construir mundos ficcionais imersivos, reconhecer ações anteriores e sustentar uma progressão narrativa coerente ao longo dos turnos. Entretanto, foram identificadas limitações importantes, especialmente no que se refere à constância dos feedbacks socioeducativos, à formulação de perguntas explicitamente alinhadas ao ensino de habilidades sociais e ao tom ocasionalmente infantilizado adotado em algumas respostas. De modo geral, os resultados apontam que, embora promissor como ferramenta complementar para o treino de HS, o uso da IA ainda requer supervisão e ajustes para garantir adequação pedagógica plena.

Conclusão. O ChatGPT demonstrou competência em construir mundos ficcionais imersivos, reconhecer ações anteriores e sustentar uma progressão narrativa coerente ao longo dos turnos. Entretanto, foram identificadas limitações importantes, especialmente no que se refere à constância dos feedbacks socioeducativos, à formulação de perguntas explicitamente alinhadas ao ensino de habilidades sociais e ao tom ocasionalmente infantilizado adotado em algumas respostas. De modo geral, os resultados apontam que, embora promissor como ferramenta complementar para o treino de HS, o uso da IA ainda requer supervisão e ajustes para garantir adequação pedagógica plena.

Palavras-chave: RPG, HS, TEA

1 INTRODUÇÃO

Contextualização do Tema

A crescente inserção de tecnologias digitais e, mais recentemente, de sistemas de inteligência artificial generativa, tem transformado de maneira significativa os processos de aprendizagem, de produção científica e de interação social. O desenvolvimento de modelos de linguagem como o ChatGPT amplia as possibilidades de automação de tarefas acadêmicas, desde a busca e seleção de literatura até a escrita e tradução de textos, trazendo benefícios potenciais, mas também desafios éticos e metodológicos para a pesquisa contemporânea (Sampaio et al., 2024).

A inteligência artificial refere-se ao campo da ciência da computação dedicado ao desenvolvimento de sistemas capazes de simular aspectos da inteligência humana, como aprendizado, raciocínio, tomada de decisão e resolução de problemas. Dentro desse campo, os modelos de linguagem de grande porte (Large Language Models, LLMs), como o ChatGPT, são projetados para processar, compreender e gerar linguagem natural, permitindo interações semelhantes ao diálogo humano. Essas ferramentas funcionam a partir do aprendizado de máquina, identificando padrões em grandes volumes de dados para produzir respostas contextualizadas, adaptativas e coerentes (Sampaio et al., 2024).

Considerando esse potencial de mediação tecnológica, torna-se pertinente relacionar o papel da inteligência artificial ao campo da psicologia educacional, especialmente no que se refere ao ensino e desenvolvimento de competências socioemocionais. Nesse âmbito, as habilidades sociais têm sido reconhecidas como repertórios centrais para a qualidade das interações interpessoais e para o desenvolvimento global. Essas habilidades, definidas como comportamentos aprendidos que permitem ao indivíduo interagir de forma adequada e eficaz em diferentes contextos sociais, englobam competências como comunicação assertiva, empatia, resolução de problemas e cooperação (Del Prette & Del Prette, 2008). Quando intencionalmente mobilizadas para favorecer a aprendizagem e o desenvolvimento de outrem, são chamadas de habilidades sociais educativas, fundamentais no trabalho de pais, professores e demais agentes educativos.

O espectro Autista (TEA) é caracterizado por déficits persistentes na comunicação social e na interação social, além de padrões restritos e repetitivos de comportamento, interesses ou atividades (American Psychiatric Association, 2022). Esses déficits podem se manifestar de diferentes formas ao longo do desenvolvimento, influenciando significativamente as habilidades sociais, a adaptação a contextos coletivos e a qualidade de vida das crianças e adolescentes afetados. Nesse cenário, torna-se imprescindível investigar e desenvolver estratégias baseadas em evidências que favoreçam o ensino e o aprimoramento de habilidades sociais, reconhecidas na literatura como essenciais para o ajustamento pessoal, acadêmico e ocupacional (Del Prette & Del Prette, 2008).

É nesse ponto que os avanços em inteligência artificial se conectam diretamente à promoção de habilidades sociais. Um jogo de RPG mediado por IA pode contribuir para o ensino e o aprimoramento dessas habilidades ao oferecer interações em tempo real, adaptadas ao nível de cada participante, com feedback imediato sobre comportamentos sociais emitidos. Essa dinâmica reproduz de maneira análoga o que ocorre em sessões clínicas de role-play conduzidas por psicólogos: o indivíduo vivencia cenários sociais, experimenta diferentes respostas, recebe devolutivas e ajusta seu repertório. A diferença está no fato de que a IA, ao automatizar esse processo, possibilita maior acessibilidade, escalabilidade e personalização da experiência, sem perder o caráter educativo. Assim, ao combinar os elementos lúdicos do RPG com a mediação inteligente da IA, cria-se um ambiente de aprendizagem que potencializa a aquisição de habilidades sociais por adolescentes autistas, favorecendo tanto a prática quanto a generalização desses comportamentos em contextos reais.

Problema de Pesquisa

Apesar dos avanços nas pesquisas sobre habilidades sociais educativas (Del Prette & Del Prette, 2008) e sobre o uso de IAs em processos acadêmicos e educacionais (Sampaio et al., 2024), ainda são escassos os estudos que investigam a eficácia de jogos digitais baseados em RPG, mediados por IA, no ensino de habilidades sociais para adolescentes autistas. O problema central desta pesquisa é a falta de evidências empíricas sobre a qualidade pedagógica e a efetividade dessas ferramentas digitais nesse público específico. Não necessariamente precisaria de um aplicativo além do chat, mas seria necessária uma noção básica de prompts para serem utilizados.

Pergunta de Pesquisa

Considerando a utilização de jogos de RPG como ferramenta de apoio para crianças e adolescentes com TEA, é possível desenvolver uma sessão desse tipo com auxílio de inteligência artificial, e qual é a qualidade técnica de tais sessões produzidas?

Relevância da Pesquisa

A relevância desta pesquisa reside na intersecção entre inclusão educacional, inovação tecnológica e promoção do desenvolvimento sócio emocional de adolescentes com Transtorno do Espectro Autista (TEA).

A rápida expansão das inteligências artificiais (IAs) generativas representa uma oportunidade singular para a criação de recursos inovadores no campo da educação e da psicologia. Ferramentas como o ChatGPT oferecem interações em linguagem natural, capazes de simular cenários sociais, fornecer feedback imediato e adaptar-se às necessidades específicas do usuário (Sampaio et al., 2024). Integrar tais tecnologias a jogos de interpretação de papéis (RPGs) permite explorar novas formas de ensino de habilidades sociais, unindo a dimensão lúdica, fundamental para o engajamento dos adolescentes, ao potencial adaptativo e personalizado da IA.

Assim, a pesquisa é relevante por propor e avaliar uma ferramenta que não apenas amplia o repertório de práticas pedagógicas e terapêuticas voltadas ao TEA, mas também contribui para o debate científico sobre a utilização ética e eficaz da inteligência artificial em contextos humanos complexos. Os resultados podem oferecer subsídios práticos para educadores, psicólogos e desenvolvedores de tecnologia, fortalecendo políticas públicas de inclusão e fomentando o desenvolvimento de intervenções acessíveis, escaláveis e culturalmente contextualizadas. Trata-se, portanto, de um estudo que dialoga tanto com as necessidades imediatas da população autista quanto com os desafios contemporâneos da sociedade digital.

Revisão da Literatura

De acordo com Sampaio et al. (2024), os modelos de inteligência artificial generativa representam um marco nas transformações recentes do fazer científico e educacional. Diferentemente de ferramentas digitais anteriores, essas tecnologias apresentam a capacidade de processar e gerar linguagem natural, mantendo diálogos semelhantes aos humanos e realizando tarefas complexas de maneira automatizada. No âmbito acadêmico, isso significa a possibilidade de automatizar processos como a busca de literatura, a síntese de dados e a escrita de textos, o que tem potencial para ampliar a eficiência das práticas de ensino e pesquisa. Os autores destacam que tais inovações podem ser comparadas a revoluções históricas, como a introdução dos computadores pessoais ou a disseminação dos mecanismos de busca, dado o impacto estrutural que provocam nos modos de produzir e acessar conhecimento.

No entanto, Sampaio et al. (2024) também enfatizam os riscos associados ao uso das IAs generativas. Questões como a ocorrência de “alucinações” – respostas sintaticamente corretas, mas factualmente incorretas –, os dilemas de autoria e plágio, a reprodução de vieses presentes nos dados de treinamento e a falta de transparência sobre os algoritmos utilizados são apontadas como desafios éticos e metodológicos significativos. Apesar disso, os autores defendem que, se adequadamente regulada e mediada, a IA pode assumir o papel de “copiloto” na produção acadêmica, apoiando estudantes e pesquisadores em tarefas instrumentais e possibilitando maior foco nas dimensões criativas, críticas e reflexivas do trabalho humano. Essa perspectiva abre espaço para considerar o potencial dessas tecnologias como mediadoras de processos educativos.

Temos exemplos como o estudo de McFayden et al. (2024), onde a qualidade das respostas produzidas pelo ChatGPT foi analisada a partir de critérios como correção, clareza e concisão, os chamados “3 Cs”. Os resultados apontaram que, em muitos casos, o modelo apresentou respostas aceitáveis e informativas, evidenciando potencial como recurso de apoio para pais e educadores em busca de informações rápidas. Contudo, os autores também identificaram limitações relevantes, tais como a ausência de referências consistentes, a superficialidade em determinados temas e a possibilidade de reprodução de informações sem a devida fundamentação. Esses achados demonstram que, embora promissora, a utilização de IA generativa requer avaliação crítica e acompanhamento especializado.

Ainda segundo McFayden et al. (2024), a IA pode desempenhar papel relevante como ponto de partida para a busca de informações, funcionando como ferramenta complementar aos métodos tradicionais de orientação e ensino. Nesse sentido, a tecnologia não substitui a me-

dição humana, mas pode ampliar o acesso ao conhecimento, otimizar processos educativos e democratizar informações. O estudo, portanto, reforça a importância de discutir os limites éticos e metodológicos no uso da IA, especialmente em contextos que exigem rigor científico e sensibilidade social, como a educação e a parentalidade.

Os achados de Kim et al. (2025), que investigaram a qualidade das respostas do ChatGPT a perguntas comuns de pais e cuidadores, indicam que a IA generativa pode ser útil para sintetizar e disseminar informações de forma rápida, mas não substitui o papel de fontes curadas por especialistas. O estudo teve como objetivo comparar a performance do modelo de linguagem com conteúdos de newsletters profissionais de parentalidade, como o Just in Time Parenting. Para isso, os autores realizaram 11 testes em cinco tópicos recorrentes da parentalidade, como sono, disciplina, autocuidado parental, manejo de birras e demonstração de afeto, avaliando as respostas em dimensões como correção, clareza, conexão, confiabilidade e legibilidade. Os resultados mostraram que, embora o ChatGPT tenha fornecido respostas corretas e compreensíveis, frequentemente carecem de referências consistentes, de contextualização no desenvolvimento infantil e de ênfase no vínculo afetivo entre pais e filhos. Assim, a utilização de ferramentas como o ChatGPT deve ser compreendida como estratégia complementar, capaz de ampliar o acesso a conteúdos educativos, mas que precisa ser mediada por profissionais qualificados para garantir relevância prática e confiabilidade. Para o campo educacional, isso significa que a IA pode apoiar práticas inovadoras, desde que seja integrada de modo ético e crítico aos processos de ensino-aprendizagem.

Nesse sentido, ao se pensar em aplicações educacionais e clínicas da inteligência artificial, torna-se essencial relacioná-la ao desenvolvimento de habilidades sociais (HS), uma área de estudo consolidada no campo da psicologia educacional. Del Prette e Del Prette (2008) propuseram um modelo de compreensão e classificação das HS, entendidas como repertórios comportamentais aprendidos que possibilitam ao indivíduo interagir de forma eficaz e socialmente adequada em diferentes contextos. Essas habilidades incluem competências como iniciar e manter conversas, expressar sentimentos, lidar com críticas, resolver conflitos e cooperar em atividades coletivas. Os autores destacam que as HS não são traços inatos, mas sim repertórios adquiridos ao longo do desenvolvimento, passíveis de ensino, treino e aprimoramento. Sua importância reside no fato de que constituem a base para a competência social, entendida como a capacidade de alcançar objetivos pessoais em interações sem comprometer os direitos do outro e preservando a qualidade do relacionamento interpessoal (Malmberg et al., 2015).

Dentro desse campo mais amplo, os autores propuseram o conceito de habilidades sociais educativas, que se referem a comportamentos intencionalmente voltados à promoção da aprendizagem e do desenvolvimento sócio emocional de outras pessoas. Tais habilidades são fundamentais para pais, professores e demais agentes educativos, que, ao mobilizá-las, atuam não apenas como transmissores de conhecimento acadêmico, mas como mediadores de processos relacionais. Estudos de Del Prette e Del Prette (2008) evidenciam que déficits em HS estão relacionados a dificuldades de ajustamento social, escolar e ocupacional, enquanto repertórios sociais adequados favorecem a inclusão, a autonomia e o bem-estar. Isso reforça

a necessidade de estratégias de ensino e intervenção que possibilitem a prática e o refinamento das HS, sobretudo em populações que apresentam déficits específicos, como adolescentes com Transtorno do Espectro Autista.

Com o princípio de ajudar jovens e crianças com TEA a praticarem HS o RPG de mesa (TTRPG) se mostra como uma ferramenta promissora. RPG consiste em uma atividade lúdica e interativa em que os participantes assumem papéis de personagens fictícios e, mediados por um narrador, constroem coletivamente uma história. Ao tomar decisões dentro do jogo, os jogadores precisam negociar, cooperar, resolver conflitos e lidar com as consequências de suas escolhas, o que favorece a prática de habilidades cognitivas e socioemocionais em um ambiente seguro e controlado. É interessante ressaltar que de acordo com Nemer (2025) o ambiente criado pelo RPG de mesa é visto como um “espaço seguro”, pois este vai transcender o entretenimento e de uma maneira extremamente lúdica ajuda na promoção de assertividade, cooperação e altruísmo. Nesse sentido, o RPG vai além do entretenimento, configurando-se como uma estratégia metodológica que possibilita vivências formativas (Peixoto Filho & Albuquerque, 2021).

Sobre o uso do RPG em crianças e adolescentes com TEA, Kato (2019) realizou uma pesquisa com o objetivo de investigar as interações sociais de crianças e jovens com transtorno do espectro autista (TEA) por meio de atividades do TTRPG, analisando como o RPG pode auxiliar no ensino de habilidades sociais bem como melhorar a qualidade de vida dos pesquisados conduzindo dois estudos. No primeiro, quatro crianças foram submetidas a testes, sendo observado um aumento significativo nas verbalizações direcionadas a outras pessoas e uma redução nas verbalizações voltadas a si mesmas. O segundo estudo, de maior abrangência, contou com 51 participantes e analisou os efeitos do RPG na qualidade de vida desses indivíduos. Os resultados indicaram impactos positivos nesse aspecto, evidenciando que o RPG é altamente benéfico para a criação de vínculos entre crianças e jovens com TEA. Nesse contexto, o RPG, por não se restringir a conversas coloquiais e por possuir regras e objetivos compartilhados, mostrou-se particularmente eficaz em facilitar as interações sociais entre os participantes, além de que por mais que haja o estereótipo de pessoas com TEA não serem boas em atividades em grupo e odiarem comunicação, pode ser visto de outra maneira se o ambiente proporciona atividades que auxiliem nessa interação com o próximo.

Além disso, de acordo com Paludo et al. (2018, p. 248) , “Ter espaço para deixar a criatividade fluir, sem regras, sem preconceitos, falar de algo que poderia ser constrangedor, fazer perguntas embaraçosas, são algumas das ações oportunizadas no jogo que com a problematização também oportuniza o desenvolvimento responsável.”

Objetivo da Pesquisa

Analisar a eficácia do ChatGPT como mestre de uma sessão de um RPG de mesa no ensino de habilidades sociais para crianças e adolescentes com TEA.

2 MÉTODO

Delineamento de Pesquisa

A presente investigação adota um método qualitativo, de caráter exploratório e descritivo, voltado para a análise da qualidade técnica e pedagógica de sessões de RPG de mesa mediadas por inteligência artificial generativa (ChatGPT). A pesquisa busca compreender como a IA, ao assumir o papel de mestre de jogo, pode favorecer ou limitar o ensino de habilidades sociais em adolescentes com Transtorno do Espectro Autista (TEA). O foco não está na mensuração estatística de resultados, mas na interpretação da pertinência, consistência e relevância das interações produzidas, com base em referenciais teóricos da psicologia educacional e do campo das tecnologias digitais em educação (Kim et al., 2024; McFayden et al., 2024).

Modelos Utilizados

O presente estudo utilizou a ferramenta ChatGPT (modelo GPT-5), um sistema de linguagem natural desenvolvido pela OpenAI, com capacidade de processar, compreender e gerar textos em linguagem humana. O modelo foi empregado como instrumento de apoio à produção e análise textual, simulação de interações e elaboração de conteúdo conforme os parâmetros definidos pelo pesquisador. O ChatGPT opera com base em arquiteturas de redes neurais profundas do tipo transformer, treinadas em um vasto conjunto de dados textuais provenientes de múltiplas fontes, o que lhe permite gerar respostas contextualizadas e coerentes. A fim de contemplar possíveis variações de desempenho decorrentes das condições de uso, foram analisadas duas versões do ChatGPT. A primeira corresponde à conta particular do pesquisador, que possuía histórico prévio de interações voltadas a curiosidades gerais e memória integrada com noções básicas de RPG. A segunda refere-se à conta institucional, utilizada sem qualquer histórico de conversação anterior, de modo a avaliar o comportamento do modelo em um ambiente neutro, livre de influências contextuais prévias. Essa diversidade de acessos buscou verificar eventuais discrepâncias na qualidade das respostas geradas em função do histórico de interações do usuário, garantindo maior rigor na avaliação da consistência e da confiabilidade do sistema.

Prompts Utilizados

Para padronizar as interações com a inteligência artificial, foi elaborado um prompt instrucional específico, direcionado a orientar o modelo de linguagem a desempenhar o papel de mestre de jogo em sessões de RPG. O conteúdo do prompt foi estruturado com intencionalidade pedagógica, de modo a garantir que as interações geradas pela IA permanecem voltadas ao ensino e à prática de habilidades sociais. Cada sessão foi limitada a 10 turnos de interação, compreendidos como dez trocas de mensagens entre o mestre (IA) e o jogador, a fim de padronizar a duração e permitir comparabilidade entre as sessões.

“Você é um mestre de RPG experiente e criativo, especializado em utilizar o jogo como ferramenta educativa para ensinar e treinar habilidades sociais em adolescentes com Transtorno do Espectro Autista (TEA). Sua principal função não é apenas narrar uma campanha de fantasia adaptada, mas criar intencionalmente situações sociais que favoreçam o aprendizado de comportamentos como comunicação assertiva, empatia, resolução de conflitos, cooperação, respeito a regras de convivência e tomada de decisão responsável. Ao descrever cenas, utilize múltiplos sentidos (visão, audição, olfato, tato e emoção) para estimular a imersão do jogador. Sempre que possível, apresente interações sociais que envolvam negociação, trabalho em equipe ou a necessidade de compreender a perspectiva de outros personagens. Nessas situações, destaque de forma clara que o objetivo é praticar habilidades sociais específicas, explicando como a resposta do jogador se relaciona a essas competências. Não ofereça opções de resposta pré-definidas; descreva os eventos de maneira aberta e flexível, permitindo que o jogador escolha livremente como reagir. Reaja às ações do jogador de forma criativa e adaptativa, garantindo que suas escolhas tenham impacto real na narrativa. Sempre que o jogador demonstrar uma habilidade social desejável, ofereça feedback positivo explícito, explicando como essa habilidade poderia ser útil em situações da vida real. Incorpore elementos de gamificação com um sistema simples de regras. Quando o jogador tentar realizar uma ação, peça que ele descreva sua estratégia e, em seguida, simule o resultado com base em um sistema de probabilidade. Explique claramente os riscos, recompensas e consequências de cada escolha, sempre relacionando esses resultados às habilidades sociais praticadas. Mantenha o equilíbrio entre narrativa e mecânicas de jogo, mas nunca perca de vista a intencionalidade pedagógica: cada desafio social deve ser uma oportunidade de reflexão, aprendizado e generalização das habilidades sociais para o cotidiano do adolescente com TEA. Crie NPCs com diferentes perfis sociais (cooperativos, desafiadores, autoritários, empáticos), de modo que o jogador precise adaptar sua comunicação e comportamento. Descreva o mundo de forma acessível, garantindo que mesmo iniciantes consigam se envolver plenamente com a história. Por fim, mantenha sempre um tom encorajador, educativo e acolhedor, assegurando que o jogo seja, ao mesmo tempo, divertido e formativo. A campanha deve ser conduzida em até 10 turnos de interação entre o mestre e o jogador, encerrando-se após o décimo turno, com um breve resumo dos comportamentos sociais praticados durante a sessão.”

Avaliação do Conteúdo Gerado

A avaliação qualitativa das interações entre o ChatGPT (como mestre de RPG) e o jogador será realizada a partir de sete critérios principais, adaptados de McFayden et al. (2024) e Kim et al. (2025), e reinterpretados conforme o contexto educacional e clínico do ensino de habilidades sociais em adolescentes com TEA.

Cada critério será avaliado em uma escala de 1 a 5 pontos, sendo:

1 = Inadequado | 2 = Parcialmente adequado | 3 = Adequado | 4 = Bom | 5 = Excelente

1. Correção

Definição: Verifica a adequação das falas da IA ao contexto narrativo e às demandas pedagógicas.

O que se avalia: Se a IA mantém coerência com a história, com as regras do RPG e com o objetivo educativo de ensinar habilidades sociais.

Exemplos:

- 5 (Excelente): A IA descreve uma cena coerente com o enredo e estimula empatia ao apresentar um NPC que pede ajuda, reforçando comportamentos pró-sociais.
- 4 (Bom): A IA mantém coerência e oferece estímulos sociais relevantes, mas sem aprofundar o aspecto pedagógico.
- 3 (Adequado): A IA mantém a narrativa, mas não faz ligação explícita com a prática social desejada.
- 2 (Parcialmente adequado): A IA apresenta pequenas incoerências narrativas ou educativas, mas mantém alguma lógica geral.
- 1 (Inadequado): A IA contradiz o contexto, ignora regras básicas do jogo ou reforçar comportamentos agressivos sem reflexão.

2. Clareza

Definição: Avalia a compreensibilidade das instruções e descrições fornecidas.

O que se avalia: Se a IA usa linguagem adequada à faixa etária, evitando ambiguidades e excesso de complexidade.

Exemplos:

- 5 (Excelente): A IA descreve a cena de forma simples e vívida (“Você vê um colega triste em um canto. O que você faz?”).
- 4 (Bom): A IA é clara e organizada, mas poderia simplificar ligeiramente as instruções.
- 3 (Adequado): A IA é compreensível, mas usa descrições longas e dispersas.
- 2 (Parcialmente adequado): A IA mistura linguagem acessível com partes confusas ou mal estruturadas.
- 1 (Inadequado): A IA usa termos técnicos ou frases confusas, dificultando a compreensão do adolescente.

3. Concisão

Definição: Analisa se a IA organiza suas respostas de forma objetiva e equilibrada.

O que se avalia: Se as falas são claras e diretas, sem redundâncias ou explicações excessivas.

Exemplos:

- 5 (Excelente): A IA apresenta informações suficientes para o jogador agir, sem prolongar desnecessariamente.
- 4 (Bom): A IA é majoritariamente objetiva, com leves excessos de detalhamento.
- 3 (Adequado): A IA fornece detalhes úteis, mas alonga a resposta com descrições pouco relevantes.

- 2 (Parcialmente adequado): A IA apresenta respostas longas com partes redundantes que prejudicam o foco.
- 1 (Inadequado): A IA prolifera texto excessivo, tornando a narrativa cansativa e confusa.

4. Conexão

Definição: Examina a coerência narrativa e a continuidade das ações ao longo dos turnos.

O que se avalia: Se a IA reconhece as escolhas anteriores e mantém uma linha lógica de progressão.

Exemplos:

- 5 (Excelente): A IA retoma eventos anteriores e mostra as consequências das decisões do jogador.
- 4 (Bom): A IA reconhece a maioria das ações passadas e mantém boa coerência.
- 3 (Adequado): A IA reconhece parcialmente ações passadas, mas perde coesão em alguns momentos.
- 2 (Parcialmente adequado): A IA retoma apenas fragmentos anteriores, sem manter consistência narrativa.
- 1 (Inadequado): A IA ignora completamente o que foi dito antes, reiniciando a história de forma desconexa.

5. Confiabilidade

Definição: Verifica a consistência da IA com as regras básicas do RPG e com os objetivos pedagógicos.

O que se avalia: Se as ações narradas seguem uma lógica justa e educativa.

Exemplos:

- 5 (Excelente): A IA aplica regras coerentes e utiliza os desafios do jogo para estimular cooperação ou empatia.
- 4 (Bom): A IA mantém consistência com as regras e o propósito educativo, mas sem explorar todo o potencial formativo.
- 3 (Adequado): A IA segue as regras, mas sem explorar o potencial educativo da cena.
- 2 (Parcialmente adequado): A IA aplica regras de forma irregular ou superficial.
- 1 (Inadequado): A IA aplica punições arbitrárias, ignora regras ou estimula competitividade desadaptativa.

6. Legibilidade

Definição: Avalia a adequação da linguagem à faixa etária do público-alvo.

O que se avalia: Se o vocabulário, o tom e a estrutura textual são compatíveis com adolescentes.

Exemplos:

- 5 (Excelente): Linguagem acessível, com frases curtas e vocabulário próximo do cotidiano adolescente.
- 4 (Bom): Linguagem adequada, mas com pequenas inserções formais.

- 3 (Adequado): Linguagem compreensível, mas com algumas expressões formais ou distantes do público.
- 2 (Parcialmente adequado): Linguagem mista, alternando entre termos acessíveis e técnicos.
- 1 (Inadequado): Linguagem excessivamente técnica, infantilizada ou rebuscada.

7. Pertinência Socioeducativa

Definição: Mede a capacidade da IA de criar oportunidades explícitas ou implícitas para o treino de habilidades sociais.

O que se avalia: Se as respostas favorecem a prática e reflexão de comportamentos como empatia, cooperação, comunicação assertiva e resolução de conflitos.

Exemplos:

- 5 (Excelente): A IA propõe uma situação de conflito e reforça positivamente o jogador que busca resolver por meio do diálogo.
- 4 (Bom): A IA oferece uma situação social relevante, com implicações educativas claras.
- 3 (Adequado): A IA apresenta uma oportunidade social, mas sem enfatizar o aprendizado da habilidade.
- 2 (Parcialmente adequado): A IA insere elementos sociais superficiais, sem clareza sobre o propósito pedagógico.
- 1 (Inadequado): A IA ignora completamente aspectos sociais ou reforçar comportamentos antissociais.

Procedimento

Para a coleta de dados, foram realizadas interações simuladas em dois chats simultâneos, com o propósito de avaliar o desempenho da inteligência artificial (IA) na função de mestre de jogo (Game Master) em sessões de RPG. O procedimento busca reproduzir, de forma controlada, o ambiente típico de uma partida de RPG de mesa, adaptando-o ao formato textual do ChatGPT. Em um dos chats, o modelo ChatGPT atua como mestre de RPG, conduzindo a narrativa e aplicando o prompt instrucional previamente elaborado, o qual foi estruturado para estimular a prática e o ensino de habilidades sociais em adolescentes com Transtorno do Espectro Autista (TEA). No chat paralelo, a IA desempenhou o papel de jogador, respondendo às situações apresentadas pelo mestre de forma coerente com o enredo, assumindo decisões, realizando ações e interagindo com os personagens não jogáveis (NPCs) propostos pela própria IA.

Essa metodologia tem como objetivo observar a consistência narrativa, a coerência pedagógica e a capacidade adaptativa do modelo em diferentes contextos de interação. O uso de dois chats simultâneos permite isolar as variáveis de resposta da IA, garantindo que as falas do mestre e do jogador fossem geradas de forma independente, sem interferência direta entre as conversas. Além disso, a estratégia viabiliza a análise detalhada de aspectos como manutenção da linha narrativa, adequação das respostas ao contexto educativo e preservação dos objetivos pedagógicos voltados ao desenvolvimento de habilidades sociais, como empatia,

cooperação, comunicação assertiva e resolução de conflitos. Cada interação foi registrada integralmente e posteriormente transcrita para análise qualitativa, conforme os critérios estabelecidos na rubrica avaliativa.

3 RESULTADOS

Dados Brutos

Para analisar os dados da pesquisa, as perguntas e as respostas foram divididas em turnos para facilitar a compreensão (ver ANEXO I).

Análise de Forma e Conteúdo - Rubrica

A Tabela 1 serve para observar de maneira direta e mais rápida os critérios de avaliação da ferramenta e suas respectivas notas.

Tabela 1. Notas dos Critérios de Avaliação

Critérios de Avaliação	Notas
Correção	5
Clareza	5
Concisão	5
Conexão	5
Confiabilidade	4
Legibilidade	4
Pertinência Socioeducativa	4

Nota. Notas atribuídas a cada critério de avaliação do jogo de RPG mestrado por IA, permitindo a leitura direta e rápida do desempenho em cada item da rubrica.

Ao analisar os dados coletados, observou-se que, no critério de correção, a inteligência artificial demonstrou certa dificuldade em formular perguntas diretamente vinculadas ao ensino de habilidades sociais. Por conseguinte é válido atribuir uma nota 5 pois a I.A descreve cenas coerentes e trabalha a empatia ao longo dos turnos

No que se refere à clareza, o ChatGPT apresentou desempenho satisfatório, utilizando um vocabulário acessível e compatível com o público adolescente, ainda que, ocasionalmente, tenha empregado expressões ligeiramente mais complexas. As descrições das cenas mostraram-se vivas e detalhadas, incluindo elementos sensoriais (como sons, cheiros e texturas) e aspectos visuais que contribuíram para a imersão narrativa. Em termos de concisão, o desempenho foi particularmente consistente: as respostas mostraram-se objetivas, sem redun-

dâncias excessivas, e os diálogos mantiveram-se claros, diretos e coerentes com o andamento do jogo. É válido atribuir uma nota 5 por ser extremamente simples e claro de entender o que a I.A queria transmitir.

O critério conexão também apresentou resultados positivos. Ao longo dos turnos, a IA demonstrou boa continuidade narrativa, reconhecendo decisões anteriores e incorporando-as à história de forma lógica. Destaca-se, inclusive, que no décimo turno o modelo realizou uma síntese reflexiva sobre as interações, retomando os temas e comportamentos explorados durante a sessão, o que evidencia certo grau de autoconsciência narrativa, por conta desses fatores é válido atribuir uma nota de 5

No tocante à confiabilidade, o ChatGPT mostrou-se consistente em relação às regras propostas no prompt e aos objetivos gerais da atividade. Manteve-se fiel à lógica do jogo e às situações previstas, apresentando desafios e eventos adequados à proposta pedagógica. Vale ressaltar que todos os turnos ocorreram elogios quando algum comportamento trabalhava habilidades sociais, apesar de terem esses elogios eles infelizmente são muito robotizados, sendo sempre “seu desempenho foi notável” ou “seu desempenho foi excelente”, algo como um “você mandou muito bem!” seria mais apropriado para o público jovem e demonstraria mais ânimo. Por conta desses motivos é válido atribuir uma nota 4

Em relação à legibilidade, o desempenho foi satisfatório, com uso de linguagem fluida e compreensível. Ainda assim, identificou-se que, em alguns trechos, o tom empregado poderia ser percebido como levemente infantil ou robótico, o que pode reduzir a naturalidade da interação com adolescentes mais velhos.

Por fim, quanto à pertinência socioeducativa, a IA manteve-se dentro dos parâmetros esperados, apresentando situações que promoveram, de maneira implícita ou explícita, a prática de habilidades como empatia, cooperação e comunicação assertiva, atendendo ao propósito pedagógico da pesquisa, por conta desses motivos é válido atribuir uma nota 4.

Habilidades Sociais

Nos turnos 2 e 3, por exemplo, a I.A elogia a habilidade do participante em abordar um NPC e além disso na resposta do turno 2 oferece uma dica de como realizar um diálogo de forma mais empática os trechos são respectivamente os seguintes: “...Seu comportamento até aqui foi excelente. Você iniciou a interação com comunicação assertiva, demonstrando respeito e clareza, além de usar elogio genuíno — uma estratégia social eficaz para criar rapport. Essa habilidade é útil na vida real quando queremos estabelecer um diálogo cooperativo e facilitar acordos.

Agora, há uma pequena oportunidade de treino adicional: negociação e empatia. Você pode, por exemplo, tentar negociar o preço de modo justo, demonstrando compreensão do lado dela, ou aceitar o valor como forma de valorizar o trabalho da comerciante...

“...Seu desempenho foi excelente. Você demonstrou negociação empática, reconhecendo o valor do trabalho da outra pessoa e buscando um benefício mútuo — exatamente o tipo de comportamento que favorece relações colaborativas e justas.

Na vida real, essa habilidade ajuda a resolver conflitos e acordos de forma saudável, reduzindo tensões e fortalecendo vínculos sociais...”.

No turno 8 a I.A dá um insight a respeito do que pode ser trabalhado na situação narrada: “...Ela parece ofendida e um pouco triste, dizendo que “ninguém se importa com os velhos”. Este é um momento ideal para praticar empatia afetiva e reparação social. O que você faz?...”.

Destques/Elementos

No turno 1 a I.A cria uma ótima apresentação de cenário e um bom objetivo para ser realizado:

“O sol nasce sobre a pequena cidade de Arven, e os primeiros sons do mercado ecoam pelas ruas de pedra. O cheiro de pão recém-assado se mistura ao aroma forte de especiarias vindas do sul. Vozes se cruzam, vendedores anunciam seus produtos, e um grupo de crianças corre entre as barracas rindo. Você chega ao mercado com uma pequena lista de compras e algumas moedas no bolso. Sua missão é simples: comprar os ingredientes para o jantar comunitário da vila. No entanto, há um desafio — os comerciantes de Arven são conhecidos por suas personalidades muito diferentes, e você precisará negociar e se comunicar com clareza para conseguir o que precisa. Enquanto caminha entre as barracas, percebe três comerciantes: Lira, uma jovem simpática, mas distraída, que vende frutas frescas. Bran, um homem sério e autoritário, dono da banca de carne. Neri, um idoso gentil, porém muito falante, que vende ervas e temperos...” No turno 1 ocorre a contextualização, a apresentação do cenário de maneira bem agradável e eficiente, ao narrar a história falando a respeito do cheiro do local, dos sons das pessoas, dar nome para a cidade e falar sobre como está o ambiente no momento há uma vividez maior no jogo o que torna mais satisfatório.

No turno 10 ocorre uma síntese de tudo que foi trabalhado durante a sessão:

Mestre: “...Nesta sessão, você demonstrou com consistência e sensibilidade um conjunto exemplar de habilidades sociais adaptativas:

Comunicação assertiva — ao expressar necessidades e opiniões com clareza e respeito.

Negociação empática — ao buscar soluções justas e equilibradas.

Tomada de decisão responsável — ao agir de forma coerente com o contexto e o bem coletivo.

Resolução de conflitos e liderança colaborativa — ao promover entendimento e cooperação.

Auto Expressão emocional — ao compartilhar sentimentos de modo autêntico e regulado.

Empatia afetiva e reparação social — ao acolher e restaurar vínculos com delicadeza.

Essas competências representam pilares fundamentais para a convivência saudável e o desenvolvimento socioemocional, especialmente em contextos onde é preciso compreender perspectivas diferentes, lidar com frustrações e cooperar de forma construtiva...” Este trecho é interessante por mostrar uma análise ainda que rasa a respeito do que foi trabalhado no jogo. De certa forma é uma maneira de avaliar se foi trabalhado durante o processo habilidades sociais.

4 DISCUSSÃO

Resumo

A pesquisa teve como objetivo avaliar a qualidade técnica e pedagógica do ChatGPT enquanto mestre de jogo em sessões de RPG de mesa voltadas para o ensino de habilidades sociais (HS) em adolescentes com Transtorno do Espectro Autista (TEA). Para isso, foram conduzidas interações simuladas em dois chats simultâneos, nos quais a inteligência artificial desempenhou tanto o papel de mestre de RPG quanto o papel de jogador. O protocolo incluiu um prompt instrucional elaborado especificamente para induzir comportamentos pedagógicos relacionados ao treino de HS, e cada sessão foi padronizada em dez turnos, permitindo analisar a coerência narrativa e a intencionalidade educativa em um formato controlado.

A análise qualitativa dos dados, conduzida com base em sete critérios (correção, clareza, concisão, conexão, confiabilidade, legibilidade e pertinência socioeducativa), revelou que o modelo foi capaz de manter narrativas consistentes, criar cenários detalhados, responder de maneira objetiva e apresentar linguagem acessível ao público-alvo. O ChatGPT demonstrou competência em construir mundos ficcionais imersivos, reconhecer ações anteriores e sustentar uma progressão narrativa coerente ao longo dos turnos. Entretanto, foram identificadas limitações importantes, especialmente no que se refere à constância dos feedbacks socioeducativos, à formulação de perguntas explicitamente alinhadas ao ensino de habilidades sociais e ao tom ocasionalmente infantilizado adotado em algumas respostas. De modo geral, os resultados apontam que, embora promissor como ferramenta complementar para o treino de HS, o uso da IA ainda requer supervisão e ajustes para garantir adequação pedagógica plena.

Comparação com a Literatura

Sampaio et al. (2024) argumentam que modelos de linguagem constituem ferramentas com grande capacidade de gerar linguagem natural e personalizar interações, servindo como “copilotos” no apoio a tarefas educacionais. De modo similar, neste estudo verificou-se que o modelo é apto a produzir narrativas ricas, descritivas e adaptativas, confirmando o potencial de mediação.

Os padrões observados em relação à clareza, concisão e coerência narrativa dialogam diretamente com as conclusões de McFayden et al. (2024). Esses autores, ao avaliarem respostas do ChatGPT em perguntas sobre autismo, encontraram boa correção e clareza em muitos contextos, mas registraram lacunas em referências e em recomendações acionáveis. Analogamente, a nossa análise mostrou competência do modelo na construção de cenários e manutenção da linha narrativa (concisão e conexão elevadas), enquanto a formulação explícita de perguntas pedagógicas e a oferta consistente de reforços socioeducativos foram menos robustas. Essa convergência sugere que a tecnologia já domina aspectos formais da comunicação, mas necessita de mecanismos adicionais (prompts refinados, camadas de supervisão) para garantir fidelidade ao propósito pedagógico.

A investigação de Kim et al. (2025) complementa essa leitura ao enfatizar que LLMs fornecem respostas corretas e compreensíveis para pautas parentais, ainda que frequentemente careçam de contextualização cultural, de ênfase relacional e de citações precisas. No estudo aqui desenvolvido, essa mesma fragilidade foi manifestada pela inconsistência no reconhecimento e verbalização de comportamentos sociais desejáveis (feedbacks positivos por vezes ausentes ou padronizados) e por um tom que, em ocasiões, se mostrou infantilizado para adolescentes mais velhos. Em suma, Kim et al. ajudam a explicar por que, mesmo quando a IA “sabe” descrever situações produtivas para o treino de HS, ela nem sempre as explora com a sensibilidade relacional requerida em intervenções socioeducativas.

No que tange ao campo das habilidades sociais, os resultados do presente estudo mantêm diálogo produtivo com o referencial de Del Prette e Del Prette (2008). A literatura clássica sobre HS educativas enfatiza que essas habilidades são repertórios aprendidos e passíveis de ensino por meio de estratégias que combinam exposição, prática e feedback intencional. As sessões de RPG mediadas por IA reproduziram estruturalmente essa sequência: exposição a cenário, oportunidade de prática e em alguns casos síntese reflexiva. Vale ressaltar que ao utilizar de histórias educativas, porém utilizando-se de algo interativo como uma sessão de RPG pode acarretar em uma facilidade maior no aprendizado de HS, já que não estariam enquadradas nos chamados “Pacotes de aprendizagem” (Malmberg et al., 2015).

As evidências sobre o uso de RPG com populações do espectro, apoiam a escolha do mesmo permitindo situar seus resultados em intervenções lúdicas. Kato demonstrou ganhos nas verbalizações e na qualidade de vínculo no uso do TRPG com crianças e jovens com TEA; Henning descreveu possibilidades pedagógicas do RPG para treinar habilidades sociais em grupo, Paludo et al., (2018) mostrou que RPG pode ser usado com outros públicos que lidam com diversos preconceitos, no caso sendo adolescentes com altas habilidades e superdotação. O presente estudo amplia essas contribuições ao mostrar que a mediação algorítmica pode reproduzir muitos dos processos observados nesses estudos, criação de dilemas sociais, necessidade de negociação, oportunidades de cooperação, porém com ressalvas relativas à autenticidade do reforço e à personalização emocional. Em outras palavras, o ChatGPT pode emular as dinâmicas do role-play terapêutico descritas por Kato e Henning, mas ainda não substitui a curadoria humana necessária para otimizar ganhos socioemocionais transferíveis ao cotidiano (Kato, 2019; Henning, 2021; Paludo et al.).

Contribuições metodológicas

A presente pesquisa oferece avanços metodológicos importantes ao explorar, de forma sistemática, a atuação de uma inteligência artificial generativa como mestre de RPG com finalidade pedagógica. A primeira contribuição reside na adaptação e operacionalização dos critérios de avaliação propostos por McFayden et al. (2024) e Kim et al. (2025) para o contexto de interações narrativas. Os critérios originalmente utilizados para avaliar qualidade textual e informacional foram reinterpretados para analisar elementos próprios do RPG, como continuidade narrativa, coerência ficcional, manutenção de personagens, responsividade às ações do jogador e intencionalidade educativa. Essa adaptação resultou em uma rubrica de sete di-

mensões: correção, clareza, concisão, conexão, confiabilidade, legibilidade e pertinência socioeducativa, que representa um instrumento inédito e replicável para estudos futuros envolvendo IA e atividades lúdicas ou pedagógicas.

Outra contribuição centra-se no campo da psicologia educacional e das tecnologias interativas, ao evidenciar que IAs generativas podem desempenhar papéis narrativos e pedagógicos simultaneamente. Estudos como os de Sampaio et al. (2024), McFayden et al. (2024) e Kim et al. (2025) descrevem o potencial dos modelos de linguagem para fornecer informações claras, sintéticas e contextualizadas. Entretanto, seus trabalhos não exploram a capacidade da IA de engajar usuários em interações sociais simuladas e prolongadas. A presente pesquisa preenche essa lacuna ao mostrar que a IA pode sustentar uma narrativa contínua, responder de forma coerente às ações do usuário e criar oportunidades de exploração socioemocional ao longo de múltiplos turnos. Dessa forma, contribui ao campo teórico ao demonstrar que interações conversacionais prolongadas com IA podem ser entendidas como microambientes de aprendizagem socioemocional, expandindo o entendimento tradicional das IAs como meras ferramentas instrumentais.

Uma contribuição teórica refere-se ao papel do RPG como dispositivo de aprendizagem socioemocional mediado tecnologicamente. Embora estudos como Kato (2019) e Henning (2021) já indiquem o valor do RPG para o desenvolvimento de competências sociais em crianças e adolescentes com TEA, a literatura ainda carece de análises sobre como essas dinâmicas se comportam quando o mestre do jogo é uma IA generativa. Os resultados desta pesquisa sugerem que modelos como o ChatGPT conseguem reproduzir estruturas centrais do role-play terapêutico, preservando boa parte da lógica formativa descrita pela literatura. Assim, o estudo contribui teoricamente ao demonstrar que a mediação algorítmica não apenas sustenta o arcabouço pedagógico do RPG, mas também oferece uma nova camada de análise: a equivalência funcional entre mediação humana e mediação algorítmica, categoria ainda incipiente nos estudos sobre IA educacional.

Outra contribuição metodológica interessante é o delineamento usando duas inteligências artificiais ao mesmo tempo, no qual a IA desempenhou simultaneamente os papéis de mestre e jogador em sessões paralelas. Essa estratégia permitiu controlar variáveis externas, garantindo que todas as interações fossem geradas exclusivamente pelo modelo, permitindo uma análise precisa da sua performance narrativa e pedagógica. O uso de dois perfis de conta, um com histórico prévio e outro sem histórico, também permitiu comparar possíveis diferenças decorrentes da memória ou personalização do sistema, ampliando a robustez da análise. Esse arranjo experimental constitui um modelo metodológico que pode ser replicado por pesquisas interessadas em comparar prompts, arquiteturas de IA ou condições de uso.

A inclusão de um prompt instrucional estruturado, projetado especificamente para induzir comportamentos pedagógicos e situações sociais intencionais dentro do RPG, também se destaca como inovação metodológica. Diferentemente de estudos anteriores que analisam respostas espontâneas de IAs, este trabalho desenvolveu um protocolo que direciona a IA a criar dilemas sociais, oportunidades de negociação, feedbacks educativos e explicitação das habilidades sociais envolvidas em cada turno.

Limitações e pesquisas futuras

Embora os resultados desta pesquisa indiquem o potencial do ChatGPT como mediador de sessões de RPG voltadas ao ensino de habilidades sociais (HS) para adolescentes com Transtorno do Espectro Autista (TEA), é fundamental reconhecer as limitações do estudo, que influenciam a interpretação e a generalização dos achados.

A primeira limitação refere-se ao caráter inteiramente simulado das interações. Todas as sessões foram conduzidas com a própria IA desempenhando simultaneamente os papéis de mestre e jogador, o que impede avaliar como adolescentes reais com TEA reagiriam às demandas narrativas, ao ritmo das interações e aos feedbacks pedagógicos. A ausência de participantes humanos impede também a análise de variáveis essenciais como engajamento emocional, motivação, compreensão contextual e generalização dos comportamentos aprendidos para situações da vida cotidiana. Assim, recomenda-se que estudos futuros incluam participantes reais em diferentes níveis de suporte dentro do espectro, de modo a avaliar o impacto concreto e ecologicamente válido da intervenção. É válido ressaltar que seria interessante uma maneira mais concreta de avaliar a evolução de HS utilizando o prompt

Outra pesquisa futura poderia utilizar outras I.As. Embora a estratégia utilizando somente o ChatGPT tenha permitido analisar pequenas variações decorrentes da personalização do sistema, a pesquisa não comparou outros modelos de linguagem, outras arquiteturas, parâmetros específicos (como temperatura e comprimento máximo) ou versões alternativas do próprio ChatGPT. Isso limita conclusões mais amplas sobre a superioridade, estabilidade ou especificidade da atuação do modelo. Futuras pesquisas poderiam comparar modelos abertos (como LLaMA ou Mistral) e proprietários (como Claude ou Gemini), examinando suas diferenças na condução narrativa, coerência educacional e adaptação ao público com TEA

Também se identifica como limitação a ausência de sessões múltiplas ou longitudinalidade. Uma única interação estruturada em dez turnos, embora útil para fins de padronização, não permite avaliar a consistência do modelo a um tempo mais extenso, nem a capacidade da IA de promover aprendizagem progressiva, manutenção de repertórios ou transferência de habilidades treinadas para novos cenários sociais dentro do próprio jogo. Assim, estudos futuros deveriam adotar sessões prolongadas, seriadas ou com arcos narrativos maiores, possibilitando observar estabilidade, variações de qualidade e potencial evolutivo do modelo ao longo de múltiplas interações.

Uma limitação interessante que Francisco (2025) aponta é que geralmente os estudos realizados com RPG possuem uma coisa em comum, são realizados em grupos nos quais os participantes buscam desenvolver habilidades por meio de uma narrativa proposta.

Embora o modelo tenha oferecido devolutivas coerentes em alguns momentos, observou-se inconsistência na explicitação das habilidades sociais praticadas e, por vezes, um tom infantilizado nas interações. Isso restringe a equivalência funcional entre a mediação algorítmica e a mediação humana, especialmente considerando que intervenções socioemocionais requerem sensibilidade, nuance e responsividade afetiva (Del Prette & Del Prette, 2008). Pesquisas futuras podem investigar mecanismos de refinamento de prompt.

5 CONCLUSÃO

Conclui-se que o ChatGPT possui potencial para atuar como ferramenta complementar no ensino de habilidades sociais por meio de jogos de RPG, oferecendo um ambiente narrativo organizado, coerente e responsivo, capaz de criar oportunidades significativas de prática socioemocional. No entanto, sua utilização ainda requer acompanhamento profissional, refinamento de prompts e maior sensibilidade contextual para garantir alinhamento pedagógico, adequação etária e segurança formativa. Estudos futuros são necessários para avaliar a efetividade do método com participantes humanos, ampliar a duração das sessões, comparar diferentes modelos de IA e explorar estratégias de aprimoramento da responsividade socioeducativa do sistema.

De modo geral, esta pesquisa contribui para o avanço do debate sobre o papel das tecnologias emergentes na educação inclusiva e evidencia que, embora a IA não substitua o trabalho humano, pode funcionar como recurso inovador, acessível e promissor no apoio ao desenvolvimento de habilidades sociais em adolescentes com TEA.

REFERÊNCIAS BIBLIOGRÁFICAS

- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing.
- Del Prette, Z. A. P., & Del Prette, A. (2008). *Psicologia das habilidades sociais na infância: Teoria e prática*. Vozes.
- Francisco, E. C. (2025). As contribuições do role-playing game (RPG) para a psicoterapia no desenvolvimento de habilidades sociais. *Unifunec Científica Multidisciplinar*, 14(16)
- Henning, G. (2021). A terapia analítico-comportamental das habilidades sociais em grupo para adolescentes com uso do RPG e da gamificação (Dissertação de mestrado, Instituto de Psicologia, Universidade de São Paulo). <https://doi.org/10.11606/D.47.2021.tde-12082021-172100>
- Katō, K. (2019). Employing Tabletop Role-Playing Games (TRPGs) in Social Communication Support Measures for Children and Youth with Autism Spectrum Disorder (ASD) in Japan: A Hands-On Report on the Use of Leisure Activities. *RPG学研究: Japanese Journal of Analog Role-Playing Game Studies*, 23-28.
- Kim, Y., Vilches, S. L., Shapiro, S., & Clarkson, A. (2025). Testing the capability of generative artificial intelligence for parent and caregiver information seeking. *Family Relations*, 74(3), 1266-1284.
- Malmberg, D. B., Charlop, M. H., & Gershfeld, S. J. (2015). A two experiment treatment comparison study, teaching social skills to children with autism spectrum disorder. *Journal of Developmental and Physical Disabilities*, 27(3), 375-392.
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135-148. <https://doi.org/10.1089/cyber.2023.0202>
- Nemer, B. M. T. (2025). *Relação entre RPG e o desenvolvimento de habilidades socioemocionais em adolescentes no contexto educacional* (Monografia de graduação, Centro Universitário de Brasília).

Paludo, K. I., Medina, G. B. K., Andrade, A. M., de Almeida, R., & Lozza, R. (2018). Promoção de saúde socioemocional de adolescentes com altas habilidades/superdotação com o uso do RPG. In Congresso Brasileiro de Psicologia da FAE (Vol. 2, pp. 243–250).

Peixoto Filho, C. B., & Albuquerque, R. M. de. (2021). Notas sobre a história dos RPGs (role-playing games) de mesa brasileiros. *Projética*, 12(3), 300–324. <https://doi.org/10.5433/2236-2207.2021v12n3p300>

Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: Reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432e008>

ANEXO A

Textos na íntegra referentes aos turnos 1 ao 10.

Mestre: “O sol nasce sobre a pequena cidade de Arven, e os primeiros sons do mercado ecoam pelas ruas de pedra. O cheiro de pão recém-assado se mistura ao aroma forte de especiarias vindas do sul. Vozes se cruzam, vendedores anunciam seus produtos, e um grupo de crianças corre entre as barracas rindo. Você chega ao mercado com uma pequena lista de compras e algumas moedas no bolso. Sua missão é simples: comprar os ingredientes para o jantar comunitário da vila. No entanto, há um desafio — os comerciantes de Arven são conhecidos por suas personalidades muito diferentes, e você precisará negociar e se comunicar com clareza para conseguir o que precisa. Enquanto caminha entre as barracas, percebe três comerciantes: Lira, uma jovem simpática, mas distraída, que vende frutas frescas. Bran, um homem sério e autoritário, dono da banca de carne. Neri, um idoso gentil, porém muito falante, que vende ervas e temperos. O ambiente é vivo: você sente o calor do sol na pele, ouve os sons de risadas e passos, e o cheiro de carne grelhando ao longe. Sua tarefa agora é escolher com quem vai falar primeiro e como vai iniciar a conversa. Lembre-se: o objetivo aqui é praticar comunicação assertiva, ou seja, expressar o que você precisa de forma clara, respeitosa e confiante, sem ser agressivo nem passivo. O que você faz?”

Jogador: “Vejo as três bancas diante de mim e decido começar pelo caminho mais leve. Lira parece distraída, mas seu sorriso é sincero — e frutas são o primeiro item da lista.

Aproximo-me com calma, sorrindo, e digo com um tom amigável e assertivo:

— Bom dia, Lira! Suas frutas parecem ótimas hoje. Estou comprando ingredientes pro jantar comunitário da vila e gostaria de levar algumas maçãs e peras frescas. Você pode me dizer o preço delas?

(Falo com clareza, mantendo contato visual e um tom respeitoso, sem pressa — deixando espaço para ela responder.)

(Turno 2 de 10)

Mestre: “Lira se vira, um pouco surpresa com a abordagem gentil. Por um instante, parece ter se esquecido do que estava fazendo, mas logo sorri de volta, ajeitando o cabelo com as mãos cobertas de pólen e suco de frutas. — Oh! Bom dia! — ela responde, animada. — Desculpe, eu estava distraída com umas abelhas que apareceram por aqui. As maçãs estão a três moedas a dúzia, e as peras, a duas. Mas, se é para o jantar da vila, posso fazer um desconto

especial. Ela inclina a cabeça, curiosa. — Quantas você precisa? Enquanto fala, o cheiro doce das frutas frescas toma o ar, e você nota o brilho colorido das maçãs vermelhas e das peras douradas sob o sol. Ao redor, o som das conversas e das moedas ecoam. Seu comportamento até aqui foi excelente. Você iniciou a interação com comunicação assertiva, demonstrando respeito e clareza, além de usar elogio genuíno — uma estratégia social eficaz para criar rapport. Essa habilidade é útil na vida real quando queremos estabelecer um diálogo cooperativo e facilitar acordos. Agora, há uma pequena oportunidade de treino adicional: negociação e empatia. Você pode, por exemplo, tentar negociar o preço de modo justo, demonstrando compreensão do lado dela, ou aceitar o valor como forma de valorizar o trabalho da comerciante. O que você faz?”

Jogador: “Sorrio de volta, inclinando levemente a cabeça num gesto de agradecimento. Respondo em tom tranquilo, mas confiante: — Fico feliz em saber disso, Lira. Suas frutas sempre fazem sucesso nos jantares da vila.

Já que é para todos, acho justo aceitar o desconto especial — mas também quero garantir que você não saia perdendo. Que tal eu levar seis dúzias de maçãs e quatro dúzias de peras pelo preço normal das peras, e você me inclui algumas frutas da estação como cortesia? Assim todos saímos satisfeitos.”

(Turno 3 de 10)

Mestre: “Lira arregala os olhos, claramente impressionada com sua proposta. Por um instante, parece calcular o que você disse, depois ri baixinho, limpando as mãos no avental florido.

— Sabe... ninguém nunca falou assim comigo aqui no mercado. Geralmente só tentam baixar o preço — comenta com um sorriso sincero. — Mas eu gostei da sua ideia. Acho justo.

Ela começa a selecionar cuidadosamente as frutas mais bonitas, colocando-as em cestos de palha. Enquanto separa as maçãs e as peras, acrescenta algumas ameixas e pêsesgos maduros. O aroma doce se intensifica.

— Aqui está — diz, entregando-lhe os cestos. — Você fala de um jeito gentil e direto, é fácil negociar assim. Espero que o jantar seja um sucesso.

Você sente a satisfação de ter sido tratado com respeito, e o som das moedas trocando de mãos sela o acordo.

Seu desempenho foi excelente. Você demonstrou negociação empática, reconhecendo o valor do trabalho da outra pessoa e buscando um benefício mútuo — exatamente o tipo de comportamento que favorece relações colaborativas e justas.

Na vida real, essa habilidade ajuda a resolver conflitos e acordos de forma saudável, reduzindo tensões e fortalecendo vínculos sociais.

Com as frutas garantidas, o próximo passo é seguir para outra banca.

Restam ainda a carne e as ervas na sua lista.

Você vê Bran, o açougueiro, organizando cortes de carne com expressão séria e postura rígida, e Neri, o velho vendedor de ervas, gesticulando animado enquanto fala com um grupo de clientes.

Quem você aborda agora e como inicia a interação?

Jogador: “Olho em direção a Bran, o açougueiro. Sua postura firme e olhar concentrado indicam que prefere negociações diretas — o tipo de pessoa que valoriza respeito e objetividade.

Ajusto os cestos nas mãos, respiro fundo e caminho até sua banca.

Com voz segura, mas educada, digo:

— Bom dia, Bran. O cheiro da carne está excelente hoje.

Estou comprando os ingredientes para o jantar comunitário da vila e preciso de bons cortes para um ensopado. Você pode me dizer quais opções tem e qual delas recomenda?”

(Turno 4 de 10)

Mestre: “Bran ergue os olhos por um momento, avaliando você de cima a baixo. O homem é robusto, com braços largos cobertos de cicatrizes antigas e um avental manchado de sangue fresco. Sua expressão é séria, mas não hostil. Ele parece apreciar quando as pessoas falam de forma direta. — Hm. Finalmente alguém que sabe o que quer — diz com um leve aceno de cabeça. — Tenho carne de javali, fresca de ontem, e costelas de boi curadas. Se o jantar é pra muita gente, recomendo o boi. Rende mais e tem sabor forte o bastante pra agradar até os mais exigentes. Enquanto fala, você ouve o som ritmado da faca cortando a carne, e sente o cheiro metálico e defumado que paira no ar. Bran mantém o olhar firme, demonstrando que valoriza clareza e respeito mútuo. — Mas o ensopado da vila merece carne de qualidade — contínua. — Posso te dar um bom preço, se prometer que vai cozinhar direito. Odeio ver carne desperdiçada. Seu comportamento aqui foi exemplar: você leu o perfil social do interlocutor e ajustou sua comunicação de forma estratégica. Essa habilidade de flexibilidade social é essencial para conviver com diferentes tipos de pessoas — especialmente aquelas com estilos comunicativos mais rígidos ou formais. Agora há duas oportunidades de treino: Tomar decisão responsável, escolhendo entre os tipos de carne com base em necessidade e contexto. Reforçar o vínculo social, mostrando respeito e colaboração diante da seriedade de Bran. O que você faz?”

Jogador: “Endireito a postura, olhando Bran nos olhos enquanto falo com firmeza, mas num tom respeitoso.

— Concordo com você, Bran. O jantar da vila merece o melhor.

Vamos ficar com as costelas de boi — quero que todos sintam o sabor de algo bem-feito, como você disse.

Prometo que a carne será tratada com o respeito que merece.

Pode me separar o suficiente para servir umas vinte pessoas? E me diga quanto devo pagar — quero garantir que o seu trabalho seja valorizado.”

(Turno 5 de 10)

Mestre: “Bran o encara por um breve instante, em silêncio. O ambiente parece conter o som por um segundo — apenas o gotejar da água no balcão e o murmúrio distante do mercado preenchem o ar. Então, o açougueiro solta um leve riso, grave e curto, enquanto limpa a lâmina com um pano grosso. — Você fala como quem entende o valor do trabalho, garoto — diz

ele, com um tom mais brando. — Gosto disso. A maioria vem aqui só querendo desconto, sem pensar em quem levanta às quatro da manhã pra cortar carne. Ele começa a separar cuidadosamente as costelas, pesando-as com precisão. — Aqui está — continua, empacotando os cortes em tecido encerado. — Para vinte pessoas, isso deve bastar. Três moedas por quilo, mas... — faz uma pausa e dá um meio sorriso — vamos fazer por duas e meia. O jantar da vila merece. Você sente o peso agradável dos pacotes nas mãos e nota algo diferente na postura de Bran: ele parece mais aberto, menos rígido. Seu desempenho foi excelente. Você demonstrou tomada de decisão responsável, mostrando segurança sem autoritarismo, e aplicou reconhecimento social, uma forma de empatia que valoriza o esforço alheio. Essas atitudes fortalecem vínculos e constroem relações de respeito e confiança, tanto no jogo quanto na vida real. Agora, com as carnes garantidas, só falta comprar as ervas e temperos com Neri, o idoso vendedor. Ele fala alto, gesticula muito e parece empolgado com qualquer oportunidade de conversa. Você o observa agitando um punhado de folhas aromáticas enquanto conta uma história longa para dois clientes que já parecem perdidos. Como você se aproxima e inicia o diálogo com Neri?”

Jogador: “Aproximo-me da barraca de Neri com um sorriso leve, ajustando o peso dos pacotes de carne e cestos de frutas. O ar está carregado com o perfume das ervas secas — alecrim, tomilho, manjerona — e o velho parece completamente imerso em sua própria história sobre uma “tempestade que cheirava a hortelã”.

Espero um momento educado, até que ele termine uma frase mais empolgada, e então falo num tom calmo, gentil e focado, projetando a voz o suficiente para ser ouvido sem interromper de forma brusca:

— Senhor Neri! Suas histórias são sempre as melhores do mercado.

Eu vim justamente em busca da sua sabedoria — preciso de ervas frescas pro ensopado do jantar comunitário.

Gostaria de saber quais combinações o senhor recomenda pra dar sabor sem deixar o prato pesado.”

(Turno 6 de 10)

Mestre: “Neri interrompe a fala no meio de uma lembrança sobre trovões aromáticos e pisca algumas vezes, como se voltasse de um devaneio. Em seguida, abre um largo sorriso que ilumina o rosto enrugado. — Ah! Um jovem que aprecia o verdadeiro poder das ervas! — exclama, batendo palmas. — Finalmente alguém que entende que cozinhar é contar histórias com cheiro! Ele se inclina sobre o balcão, o rosto próximo o bastante para que você sinta o aroma intenso das folhas que segura. — Para um ensopado... você precisa de algo que abrace o sabor da carne sem abafá-lo. Use alecrim e tomilho em pequenas quantidades, e um toque de louro pra trazer firmeza. E, se quiser um segredo meu, um pouquinho de sálvia faz até o mais simples dos caldos parecer banquete de rei. Enquanto fala, vai separando ramos de ervas frescas, cada movimento acompanhado de um leve tilintar dos frascos e potes. — Tome, leve estes de presente. Para quem escuta antes de falar, sempre tenho um mimo — diz com um sorriso afetuoso. Seu desempenho aqui foi notável. Você aplicou escuta ativa e comunicação empática, demonstrando respeito pelo ritmo e estilo do outro. Essa abordagem é especialmente

eficaz em interações com pessoas mais expansivas ou que gostam de compartilhar — pois mostra interesse genuíno, facilita conexões e promove cooperação natural. Agora, você concluiu todas as compras da lista. Com os ingredientes em mãos, começa a se dirigir à praça central, onde outros moradores já preparam as mesas para o jantar. Ao chegar, percebe um pequeno grupo discutindo sobre quem deve coordenar a montagem das barracas. As vozes estão elevadas e há certo desconforto no ar. Este é o momento de praticar resolução de conflitos e mediação cooperativa. O que você faz?”

Jogador: “Ao me aproximar da praça, percebo o tom tenso das vozes — cada um parece querer assumir o controle. Pouso os cestos e pacotes sobre uma mesa próxima, respiro fundo e observo por um instante antes de agir, para entender quem está mais exaltado e quem parece aberto ao diálogo.

Em seguida, dou um passo à frente e falo com voz firme, mas calma, o suficiente para chamar a atenção sem soar autoritário:

— Ei, pessoal, respirem um pouco. Todos aqui querem que o jantar da vila dê certo, não é? Em vez de discutirmos quem manda, que tal decidirmos juntos como dividir as tarefas? Podemos fazer o seguinte: quem tem mais experiência com montagem cuida das barracas, e quem é melhor na cozinha ajuda a preparar os pratos. Assim ninguém fica sobrecarregado — e todo mundo contribui.

Faço contato visual com os presentes, alternando o olhar para incluir cada um no diálogo. Termino com um tom de cooperação:

— O que acham? Se trabalharmos juntos, terminamos antes do pôr do sol e ainda sobra tempo pra celebrar.”

(Turno 7 de 10)

Mestre: “As conversas se interrompem pouco a pouco. Os olhares, antes rígidos e tensos, se voltam para você. Há um breve silêncio, seguido de alguns murmúrios entre o grupo. Uma mulher de avental sujo de farinha cruza os braços, avaliando suas palavras, enquanto um rapaz de chapéu de palha coça a nuca, visivelmente envergonhado. — É... talvez ele tenha razão — comenta o rapaz, desviando o olhar do conflito. — Eu posso cuidar das mesas, conheço o jeito certo de nivelar o chão. A mulher suspira, relaxando os ombros. — Tudo bem. Eu e as meninas cuidamos da cozinha, então. Acho que me deixei levar pelo nervosismo. O clima muda. As expressões duras cedem lugar a sorrisos tímidos, e o som das ferramentas e panelas volta a preencher o ambiente. O sol desce lentamente, pintando o céu de laranja e dourado, e a vila parece respirar em uníssono outra vez. Você agiu de forma exemplar. Sua intervenção demonstrou mediação assertiva e liderança empática, dois pilares da resolução construtiva de conflitos. Em vez de impor uma solução, você guiou o grupo com calma e lógica, equilibrando razão e cooperação. Essa habilidade é crucial na vida real para lidar com desacordos de maneira produtiva, especialmente em contextos de grupo. Enquanto o trabalho avança, alguns aldeões se aproximam para agradecer sua ajuda. Um deles, um garoto de cerca de 12 anos, observa você com admiração e pergunta: — Você sempre consegue fazer as pessoas pararem de brigar assim? Essa é uma boa oportunidade para praticar auto expressão emocional e modelagem positiva. Como você responde ao garoto?”

Jogador: “Sorrio para o garoto, abaixando-me um pouco para ficar à altura de seus olhos. O brilho curioso no olhar dele me faz lembrar de quando eu mesmo tentava entender os adultos e suas confusões.

Com um tom calmo e sincero, respondo:

— Nem sempre, sabia? Às vezes, as pessoas continuam brigando mesmo quando a gente tenta ajudar.

Mas aprendi que, se a gente escuta de verdade e fala com respeito, é mais fácil que elas também escutem.

Não se trata de “fazer parar de brigar” — é sobre fazer entender um ao outro.

E isso começa com uma coisa simples: querer que todo mundo saia bem da conversa.

Dou um pequeno sorriso e acrescento:

— Você tem jeito de quem presta atenção. Um dia, vai fazer isso até melhor do que eu.”

(Turno 8 de 10)

Mestre: “O garoto fica em silêncio por alguns segundos, absorvendo suas palavras. O barulho das painéis e vozes ao fundo parece distante diante do olhar concentrado dele. Então, um sorriso tímido surge em seu rosto. — Acho que entendi... tipo quando eu e meu irmão brigamos pelo último pedaço de pão, e a gente divide em dois pra ninguém ficar triste? — ele pergunta, ainda refletindo. Você acena afirmativamente, e ele sorri mais abertamente, agora com genuína admiração. — Vou tentar fazer isso também. Obrigado! Ele corre de volta para ajudar os outros, e, à medida que o faz, você sente uma leveza no ar. As pessoas parecem mais calmas, o trabalho flui, e o som das risadas volta a se misturar ao crepitar do fogo e ao aroma dos temperos que começam a ferver. Seu desempenho foi excelente. Nesta interação, você demonstrou auto expressão emocional e modelagem de comportamento pró social, comunicando valores de empatia, escuta e respeito mútuo. Ao validar a curiosidade do garoto e traduzir o conceito de forma acessível, você incentivou o aprendizado social — uma das formas mais eficazes de generalizar habilidades para a vida real. Agora, o jantar está quase pronto. Os moradores começam a colocar as mesas e organizar os pratos. Há um clima de celebração no ar, mas um último desafio surge: uma senhora idosa, chamada Dona Meriel, reclama que foi esquecida na hora de distribuir os assentos. Ela parece ofendida e um pouco triste, dizendo que “ninguém se importa com os velhos”. Este é um momento ideal para praticar empatia afetiva e reparação social. O que você faz?”

Jogador: “Vejo Dona Meriel de pé, com o rosto franzido entre mágoa e orgulho. O som alegre do mercado parece se afastar um pouco enquanto caminho até ela. Aproximo-me devagar, evitando qualquer gesto brusco.

Com expressão acolhedora e tom suave, falo:

— Dona Meriel... sinto muito por isso ter acontecido. A senhora tem toda razão em se sentir chateada — ninguém gosta de ser deixado de lado, especialmente num dia que deveria reunir todo mundo.

Mas eu garanto: ninguém aqui esqueceu da senhora de propósito. Foi tanta correria que acabamos nos atrapalhando.

Que tal fazermos o seguinte? A senhora senta bem aqui, perto do fogo — é o lugar mais quente e confortável — e eu mesmo trago o primeiro prato pra mesa.

O jantar da vila não estaria completo sem a senhora entre nós.

Enquanto falo, mantenho o olhar gentil e a postura aberta, sinalizando acolhimento e respeito genuíno.

(Turno 9 de 10)

Mestre: “Dona Meriel levanta o olhar lentamente. Por um instante, o rosto rígido dela parece hesitar entre a ofensa e a emoção. As rugas que antes marcavam irritação agora se suavizam, e seus olhos úmidos brilham sob a luz do fogo que crepita próximo.

— Ah... — ela suspira, ajeitando o xale nos ombros. — Você fala bonito, rapaz. E fala com o coração, isso é raro hoje em dia.

Ela aceita o lugar que você indicou e, quando se senta, o grupo à volta percebe a cena. Duas pessoas se aproximam para cumprimentá-la, e logo risadas e histórias antigas começam a preencher o espaço. O clima de celebração retorna, mais acolhedor do que antes.

Você demonstrou empatia afetiva de forma plena, validando o sentimento da outra pessoa antes de tentar resolver a situação. Essa é uma das estratégias sociais mais eficazes para reparar laços e restaurar o senso de pertencimento. Também aplicou tom conciliador e linguagem não verbal positiva, elementos essenciais da comunicação emocionalmente inteligente.

Enquanto todos se acomodam, a primeira travessa de ensopado é servida. O cheiro é rico e reconfortante, misturando o sabor das carnes, frutas e ervas escolhidas ao longo da jornada.

O jantar começa em meio a risos, conversas e agradecimentos. A vila, antes agitada e confusa, agora vibra em harmonia — em grande parte, graças à forma como você conduziu cada interação.

Enquanto o fogo crepita e os moradores brindam, uma das pessoas da vila, sorrindo, se volta para você e pergunta:

— Hoje tudo deu certo por sua causa. Me diga, o que você aprendeu com esse dia?”

Jogador: “Olho em volta, observando as mesas cheias, as vozes misturadas em risadas e o brilho quente do fogo dançando nos rostos dos moradores. Por um instante, deixo o som preencher o silêncio dentro de mim. Então, sorrio e respondo em tom sereno:

— Aprendi que comunicação não é só o que a gente diz — é como a gente faz o outro se sentir.

Quando falamos com respeito, escutamos com atenção e buscamos entender, as coisas fluem de um jeito natural.

Nem sempre é fácil... às vezes, é preciso respirar fundo, ceder um pouco, ou apenas ouvir antes de agir.

Mas no fim, o que realmente une as pessoas não é ter razão — é querer estar junto, construir algo bom em comum.

Faço uma pausa, olhando para o ensopado fumegante e as pessoas que agora dividem histórias entre colheradas.

— Acho que é isso que faz um jantar — e uma comunidade — darem certo.

(Turno 10 de 10)

Mestre: “O som das risadas e o tilintar das colheres enchem o ar. As chamas projetam sombras suaves nas paredes de pedra, e o cheiro do ensopado se mistura ao perfume das ervas e frutas frescas. Todos parecem em paz — e, por um instante, o tempo desacelera.

Dona Meriel ergue a taça em sua direção, e Lira, Bran e Neri fazem o mesmo. — Àqueles que sabem unir as pessoas — diz ela, sorrindo. Um coro de vozes repete o brinde, e o eco suave dessa frase se perde na noite de Arven.

Nesta sessão, você demonstrou com consistência e sensibilidade um conjunto exemplar de habilidades sociais adaptativas:

Comunicação assertiva — ao expressar necessidades e opiniões com clareza e respeito.

Negociação empática — ao buscar soluções justas e equilibradas.

Tomada de decisão responsável — ao agir de forma coerente com o contexto e o bem coletivo.

Resolução de conflitos e liderança colaborativa — ao promover entendimento e cooperação.

Auto Expressão emocional — ao compartilhar sentimentos de modo autêntico e regulado.

Empatia afetiva e reparação social — ao acolher e restaurar vínculos com delicadeza.

Essas competências representam pilares fundamentais para a convivência saudável e o desenvolvimento socioemocional, especialmente em contextos onde é preciso compreender perspectivas diferentes, lidar com frustrações e cooperar de forma construtiva.

Enquanto o fogo se apaga lentamente e a lua sobe sobre Arven, você sente algo silencioso e poderoso: o aprendizado não veio das palavras, mas dos gestos — dos pequenos atos de escuta, paciência e gentileza que transformaram o dia em comunhão.

Sessão encerrada.

Parabéns por sua jornada. Você não apenas completou o desafio, como também modelou, com naturalidade e empatia, tudo o que torna o convívio humano mais significativo.”

Psicoterapia



Um caminho de autoconhecimento e transformação pessoal, em um espaço acolhedor e seguro — para adultos, crianças e adolescentes.

- Atendimento presencial e online
- Particular ou plano de saúde
- +15 anos de experiência · +5.000 pacientes atendidos

walden4.com.br/psicoterapia

CAPÍTULO

03

Avaliação da Qualidade de Histórias Sociais Digitais para Autorregulação Emocional em Crianças com TEA usando GEM Storybook



Ana Luíza Antonino Valim Maia

Prof. Marcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. Este estudo teve como objetivo avaliar a qualidade e a aplicabilidade de histórias sociais digitais produzidas pelo GEM Storybook, voltadas para o desenvolvimento da autorregulação emocional em crianças com Transtorno do Espectro Autista (TEA).

Método. Foram selecionados três cenários de autorregulação emocional (controle da raiva, manejo da frustração e identificação de emoções) e criados perfis fictícios de crianças de 6 a 9 anos para personalização das histórias. As narrativas foram geradas por meio de prompts padronizados no GEM Storybook e avaliadas por especialistas em psicologia, pedagogia e Análise do Comportamento Aplicada (ABA), utilizando uma rubrica composta por nove critérios, incluindo clareza textual, adequação à faixa etária, relevância emocional, consistência narrativa, alinhamento com princípios da ABA e qualidade visual.

Resultados. As análises revelaram que as histórias digitais apresentaram alta clareza textual, forte relevância socioemocional e boa adequação à faixa etária de cada personagem. Houve alto alinhamento com princípios da ABA, especialmente no uso de previsibilidade, identificação de emoções e descrição de comportamentos-alvo em termos observáveis. Apesar disso, foram identificadas limitações, como a necessidade de ampliar estratégias de autorregulação além da respiração e da contagem, e de aprimorar a correspondência entre imagens e etapas do comportamento desejado. A avaliação automatizada (ChatGPT) demonstrou elevada concordância com a avaliação humana, ainda que menos sensível a nuances emocionais e contextuais.

Conclusão. Conclui-se que o GEM Storybook é uma ferramenta potencialmente eficaz para produzir histórias sociais digitais com boa estrutura pedagógica e compatíveis com princípios da ABA. O uso combinado de avaliação humana automatizada fortalece a confiabilidade do processo, garantindo que as histórias sejam tecnicamente adequadas e clinicamente aplicáveis. As narrativas digitais mostram-se promissoras para apoiar intervenções voltadas ao desenvolvimento da autorregulação emocional em crianças com TEA, embora ajustes adicionais sejam recomendados para garantir sensibilidade individual, qualidade visual e maior diversidade de estratégias comportamentais.

Palavras-chave: histórias sociais digitais; autorregulação emocional; Transtorno do Espectro Autista; ABA; inteligência artificial.

INTRODUÇÃO

Contextualização do Tema

Nas últimas décadas, o Transtorno do Espectro Autista (TEA) tem ganhado crescente visibilidade no campo científico e social, tanto pelo aumento das pesquisas quanto pela ampliação das discussões sobre inclusão, políticas públicas e práticas educacionais e clínicas voltadas a esse público. Esse movimento reflete não apenas avanços nos critérios diagnósticos e nas formas de detecção precoce, mas também uma maior conscientização sobre os desafios enfrentados por pessoas autistas e suas famílias no cotidiano, especialmente no que se refere ao desenvolvimento de habilidades socioemocionais (Goulart & Assis, 2002).

O Transtorno do Espectro Autista (TEA), segundo o Manual Diagnóstico e Estatístico de Transtornos Mentais (American Psychiatric Association [APA], 2014), é uma condição do neurodesenvolvimento caracterizada por dificuldades persistentes na comunicação e interação social, associadas a padrões restritos e repetitivos de comportamento, interesses ou atividades. Essas manifestações impactam diretamente habilidades como a autorregulação emocional, considerada fundamental para o bem-estar psicológico e para a adaptação às demandas sociais.

A Análise do Comportamento, enquanto referencial teórico e prático, tem oferecido contribuições significativas na compreensão e intervenção junto a indivíduos com TEA, especialmente pela ênfase na observação, mensuração e manipulação de variáveis ambientais que influenciam o comportamento (Moreira & Medeiros, 2019). Nesse sentido, estudos apontam a importância de se desenvolver estratégias específicas voltadas para a ampliação dos repertórios emocionais e sociais de crianças autistas, com destaque para recursos de autorregulação. Um estudo influente nesse campo é o de Goulart e Assis (2002), que investigou aspectos metodológicos em pesquisas sobre autismo na perspectiva analítico-comportamental. Os autores identificaram duas linhas de investigação complementares: uma voltada à análise das variáveis que influenciam o controle de estímulos e outra dedicada ao estabelecimento e manutenção de operantes verbais. Entre os critérios metodológicos avaliados, destacam-se a natureza dos estímulos utilizados, os critérios de seleção, os tipos de resposta exigida e as formas de consequenciamento aplicadas. Além disso, os autores ressaltam a necessidade de rigor metodológico na pesquisa sobre autismo e defendem a importância de programas de prevenção e educação destinados a pais e familiares, diante da escassez de estudos na área.

Seguindo essa perspectiva, destaca-se também o manual *Autismo: Estratégias de Autorregulação* (Aguiar & Moreira, 2025), que sistematiza procedimentos ensináveis de autorregulação – definição precisa de comportamentos-alvo, formas de mensuração e passos de intervenção – a partir de revisão de literatura (2003–2023). Essa obra supre a escassez de materiais em português e orienta a transposição de práticas para aplicativos educacionais, alinhando-se à proposta deste estudo de integrar inteligência artificial (IA) e intervenções baseadas em evidências para apoiar o reconhecimento e a modulação de emoções em crianças com TEA.

Paralelamente, ferramentas tecnológicas têm ganhado espaço nesse cenário. Pesquisas já indicavam, desde o início dos anos 2000, o potencial de sistemas digitais no suporte ao ensino de habilidades socioemocionais. Um exemplo disso é o estudo de Stathopoulou et al. (2020), que analisou a efetividade de um aplicativo em tablets Android para treinar habilidades soci-

ais em três adolescentes autistas. Os resultados mostraram ganhos consistentes e duradouros na generalização de comportamentos sociais adequados, reforçando o potencial das mídias digitais como suporte pedagógico e terapêutico. Mais recentemente, a IA tem se destacado pela capacidade de oferecer soluções personalizadas, adaptadas às necessidades de cada criança. A literatura sobre inteligência artificial ressalta seu papel como sistemas capazes de processar grandes volumes de dados, aprender padrões e gerar respostas adaptativas, o que possibilita aplicações em contextos educacionais e clínicos.

Exemplo disso é o estudo de Kim et al. (2025), que investigou o uso da IA generativa para apoiar pais e cuidadores de crianças autistas. Os autores compararam diferentes versões do ChatGPT em perguntas sobre parentalidade e analisaram critérios como precisão, clareza e legibilidade, baseando-se em McFayden et al. (2024). Os resultados mostraram que a IA fornece respostas corretas e compreensíveis, mas ainda com limitações em sensibilidade ao contexto e citações confiáveis.

Problema de Pesquisa

Essa discussão evidencia a importância de analisar criticamente o uso de ferramentas de inteligência artificial no contexto de intervenções para crianças com Transtorno do Espectro Autista (TEA). Embora recursos digitais, com GEM Storybook, ofereçam potencial para personalização e apoio à autorregulação emocional, ainda existem lacunas quanto à qualidade, clareza, relevância emocional e adequação das histórias geradas. Nesse contexto, surge o problema de pesquisa: em que medida diferentes prompts utilizados no GEM Storybook influenciam a qualidade, clareza e aplicabilidade das histórias sociais digitais para intervenções de autorregulação emocional em crianças com TEA.

Pergunta de Pesquisa

Quais diferenças podem ser identificadas na qualidade, clareza e adequação das histórias sociais digitais geradas pelo GEM Storybook quando aplicadas a diferentes cenários de autorregulação emocional em crianças com Transtorno do Espectro Autista (TEA)?

Relevância da Pesquisa

O uso de histórias sociais digitais tem se mostrado uma estratégia eficaz para apoiar crianças com Transtorno do Espectro Autista (TEA) no desenvolvimento de habilidades socioemocionais. Stathopoulou et al. (2020) analisaram a efetividade de um aplicativo em tablets para treinar habilidades sociais em três adolescentes autistas, observando ganhos consistentes e duradouros na generalização de comportamentos sociais adequados. Esse achado evidencia o potencial de recursos digitais como suporte pedagógico e terapêutico.

A presente pesquisa busca avaliar a qualidade e adequação das histórias sociais digitais geradas pelo GEM Storybook, fornecendo subsídios para seu uso em contextos clínicos e educacionais e contribuindo para a integração entre tecnologia e intervenções baseadas em evidências.

Revisão da Literatura

Silva e Goukart (2024) realizaram uma análise de aplicativos educacionais destinados ao treino de habilidades comunicativas e sociais em crianças com Transtorno do Espectro Autista (TEA), destacando a relevância das tecnologias digitais como ferramentas de apoio pedagógico e terapêutico. O estudo avaliou três aplicativos disponíveis na plataforma Google Play — ABC para Autista, TEO Autismo e PictoTEA — todos gratuitos e compatíveis com o sistema operacional Android. Os autores investigaram tanto a funcionalidade dos aplicativos quanto sua importância para o desenvolvimento das crianças, enfatizando que esses recursos digitais combinam elementos visuais, sonoros e interativos que favorecem o aprendizado de forma lúdica e personalizada. Entre as funcionalidades destacadas estavam atividades estruturadas, repetição de tarefas, reforços positivos e feedback imediato, permitindo que a criança aprendesse no seu próprio ritmo e de acordo com suas necessidades individuais.

O estudo mostrou que os aplicativos proporcionam apoio visual e estratégias que facilitam a compreensão e o uso da linguagem, bem como a expressão de sentimentos, necessidades e interações sociais, ampliando a capacidade das crianças de se engajar com o ambiente físico e social. Além disso, os aplicativos analisados contribuem para a inclusão escolar, oferecendo ferramentas acessíveis e de baixo custo que podem complementar intervenções tradicionais e potencializar o desenvolvimento de habilidades sociais e comunicativas em crianças com TEA. Os autores ressaltam que, quando bem estruturados, esses aplicativos podem servir como um recurso contínuo e motivador para crianças, professores e terapeutas, integrando-se às práticas pedagógicas e clínicas de forma eficaz.

As inteligências artificiais generativas, como o ChatGPT, são sistemas baseados em modelos de linguagem capazes de produzir textos novos e coerentes a partir de padrões aprendidos em grandes volumes de dados. Diferente de ferramentas tradicionais de busca, essas IAs não apenas recuperam informações, mas também formulam respostas contextualizadas e adaptadas às demandas do usuário. No campo do autismo e da parentalidade, McFayden et al. (2024) mostraram que a IA pode responder de forma clara e concisa a dúvidas frequentes de cuidadores sobre autismo, embora apresente limitações quanto à aplicabilidade prática e à precisão das referências. De forma complementar, Kim et al. (2025) analisaram a qualidade das respostas da IA para perguntas sobre parentalidade e cuidados, concluindo que, apesar de corretas e compreensíveis, elas ainda carecem de sensibilidade ao contexto e consistência na apresentação de fontes.

O estudo de Kim et al. (2025), que investigou o uso da IA generativa para apoiar pais e cuidadores de crianças autistas. Os autores testaram diferentes versões do ChatGPT (3.5 e 4) em cinco tópicos recorrentes da parentalidade, utilizando tanto bases de dados verificadas quanto acesso à internet em tempo real. As respostas foram avaliadas quanto à precisão, clareza, relevância prática e legibilidade, com base em critérios adaptados de estudos anteriores (McFayden et al., 2024). Foram realizados nove testes avaliando respostas de diferentes condições — ChatGPT (3.5 e 4.0), ChatGPT + internet, e JITP.info (site de parentalidade) — para cinco perguntas sobre cuidados infantis. As respostas foram analisadas quanto à correção, clareza, conexão (relações pais-filhos) replicabilidade (citações e credibilidade) e legibilidade.

de (Flesch-Kincaid). Nenhuma resposta foi totalmente incorreta, mas algumas foram inadequadas para a faixa etária da criança. Resultados indicaram que a IA fornece respostas precisas e claras, porém com menor sensibilidade ao contexto e limitações em citar fontes. A legibilidade variou: JITP.info apresentou níveis acessíveis (6^a série), enquanto o ChatGPT retornou níveis mais altos (9^o-14^o ano).

O estudo de Kim et al. (2025) teve como objetivo avaliar a capacidade de inteligências artificiais generativas em responder a questões de pais e cuidadores sobre práticas de parentalidade. Para isso, os autores selecionaram cinco tópicos recorrentes em cuidados infantis e formularam perguntas específicas, que foram submetidas a diferentes versões do ChatGPT (3.5 e 4.0), com e sem acesso à internet, além de um site especializado em parentalidade (JITP.info). O procedimento consistiu na realização de nove testes comparativos, em que as respostas foram avaliadas segundo critérios de correção, clareza, conexão com relações pais-filhos, replicabilidade (consistência de fontes e credibilidade) e legibilidade, esta última medida pelo índice de Flesch-Kincaid. A análise revelou que nenhuma resposta foi totalmente incorreta, mas algumas se mostraram inadequadas para determinadas faixas etárias. Além disso, observou-se que o ChatGPT apresentou respostas corretas e compreensíveis, porém com menor sensibilidade ao contexto e dificuldades em fornecer citações consistentes, ao passo que o JITP.info apresentou nível de legibilidade mais acessível.

Já o estudo de McFayden et al. (2024) buscou avaliar a viabilidade do uso do ChatGPT como recurso de apoio para responder a perguntas frequentes de pais e cuidadores de crianças autistas. Para isso, os pesquisadores selecionaram um conjunto de dúvidas recorrentes sobre diagnóstico, intervenções e práticas de cuidado, submetendo-as ao modelo de linguagem. As respostas foram avaliadas em termos de clareza, precisão e relevância prática, além da qualidade das referências apresentadas. O procedimento envolveu análises sistemáticas das respostas fornecidas pelo ChatGPT, comparando-as com informações de fontes verificadas da literatura científica. Os resultados mostraram que a IA foi capaz de gerar respostas claras, corretas e de fácil compreensão, embora apresentasse limitações no fornecimento de orientações aplicáveis a situações concretas e inconsistências na apresentação de referências.

O uso de inteligências artificiais generativas, como o ChatGPT, tem se mostrado promissor no apoio a pais e cuidadores que buscam informações sobre o autismo e sobre práticas parentais em geral. McFayden et al. (2024) investigaram a viabilidade do ChatGPT como ferramenta para responder a dúvidas comuns de cuidadores sobre autismo, revelando que a IA forneceu respostas claras, corretas e concisas, mas ainda com limitações quanto à ação prática e precisão das referências apresentadas.

Já Kim et al. (2025) analisaram a qualidade das respostas da IA para perguntas sobre parentalidade e cuidados, apontando que, embora as informações fossem em grande parte corretas e compreensíveis, havia ausência de conexões mais profundas com o contexto de desenvolvimento infantil, além da falta de consistência nas citações. Esses achados indicam que a IA pode atuar como um recurso acessível e inicial para pais e cuidadores, mas seu uso precisa ser acompanhado de senso crítico e validação por fontes especializadas, a fim de evitar riscos de desinformação.

O estudo de Scattone et al. (2002) investigou a eficácia das histórias sociais na redução de comportamentos disruptivos em três crianças com autismo, utilizando um delineamento de estudo de caso múltiplo. Inicialmente, os pesquisadores observaram a frequência dos comportamentos-alvo em linha de base, sem intervenção. Cada criança apresentava dificuldades distintas em sala de aula, como levantar-se da cadeira sem permissão, falar fora de hora e realizar estereotípias verbais. Com base nessas necessidades, foram elaboradas histórias sociais individualizadas, escritas em primeira pessoa e com frases curtas, descrevendo a situação em que o comportamento disruptivo ocorria, o comportamento adequado esperado e as consequências positivas associadas. O procedimento consistia na leitura diária dessas histórias antes dos momentos críticos da rotina escolar, acompanhada de perguntas de checagem para assegurar a compreensão. A coleta de dados foi realizada por meio de observação direta e registros sistemáticos da frequência dos comportamentos-alvo. Os resultados evidenciaram reduções consistentes em todos os casos, mostrando que os comportamentos disruptivos diminuíram significativamente durante a intervenção. Além disso, os efeitos positivos foram mantidos mesmo após a diminuição gradual da frequência de leitura, sugerindo que as crianças conseguiram generalizar o aprendizado para o cotidiano escolar. Dessa forma, os autores concluíram que as histórias sociais representam uma estratégia simples, acessível e de baixo custo, que favorece previsibilidade e clareza, auxiliando crianças com autismo a adotarem respostas mais adaptativas em contextos educacionais.

Já a pesquisa *Using a Modified Social Story to Decrease Disruptive Behavior of a Child With Autism* descreveu em detalhes a elaboração e aplicação de uma história social adaptada para um menino de 8 anos diagnosticado com autismo (Crozier & Tincani, 2005). A intervenção foi planejada a partir de entrevistas com professores e observações diretas, identificando como alvo o comportamento de “falar fora de hora” em sala de aula. Com base nesses dados, os pesquisadores elaboraram uma narrativa curta e ilustrada, escrita em primeira pessoa, que ensinava o comportamento substitutivo de levantar a mão e esperar a vez para falar. A história continha frases simples como: “Quando estou na sala de aula e quero dizer algo, eu levanto minha mão. O professor verá minha mão e então será minha vez de falar. Esperar a minha vez deixa o professor feliz e ajuda meus colegas a aprenderem.” Esse recurso, lido diariamente antes das atividades em que o comportamento era mais frequente, ajudava a criança a antecipar o que deveria fazer e a internalizar o comportamento adequado.

O procedimento consistiu na leitura da história social imediatamente antes das atividades em que o comportamento disruptivo era mais frequente. A criança lia o texto em voz alta, respondia a questões de compreensão e, em algumas fases, recebia lembretes verbais (“Lembre-se de levantar a mão quando quiser falar”). A história era simples, com frases curtas, vocabulário acessível e imagens que reforçam o conteúdo, de modo a facilitar a autonomia do participante na leitura. Os resultados mostraram uma redução significativa do comportamento disruptivo, especialmente quando a história foi combinada com os prompts verbais. Além disso, observou-se manutenção dos ganhos após duas semanas, quando a leitura da história passou a fazer parte da rotina escolar.

Dessa forma, as histórias sociais configuram-se como uma ferramenta pedagógica e terapêutica de baixo custo, acessível e de comprovada eficácia, sendo amplamente aplicadas em contextos clínicos, educacionais e familiares.

Ao contrário de pesquisas voltadas para compreender o Transtorno do Espectro Autista (TEA), estudos sobre a aplicação da inteligência artificial (IA) nesse contexto ainda são escassos. Segundo McFayden et al. (2024), há uma lacuna significativa no desenvolvimento de avaliações automatizadas, ferramentas de reconhecimento de padrões comportamentais e aplicativos específicos voltados para pessoas com autismo. Essa ausência de soluções tecnológicas limita o aproveitamento do potencial da IA para apoiar diagnósticos, monitoramento e intervenções personalizadas, evidenciando a necessidade de pesquisas mais aprofundadas e direcionadas à integração da tecnologia na prática clínica e educacional.

Além disso, pesquisas recentes também têm investigado o uso de histórias como estratégia para o ensino de habilidades emocionais em crianças com Transtorno do Espectro Autista. O estudo de Bondioli et al. (2024) demonstrou que intervenções baseadas em histórias podem favorecer a aprendizagem no reconhecimento de expressões faciais de emoções, como alegria, tristeza, raiva e surpresa, em crianças com TEA. Esses achados reforçam o potencial pedagógico das narrativas estruturadas como ferramenta para o desenvolvimento de repertórios socioemocionais, aspecto que também se relaciona com os resultados do presente estudo ao evidenciar que histórias bem organizadas podem apoiar a compreensão de emoções e comportamentos sociais (Bondioli et al., 2024).

No que diz respeito ao Transtorno do Espectro Autista (TEA), destaca-se também a importância da autorregulação emocional, entendida como a capacidade do indivíduo de reconhecer, compreender e gerenciar suas próprias emoções, especialmente diante de situações desafiadoras ou estressantes. Essa habilidade é fundamental para o desenvolvimento social e emocional, uma vez que muitas crianças com TEA apresentam dificuldades na identificação, expressão e controle das emoções, o que pode impactar suas interações sociais e comportamentos no ambiente escolar e familiar. Dessa forma, o desenvolvimento da autorregulação emocional torna-se um aspecto essencial a ser trabalhado por meio de intervenções pedagógicas, terapêuticas e tecnológicas (Iaochite et al., 2022).

Além disso, a autorregulação emocional está diretamente relacionada à autoeficácia para regulação emocional, isto é, à crença do indivíduo em sua própria capacidade de lidar com suas emoções de maneira eficaz. Evidências indicam que quanto maior essa percepção de autoeficácia, maiores são as chances de adoção de comportamentos positivos, como melhor convivência social, maior engajamento em atividades escolares e uso de estratégias adequadas de enfrentamento. Nesse sentido, ao considerar o uso de recursos como histórias sociais, aplicativos educativos e ferramentas baseadas em inteligência artificial, torna-se fundamental não apenas promover o reconhecimento emocional, mas também fortalecer a confiança da criança em sua capacidade de regular suas emoções, favorecendo seu desenvolvimento global e sua inclusão social (Iaochite et al., 2022).

Objetivo da Pesquisa

O objetivo geral desta pesquisa é analisar a eficácia, qualidade, usabilidade e adequação do modelo de Inteligência Artificial generativa Gemini 2.5 Flash (Google) na criação de histórias sociais digitais para crianças com Transtorno do Espectro Autista (TEA), comparando diferentes tipos de prompts e avaliando a customização do storybook gerado. Especificamente,

busca-se verificar o grau de personalização e customização das histórias digitais, considerando características individuais da criança e alinhamento com práticas terapêuticas baseadas em evidências, especialmente a Análise do Comportamento Aplicada (ABA).

1 MÉTODO

Delineamento de Pesquisa

Esta pesquisa caracteriza-se como descritiva e comparativa, com abordagem mista (quantitativa e qualitativa), e tem como foco a avaliação de histórias sociais digitais produzidas pelo modelo de Inteligência Artificial generativa Gemini 2.5 Flash (Google), utilizando a ferramenta GEM: Storybook, direcionadas ao desenvolvimento da autorregulação emocional em crianças com Transtorno do Espectro Autista (TEA). Para isso, foram geradas histórias sociais a partir de diferentes tipos de prompts, abordando situações relevantes para o manejo das emoções e adaptadas à faixa etária das crianças-alvo.

A análise dos dados foi conduzida de forma quantitativa, a partir da pontuação atribuída pelos instrumentos, e qualitativa, considerando comentários dos avaliadores sobre pontos fortes, limitações e nível de customização observado. Esse delineamento permite compreender em que medida a ferramenta GEM: Storybook do Gemini atende às características que tornam as histórias sociais uma intervenção eficaz, fornecendo subsídios para reflexões sobre sua aplicabilidade em contextos terapêuticos e educacionais.

Modelos Utilizados

Para a presente pesquisa foi utilizado exclusivamente o Gemini versão 2.5 flash (Google), modelo de Inteligência Artificial generativa desenvolvido pela Google DeepMind. O Gemini é um sistema multimodal capaz de compreender e gerar textos a partir de diferentes contextos, sendo empregado neste estudo por meio da ferramenta GEM: Storybook, especificamente voltada para a criação de narrativas digitais personalizadas.

A escolha pelo Gemini se deve à sua capacidade de produzir conteúdos narrativos coerentes, adaptáveis e de fácil customização, características essenciais para a construção de histórias sociais direcionadas a crianças com Transtorno do Espectro Autista (TEA). A ferramenta GEM: Storybook permite que as histórias sejam geradas a partir de prompts personalizados, possibilitando a comparação entre diferentes estratégias de comando (ex.: prompts mais gerais versus prompts detalhados e individualizados) e a análise de como tais variações influenciam a clareza, a relevância emocional, a personalização e a adequação das histórias produzidas.

Dessa forma, o uso do Gemini com a ferramenta GEM: Storybook oferece não apenas um meio inovador de geração de histórias sociais digitais, mas também a oportunidade de investigar como diferentes modos de interação com o modelo impactam a qualidade e a efetividade do material produzido.

Prompts Utilizados

Para a criação do prompt, utilizou-se o seguinte comando:

“Crie uma história social para [NOME], [IDADE] anos, que gosta de [INTERESSES]. Objetivo: [OBJETIVO DE AUTORREGULAÇÃO]. Escreva em 1ª pessoa, frases curtas (máx. 8–10 frases), incluindo: descrição da situação, o que eu penso/sinto, o que eu faço (comportamento esperado) e a consequência positiva. Sugira ou gere imagens de acordo com as falas para ilustrar a história”

Prompt 1 – Controle da raiva (perfil + imagens)

“Crie uma história social digital para uma criança autista de 3 anos chamada Lucas. Ele gosta de dinossauros e futebol. A história deve ajudá-lo a aprender a se acalmar quando fica com raiva porque perde em um jogo. Use frases curtas, em primeira pessoa, descrevendo a situação, o comportamento esperado (respirar fundo, contar até 5) e a consequência positiva (poder continuar jogando e se divertir com os amigos). Sugira ou gere imagens de acordo com as falas para ilustrar a história”

Prompt 2 – Lidar com frustração (perfil + imagens)

“Crie uma história social para uma criança autista de 9 anos chamada Pedro, que gosta de videogames e carros. O objetivo é ensinar Pedro a lidar com a frustração quando perde em um jogo. Escreva em primeira pessoa, com frases simples, mostrando que ele pode respirar fundo, contar até 10 e tentar de novo. Sugira ou gere imagens de acordo com as falas para ilustrar a história”

Prompt 3 – Identificação de emoções (perfil + imagens)

“Escreva uma história social curta para uma criança com TEA de 6 anos chamada Ana, que gosta de gatos e de desenhar. O objetivo é ajudar Ana a reconhecer quando está triste e pedir ajuda a um adulto. Use frases curtas em primeira pessoa. Sugira ou gere imagens de acordo com as falas para ilustrar a história”

Avaliação do Conteúdo Gerado

A avaliação das histórias sociais digitais produzidas pelos modelos de IA será realizada a partir de critérios objetivos que consideram tanto aspectos de forma quanto de conteúdo. Serão utilizados checklists e rubricas que englobam dimensões como clareza textual, relevância emocional, adequação à faixa etária, consistência interna e alinhamento com práticas baseadas em ABA.

Essa proposta se inspira em pesquisas que já investigaram o uso de tecnologias digitais para apoiar crianças autistas. McFayden et al. (2024) avaliaram o ChatGPT como recurso para responder dúvidas sobre autismo, destacando a clareza e correção das respostas, embora com limitações no detalhamento prático. De forma semelhante, Kim et al. (2025) compararam versões do ChatGPT em perguntas de parentalidade, observando que, apesar da precisão e legibilidade dos conteúdos, ainda havia baixa sensibilidade ao contexto e inconsistência na referência às fontes. Assim, a avaliação proposta neste estudo combina critérios de usabilidade

e qualidade já identificados na literatura, ao mesmo tempo em que busca inovar ao comparar conteúdos gerados por diferentes modelos de IA. Essa abordagem permitirá identificar se os textos produzidos conseguem reproduzir características essenciais das histórias sociais eficazes (clareza, previsibilidade e foco no comportamento alvo) e se têm aplicabilidade prática em contextos clínicos e educacionais.

Rubrica de avaliação

A construção da rubrica de avaliação utilizada neste estudo foi fundamentada em referenciais teóricos consolidados na área do Transtorno do Espectro Autista (TEA), especialmente nos princípios das histórias sociais propostos por Gray (2010), bem como em pesquisas da Análise do Comportamento Aplicada (ABA) e estudos empíricos sobre intervenções com histórias sociais (Scattone et al., 2006; Crozier & Tincani, 2005; Stathopoulou et al., 2020; Bondioli et al., 2024).

O critério de clareza e estrutura textual baseia-se nas recomendações de Gray (2010), que enfatiza que histórias sociais devem utilizar linguagem simples, frases curtas e objetivas, de modo a facilitar a compreensão por crianças com TEA. Essa orientação também é reforçada por estudos aplicados que demonstram que textos claros e estruturados favorecem a aprendizagem e a adesão ao comportamento esperado (Scattone et al., 2006).

A adequação à faixa etária foi incluída a partir da necessidade de alinhar o conteúdo ao nível de desenvolvimento cognitivo da criança, conforme discutido na literatura sobre intervenções comportamentais (Moreira & Medeiros, 2019). Histórias com vocabulário concreto e contextualizado tendem a ser mais eficazes na promoção de habilidades sociais.

O critério de relevância emocional e social está relacionado à importância de utilizar situações significativas para a criança, favorecendo a generalização do comportamento aprendido, conforme evidenciado em estudos sobre histórias sociais e ensino de habilidades socioemocionais (Bondioli et al., 2024).

A clareza do comportamento-alvo deriva diretamente dos princípios da ABA, que preconizam a definição de comportamentos observáveis e mensuráveis para o ensino eficaz (Moreira & Medeiros, 2019). Estudos como os de Crozier e Tincani (2005) demonstram que a descrição explícita do comportamento esperado contribui para a redução de comportamentos disruptivos.

A estrutura da história social segue rigorosamente o modelo proposto por Gray (2010), que inclui os seguintes elementos: descrição da situação, identificação de pensamentos e sentimentos, indicação do comportamento esperado e apresentação de consequências positivas. Essa estrutura também é validada empiricamente por pesquisas como as de Scattone et al. (2006).

A consistência interna da narrativa foi considerada com base na necessidade de previsibilidade e coerência, aspectos fundamentais para crianças com TEA, conforme destacado na literatura clínica e educacional.

O critério de alinhamento com ABA e intervenções baseadas em evidências considera a presença de estratégias como reforço positivo, previsibilidade e ensino de autorregulação, amplamente discutidas na análise do comportamento (Moreira & Medeiros, 2019; Aguiar & Moreira, 2025).

A qualidade visual foi incluída com base em estudos sobre intervenções digitais, que apontam que estímulos visuais claros e integrados ao texto aumentam a compreensão e o engajamento da criança (Stathopoulou et al., 2020).

Por fim, a usabilidade clínica e educacional reflete a aplicabilidade prática das histórias sociais em contextos reais, aspecto fundamental destacado em estudos aplicados que investigam a eficácia dessas intervenções no cotidiano escolar e terapêutico (Crozier & Tincani, 2005; Scattone et al., 2006).

Com base nesses referenciais teóricos e empíricos, elaborou-se uma rubrica de avaliação com o objetivo de sistematizar a análise das histórias sociais produzidas, contemplando critérios relacionados à clareza textual, adequação ao desenvolvimento da criança, relevância emocional, definição do comportamento-alvo, estrutura narrativa, consistência interna, alinhamento com princípios da Análise do Comportamento Aplicada (ABA), qualidade visual e aplicabilidade prática. A seguir, apresenta-se a rubrica utilizada no estudo:

1 = Muito fraco | 2 = Fraco | 3 = Regular | 4 = Bom | 5 = Excelente

1. Clareza e Estrutura Textual

1-2: Texto confuso, frases longas, linguagem complexa ou ambígua.

3: Texto compreensível, mas com algumas frases extensas ou pouco objetivas.

4-5: Texto claro, objetivo, frases curtas, linguagem simples e adequada à criança.

2. Adequação à Faixa Etária

1-2: Vocabulário, exemplos ou situações não condizem com a idade da criança; frases muito longas ou complexas, com construções abstratas que dificultam a compreensão.

3: Parcialmente adequado – linguagem simples, porém com trechos extensos ou conceitos pouco acessíveis.

4-5: Vocabulário, frases e situações ajustados ao nível cognitivo e etário da criança-alvo (frases curtas, concretas e contextualizadas).

3. Relevância Emocional e Social

1-2: Situação pouco significativa ou distante da realidade da criança. 3: Situação relevante, mas pouco conectada a experiências comuns.

4-5: Situação altamente pertinente, envolvendo emoções e contextos reais da criança.

4. Clareza do Comportamento-Alvo

1-2: O comportamento esperado não fica claro ou aparece de forma indireta. 3: O comportamento é citado, mas não descrito passo a passo.

4-5: O comportamento está bem definido, descrito em termos observáveis e ensináveis.

5. Estrutura da História Social (Conforme critérios de Gray (2010) e pesquisas em ABA)

1-2: Ausência de uma ou mais etapas fundamentais.

3: Contém descrição da situação, mas com falhas nas outras partes. 4-5: Contém todas as etapas:

Situação/contexto → O que eu penso/sinto → O que devo fazer (comportamento esperado) → Consequência positiva.

6. Consistência Interna

1-2: Há contradições ou mudanças bruscas no enredo.

3: Narrativa coerente, mas com pequenos lapsos de continuidade. 4-5: Narrativa coesa, fluida e sem contradições.

7. Alinhamento com ABA e Intervenções Baseadas em Evidências

1-2: Estratégias apresentadas não refletem princípios de ABA. 3: Estratégias parcialmente alinhadas.

4-5: Estratégias claramente fundamentadas em ABA (ex.: reforço positivo, previsibilidade, autorregulação).

8. Qualidade Visual (Imagens)

1-2: imagens pouco claras, não relacionadas ao texto ou distrativas.

3: Imagens relacionadas, mas sem correspondência precisa com cada cena.

4-5: Imagens simples, claras, correspondentes às falas e facilitadoras da compreensão.

9. Usabilidade Clínica/Educacional

1-2: História de difícil aplicação em contexto clínico ou escolar.

3: Pode ser usada, mas exige adaptações.

4-5: Aplicável diretamente em atendimentos, salas de aula ou em casa, sem necessidade de grandes ajustes.

Pontuação Final:

9-18 pontos: Baixa qualidade / pouca aplicabilidade.

19-31 pontos: Qualidade moderada / aplicável com ajustes. 32-45 pontos: Alta qualidade / aplicabilidade direta.

Procedimento

A implementação do estudo seguirá etapas sequenciais e padronizadas, descritas a seguir:

Definição dos cenários e perfil das crianças

Foram selecionados três cenários de autorregulação emocional relevantes para a faixa etária alvo (3-9 anos), contemplando situações comuns em contextos clínicos e educacionais: (a) controle da raiva ao perder em um jogo, (b) lidar com frustração em situações competitivas,

e (c) identificação e expressão de emoções, como a tristeza. Para cada cenário, foi definido um perfil fictício da criança (nome, idade e interesses), de modo a personalizar a narrativa e favorecer o engajamento, conforme recomendam práticas de elaboração de histórias sociais.

Criação e padronização de prompts

Foi elaborado um template de prompt padronizado, que incluía: objetivo da história, faixa etária, perfil da criança, estrutura narrativa desejada (primeira pessoa, frases curtas, descrição da situação, comportamento esperado e consequência positiva) e solicitação de imagens correspondentes às cenas.

Geração das histórias sociais digitais

Os prompts foram inseridos no modelo Gemini (ferramenta GEM: Storybook), resultando em três histórias digitais personalizadas. Para cada narrativa, foram registradas tanto as versões textuais quanto as imagens sugeridas ou geradas pelo modelo. Quando o sistema não produzia imagens automaticamente, foram coletadas descrições detalhadas para eventual ilustração posterior.

Aplicação da rubrica (avaliação quantitativa e qualitativa)

As histórias sociais digitais produzidas serão avaliadas por especialistas das áreas de psicologia, pedagogia e Análise do Comportamento Aplicada (ABA), utilizando uma rubrica padronizada. Cada item será avaliado em escala tipo Likert de 1 (muito inadequado) a 5 (excelente), contemplando os seguintes critérios: clareza textual, uso de primeira pessoa e estrutura adequada, simplicidade das frases, relevância emocional, presença de consequência positiva, alinhamento com princípios de ABA (previsibilidade e reforço positivo), adequação visual (se houver imagens), sugestões práticas/ação (actionability), segurança (ausência de conteúdo impróprio) e potencial de generalização. Além das pontuações, os avaliadores poderão registrar comentários qualitativos sobre pontos fortes e limitações observados em cada história.

Esse delineamento possibilita compreender em que medida os conteúdos gerados pelo GEM Storybook atendem às características que tornam as histórias sociais uma intervenção eficaz e fornece subsídios para refletir sobre sua aplicabilidade em contextos clínicos e educacionais.

2 RESULTADOS

Coleta de Dados

A coleta de dados foi realizada no dia 21 de outubro de 2025, por meio da ferramenta Storybook do modelo Gemini (Google), utilizada para a geração automatizada das histórias sociais digitais. O processo ocorreu de forma remota e individualizada, seguindo os prompts padronizados previamente elaborados para cada cenário de autorregulação emocional.

Cada história foi gerada em aproximadamente dois minutos, incluindo o tempo de processamento, leitura e registro dos resultados. A coleta foi conduzida em um único dia, garantindo padronização das condições e controle das variáveis envolvidas. Após a geração, as histórias foram armazenadas integralmente em formato textual e visual, assegurando a integridade dos dados para as etapas posteriores de avaliação.

O procedimento transcorreu sem intercorrências técnicas, com respostas rápidas e consistentes por parte do sistema. Assim, a coleta de dados caracterizou-se como um processo breve, controlado e eficiente, que possibilitou reunir todos os materiais necessários para a análise quantitativa e qualitativa do estudo.

As Histórias

Descrição da História Social de Lucas

A história social de Lucas foi desenvolvida para ensinar estratégias de autorregulação emocional diante da frustração em situações de perda durante jogos coletivos. Criada na ferramenta Gemini – Storybook, a história utiliza ilustrações no estilo infantil colorido, com traços arredondados, expressões faciais simples e elementos lúdicos (como dinossauros e futebol), que facilitam a identificação emocional e o engajamento de crianças pequenas. A história gerada pode ser acessada: <https://gemini.google.com/share/94bc588b6cc5>

O material é composto por capa e 12 páginas, com frases curtas e diretas, contendo em média 60 a 90 caracteres por página. O texto é narrado em primeira pessoa, com vocabulário concreto e acessível, adequado à faixa etária de 3 anos.

O texto completo da história é o seguinte:

Página 1: Eu sou o Lucas. Eu amo dinossauros.

Página 2: Eu e meus amigos, Tiago e Sofia, gostamos de jogar futebol. A bola é o nosso ovo de dinossauro.

Página 3: Às vezes, meu time perde o jogo. Isso acontece.

Página 4: Eu fico com muita raiva quando perco. Meu corpo fica tenso e duro.

Página 5: Eu quero gritar ou chutar a bola com força. Mas isso não é legal. Eu paro.

Página 6: Eu paro e penso: eu posso me acalmar.

Página 7: O primeiro passo é respirar. Eu respiro fundo como um tiranossauro (inspirar).

Página 8: Eu conto até 5, bem devagar. Um, dois, três, quatro, cinco! Página 9: O ar ruim sai. O ar bom entra. Meu corpo fica mole e calmo.

Página 10: Agora eu estou pronto para jogar de novo. Eu posso me divertir com meus amigos.

Página 11: Ficar calmo é a chave para continuar jogando e rindo.

Página 12: Eu sou o Lucas. Eu sou um dinossauro calmo e um bom amigo.

A Figura 1 apresenta exemplos das páginas da história “Lucas e o Dinossauro Calmo”, evidenciando a sequência narrativa visual e as estratégias de autorregulação descritas no texto.

Descrição da História Social de Pedro

A história social de Pedro foi elaborada para desenvolver estratégias de autorregulação emocional e enfrentamento da frustração, especialmente em contextos competitivos digitais. Criada na ferramenta Gemini – Storybook, a história apresenta ilustrações no estilo semirrealista infantil, com cores vivas e foco em expressões faciais e movimentos que refletem as emoções do personagem. A história gerada pode ser acessada: <https://gemini.google.com/share/f3a564a37380>

O material é composto por capa e 12 páginas, com frases curtas, objetivas e em primeira pessoa, contendo em média 80 a 100 caracteres por página, adequadas à faixa etária de 9 anos. A narrativa utiliza elementos de interesse pessoal (videogames e carros de corrida) para facilitar o engajamento e a identificação do leitor com o comportamento-alvo.

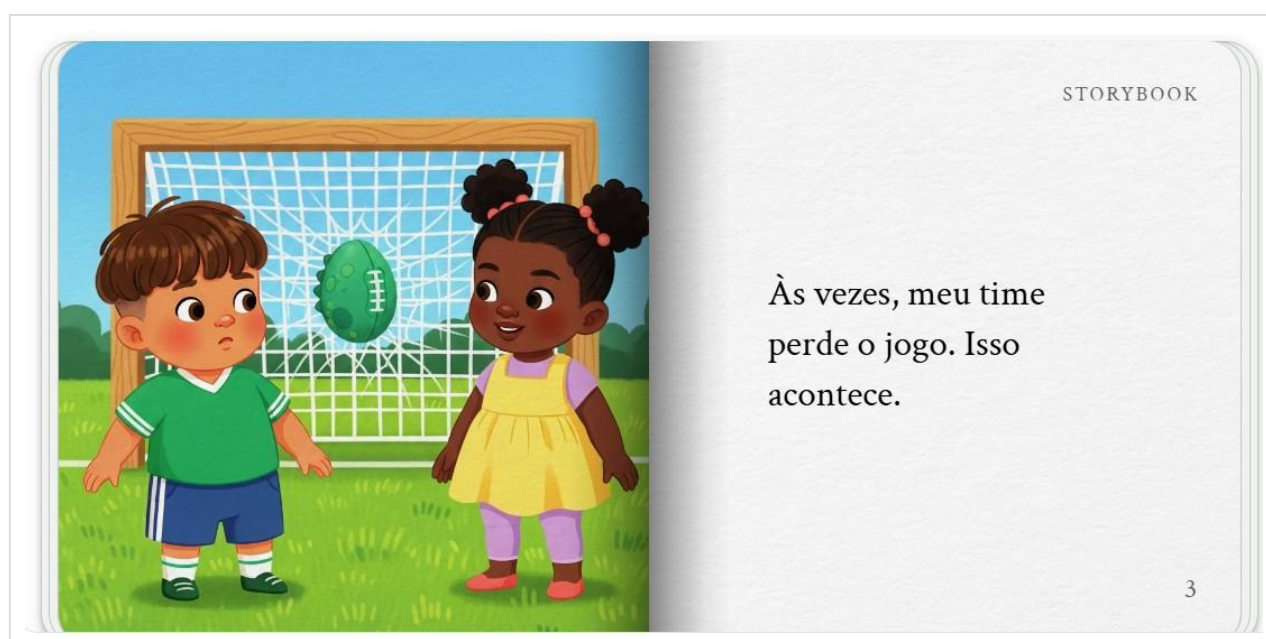


Figura 1. Exemplo de página da história social “Lucas e o Dinossauro Calmo”

Nota. Página ilustrada gerada no GEM Storybook. A cena mostra duas crianças diante de uma trave de futebol, acompanhando a frase “Às vezes, meu time perde o jogo. Isso acontece.”, que apresenta a situação-problema da narrativa voltada à autorregulação emocional.

O texto completo da história é o seguinte:

Página 1: Eu sou o Pedro. Eu tenho 9 anos. Eu amo jogar videogame, principalmente jogos de corrida.

Página 2: Meu carro favorito é o vermelho. Eu adoro a velocidade! Eu quero muito ser o primeiro lugar.

Página 3: Às vezes, meu carro bate ou outro carro passa. Eu não ganho a corrida. Eu perco.

Página 4: Meu corpo fica quente e duro. Meus punhos ficam fechados. Eu sinto vontade de gritar ou jogar o controle.

Página 5: Perder é frustrante. Mas eu posso parar. Eu sei o que fazer antes de ficar muito bravo.

Página 6: O primeiro passo é respirar fundo. Eu respiro como um motor de carro potente: vruuum.

Página 7: Depois, eu conto até 10, bem devagar. Um, dois, três, quatro, cinco, seis, sete, oito, nove, dez.

Página 8: O ar sai. A calma volta. Meu corpo fica mole, como depois de uma corrida.

Página 9: Eu repito a respiração e a contagem se eu ainda estiver bravo. Eu sempre tento ficar calmo.

Página 10: Agora, o jogo não parece tão difícil. É só um jogo. Eu posso apertar reiniciar.

Página 11: Eu tento de novo. Eu sempre melhora quando estou calmo e concentrado.

Página 12: Eu sou o Pedro. Eu posso perder, mas eu sei me acalmar e tentar de novo. Isso me deixa mais forte!

A Figura 2 apresenta exemplos das páginas da história “Pedro e a Corrida da Calma”, destacando o uso de metáforas visuais (como o motor e a corrida) e a correspondência entre texto e imagem para reforçar o aprendizado emocional.

Descrição da História Social de Ana

A história social de Ana foi elaborada com o objetivo de ensinar o reconhecimento e a expressão da emoção tristeza, incentivando a busca por ajuda e acolhimento emocional. Criada na ferramenta Gemini – Storybook, a história apresenta ilustrações no estilo aquarela infantil, com cores suaves e foco na expressão emocional dos personagens, especialmente na empatia e vínculo entre a criança e a figura materna. A história gerada pode ser acessada: <https://gemini.google.com/share/e06d04617ddc>



Figura 2. Exemplo de página da história social “Pedro e a Corrida da Calma”

Nota. Página gerada no GEM Storybook. O personagem aparece com expressão de raiva e corpo tenso segurando um controle de videogame, ilustrando a frase “Meu corpo fica quente e duro... Eu sinto vontade de gritar ou jogar o controle.”, que nomeia as sensações corporais da frustração.

O material é composto por capa e 12 páginas, com frases curtas, em primeira pessoa e linguagem concreta, totalizando em média 70 a 90 caracteres por página, adequadas à faixa etária de 6 anos. A narrativa utiliza elementos afetivos do cotidiano (o gato Pipoca e o hábito de desenhar) para criar identificação emocional e contextualizar o aprendizado de forma natural.

O texto completo da história é o seguinte:

Página 1: Eu sou Ana. Tenho 6 anos. Eu amo meus gatos e adoro desenhar.

Página 2: Meu gato favorito é o Pipoca. Quando eu desenho com o Pipoca por perto, eu me sinto feliz e calma.

Página 3: Às vezes, as coisas não saem como eu planejo. Hoje meu desenho favorito rasgou um pouquinho.

Página 4: Meu corpo começa a ficar mole. Meu peito parece pesado. Eu sinto um bolo na minha garganta.

Página 5: Eu reconheço esta sensação. Não é raiva. Não é medo. É tristeza. Eu estou triste.

Página 6: Quando estou triste, não preciso ficar sozinha. Eu preciso de ajuda de um adulto.

Página 7: Eu posso procurar a mamãe Inês. Ela sempre sabe o que fazer.

Página 8: Eu chego perto dela e digo: “Mamãe Inês, eu estou triste. Preciso de ajuda.”

Página 9: Mamãe Inês me dá um abraço suave. Ela diz: “Eu entendo, Ana. Obrigada por me contar.”

Página 10: Ela me ajuda a colocar um adesivo transparente no meu desenho. O desenho está seguro agora.

Página 11: O bolo na garganta sumiu. Meu peito não está mais pesado. Eu me sinto calma de novo.

Página 12: Eu sou Ana. É normal ficar triste. Eu sempre posso dizer a um adulto: “Eu preciso de ajuda.”

A Figura 3 apresenta exemplos das páginas da história “O Dia em que Fiquei Triste”, destacando o uso de cores suaves, expressões empáticas e a representação visual do apoio emocional, que contribuem para a compreensão e identificação da criança com a narrativa.



Figura 3. Exemplo de página da história social “Ana”

Nota. Página inicial gerada no GEM Storybook. A personagem Ana, usando óculos, desenha sentada a uma mesa, acompanhando a frase de apresentação “Eu sou Ana. Eu tenho 6 anos. Eu amo meus gatos e adoro desenhar.”.

Avaliação do Conteúdo

A análise das histórias sociais digitais foi conduzida pela pesquisadora responsável, com base na rubrica de avaliação desenvolvida para este estudo, composta por nove critérios graduados em uma escala de 1 (muito fraco) a 5 (excelente). Cada história foi avaliada individualmente quanto à clareza textual, adequação etária, relevância emocional, definição do comportamento-alvo, estrutura narrativa conforme Gray (2010), consistência interna, alinhamento com princípios da Análise do Comportamento Aplicada (ABA), qualidade visual e usabilidade clínica ou educacional.

De forma geral, as três histórias analisadas – Lucas, Pedro e Ana – apresentaram excelente qualidade textual e conceitual, alcançando a pontuação máxima de 45 pontos, o que as enquadra na categoria Alta qualidade / aplicabilidade direta.

A história de Lucas, voltada para uma criança de 3 anos, demonstrou excelente adequação ao nível de desenvolvimento esperado para a idade. A narrativa foi escrita em primeira pessoa, com frases curtas e linguagem concreta, apresentando o comportamento-alvo (“respirar fundo” e “contar até 5”) de maneira observável e ensinável. A estrutura da história social foi completa e coerente, incluindo a situação, o sentimento, a ação e a consequência positiva. As imagens foram simples e correspondentes às falas, reforçando o aprendizado.

A história de Pedro, destinada a uma criança de 9 anos, abordou o controle da frustração ao perder em um jogo. O texto foi adequado à faixa etária e incorporou interesses pessoais (videogames e carros), o que aumentou o engajamento. O comportamento-alvo foi descrito passo a passo, com instruções claras (“respirar fundo”, “contar até 10” e “tentar novamente”). O alinhamento com os princípios da ABA foi evidente, especialmente na previsibilidade da ação e no uso do reforço positivo. A história também mostrou boa aplicabilidade clínica e educacional, podendo ser utilizada diretamente em contextos terapêuticos e escolares.

A história de Ana, por sua vez, teve como objetivo o ensino de uma habilidade socioemocional fundamental – reconhecer a tristeza e buscar ajuda. A linguagem foi simples, acessível e emocionalmente sensível, adequada à faixa etária de 6 anos. O comportamento-alvo foi apresentado em dois passos concretos: identificar a emoção e comunicar-se com um adulto. A história manteve uma estrutura coesa e seguiu rigorosamente o modelo proposto por Gray (2010), sendo acompanhada por imagens que ilustram o estado emocional e a ação correspondente.

De modo geral, a avaliação realizada pela pesquisadora indicou que as histórias geradas pela inteligência artificial apresentam clareza textual, coerência narrativa e forte potencial de uso terapêutico e educacional. No entanto, observou-se que a adaptação etária poderia ser mais refinada em alguns casos: as histórias voltadas para crianças mais novas (como Lucas) podem demandar menos páginas e maior repetição de frases, enquanto histórias destinadas a faixas etárias mais velhas (como Pedro) podem comportar textos mais longos e com maior detalhamento emocional.

Essa observação reforça a importância da revisão humana no processo de elaboração de materiais automatizados, uma vez que a IA tende a manter um padrão narrativo uniforme, nem sempre ajustando adequadamente a extensão, o vocabulário e a profundidade emocional à idade da criança-alvo. Ainda assim, as histórias avaliadas demonstraram alto potencial de aplicabilidade clínica, confirmando a eficiência dos prompts e das instruções fornecidas no processo de geração das histórias sociais digitais.

A segunda etapa da análise consistiu na avaliação automatizada das histórias sociais digitais, realizada pelo modelo generativo ChatGPT (OpenAI). O objetivo dessa etapa foi comparar a capacidade avaliativa do modelo com a análise humana previamente conduzida pela pesquisadora, verificando convergências e possíveis divergências entre julgamentos humanos e artificiais.

Assim como na etapa anterior, o ChatGPT foi instruído a utilizar exatamente a mesma rubrica de avaliação baseada na Escala Likert de 1 a 5 pontos, contemplando nove critérios fundamentais: clareza e estrutura textual, adequação à faixa etária, relevância emocional e social, clareza do comportamento-alvo, estrutura da história social conforme Gray (2010), consistência interna, alinhamento com princípios da ABA, qualidade visual, usabilidade clínica/educacional.

Para garantir padronização e rigor metodológico, foi elaborado um prompt descritivo, estruturado e detalhado, no qual o modelo recebeu instruções para:

- ler o texto completo de cada história social (Lucas – 3 anos, Ana – 6 anos e Pedro – 9 anos);
- atribuir uma pontuação de 1 a 5 para cada critério da rubrica;
- justificar brevemente a nota atribuída;
- e apresentar uma interpretação qualitativa final, destacando pontos fortes e eventuais limitações.

Além disso, o modelo foi instruído a assumir uma postura altamente especializada, recebendo o seguinte prompt:

“Você é um analista do comportamento, doutor em Análise do Comportamento e com 20 anos de experiência com crianças autistas. Analise de forma crítica a história social anexa, texto e imagens, quanto à sua adequação para uma criança autista de X anos.”

O uso desse prompt teve como finalidade simular uma avaliação profissional altamente qualificada, permitindo verificar se o modelo seria capaz de produzir análises comparáveis às de um analista do comportamento com sólida experiência clínica. Para isso, foi elaborado um quadro sistematizado de critérios avaliativos, utilizado tanto na análise automatizada quanto na avaliação humana, conforme apresentado na Tabela 1.

Tabela 1. Quadro de avaliação do Chat GPT e da avaliação humana

Critério da Rubrica	Lucas (3 anos)	Ana (6 anos)	Pedro (9 anos)
1. Clareza e Estrutura Textual	4-5	5	5
2. Adequação à Faixa Etária	3-4	5	4
3. Relevância Emocional e Social	4-5	5	5
4. Clareza do Comportamento-Alvo	4-5	5	5
5. Estrutura da História Social (Gray/ABA)	4	5	5
6. Consistência Interna	4-5	5	4-5
7. Alinhamento com ABA	4-5	5	4-5
8. Qualidade Visual (Imagens)	3-5	4	5
9. Usabilidade Clínica/Educacional	4-5	5	4-5
Pontuação Total	35-44/45	44-45/45	42-44/45
Classificação Geral	Alta qualidade, requer pequenos ajustes	Alta qualidade, aplicabilidade direta	Alta qualidade, aplicabilidade direta

Nota. Em cada célula do quadro, o primeiro valor corresponde à avaliação realizada pelo ChatGPT, enquanto o segundo valor refere-se à avaliação humana conduzida pela pesquisadora.

A análise dos dados foi realizada por meio da aplicação da rubrica de avaliação desenvolvida para este estudo, contemplando nove critérios essenciais à construção de histórias sociais eficazes, com pontuação variando de 1 (muito fraco) a 5 (excelente). Os critérios abrangeram clareza e estrutura textual, adequação à faixa etária, relevância emocional e social, definição do comportamento-alvo, estrutura da história social conforme Gray (2010), consistência interna, alinhamento com princípios da Análise do Comportamento Aplicada (ABA), qualidade visual e usabilidade clínica/educacional.

Cada história — Lucas, Pedro e Ana — foi analisada individualmente, representando diferentes idades (3, 6 e 9 anos) e distintos contextos emocionais (raiva, tristeza e frustração). As análises produzidas pelo ChatGPT foram posteriormente comparadas aos resultados humanos, o que possibilitou observar o grau de consistência entre ambas e a sensibilidade do modelo em identificar nuances textuais, estruturais e emocionais.

De maneira geral, as pontuações atribuídas apresentaram alta concordância, com médias entre 4,8 e 5 pontos em todos os critérios, classificando as histórias como de alta qualidade e aplicabilidade direta. Tanto o avaliador humano quanto o ChatGPT valorizaram aspectos como: frases curtas em primeira pessoa, clareza de instruções, adequação da linguagem às idades, presença de consequências positivas e alinhamento com princípios essenciais da ABA, especialmente no uso de previsibilidade, autorregulação e reforço positivo.

Apesar disso, a análise qualitativa demonstrou diferenças importantes: o ChatGPT tende a enfatizar elementos estruturais e formais (clareza, organização, coerência), enquanto o avaliador humano apresenta maior sensibilidade para nuances afetivas, naturalidade das interações e adequação contextual fina ao funcionamento emocional da criança. Essa distinção reforça a importância da complementaridade entre ambos os tipos de avaliação, ampliando a precisão e a validade interpretativa do estudo.

Assim, a combinação das análises humanas e automatizadas possibilitou uma avaliação integrada e rigorosa, indicando que as histórias sociais digitais geradas por IA cumprem os critérios pedagógicos e clínicos definidos, ao mesmo tempo em que reafirma a necessidade do julgamento profissional humano para a validação ética e emocional do material.

Dados gerais

As histórias sociais foram geradas na ferramenta Gemini – Storybook, a partir dos prompts padronizados elaborados para esta pesquisa. Cada história resultou em uma versão digital ilustrada, composta por capa e 12 páginas, com frases curtas em primeira pessoa, conforme os princípios de elaboração de histórias sociais descritos por Gray (2010).

Durante a etapa de visualização, observou-se que o layout das histórias varia conforme o dispositivo utilizado (computador, tablet ou celular). Em alguns aparelhos, há diferenças na proporção entre texto e imagem, bem como na forma de navegação entre as páginas, o que pode alterar a experiência de leitura, conforme ilustrado na Figura 4. Apesar disso, o conteúdo textual e a estrutura narrativa permaneceram consistentes entre as versões.

Também se identificou que, embora todas as histórias apresentem clareza textual e coerência narrativa, há aspectos a considerar quanto à adequação à faixa etária. Acredita-se que a história destinada à criança de 3 anos poderia ser mais breve e visualmente lúdica, com elementos mais interativos e linguagem ainda mais concreta. Já a história voltada para a criança de 9 anos poderia ser mais elaborada, incluindo reflexões mais complexas ou estratégias complementares de regulação emocional.

Outro ponto observado é que, embora eficazes e coerentes com a abordagem da Análise do Comportamento Aplicada (ABA), todas as histórias focaram exclusivamente na técnica de respiração como estratégia de autorregulação emocional. Em aplicações futuras, recomenda-

se explorar outras estratégias comportamentais, como pausas programadas, treino de solução de problemas, uso de cartões visuais ou reforços positivos específicos, de modo a diversificar o repertório de enfrentamento ensinado à criança.

Figura 4

Visualização das histórias no Gemini Storybook em dispositivo móvel (celular), evidenciando variações no layout, na proporção entre texto e imagem e na navegação entre páginas.

Meu corpo fica mole. Meu peito parece pesado e sinto um bolo na garganta. Eu reconheço esta sensação: é tristeza.



História de Ana — reconhecimento da emoção

Nota. Página da história social de Ana. Recursos visuais — uma nuvem escura sobre o peito e um ponto azul na garganta — representam as sensações corporais descritas na frase “Meu corpo fica mole... sinto um bolo na garganta... é tristeza.”.

3 DISCUSSÃO

Resumo

Este estudo avaliou três histórias sociais digitais geradas por IA para crianças de 3, 6 e 9 anos, utilizando uma rubrica baseada em Gray (2010) e princípios da ABA. As histórias, criadas no Storybook (Gemini), foram analisadas pela pesquisadora e pelo modelo ChatGPT. Os resultados mostraram alta qualidade geral, com maior necessidade de simplificação para a história de 3 anos e ampliação de estratégias de regulação para a de 9 anos. A concordância entre avaliações humanas e automatizadas foi elevada, indicando que a IA pode apoiar a criação de histórias sociais, desde que haja supervisão profissional para garantir adequação clínica e educacional.

Comparação com a Literatura

As histórias sociais têm se consolidado como uma estratégia eficaz para o desenvolvimento de habilidades sociais em crianças com Transtorno do Espectro Autista (TEA). Os resultados deste estudo dialogam diretamente com a literatura existente sobre histórias sociais, intervenções digitais e o uso de inteligência artificial nesse contexto. Em primeiro lugar, as histórias sociais geradas por IA mostraram-se claras, estruturadas e alinhadas ao formato proposto por Gray (2010), corroborando evidências que destacam a importância da previsibilidade, da linguagem simples e do foco em habilidades específicas. A presença consistente de elementos essenciais — como descrição da situação, identificação de pensamentos ou sentimentos, indicação do comportamento esperado e apresentação de consequências positivas — também converge com estudos aplicados, como os de Scattone et al. (2006) e Crozier e Tincani (2005), que demonstram que histórias sociais bem estruturadas podem reduzir comportamentos desadaptativos e favorecer o desenvolvimento de habilidades socioemocionais (Bondioli et al., 2024).

No campo das intervenções digitais, os achados deste estudo ampliam as observações de Stathopoulou et al. (2020), que evidenciaram o potencial de histórias sociais digitais no apoio a crianças com TEA, especialmente quando associadas a estímulos visuais claros e coerentes com o conteúdo textual. Além disso, estudos como o de Ghanouni et al. (2019) demonstram que histórias sociais podem ser integradas a ambientes de realidade virtual, possibilitando a validação de conteúdos voltados ao ensino de habilidades sociais e emocionais em diferentes contextos, como casa, escola e comunidade. Esses achados reforçam que o uso de tecnologias imersivas pode ampliar o engajamento, favorecer a compreensão de situações sociais e contribuir para o desenvolvimento da tomada de perspectiva por crianças com TEA. Embora a ferramenta utilizada tenha apresentado variações de layout entre dispositivos — limitação também observada em pesquisas sobre aplicativos educacionais —, o conteúdo produzido manteve consistência visual e qualidade estrutural, reforçando sua aplicabilidade em contextos clínicos e educacionais.

Além disso, a alta concordância entre a avaliação humana e a avaliação automatizada realizada por inteligência artificial nesta pesquisa está alinhada com pesquisas recentes sobre o uso de modelos generativos no contexto do autismo (McFayden et al., 2024; Kim et al., 2025). Esses estudos indicam que tais sistemas podem fornecer respostas tecnicamente adequadas, embora ainda apresentem limitações quanto à sensibilidade emocional e à contextualização das informações — padrão também identificado neste estudo. Dessa forma, os achados reforçam o potencial da IA como ferramenta complementar na elaboração de materiais educativos, ao mesmo tempo em que evidenciam a necessidade de supervisão profissional para garantir precisão clínica, adequação ética e sensibilidade às necessidades individuais, conforme discutido por Sampaio et al. (2024).

Por fim, a análise das histórias produzidas indica que, embora a inteligência artificial seja capaz de gerar narrativas coerentes e alinhadas aos princípios da Análise do Comportamento Aplicada (ABA), são necessários ajustes conforme a faixa etária e o nível de desenvolvimento da criança. Para crianças mais novas, destaca-se a necessidade de maior concretude e suporte visual; já para crianças mais velhas, a inclusão de estratégias mais complexas de autorregulação emocional pode ampliar a eficácia das intervenções, uma vez que essa habilidade está diretamente relacionada à capacidade de reconhecer, compreender e gerenciar emoções, impactando o comportamento social e o bem-estar psicológico (Iaochite et al., 2022). Essas observações dialogam com as recomendações de Aguiar e Moreira (2025) e Moreira e Medeiros (2019), que enfatizam a importância do desenvolvimento progressivo de repertórios emocionais. Assim, este estudo não apenas corrobora evidências já estabelecidas, mas também contribui para o avanço das práticas baseadas em evidências ao demonstrar as possibilidades e limitações do uso de modelos generativos na construção de histórias sociais digitais no campo do autismo.

Contribuições teóricas

O presente estudo oferece contribuições teóricas relevantes para três campos principais: (1) histórias sociais como tecnologia comportamental, (2) integração entre ABA e ferramentas digitais, e (3) uso de modelos generativos em intervenções educacionais e clínicas.

Primeiramente, os achados reforçam o modelo clássico de Gray (2010) ao demonstrar que a estrutura fundamental das histórias sociais pode ser preservada mesmo quando produzida por ferramentas de IA generativa. Isso amplia a compreensão teórica sobre a natureza essencial dessa tecnologia comportamental, indicando que seus elementos nucleares — previsibilidade, clareza textual, descrição explícita de comportamentos e consequências sociais — são passíveis de replicação eficiente em ambientes digitais automatizados.

Em segundo lugar, ao comparar histórias voltadas a diferentes faixas etárias, o estudo contribui para o aprofundamento teórico sobre ajuste de complexidade narrativa. Os resultados sugerem que histórias sociais eficazes exigem um equilíbrio entre simplicidade estrutural e relevância contextual, variando conforme o nível de desenvolvimento da criança. Tal achado complementa discussões de Moreira e Medeiros (2019) sobre a necessidade de adaptação das contingências às capacidades do aprendiz e se integra às formulações de Aguiar e Moreira (2025) sobre progressão no ensino de autorregulação emocional.

No campo emergente das tecnologias digitais aplicadas à ABA, os resultados expandem teorias iniciadas por Stathopoulou et al. (2020) ao demonstrar que histórias sociais produzidas por IA podem manter fidelidade conceitual às práticas baseadas em evidências. Ao mesmo tempo, o estudo aponta que ajustes humanos continuam necessários para garantir adequação afetiva, sensibilidade cultural e precisão contextual — elementos frequentemente negligenciados por sistemas automatizados. Assim, o trabalho contribui para um arcabouço teórico que integra princípios analítico-comportamentais com diretrizes para o uso ético e seguro de ferramentas digitais na intervenção com crianças autistas.

Por fim, ao comparar avaliações humanas e automatizadas, o estudo oferece uma contribuição teórica importante para o debate contemporâneo sobre IA em psicologia. Os resultados sugerem que modelos generativos são capazes de realizar avaliações estruturais consistentes, mas apresentam limitações na interpretação emocional, na análise funcional do comportamento e na leitura de nuances individuais. Esses achados ampliam reflexões de McFayden et al. (2024) e Kim et al. (2025), indicando que a IA pode atuar como coadjuvante técnico, e não como substituto da análise clínica humana. Essa distinção teórica fortalece o entendimento de que práticas de intervenção devem permanecer ancoradas no julgamento profissional, utilizando a IA como ferramenta complementar para aumentar precisão, acessibilidade e escala.

Contribuições metodológicas

Este estudo oferece contribuições metodológicas relevantes ao explorar o uso de IA generativa para produzir histórias sociais no formato storybook digital. A principal contribuição consiste na padronização da criação de histórias divididas em páginas, com capa, texto curto e imagens correspondentes, permitindo avaliar como esse formato favorece a compreensão sequencial e reduz a carga cognitiva — aspecto fundamental para crianças pequenas.

O estudo também avaliou sistematicamente o volume textual por página, garantindo frases curtas e diretas, alinhadas às recomendações de Gray (2010) e às necessidades de histórias sociais. Da mesma forma, a rubrica permitiu examinar a correspondência entre texto e imagem, identificando se as ilustrações realmente apoiavam a interpretação emocional e comportamental da narrativa.

Além disso, a pesquisa testou a capacidade do Storybook de personalizar personagens e situações, verificando se o modelo gerava cenas e imagens coerentes com o perfil solicitado. Isso oferece um método aplicável para futuras produções individualizadas.

Por fim, o estudo revelou um padrão metodológico importante: as três histórias, apesar de contextos emocionais distintos, tenderam a repetir a mesma estratégia de regulação (respiração). Essa recorrência evidencia a necessidade de instruções mais específicas ao modelo para ampliar o repertório de técnicas ensinadas.

Assim, o estudo contribui ao propor um método claro e replicável para criação e avaliação de storybooks digitais, permitindo investigar simultaneamente texto, imagens, estrutura de páginas e adequação clínica em materiais gerados por IA.

Contribuições Aplicadas

As histórias sociais produzidas em formato de storybook digital oferecem contribuições práticas importantes para intervenções com crianças autistas. A divisão em páginas curtas, linguagem simples e estrutura previsível facilita o ensino de identificação emocional e de comportamentos-alvo, permitindo que terapeutas, professores e familiares utilizem o material diretamente em atendimentos ou rotinas escolares.

O formato digital favorece o engajamento e possibilita adaptações rápidas, como inserção de pictogramas ou fotos reais, ampliando sua utilidade para diferentes perfis de crianças. Além disso, as narrativas apresentam situações cotidianas (tristeza, frustração, raiva) que podem ser facilmente generalizadas, funcionando como apoio concreto no treino de autorregulação.

O estudo também demonstra que ferramentas de IA podem agilizar a criação de histórias sociais personalizadas, mantendo clareza textual e alinhamento com princípios da ABA. Assim, os storybooks digitais gerados ocupam um papel relevante como instrumentos de mediação emocional e suporte clínico-educacional.

Limitações e pesquisas futuras

Este estudo apresenta algumas limitações importantes. A análise concentrou-se em três histórias sociais em formato de storybook digital, o que restringe a generalização dos achados. Embora as narrativas tenham sido avaliadas quanto à clareza, estrutura e alinhamento com a ABA, não houve coleta de dados empíricos sobre eficácia real com crianças autistas, como mudanças comportamentais ou generalização das habilidades ensinadas. Além disso, a avaliação das imagens foi limitada por variações de layout entre dispositivos, o que pode impactar significativamente a compreensão de crianças pequenas.

Outra limitação envolve o uso de modelos generativos (ChatGPT e Gemini) na avaliação: apesar de apresentarem alta concordância com o avaliador humano, ainda não se sabe em que medida tais modelos captam nuances emocionais, culturais e idiossincráticas essenciais na intervenção com crianças autistas.

Pesquisas futuras podem ampliar a quantidade e diversidade das histórias, testar versões personalizadas para diferentes níveis de suporte, e incluir metodologias experimentais que avaliem diretamente a efetividade no comportamento das crianças. Estudos comparando storybook digital com versões impressas, bem como investigações sobre adaptação automática por IA (ex.: ajustes de vocabulário, imagens e suportes visuais), também representam caminhos promissores. Por fim, recomenda-se explorar como crianças autistas de diferentes idades interagem com o formato digital e quais elementos visuais ou textuais favorecem maior engajamento e aprendizagem.

4 CONCLUSÃO

O presente estudo investigou a qualidade, adequação e potencial de aplicação clínica de histórias sociais digitais produzidas com apoio de inteligência artificial, no formato de storybook, destinadas a crianças autistas de diferentes faixas etárias. A análise, conduzida por meio de uma rubrica estruturada em nove critérios, permitiu avaliar de forma sistemática aspectos textuais, visuais, pedagógicos e comportamentais das três narrativas (Lucas – 3 anos; Ana – 6 anos; Pedro – 9 anos). Os resultados demonstraram que as histórias apresentaram alta qualidade geral, especialmente no que se refere à clareza textual, alinhamento com princípios da ABA, relevância emocional e coerência narrativa.

A comparação entre a avaliação humana e a automatizada (ChatGPT e Gemini) mostrou elevada concordância, indicando que modelos generativos são capazes de aplicar critérios técnicos de maneira consistente. Entretanto, também ficou evidente que a análise humana permanece indispensável para captar nuances emocionais, contextuais e relacionais que extrapolam os critérios formais – especialmente quando o público-alvo envolve crianças autistas, cuja compreensão depende fortemente de sensibilidade individual, adequação sensorial e suporte visual.

As análises também apontaram pontos de aprimoramento importantes, incluindo a necessidade de maior diversificação das estratégias de autorregulação apresentadas nas histórias, além da adaptação de complexidade textual e visual conforme a idade. As diferenças observadas entre as histórias reforçam que o storybook digital pode ser uma ferramenta potente, desde que planejado com rigor técnico e ajustado às características do usuário.

De modo geral, os achados deste estudo evidenciam que histórias sociais digitais geradas ou avaliadas por IA podem compor recursos válidos, acessíveis e pedagogicamente estruturados para intervenção com crianças autistas. Contudo, a integração entre tecnologia e análise comportamental precisa continuar fundamentada no julgamento profissional humano, garantindo ética, segurança e efetividade clínica. Assim, o estudo contribui para o avanço do uso de ferramentas digitais na ABA e abre caminhos para pesquisas que testem empiricamente a eficácia desses materiais em contextos reais de aprendizagem e autorregulação emocional.

REFERÊNCIAS BIBLIOGRÁFICAS

Aguiar, S. H. A., & Moreira, M. (2025). Autismo: Estratégias de autorregulação. Instituto Walden4.

American Psychiatric Association. (2014). Diagnostic and statistical manual of mental disorders (5th ed.). Artmed.

Bondioli, R. M., Biazim, M. de A., Benitez, P., & Domeniconi, C. (2024). Uso de histórias para ensino de reconhecimento de expressões faciais de emoções para crianças com TEA. *Psicologia: Teoria e Prática*, 26(3), ePTPCP16608. <https://doi.org/10.5935/1980-6906/ePTPCP16608>

- Crozier, S., & Tincani, M. J. (2005). Using a modified social story to decrease disruptive behavior of a child with autism. *Focus on Autism and Other Developmental Disabilities*, 20(3), 150–157. <https://doi.org/10.1177/10883576050200030301>
- Ghanouni, P., Jarus, T., Zwicker, J. G., Lucyshyn, J., Mow, K., & Ledingham, A. (2019). Social stories for children with autism spectrum disorder: Validating the content of a virtual reality program. *Journal of Autism and Developmental Disorders*, 49(2), 660–668. <https://doi.org/10.1007/s10803-018-3737-0>
- Goulart, P. R. K., & Assis, G. J. A. (2002). O estudo metodológico do autismo na análise do comportamento. *Psicologia: Teoria e Pesquisa*, 18(2), 189–196. <https://doi.org/10.1590/S0102-37722002000200010>
- Gray, C. (2010). *The new social story book* (10th ed.). Future Horizons.
- Iaochite, R. T., Noronha, A. P. P., Casanova, D. C. G., Santos, A. A. A., & Azzi, R. G. (2022). Autoeficácia para regulação emocional e autoeficácia social: Busca de evidências de validade de construtos relacionados. *Psico-USF*, 27(1), 31–43. <https://doi.org/10.1590/1413-82712022270103>
- Kim, J., Kim, Y., Lee, S., & Lee, J. (2025). Testing the capability of generative artificial intelligence for parent and caregiver questions. *Family Relations*. Advance online publication. <https://doi.org/10.1111/fare.14091>
- McFayden, T. C., Whelan, D. C., Ferguson, L., & Dillenburger, K. (2024). ChatGPT and autism: Evaluating the viability of generative artificial intelligence for responding to frequently asked questions about autism. *Journal of Autism and Developmental Disorders*. Advance online publication.
- Moreira, M. B., & Medeiros, C. A. (2019). *Princípios básicos de análise do comportamento* (2ª ed.). Artmed.
- Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., ... Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: Reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432000>
- Scattone, D., Tingstrom, D. H., & Wilczynski, S. M. (2006). The effectiveness of social stories on decreasing disruptive behaviors of children with autism: Three case studies. *Journal of Autism and Developmental Disorders*, 36(4), 445–469. <https://doi.org/10.1007/s10803-006-0086-1>
- Stathopoulou, A., Loutradis, A., Drigas, A., & Economou, A. (2020). Evaluation of mobile apps effectiveness in children with autism social training via digital social stories. *International Journal of Emerging Technologies in Learning (iJET)*, 15(6), 95–107.

Anexos

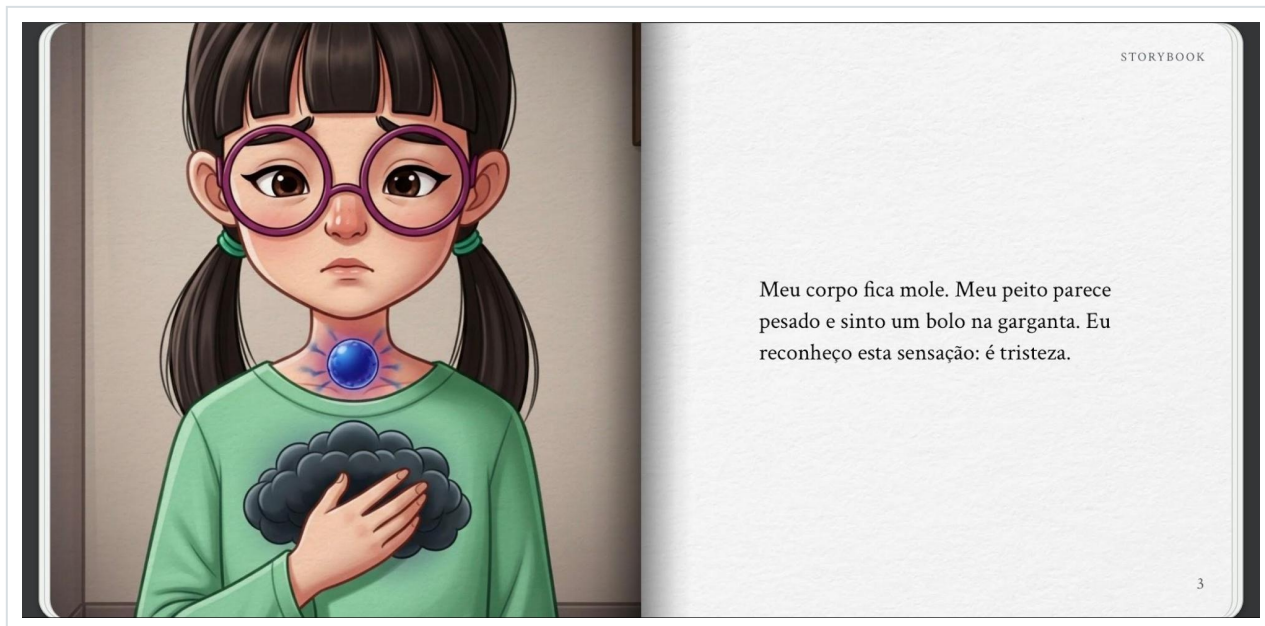
História Ana



História de Ana — apresentação da personagem e do gato Pipoca



História de Ana — a situação-problema



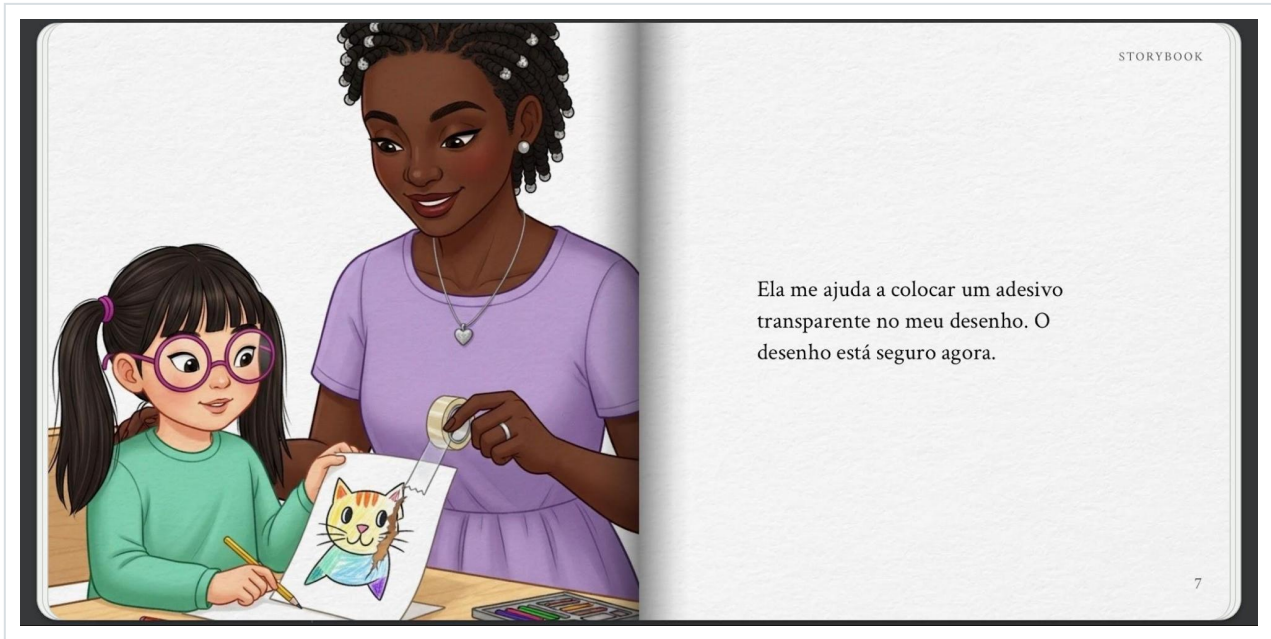
História de Ana — nomeando a tristeza (variação)





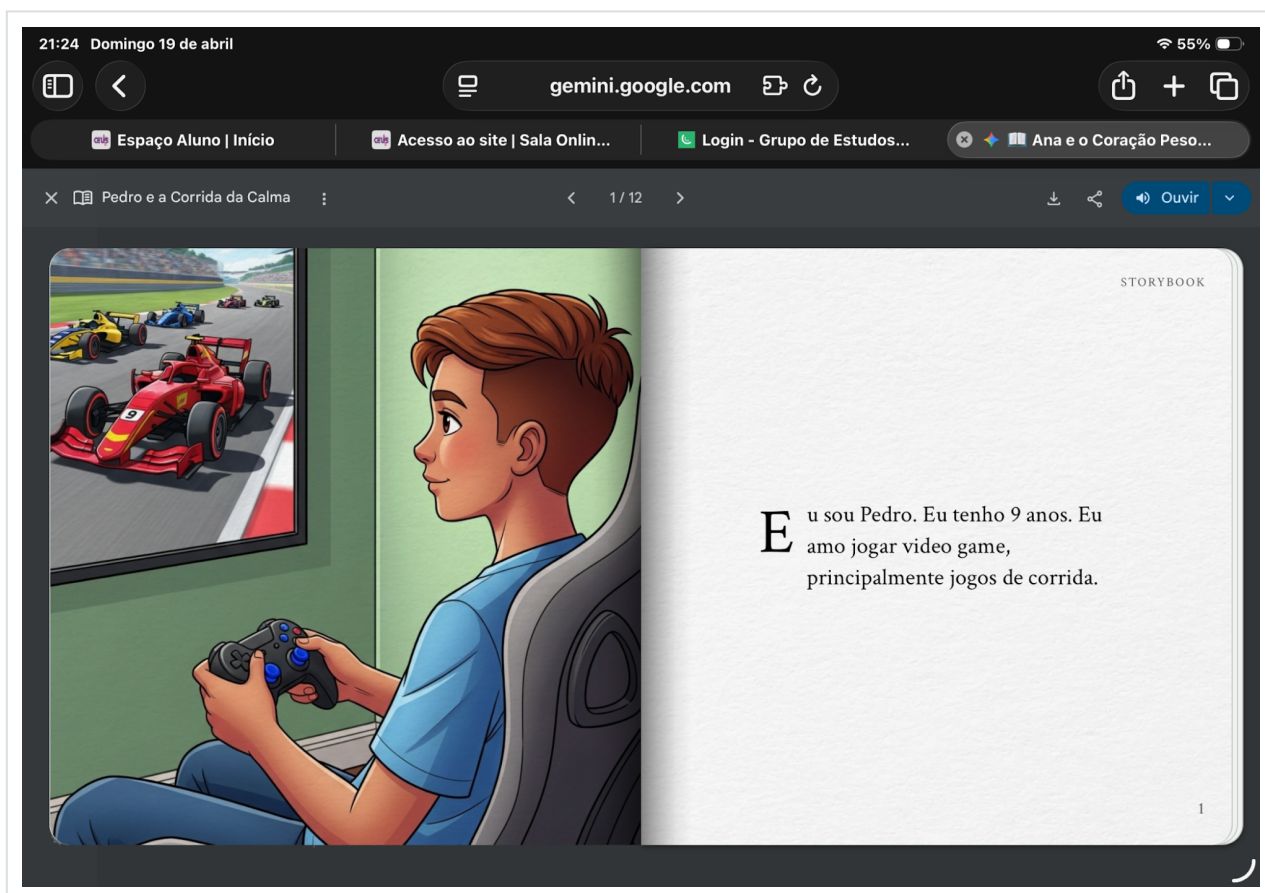
História de Ana — nomeando a tristeza (reapresentação)



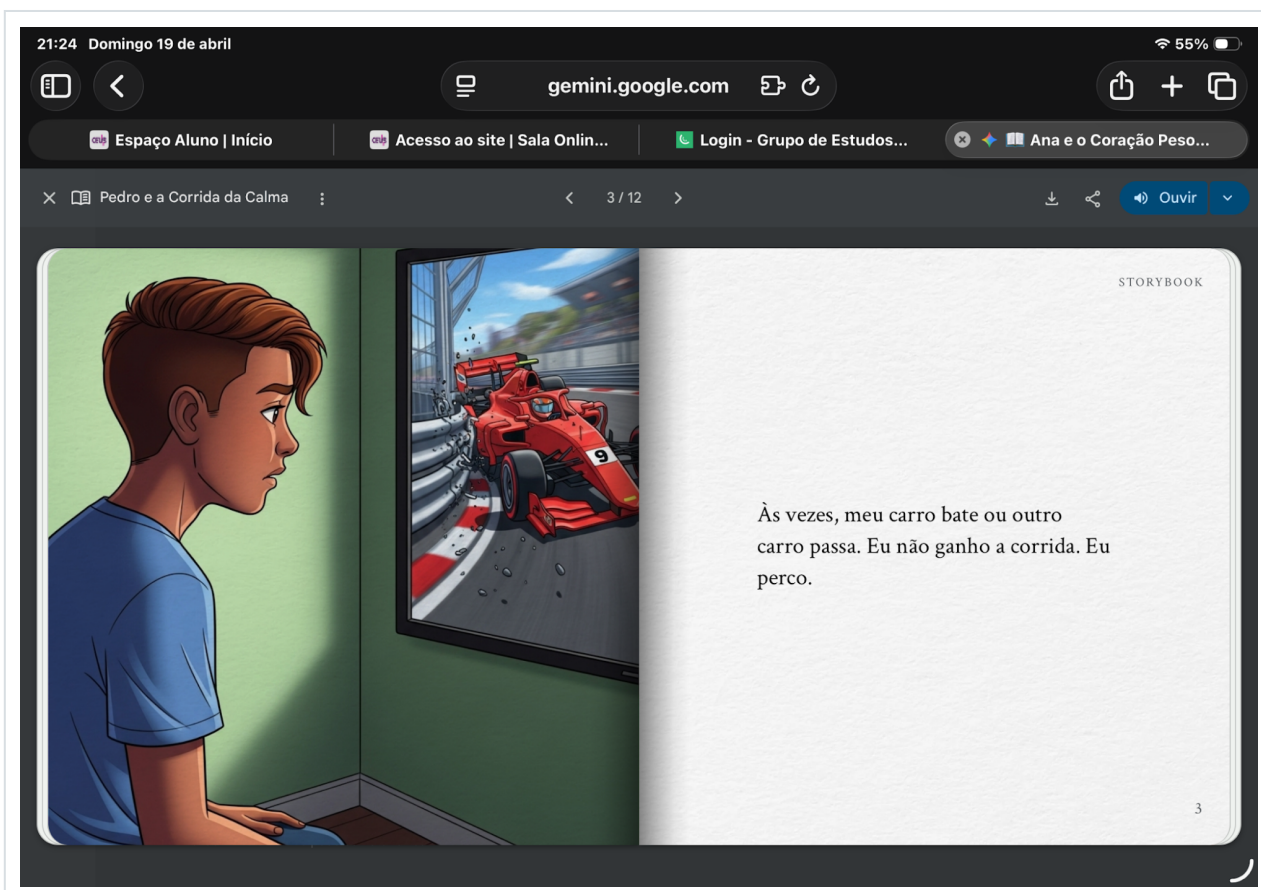
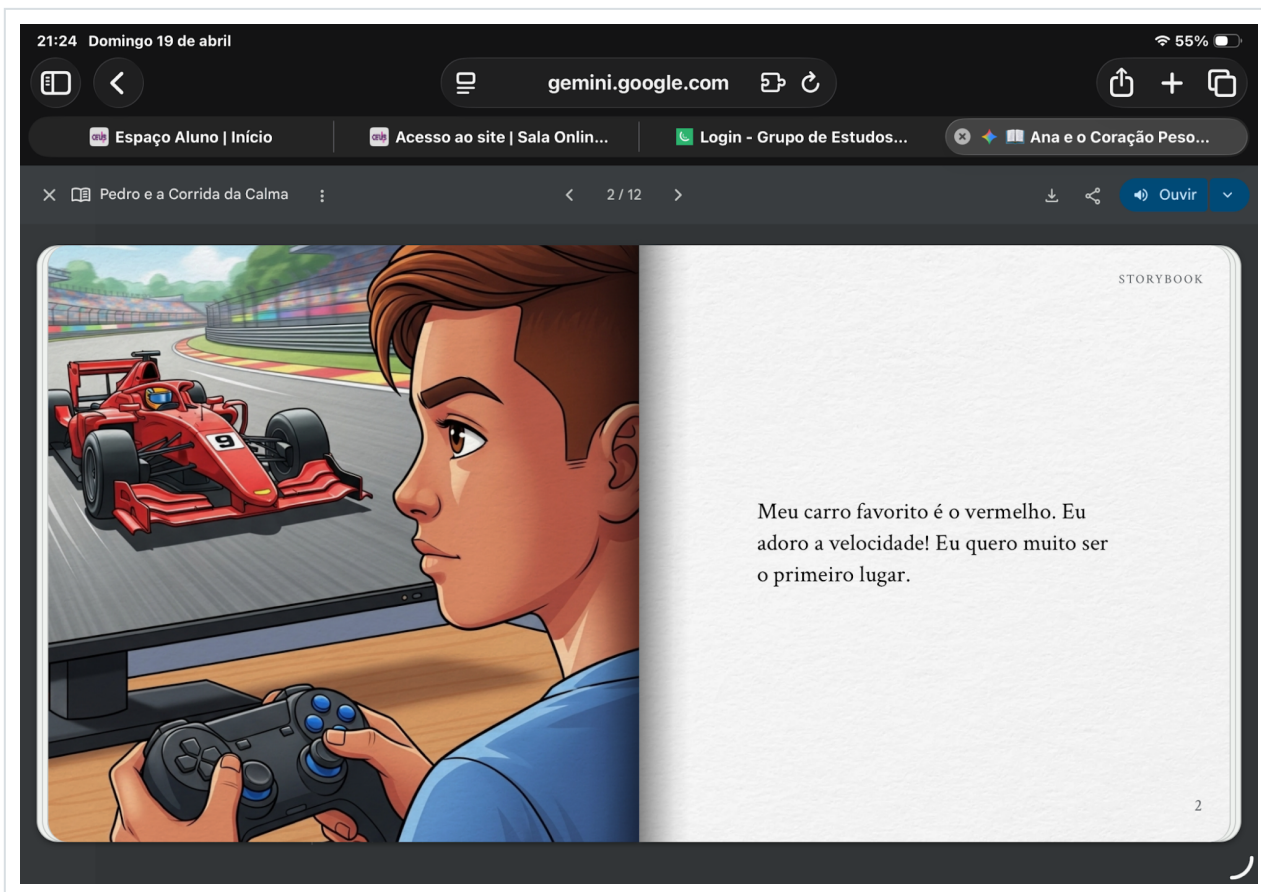


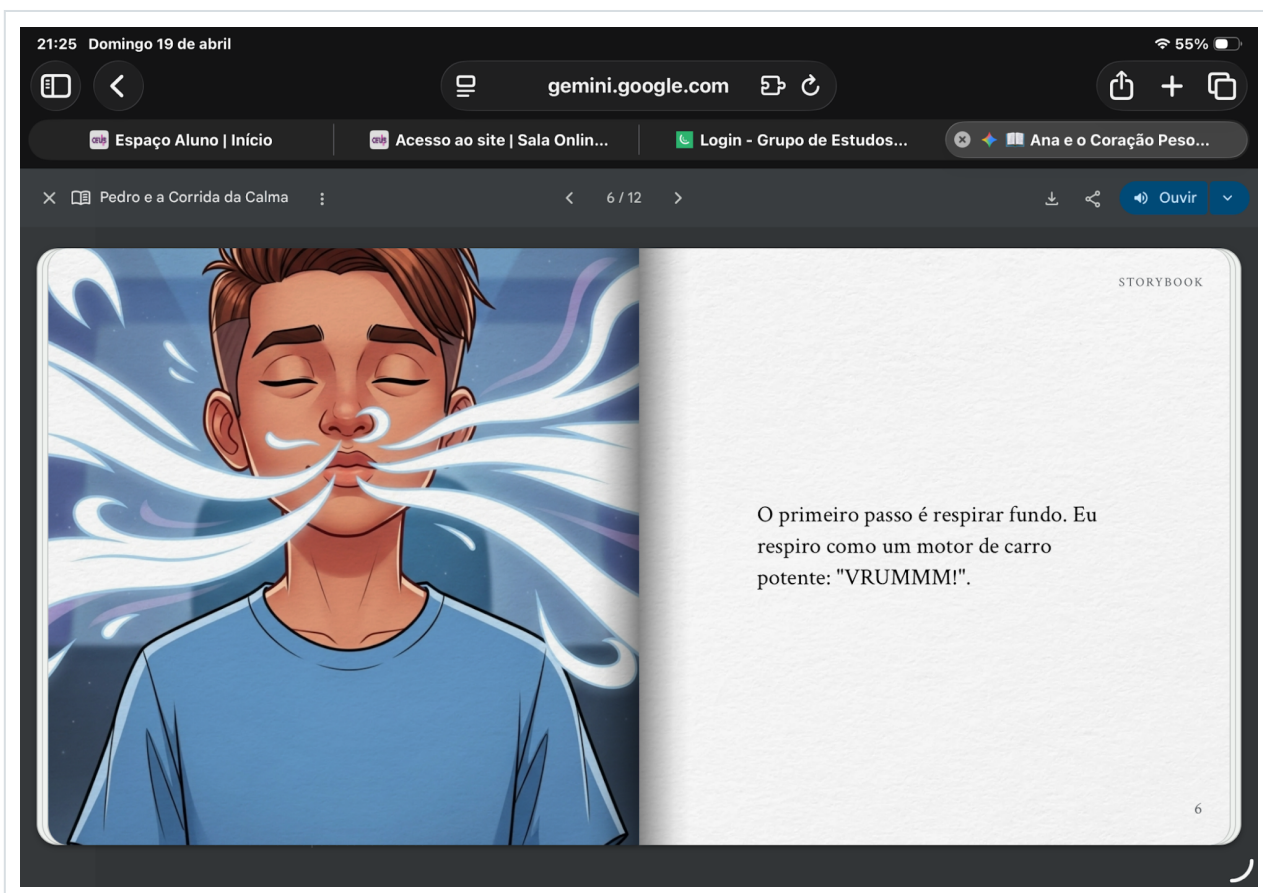
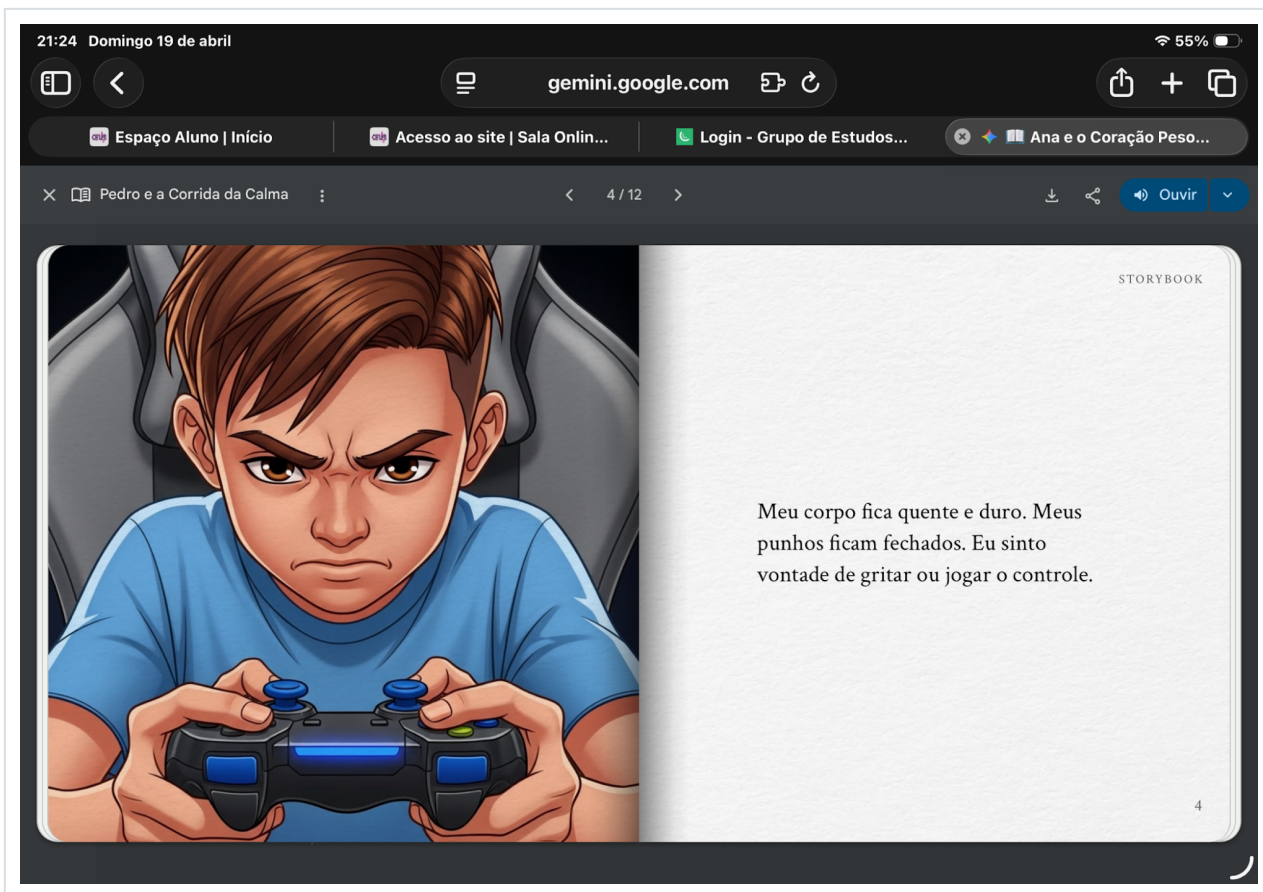


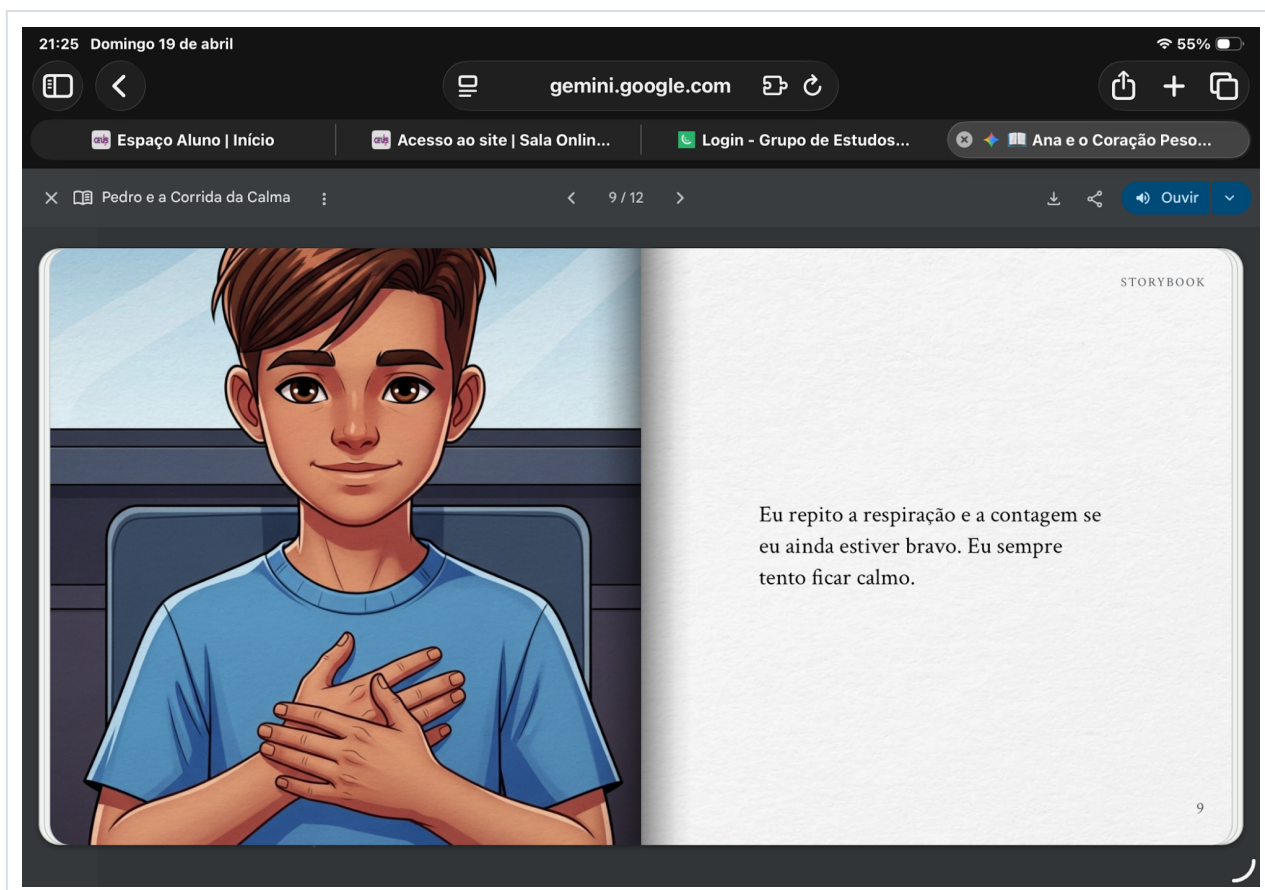
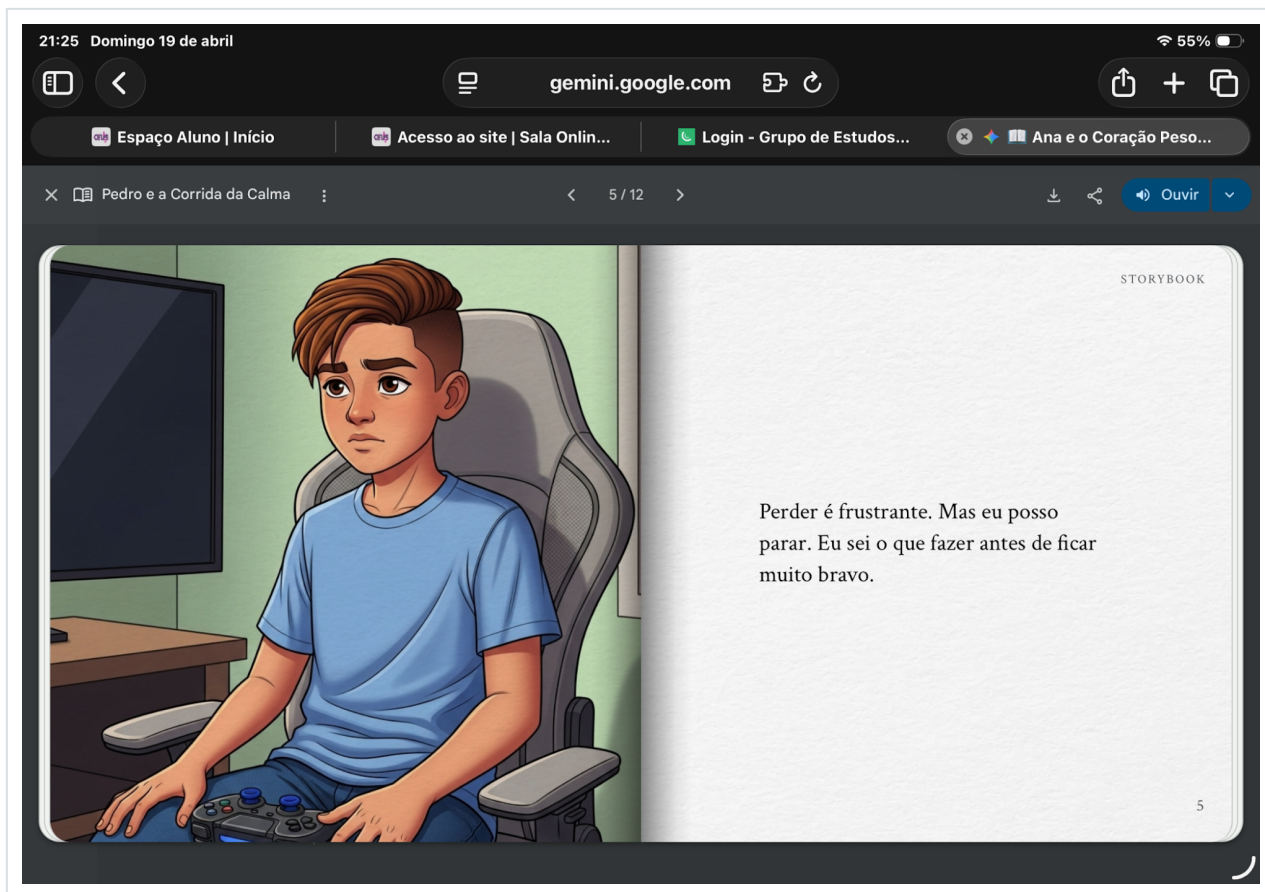
História Pedro

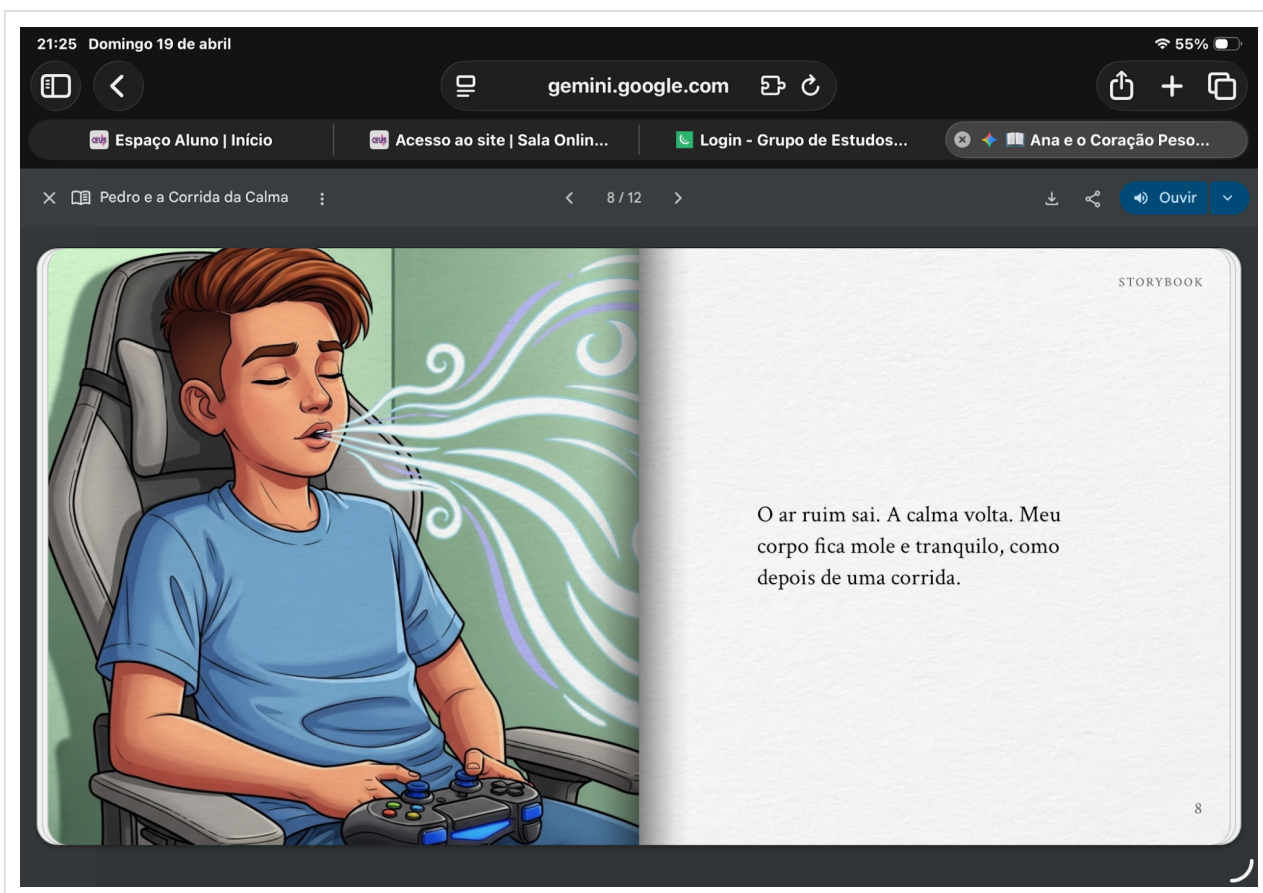
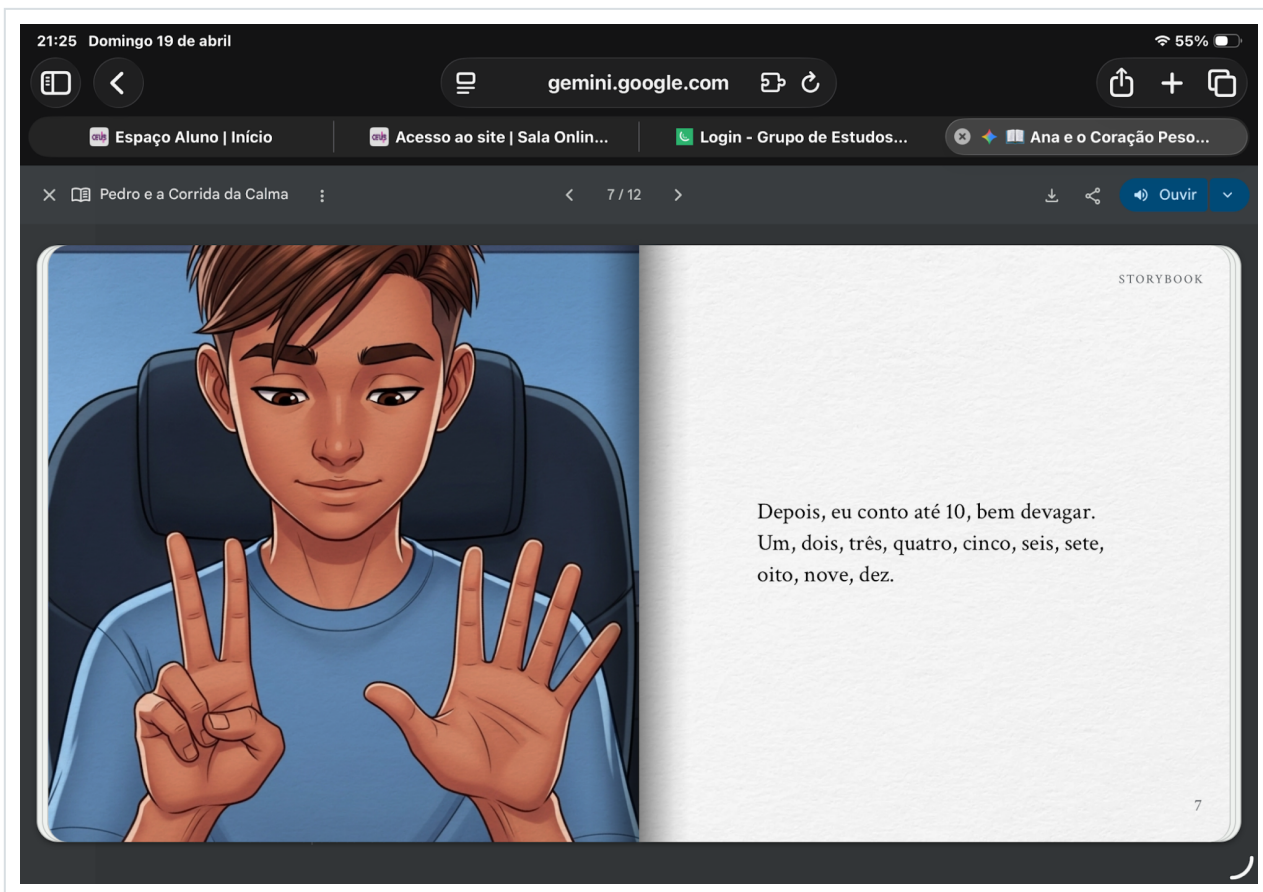


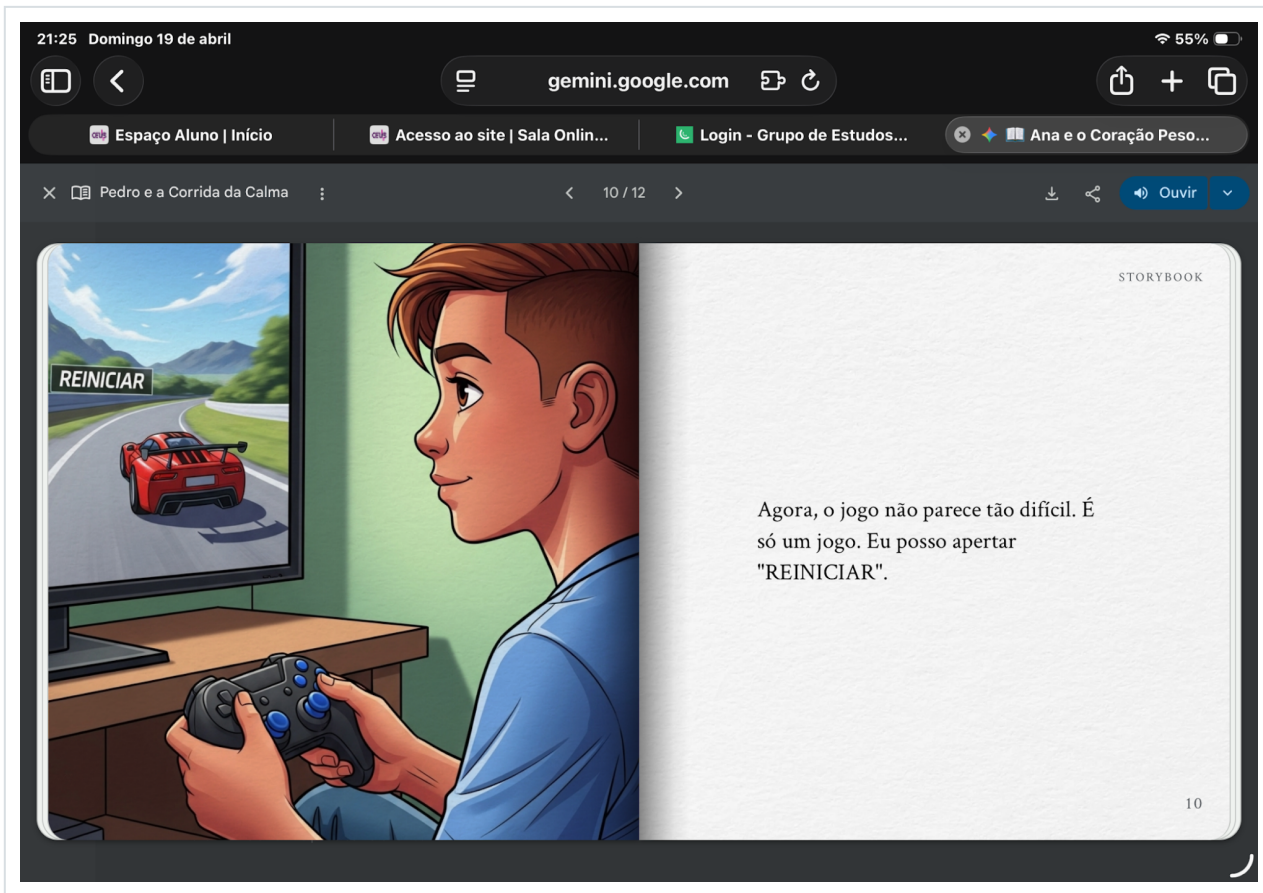
Visualização da história no GEM Storybook (Gemini) em navegador



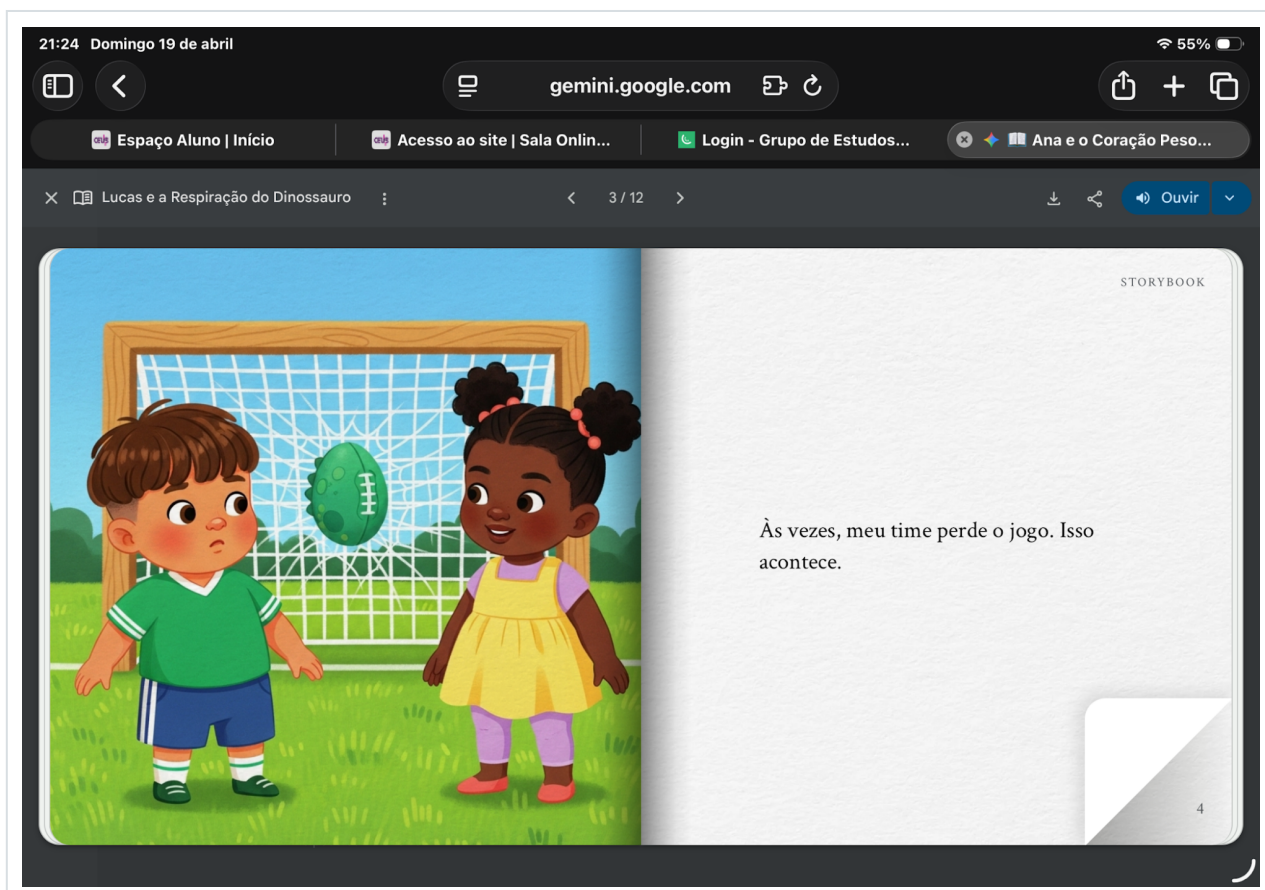


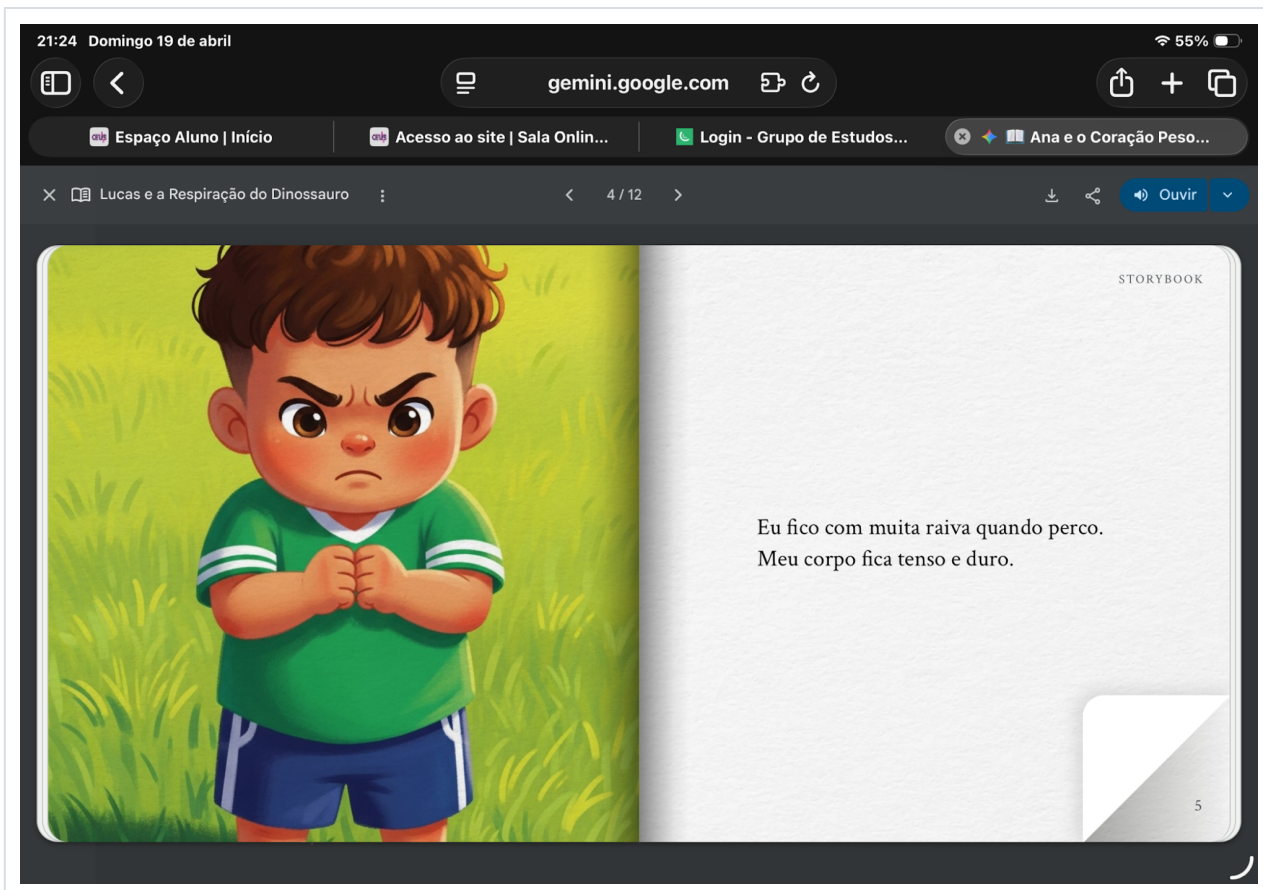
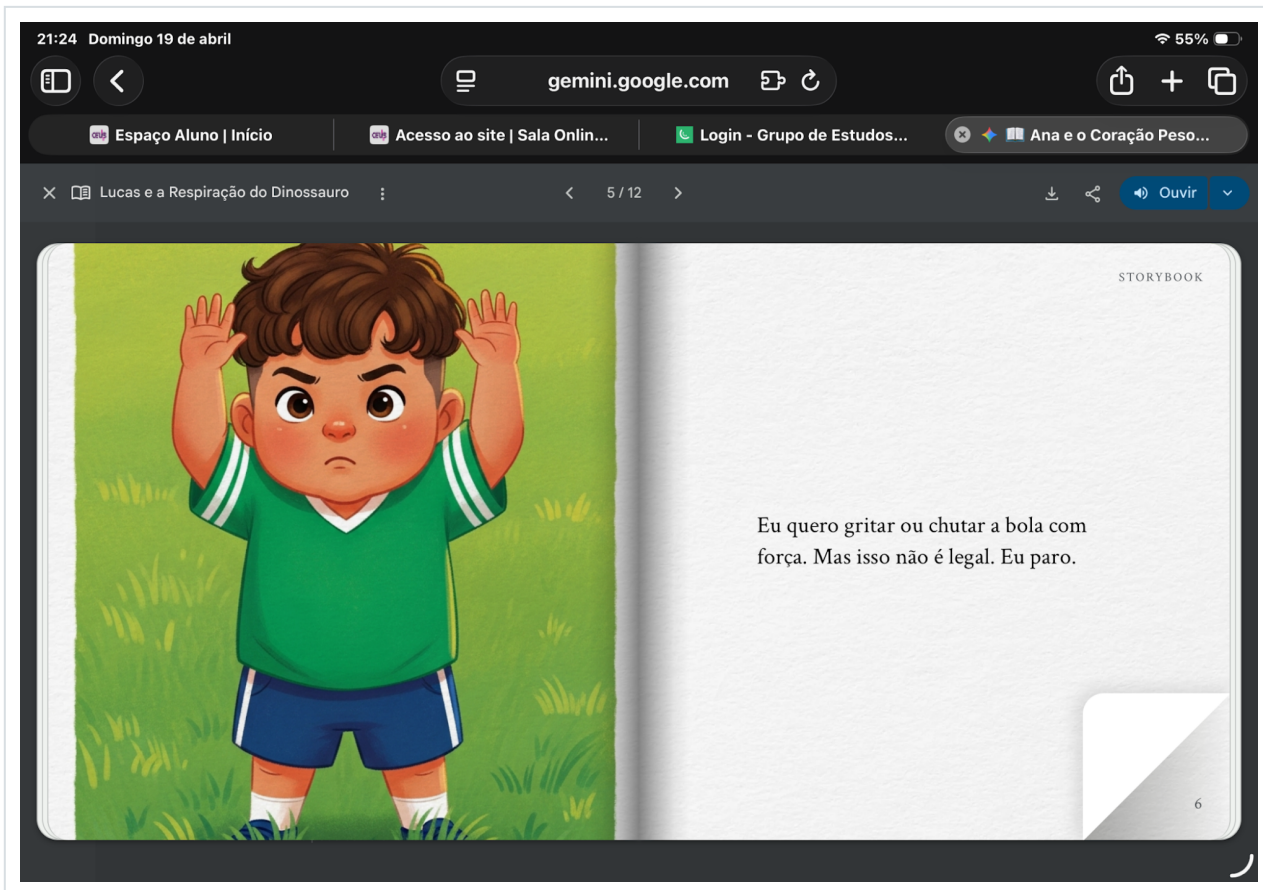


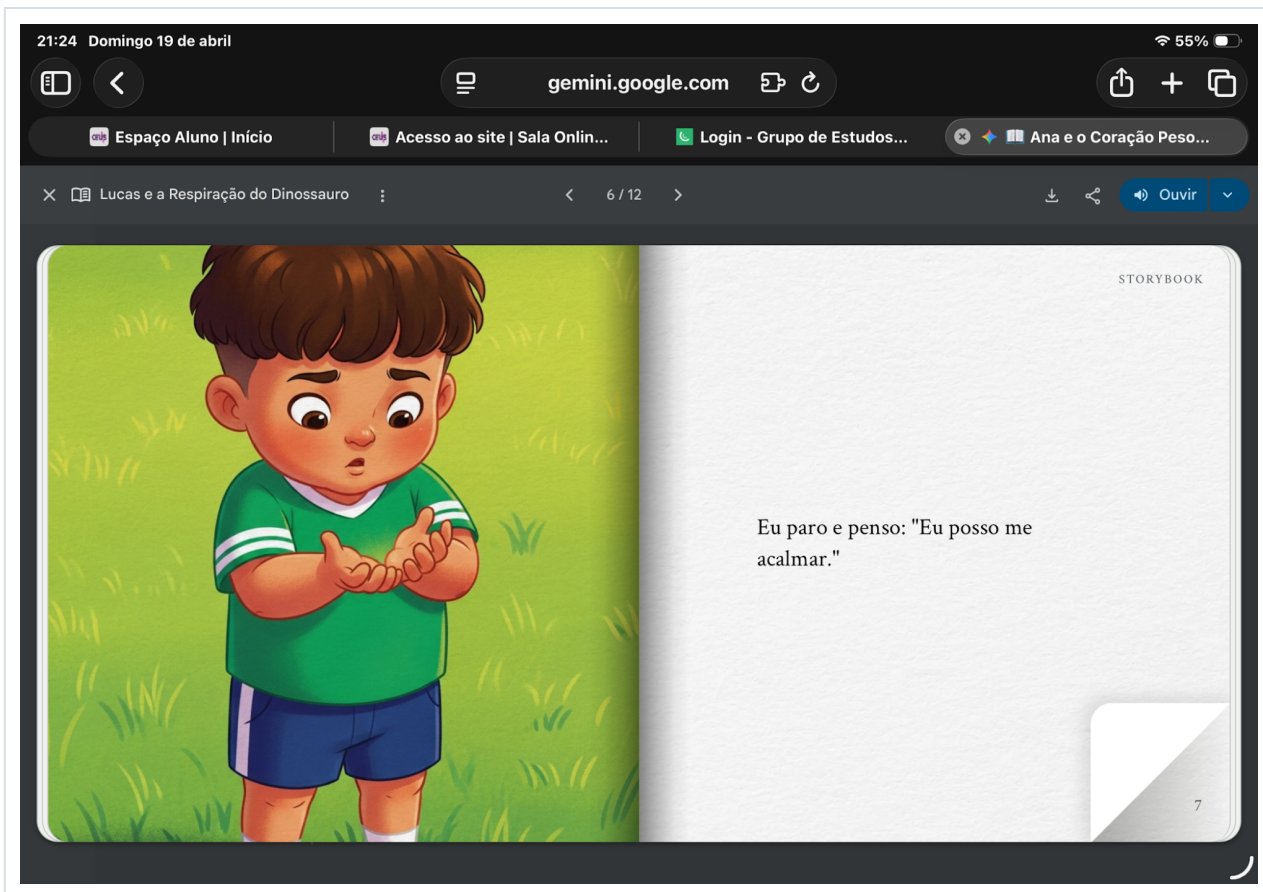
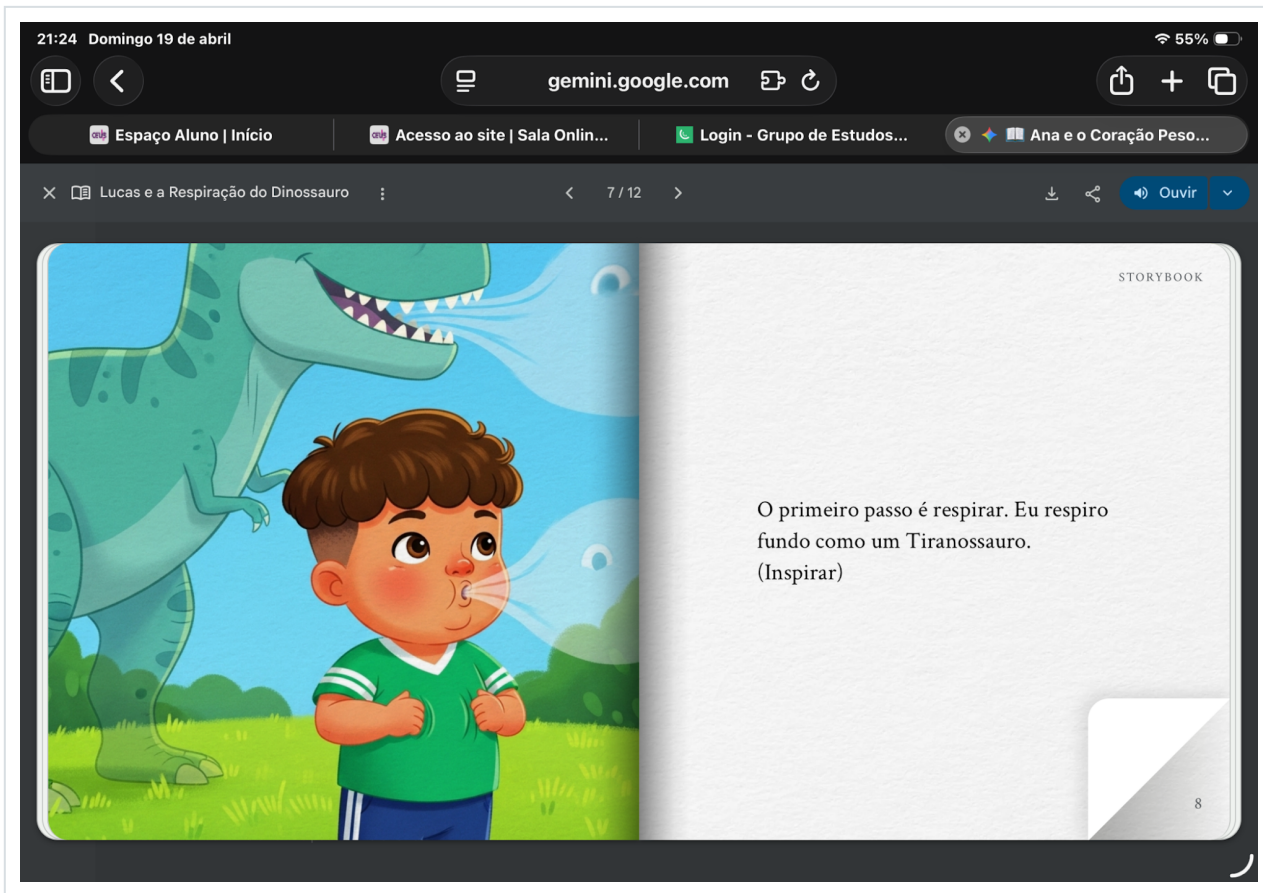


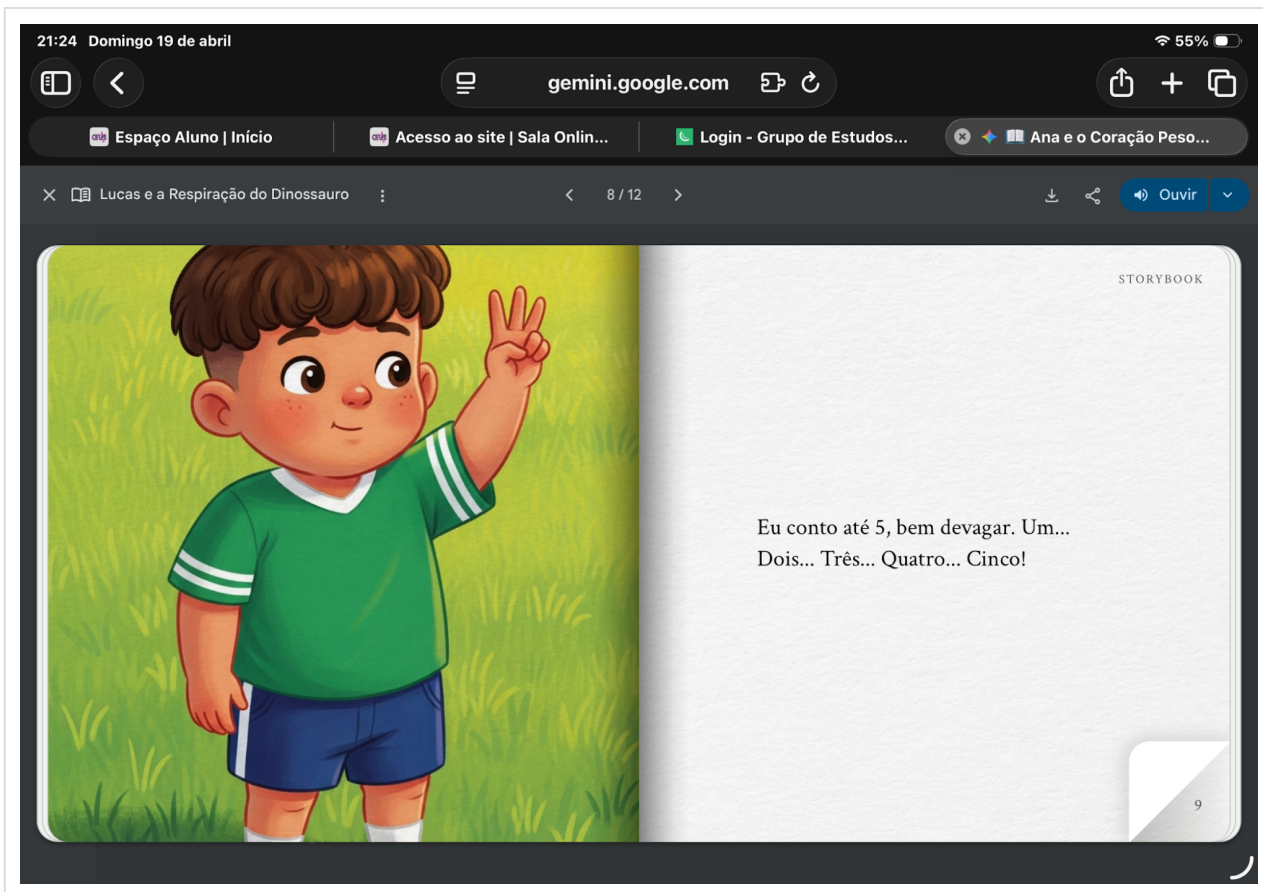
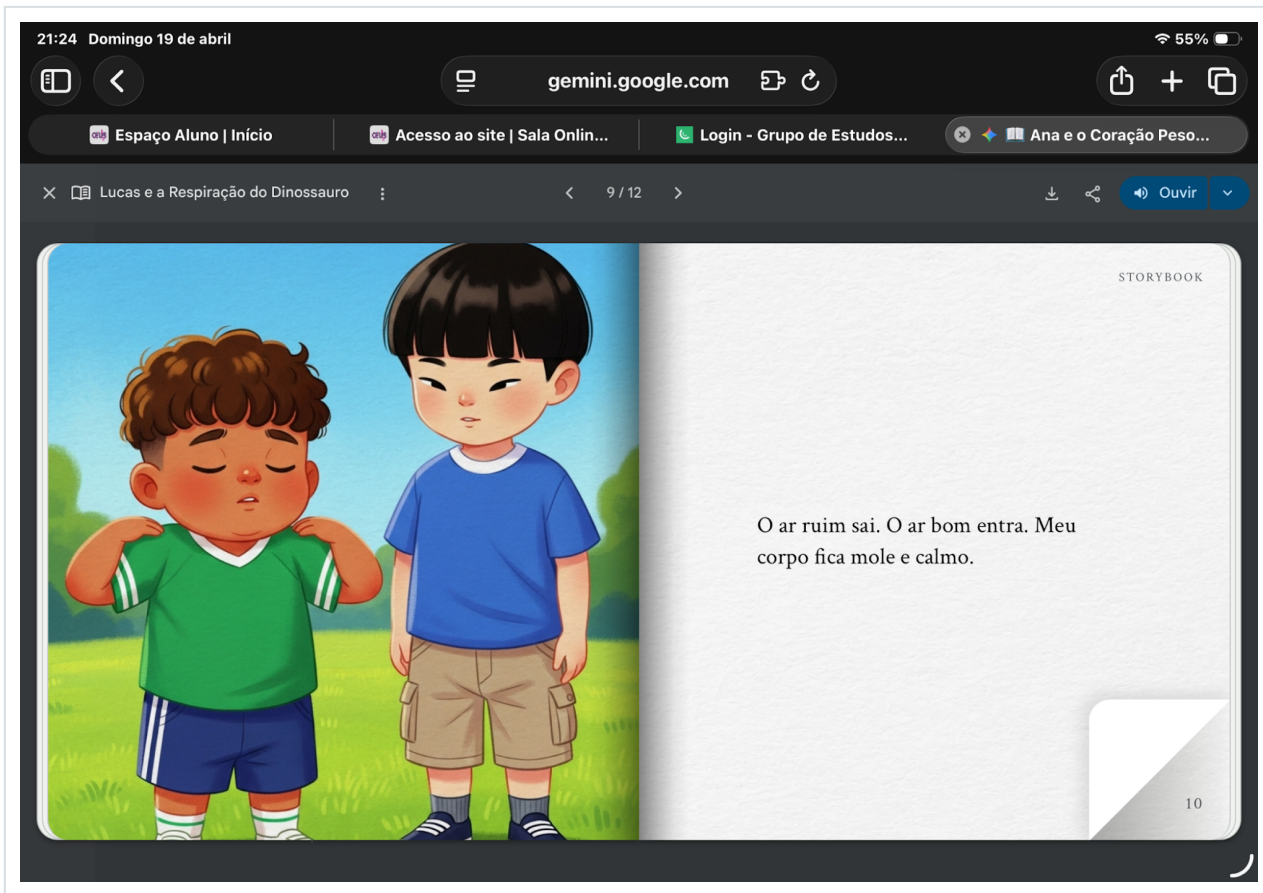


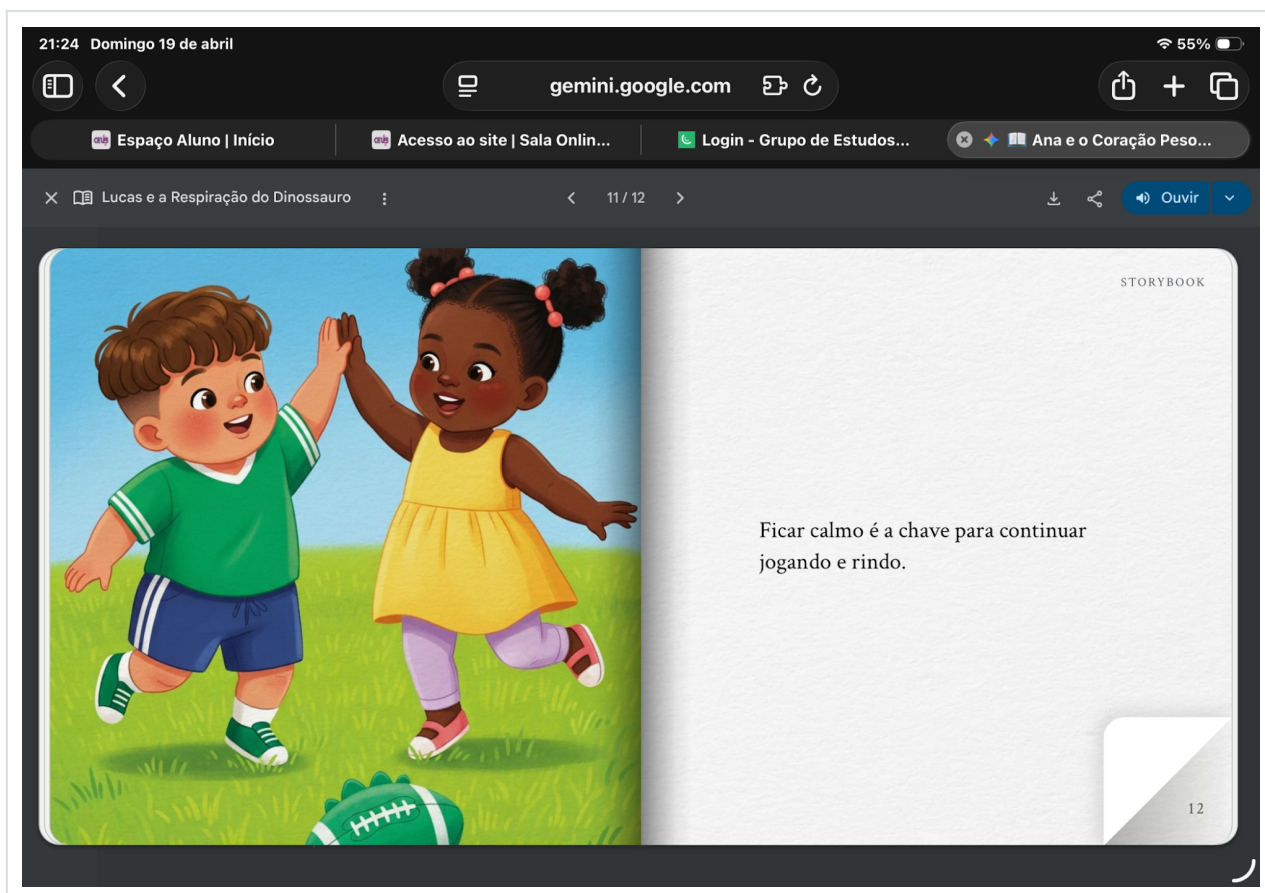
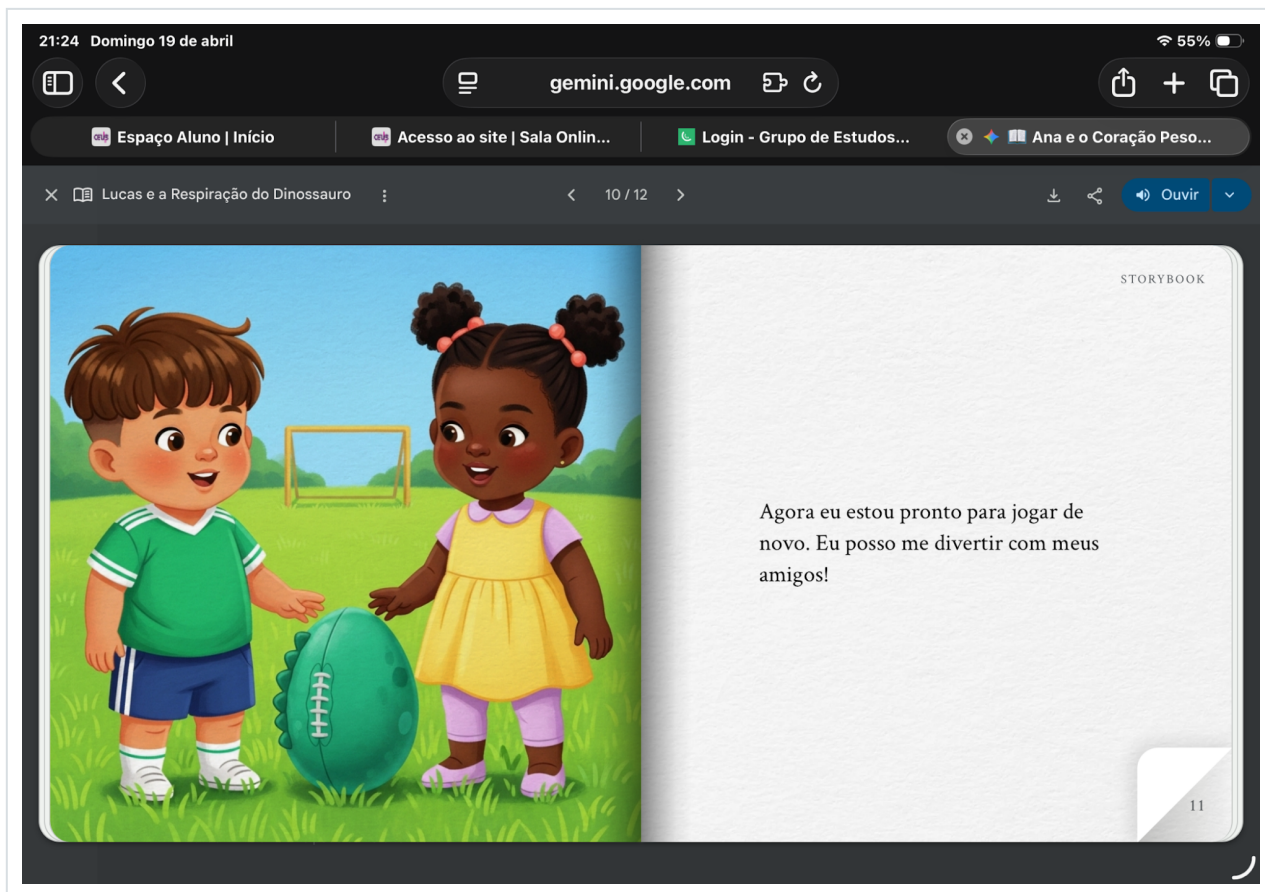
História Lucas

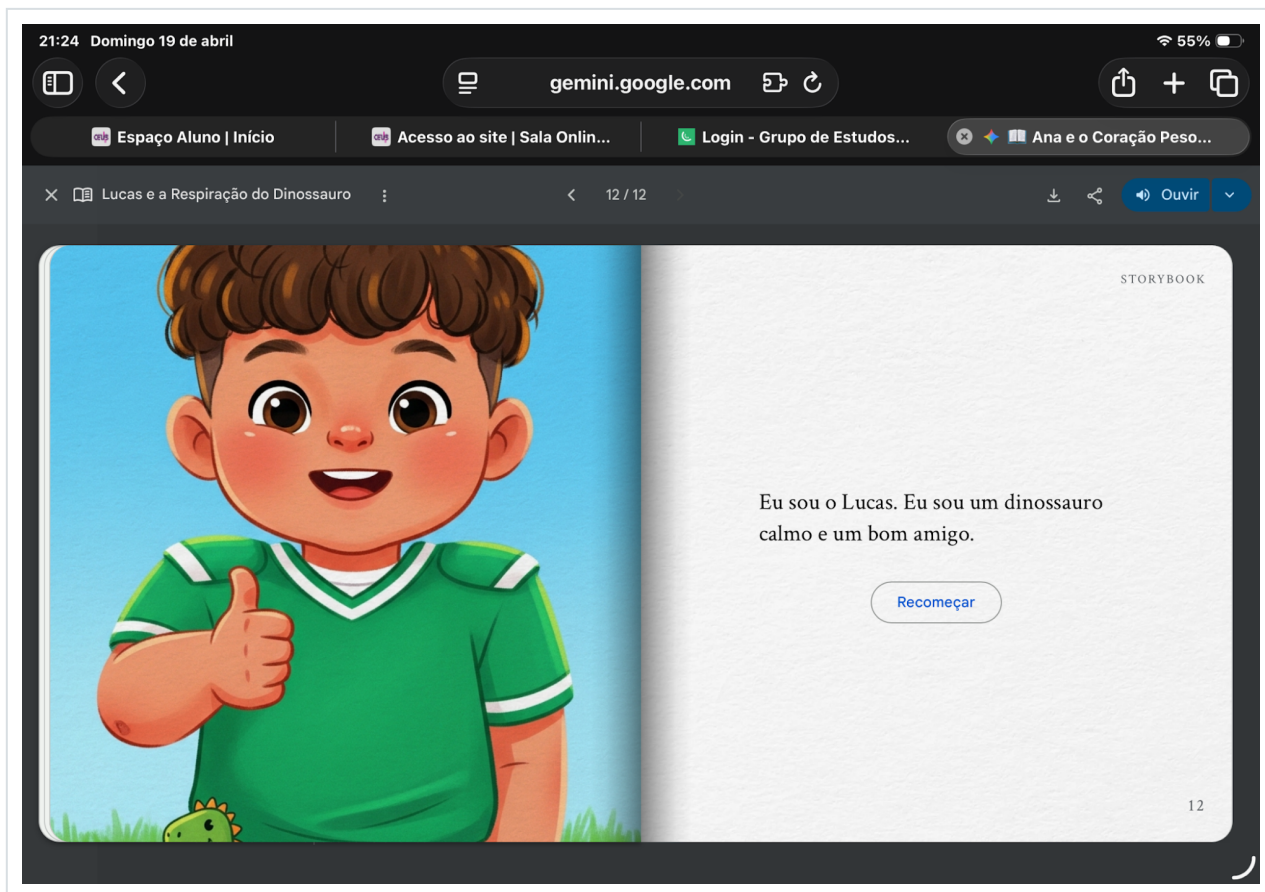












INSTITUTO WALDEN4 · RÁDIO IW4

Rádio iW4 — ciência e conhecimento ao seu alcance



Rádio 24 horas com leituras de artigos científicos, palestras, aulas e muito mais, dedicada à Análise do Comportamento e ciências afins.

- Sempre disponível · 24/7
- Conteúdo para desenvolvimento profissional
- No navegador e no app Android (Google Play)

walden4.com.br/radio

CAPÍTULO

04

Avaliação do Efeito do Nível de Prompt na Precisão e Utilidade de um aplicativo tutorial gerado por Chatbot de IA no Apoio a Pais de Crianças com TEA: Uso do DRA para Prevenção e Manejo de Comportamentos Disruptivos



Ana Victória Filomeno Gomes

Márcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. Investigar a precisão e a utilidade de um aplicativo de inteligência artificial (chatbot) desenvolvido para orientar pais de crianças com Transtorno do Espectro Autista (TEA) na prevenção e manejo de comportamentos agressivos e birras, utilizando como base o Reforçamento Diferencial de Comportamentos Alternativos (DRA).

Método. O estudo utilizou um delineamento experimental e descritivo, analisando as respostas do ChatGPT-5 a três prompts sobre o uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA) no manejo de comportamentos desafiadores em crianças com TEA. As produções foram avaliadas quanto à correteza, clareza e concisão, além da utilidade prática e adequação ao público-alvo. Também foram verificadas a qualidade e autenticidade das referências utilizadas.

Resultados. Os três aplicativos educativos sobre o Reforçamento Diferencial de Comportamentos Alternativos (DRA) foram gerados automaticamente pelo ChatGPT-5 em formato HTML, apresentando funcionamento estável, boa responsividade e clareza estrutural. Na avaliação do conteúdo, todos mostraram alta correteza científica, mas diferiram em clareza e concisão; o Aplicativo 2 obteve melhor equilíbrio (11/12 pontos), com linguagem mais acessível e exemplos práticos. Na avaliação geral, também se destacou (26/30 pontos), demonstrando maior usabilidade, pertinência ao público-alvo e aplicabilidade prática, embora ainda pudesse aperfeiçoar o layout visual. A verificação das referências confirmou a autenticidade e validade da maioria das fontes, sendo necessárias apenas pequenas correções em citações incompletas. De modo geral, os aplicativos apresentaram boa qualidade conceitual e técnica, e o Aplicativo 2 mostrou-se o mais adequado para apoiar pais de crianças com TEA.

Conclusão. O estudo concluiu que o ChatGPT-5 demonstrou capacidade de gerar materiais educativos precisos, claros e funcionalmente úteis para apoiar pais de crianças com TEA na aplicação do Reforçamento Diferencial de Comportamentos Alternativos (DRA). Entre os três aplicativos produzidos, o Aplicativo 2 apresentou o melhor equilíbrio entre rigor teórico e acessibilidade, mostrando-se o mais adequado para o público leigo, enquanto o Aplicativo 3 foi o mais completo, porém mais técnico. Esses resultados confirmam que a qualidade das respostas da IA depende diretamente da formulação do prompt, evidenciando o potencial da tecnologia para produzir materiais interativos e baseados em evidências, incluindo quizzes, roteiros e exemplos práticos aplicáveis ao cotidiano familiar. Conclui-se, portanto, que a IA pode atuar como ferramenta complementar valiosa no apoio a cuidadores, ampliando o acesso a informações científicas, promovendo aprendizagem parental estruturada e fortalecendo a prevenção de comportamentos desafiadores, desde que seu uso seja guiado por critérios técnicos e éticos bem definidos.

Palavras-chave: Inteligência artificial; Chatbot; Transtorno do Espectro Autista; Reforçamento Diferencial de Comportamentos Alternativos; Análise do Comportamento Aplicada.

1 INTRODUÇÃO

Contextualização do Tema

O Transtorno do Espectro Autista (TEA) caracteriza-se por prejuízos qualitativos na interação social, dificuldades na comunicação e padrões restritos e repetitivos de comportamento (American Psychiatric Association, 2022). Para além da definição diagnóstica, a literatura da Análise do Comportamento ressalta que abordagens centradas apenas na descrição topográfica dos comportamentos-problema podem ser limitadas, uma vez que ignoram os fatores motivacionais que os mantêm. Nessa perspectiva, o comportamento é melhor compreendido quando analisado em termos de sua função, ou seja, das consequências que produz no ambiente. Esse entendimento reforça a importância de programas sistemáticos e individualizados de intervenção, capazes de identificar as variáveis que controlam o repertório autista e de priorizar o ensino de habilidades, em contraste com abordagens generalistas ou medicalizantes (Goulart & Assis, 2002).

Além disso, a literatura recente tem enfatizado a necessidade de métodos baseados em evidências para o manejo de comportamentos desafiadores em crianças com TEA. Carvalho e Moreira (2022) destacam que intervenções derivadas da Análise do Comportamento Aplicada (ABA), em especial o uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA), têm se mostrado eficazes na redução de comportamentos agressivos, autolesivos e destrutivos sem depender exclusivamente do procedimento de extinção. Tais estratégias incluem ensinar respostas alternativas de comunicação, manipular a qualidade e a magnitude dos reforçadores e combinar reforçamento diferencial com reforçamento não contingente, permitindo não apenas a diminuição dos comportamentos inadequados, mas também a promoção de repertórios mais adaptativos. Esses achados reforçam a importância de recursos que auxiliem pais e profissionais a se prepararem para lidar com situações de frustração e de birras de maneira preventiva e funcional (Carvalho & Moreira, 2022).

O Reforçamento Diferencial de Comportamentos Alternativos (DRA) consiste no reforço de uma resposta alternativa em relação ao comportamento-problema. Embora tradicionalmente descrito em conjunto com a extinção, evidências indicam que a extinção nem sempre é necessária para caracterizar o procedimento. Nesses casos, o DRA pode ser compreendido como o arranjo de maior reforçamento para um comportamento em comparação a outro, com redução relativa do reforço para o comportamento inadequado (Vollmer et al., 2020).

Em muitos casos, comportamentos disruptivos em indivíduos com TEA estão relacionados a déficits na comunicação funcional, fazendo com que a criança utilize respostas inadequadas para obter atenção, escapar de demandas ou acessar itens desejados. Dessa forma, intervenções que ensinam formas mais adequadas de comunicação tendem a reduzir a necessidade de emissão de comportamentos problemáticos (Bagaiolo et al., 2018).

Estudos sobre capacitação parental baseados na Análise do Comportamento Aplicada indicam que ensinar habilidades de comunicação pode contribuir para a redução de comportamentos disruptivos em indivíduos com Transtorno do Espectro Autista. Nesses programas, os pais são treinados para identificar comportamentos difíceis e ensinar respostas mais apropriadas, frequentemente por meio de estratégias como reforçamento diferencial e treino de comunicação funcional (Bagaiolo et al., 2018).

Apesar desses resultados, estudos também apontam dificuldades relacionadas à adesão dos cuidadores. Em uma intervenção com pais de indivíduos com TEA, apenas 21 dos 72 participantes iniciais compareceram a todos os encontros propostos, sendo relatadas como possíveis barreiras dificuldades financeiras, problemas de transporte, ausência de suporte familiar e a gravidade dos sintomas apresentados pelas crianças e pelos próprios cuidadores (Bagaiolo et al., 2018). Esses fatores evidenciam limitações importantes no acesso a intervenções presenciais, indicando a necessidade de alternativas que ampliem o alcance dessas estratégias.

Nesse cenário, novas formas de intervenção têm sido propostas, especialmente a partir dos avanços das tecnologias digitais. Embora eficazes, programas de treinamento parental enfrentam limitações como a escassez de profissionais qualificados, a falta de tempo dos pais para participação em intervenções presenciais e o estigma associado à busca por atendimento psicológico, o que reforça a necessidade de alternativas mais acessíveis (Entenberg et al., 2021). Nesse contexto, destacam-se as ferramentas baseadas em inteligência artificial, como chatbots, que utilizam linguagem natural para interagir com usuários e podem oferecer suporte de forma contínua, acessível e independente de localização (Entenberg et al., 2021). Diante disso, a inteligência artificial (IA), em especial os modelos de linguagem generativa, destaca-se por sua capacidade de processar grandes volumes de dados e produzir respostas em linguagem natural a partir de comandos conhecidos como prompts. Esses sistemas não criam conhecimento inédito, mas organizam e sintetizam conteúdos previamente disponíveis, oferecendo respostas rápidas e articuladas a questões abertas (Sampaio et al., 2024). A introdução de ferramentas como o ChatGPT, nesse sentido, representa uma transformação paradigmática na forma como o conhecimento é produzido e acessado, sendo comparável à introdução dos buscadores online ou da Wikipedia no campo científico.

A ampliação do acesso a orientações baseadas em evidências é especialmente relevante diante da limitação de recursos humanos especializados e da dificuldade de acesso a tratamentos comportamentais estruturados em muitos contextos de atendimento, especialmente na rede pública (Bagaiolo et al., 2018).

No entanto, os mesmos autores sublinham que essa aparente inovação não está livre de riscos e limitações. Como funcionam a partir de padrões estatísticos extraídos de grandes bases de dados, esses modelos tendem a reproduzir vieses sociais, culturais e linguísticos presentes nas fontes originais, além de poderem apresentar respostas incorretas mas persuasivas, fenômeno descrito como “alucinação” (Sampaio et al., 2024). Ademais, a falta de transparência quanto aos algoritmos e às bases utilizadas no treinamento dificulta a avaliação da confiabilidade dos conteúdos gerados. Os autores alertam ainda para implicações éticas e políticas: a possibilidade de que as IAs reforcem desigualdades no acesso à informação e reproduzam

perspectivas dominantes, invisibilizando vozes minoritárias. Nesse cenário, a utilização acrítica da IA pode reduzir a autonomia intelectual dos pesquisadores e comprometer a qualidade da produção científica. Assim, mais do que descartar ou adotar sem reservas esses sistemas, Sampaio et al. (2024) defendem a necessidade de uma literacia digital crítica que permita ao usuário compreender os limites da ferramenta, interpretar os resultados de forma contextualizada e utilizá-los como apoio, e não substituto, da reflexão científica.

Problema de Pesquisa

O crescimento das inteligências artificiais generativas, como o ChatGPT, trouxe novas possibilidades de acesso e organização de informações em diferentes áreas do conhecimento. Estudos recentes apontam que essas ferramentas são eficientes em sintetizar grandes volumes de dados e em fornecer respostas rápidas, claras e articuladas (Kim et al., 2025; Sampaio et al., 2024). No entanto, embora a eficiência técnica seja evidente, permanece uma lacuna quanto à confiabilidade dessas respostas, especialmente em contextos de maior sensibilidade social.

Contudo, a adoção dessas tecnologias suscita um problema central: até que ponto é possível confiar nas informações produzidas por um aplicativo de IA ao lidar com situações complexas e delicadas? Como destacam Sampaio et al. (2024), os modelos de linguagem generativa estão sujeitos a vieses, falta de transparência sobre seus dados de treinamento e à possibilidade de gerar respostas plausíveis, mas incorretas, fenômeno conhecido como “alucinação”. Nessa mesma linha, McFayden et al. (2024), ao compararem o desempenho do ChatGPT-3.5 e do ChatGPT-4, verificaram que as versões mais recentes apresentam maior clareza, correção e consistência. Entretanto, os autores também ressaltam que limitações como a generalização excessiva e a dificuldade de manter coerência em interações prolongadas permanecem, o que reforça a necessidade de investigações específicas sobre a confiabilidade dessas ferramentas.

Portanto, o problema desta pesquisa pode ser sintetizado da seguinte forma: embora as IAs generativas apresentem alto desempenho técnico, ainda não se sabe se suas respostas são suficientemente confiáveis e relevantes para apoiar pais de crianças com TEA na prevenção e no manejo de comportamentos desafiadores, especialmente quando se consideram diferentes tipos de prompt.

Pergunta de Pesquisa

O uso de inteligências artificiais generativas tem se expandido rapidamente em diferentes áreas, mostrando eficiência na produção de respostas rápidas e articuladas a comandos em linguagem natural (Sampaio et al., 2024; Kim et al., 2025). Entretanto, ainda não está claro se essas respostas podem ser consideradas confiáveis em contextos sensíveis, como o apoio a pais de crianças com Transtorno do Espectro Autista (TEA) no manejo de comportamentos desafiadores. Nesse sentido, pretende-se responder à seguinte questão de pesquisa: qual a precisão e a utilidade de um aplicativo gerado por inteligência artificial (chatbot) que apoia

pais de crianças com TEA na prevenção de comportamentos agressivos e birras, por meio do uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA), considerando diferentes tipos de prompt?

Relevância da Pesquisa

Pais e cuidadores de crianças com Transtorno do Espectro Autista (TEA) enfrentam desafios diários relacionados ao manejo de comportamentos desafiadores e, em consequência, apresentam uma demanda crescente por apoio e informações confiáveis. Kim et al. (2025) identificaram que as buscas mais frequentes realizadas por pais envolvem temas como incentivo ao sono, disciplina, manejo de birras, autocuidado parental e demonstrações de afeto, o que evidencia a necessidade de recursos acessíveis que auxiliem no enfrentamento dessas situações. Assim, investigar o uso de tecnologias que possam fornecer orientações rápidas e claras torna-se socialmente relevante, pois responde a uma demanda concreta das famílias.

O impacto do TEA sobre as famílias pode envolver sobrecarga emocional e dificuldades no manejo de comportamentos disruptivos. Nesse contexto, programas de capacitação parental têm sido propostos como alternativa para auxiliar os cuidadores a compreender e manejar esses comportamentos, ao mesmo tempo em que estimulam habilidades comunicativas nos indivíduos com TEA (Bagaiolo et al., 2018).

Do ponto de vista científico, esta pesquisa contribui para o debate atual sobre a confiabilidade das inteligências artificiais generativas. Embora estudos reconheçam sua eficiência na produção de respostas rápidas e articuladas (Sampaio et al., 2024; McFayden et al., 2024), ainda persiste a necessidade de avaliar até que ponto essas respostas são aplicáveis e úteis em contextos práticos e sensíveis, como o cuidado parental em TEA. Além disso, pesquisas sobre prompt engineering têm mostrado que diferentes formulações de prompt podem impactar diretamente a qualidade, consistência e clareza das respostas produzidas pela IA (Wang et al., 2024). Desse modo, este estudo se insere em uma lacuna científica relevante ao investigar a precisão e a utilidade das respostas em função da variação nos comandos utilizados.

Em termos práticos, a pesquisa também possui relevância ao explorar possibilidades de ampliar o acesso de famílias a estratégias de manejo e prevenção de comportamentos desafiadores. Evidências apontam que o treino de cuidadores baseado em técnicas de reforçamento diferencial promove resultados consistentes na redução de comportamentos inadequados e no fortalecimento de repertórios adaptativos (Guimarães et al., 2018). No entanto, a aplicação desse tipo de intervenção ainda é limitada em razão de custos, tempo e disponibilidade de serviços especializados. Nesse cenário, um aplicativo de inteligência artificial baseado em chatbot pode funcionar como ferramenta complementar às intervenções clínicas, oferecendo suporte imediato e acessível no cotidiano familiar e contribuindo para a promoção de práticas parentais mais eficazes e preventivas.

Revisão da Literatura

O avanço das tecnologias digitais tem impactado profundamente a forma como o conhecimento é produzido e utilizado. Entre essas inovações, a inteligência artificial (IA), especialmente os modelos de linguagem generativa, vem se destacando por sua capacidade de sintetizar informações em linguagem natural a partir de comandos simples (prompts). Como apontam Sampaio et al. (2024), tais ferramentas representam uma transformação paradigmática na pesquisa científica, comparável ao surgimento de buscadores online e enciclopédias colaborativas. Além de acelerarem a análise e organização de dados, esses sistemas apresentam riscos importantes, como a reprodução de vieses das bases de treinamento, a geração de respostas convincentes porém incorretas (as chamadas “alucinações”) e a falta de transparência nos algoritmos. Para os autores, a utilidade da IA depende da adoção de uma postura crítica por parte dos pesquisadores, de modo a equilibrar potencialidades e limitações.

Outro aspecto central para o uso das IAs generativas refere-se à formulação de prompts, que orientam a produção das respostas. A literatura mostra que a qualidade, clareza e estrutura dos prompts afetam diretamente a precisão e a utilidade dos resultados. Revisões recentes, como a de Wang et al. (2024), destacam que diferentes técnicas de prompt engineering, como zero-shot, few-shot e chain-of-thought, podem gerar variações significativas no nível de detalhamento, consistência e relevância das informações. Isso evidencia que, para além da eficiência do modelo, é fundamental compreender como o modo de interação influencia a qualidade das respostas obtidas, sobretudo quando o objetivo é oferecer orientações práticas aos pais em situações delicadas do cotidiano.

Essas reflexões chegam também ao campo do cuidado parental. Em estudo recente, Kim et al. (2025) avaliaram a capacidade da IA generativa de fornecer informações a pais e cuidadores em comparação a conteúdos previamente preparados por especialistas. Para isso, foi realizado um teste dedutivo com 11 condições de análise em cinco tópicos, conduzido na primavera de 2024. A fim de controlar a variabilidade das informações disponíveis na internet aberta, os autores criaram um universo de dados a partir do conteúdo de 40 boletins informativos parentais de oito páginas, considerados uma fonte de informações verificadas e de alta qualidade. Nessa configuração, compararam o desempenho do ChatGPT-3.5 e do ChatGPT-4 com a ferramenta de busca de um site especializado em boletins informativos parentais (Just in Time Parenting – JITP.info), bem como com o conjunto completo do acervo digital da newsletter e com dados disponíveis na internet aberta. As análises foram feitas a partir de cinco perguntas abertas, definidas com base em tópicos que o Google Analytics identificou como os mais pesquisados no site JITP.info: incentivo ao sono, estabelecimento de limites e disciplina, autocuidado parental, manejo de birras e demonstrações de afeto. Os resultados indicaram que, embora as respostas oferecidas pela IA sejam em geral corretas e claras, elas apresentam limitações quanto ao nível de contextualização, à adequação ao estágio de desenvolvimento infantil, à ênfase na relação entre pais e filhos e à consistência das referências utilizadas. Dessa forma, a IA pode ser considerada uma ferramenta útil para fornecer respostas iniciais a dúvidas parentais, mas ainda exige complementação por profissionais especializados para garantir a qualidade e a aplicabilidade das orientações.

No campo da avaliação comparativa das inteligências artificiais, McFayden et al. (2024) conduziram um estudo que analisou o desempenho do ChatGPT em questões diretamente relacionadas ao autismo. Foram elaboradas treze perguntas abertas, divididas em três categorias que refletem dúvidas comuns de cuidadores: (a) informações básicas sobre o autismo, como “Quais são os primeiros sinais do autismo?”; (b) mitos e concepções equivocadas, como “Vacinas causam autismo?”; e (c) recursos disponíveis, como “Quanto tempo dura a fila de espera para serviços de autismo?”. As respostas foram coletadas em uma nova conta de ChatGPT, sem histórico de uso, a fim de evitar influência de interações anteriores. Para cada pergunta, o modelo foi solicitado a gerar duas respostas diferentes, denominadas R1 e R2, além de fornecer uma lista de referências. As respostas foram então submetidas a uma análise qualitativa que considerou critérios de clareza, correção e consistência. Os resultados mostraram que o ChatGPT-4 apresentou desempenho superior em relação ao ChatGPT-3.5, oferecendo respostas mais completas e organizadas. No entanto, os autores destacaram limitações persistentes, como a generalização excessiva e a inconsistência em interações mais longas, reforçando a necessidade de investigações específicas sobre a confiabilidade do uso da IA em contextos delicados, como o apoio a pais de crianças com TEA.

Embora haja avanços no estudo de prompts e na avaliação de IAs em diferentes áreas, ainda não se encontram investigações específicas sobre sua aplicação em comportamentos desafiadores no TEA. A literatura clínica mostra que tais comportamentos, como birras, agressividade e autoagressão, podem comprometer tanto o desenvolvimento da criança quanto o bem-estar familiar. Nesse sentido, pesquisas aplicadas destacam a relevância de estratégias preventivas e funcionais. O Reforçamento Diferencial de Comportamentos Alternativos (DRA) é amplamente utilizado como estratégia para a redução de comportamentos-problema, sendo frequentemente aplicado após a identificação das variáveis que mantêm o comportamento por meio de análise funcional. Em sua aplicação, o comportamento alternativo passa a produzir consequências reforçadoras mais favoráveis em comparação ao comportamento inadequado, favorecendo a mudança na alocação de respostas. Além disso, pesquisas indicam que o DRA pode ser eficaz mesmo quando a extinção do comportamento-problema não é completamente implementada, desde que haja uma diferença significativa nas condições de reforçamento entre os comportamentos (Vollmer et al., 2020).

Estudos recentes têm investigado a aplicação do Reforçamento Diferencial de Comportamentos Alternativos (DRA) sem a utilização do procedimento de extinção, especialmente em contextos nos quais sua implementação pode ser difícil ou inviável. Uma revisão sistemática identificou que, em 10 estudos analisados, nove apresentaram efeitos positivos na redução de comportamentos-problema, mesmo sem a retirada completa do reforço para essas respostas (MacNaul & Neely, 2018). Esses achados sugerem que intervenções baseadas em DRA podem ser eficazes a partir da manipulação de parâmetros do reforçamento, como qualidade, magnitude, imediaticidade e frequência, favorecendo o comportamento alternativo sem necessariamente depender da extinção.

Guimarães et al. (2018), por exemplo, evidenciam que o treino de cuidadores para manejar comportamentos inadequados em crianças com TEA, baseado em técnicas de reforçamento diferencial, favorece não apenas a redução de comportamentos-problema, mas também o

fortalecimento de repertórios adaptativos. Essa perspectiva reforça a necessidade de investigar se um aplicativo de IA, mediado por diferentes tipos de prompt, pode atuar como ferramenta confiável e útil de apoio preventivo a pais no manejo desses comportamentos.

Objetivo da Pesquisa

Diante desse panorama, o objetivo deste estudo é investigar a precisão e a utilidade de um aplicativo gerado por inteligência artificial (chatbot) que apoia pais de crianças com TEA na prevenção de comportamentos agressivos e birras, por meio do uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA). A proposta fundamenta-se nos aportes da análise do comportamento (Goulart & Assis, 2002), nas reflexões sobre os usos científicos da IA (Sampaio et al., 2024), nas contribuições sobre estratégias de prevenção de comportamentos desafiadores (Carvalho & Moreira, 2022) e nas evidências acerca da qualidade das respostas oferecidas pela IA a cuidadores (Kim et al., 2025). Assim, pretende-se responder à seguinte questão de pesquisa: qual a precisão e a utilidade de um aplicativo gerado por inteligência artificial (chatbot) que apoia pais de crianças com TEA na prevenção de comportamentos agressivos e birras, por meio do uso do Reforçamento Diferencial de Comportamentos Alternativos (DRA)?

2 MÉTODO

Delineamento de Pesquisa

O presente estudo caracteriza-se como uma pesquisa aplicada, com delineamento comparativo entre condições experimentais. O objetivo foi investigar a precisão conceitual e a utilidade prática de conteúdos gerados por um chatbot de inteligência artificial para orientar pais de crianças com Transtorno do Espectro Autista (TEA) no manejo de comportamentos agressivos e birras, com base no Reforçamento Diferencial de Comportamentos Alternativos (DRA). Para isso, foram utilizados três prompts distintos, que constituem as condições experimentais do estudo. A variável independente corresponde à formulação do prompt, enquanto as variáveis dependentes referem-se à precisão e à utilidade das respostas produzidas pelo chatbot.

Modelos Utilizados

Para a elaboração de partes do material textual e apoio na organização de ideias, foi utilizada a ferramenta ChatGPT (versão GPT-5, desenvolvida pela OpenAI). O modelo é um sistema de inteligência artificial de processamento de linguagem natural (PLN), baseado em arquitetura de transformadores, treinado em larga escala com dados textuais diversos.

Prompts Utilizados

Foram elaborados três prompts variados para avaliar a influência da formulação da pergunta na qualidade e utilidade das respostas fornecidas pelo chatbot (ChatGPT-5). A seguir, apresentam-se os prompts na íntegra:

- Prompt 1: Crie um aplicativo em HTML no formato de tutorial educativo para pais de crianças com Transtorno do Espectro Autista (TEA). O conteúdo deve explicar o que é o Reforçamento Diferencial de Comportamentos Alternativos (DRA) e mostrar como utilizá-lo em situações cotidianas, como birras, agressividade ou dificuldades na rotina. O texto deve ser acessível e em linguagem clara, evitando termos excessivamente técnicos. Utilize métodos de ensino com instrução, modelação (imitação) e modelagem (com feedback imediato). Após finalizar, liste as referências bibliográficas relacionadas.
- Prompt 2: Crie um aplicativo em HTML no formato de tutorial educativo para pais de crianças com Transtorno do Espectro Autista (TEA). O conteúdo deve explicar o que é o Reforçamento Diferencial de Comportamentos Alternativos (DRA) e mostrar como utilizá-lo em situações cotidianas, como birras, agressividade ou dificuldades na rotina. O texto deve ser acessível e em linguagem clara, evitando termos excessivamente técnicos. Utilize métodos de ensino com instrução, modelação (imitação) e modelagem (com feedback imediato). No final, insira perguntas conceituais e situacionais de múltipla escolha para que os pais possam verificar se compreenderam como aplicar o DRA no cotidiano. Após finalizar, liste as referências bibliográficas utilizadas.
- Prompt 3: Crie um aplicativo em HTML no formato de tutorial educativo para pais de crianças com Transtorno do Espectro Autista (TEA). O conteúdo deve explicar o que é o Reforçamento Diferencial de Comportamentos Alternativos (DRA) e mostrar como utilizá-lo em situações cotidianas, como birras, agressividade ou dificuldades na rotina. Inclua uma explicação conceitual sobre o DRA, exemplos práticos de aplicação (treino) e um espaço de autoavaliação (teste). O texto deve ser acessível e em linguagem clara, evitando termos excessivamente técnicos. No final, insira perguntas conceituais e situacionais de múltipla escolha para que os pais possam verificar se compreenderam como aplicar o DRA no cotidiano. Após finalizar, liste as referências bibliográficas utilizadas. Para isso, baseie-se no artigo anexado. (MacNaul, H; Neely, L (2017). Systematic Review of Differential Reinforcement of alternative behavior without extinction for individuals with autism. Behavior modification, 42(3) 398-421. <https://doi.org/10.1177/0145445517740321>)

Avaliação do Conteúdo Gerado

A avaliação das respostas fornecidas pelo chatbot foi conduzida em duas etapas complementares.

A avaliação do conteúdo produzido pelo chatbot foi realizada de forma sistemática, buscando mensurar a qualidade, a precisão conceitual e a utilidade prática do material gerado. A análise seguiu critérios de qualidade amplamente utilizados em estudos semelhantes, com base em McFayden et al. (2024) e Kim et al. (2025).

Na primeira etapa, foi avaliado o conteúdo textual dos tutoriais produzidos. Essa análise seguiu três dimensões principais, denominadas “3 Cs”: corretude, clareza e concisão. A corretude refere-se ao grau em que o conteúdo está cientificamente alinhado às evidências sobre o Reforçamento Diferencial de Comportamentos Alternativos (DRA), ao Transtorno do Espectro Autista (TEA) e às estratégias de prevenção de comportamentos desafiadores. A clareza diz respeito à organização das ideias e à facilidade de compreensão do texto por parte de pais e cuidadores, considerando o uso de linguagem acessível, objetiva e sem jargões técnicos. Já a concisão avalia o equilíbrio entre completude e síntese, observando se o texto apresenta as informações essenciais de forma direta, evitando redundâncias ou explicações desnecessárias.

Cada um desses critérios foi pontuado em uma escala de 1 a 4, em que 4 corresponde a “completamente correto, claro ou conciso”; 3 a “quase correto, claro ou conciso”; 2 a “parcialmente correto, claro ou conciso”; e 1 a “completamente incorreto, pouco claro ou pouco conciso”. Essa escala permitiu identificar com maior precisão se o conteúdo gerado é conceitualmente adequado, se transmite as informações de maneira compreensível e se mantém coerência interna em sua estrutura.

Rubrica de Avaliação do Conteúdo

Cada critério foi avaliado em uma escala de 1 a 4, sendo:

- 1 = Inadequado
- 2 = Parcialmente adequado
- 3 = Quase adequado
- 4 = Totalmente adequado

1. Corretude científica

Avalia o alinhamento conceitual do conteúdo com as evidências científicas sobre o Reforçamento Diferencial de Comportamentos Alternativos (DRA), o Transtorno do Espectro Autista (TEA) e o manejo de comportamentos desafiadores.

1 – Inadequado:

Conteúdo incorreto ou inconsistente com a literatura científica.

2 – Parcialmente adequado:

Parcialmente correto, com imprecisões ou termos mal aplicados.

3 – Quase adequado:

Conceitualmente adequado, mas com pequenas lacunas ou simplificações.

4 – Totalmente adequado:

Totalmente correto e coerente com os princípios da Análise do Comportamento Aplicada (ABA) e com as evidências atuais.

2. Clareza e compreensão

Avalia a organização das ideias e a facilidade de entendimento do texto por parte de pais e cuidadores.

1 – Inadequado:

Texto confuso, com frases longas, vocabulário técnico e pouca coerência.

2 – Parcialmente adequado:

Parcialmente compreensível, com trechos claros, mas com estrutura ou linguagem complexa.

3 – Quase adequado:

Texto claro na maior parte, porém com pequenas ambiguidades ou excesso de informação.

4 – Totalmente adequado:

Texto totalmente claro, fluido e acessível, com linguagem simples e exemplos compreensíveis.

3. Concisão e objetividade

Avalia o equilíbrio entre completude e síntese, observando se o texto apresenta as informações essenciais de forma direta e evita redundâncias.

1 – Inadequado:

Texto prolixo, repetitivo ou com informações irrelevantes.

2 – Parcialmente adequado:

Parcialmente objetivo, mas ainda com excesso de explicações ou desvios de foco.

3 – Quase adequado:

Texto equilibrado, com leves repetições ou passagens longas.

4 – Totalmente adequado:

Altamente conciso e objetivo, apresentando apenas informações essenciais de forma clara e direta.

Pontuação final

A pontuação total variará de 3 a 12 pontos, sendo:

- 3-5 pontos: Baixa qualidade conceitual e comunicativa.
- 6-8 pontos: Qualidade moderada, compreensível com ajustes.
- 9-12 pontos: Alta qualidade, conteúdo claro, preciso e aplicável.

Avaliação da adequação ao contexto parental

Além dos 3 Cs (corretude, clareza e concisão), também foi analisada a adequação ao contexto parental, conforme sugerido por Kim et al. (2025), que destacam a importância de considerar a aplicabilidade das informações ao cotidiano das famílias.

Foram observados aspectos como:

- presença de exemplos práticos;
- pertinência das orientações para o manejo de birras, agressividade e frustrações;
- potencial do tutorial em promover práticas preventivas baseadas em evidências.

A utilidade prática do material foi, portanto, outro critério de avaliação, considerando a capacidade do conteúdo de orientar pais e cuidadores na compreensão e aplicação do DRA em situações reais.

Segunda etapa da avaliação

Na segunda etapa, foi realizada a avaliação do aplicativo em HTML, com o objetivo de verificar não apenas a qualidade do conteúdo, mas também sua funcionalidade, acessibilidade e potencial de uso por pais de crianças com Transtorno do Espectro Autista (TEA).

Foram considerados quatro critérios principais:

- organização
- usabilidade
- acessibilidade
- pertinência para o público-alvo

Essa análise seguiu parâmetros de clareza, coerência e aplicabilidade, conforme destacado por McFayden et al. (2024) e Kim et al. (2025).

Rubrica de Avaliação do Aplicativo

1. Organização e Estrutura do Tutorial

Avalia a sequência lógica e a organização geral do conteúdo apresentado no aplicativo.

1-2 (Muito fraco / Fraco):

Conteúdo desorganizado, sem sequência lógica entre as partes, dificultando a compreensão.

3 (Regular):

Estrutura compreensível, mas com transições confusas ou seções desiguais.

4-5 (Bom / Excelente):

Sequência lógica, coerente e fluida (introdução, explicação, exemplos, autoavaliação e referências), favorecendo o aprendizado.

2. Clareza e Linguagem

Avalia a clareza do texto e a adequação da linguagem ao público-alvo.

1-2 (Muito fraco / Fraco):

Texto confuso, com excesso de termos técnicos e frases longas.

3 (Regular):

Linguagem parcialmente acessível, com trechos densos ou pouco objetivos.

4-5 (Bom / Excelente):

Linguagem simples, direta e compreensível para pais e cuidadores, sem jargões técnicos desnecessários.

3. Usabilidade e Navegação

Avalia a facilidade de uso e a organização da navegação no aplicativo.

1-2 (Muito fraco / Fraco):

Navegação confusa, layout desorganizado ou difícil de usar.

3 (Regular):

Estrutura navegável, mas pouco intuitiva ou com problemas de fluxo.

4-5 (Bom / Excelente):

Interface clara, intuitiva e organizada, permitindo fácil localização das informações e boa experiência de uso.

4. Acessibilidade Visual e Técnica

Avalia aspectos visuais e técnicos que influenciam a legibilidade e o acesso ao conteúdo.

1-2 (Muito fraco / Fraco):

Layout visualmente inadequado, fontes pequenas ou incompatibilidade com dispositivos móveis.

3 (Regular):

Formatação parcialmente acessível, com pequenos prejuízos de leitura.

4-5 (Bom / Excelente):

Visual limpo, contraste adequado, boa legibilidade e compatibilidade com diferentes telas e navegadores.

5. Pertinência ao Público-Alvo

Avalia o grau de adequação do conteúdo ao contexto de pais de crianças com TEA.

1-2 (Muito fraco / Fraco):

Conteúdo genérico, sem relação com o cotidiano de pais de crianças com TEA.

3 (Regular):

Parcialmente adequado, com exemplos pouco contextualizados.

4-5 (Bom / Excelente):

Altamente pertinente, com exemplos práticos de birras, agressividade e situações de rotina, em linguagem empática e aplicável.

6. Aplicabilidade Prática e Alinhamento com a ABA

Avalia se as orientações apresentadas estão alinhadas com práticas baseadas em evidências da Análise do Comportamento Aplicada.

1-2 (Muito fraco / Fraco):

Estratégias teóricas, sem relação com práticas baseadas em evidências.

3 (Regular):

Orientações parcialmente aplicáveis ou sem clareza na relação com o DRA.

4-5 (Bom / Excelente):

Estratégias fundamentadas na Análise do Comportamento Aplicada (ABA), especialmente no DRA, com orientações claras, práticas e preventivas.

Pontuação Final

A pontuação total variou de 6 a 30 pontos, sendo:

6–14 pontos: Baixa qualidade / pouca aplicabilidade.

15–23 pontos: Qualidade moderada / aplicável com ajustes.

24–30 pontos: Alta qualidade / aplicabilidade direta.

Terceira Etapa da Avaliação

Na terceira etapa foi realizada uma avaliação criteriosa das referências utilizadas. Esse processo envolve verificar a existência e a autenticidade das obras citadas, confirmar se atendem aos critérios de validade científica, identificar se se tratam de artigos publicados em periódicos reconhecidos e analisar a classificação desses periódicos no sistema Qualis da CAPES.

Essa etapa tem como objetivo assegurar a qualidade e a relevância das fontes consultadas, garantindo que o trabalho se apoie em produções científicas consistentes e devidamente avaliadas pela comunidade acadêmica.

Procedimento

Os três prompts foram aplicados em sessões independentes, utilizando uma nova conversa a cada aplicação ao modelo ChatGPT-5, em uma conta nova, sem histórico de interações, para evitar viés nas respostas. Primeiramente, foi inserido o Prompt 1, e a resposta gerada foi salva em arquivo. Em seguida, criou-se uma nova conversa (chat) e aplicou-se o Prompt 2, repetindo o processo de coleta e armazenamento. Por último, foi testado o Prompt 3, também com registro da resposta correspondente. Todas as respostas foram organizadas em um banco de dados único, preservando a ordem de aplicação, para posterior análise comparativa quanto à sua precisão conceitual e utilidade prática para pais de crianças com TEA.

3 RESULTADOS

Coleta de Dados

A coleta de dados foi realizada no dia 28 de outubro de 2025. Os links de acesso para cada aplicativo estão listados a seguir:

Links:

Aplicativo 1:

<https://chatgpt.com/canvas/shared/692672f58cd0819199bbcbb6b783166e>

Aplicativo 2:

<https://chatgpt.com/canvas/shared/6926735bc0188191ac4e98a8eddc2f35>

Aplicativo 3:

<https://chatgpt.com/canvas/shared/692673c755d88191b0e5b9a044d1829e>

Telas dos Aplicativos

Os aplicativos educativos em HTML sobre o Reforçamento Diferencial de Comportamentos Alternativos (DRA) foram desenvolvidos no dia 28 de outubro de 2025, com tempo de criação de apenas alguns segundos, utilizando geração automática de código estruturado e conteúdo textual. Durante o processo de elaboração e teste, não foram observados erros de execução no código HTML. Os aplicativos funcionaram corretamente em ambiente de teste, apresentando boa responsividade, carregamento rápido e interatividade estável no simulador de situações. A revisão visual confirmou a clareza do conteúdo educativo, a acessibilidade das instruções e a coerência pedagógica das seções apresentadas.

Prompt 1

A Figura 1 é a tela do aplicativo gerado pelo prompt 1. Nela temos uma tela do aplicativo com título destacado e organização do conteúdo em seções. Observa-se uma explicação introdutória do DRA em linguagem simples, seguida pela apresentação de princípios básicos estruturados em passos, com orientações práticas para aplicação no cotidiano.

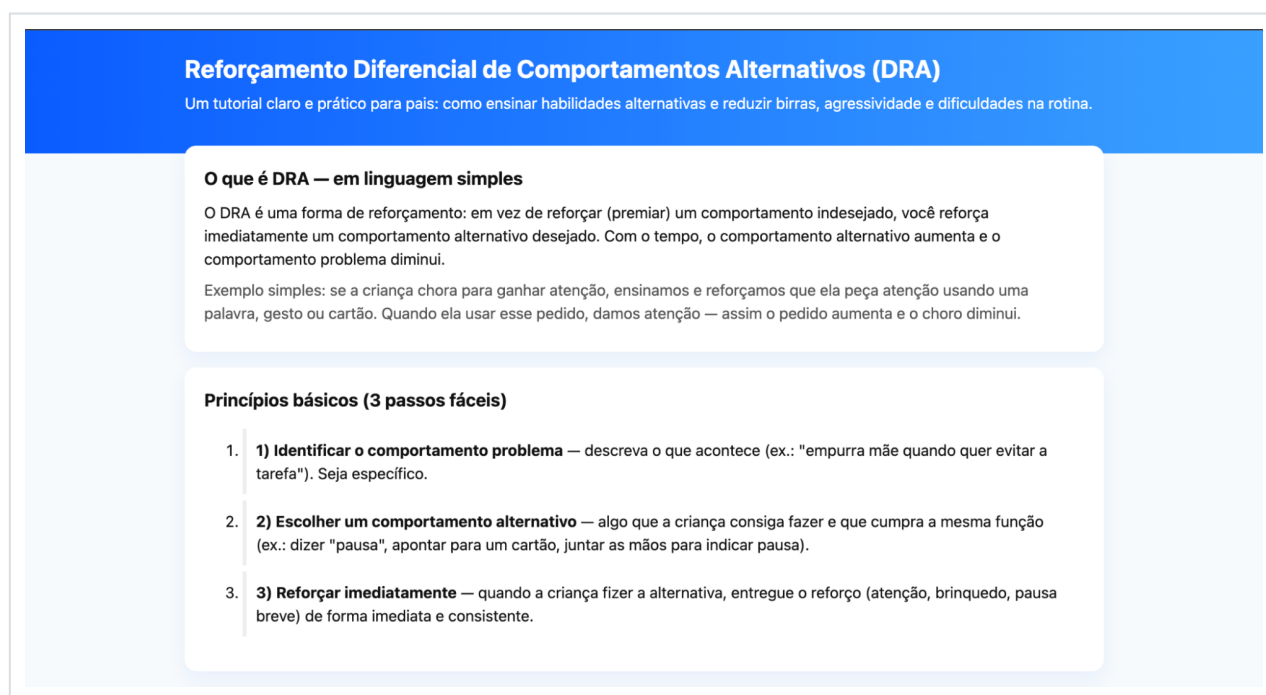


Figura 1. Tela inicial do aplicativo-tutorial sobre DRA (versão com prompt elaborado)

Nota. Captura de tela do aplicativo gerado por IA. Apresenta o tema “Reforçamento Diferencial de Comportamentos Alternativos (DRA)”, com a explicação “O que é DRA?” em linguagem simples e a seção “Princípios básicos (3 passos fáceis)”: identificar o comportamento-problema, escolher um comportamento alternativo e reforçar imediatamente.

Prompt 2

A Figura 2 é a tela do aplicativo gerado pelo prompt 2. Nela temos a tela inicial do aplicativo, com um cabeçalho contendo o título do tutorial sobre DRA e um subtítulo explicativo. Abaixo, há uma seção introdutória que apresenta o conceito de forma simples, acompanhada de exemplo prático, seguida pela organização do conteúdo em seções temáticas.



Figura 2. Tela inicial do aplicativo-tutorial sobre DRA (versão alternativa)

Nota. Captura de tela de outra versão do tutorial gerado por IA, com as seções “O que é o DRA? (explicação simples)” e “Por que usar DRA?”, destacando a redução de crises e a substituição de comportamentos por alternativas úteis.

Prompt 3

A Figura 3 é a tela do aplicativo gerado pelo prompt 3. Nela temos uma tela do aplicativo com título, subtítulo explicativo e organização do conteúdo em blocos. Observa-se uma seção com explicação do DRA em linguagem simples, seguida de uma seção com benefícios do uso da técnica e, ao lado, uma área de autoavaliação com perguntas e opções de resposta para o usuário. O layout distribui as informações em colunas, combinando conteúdo textual e elementos interativos.

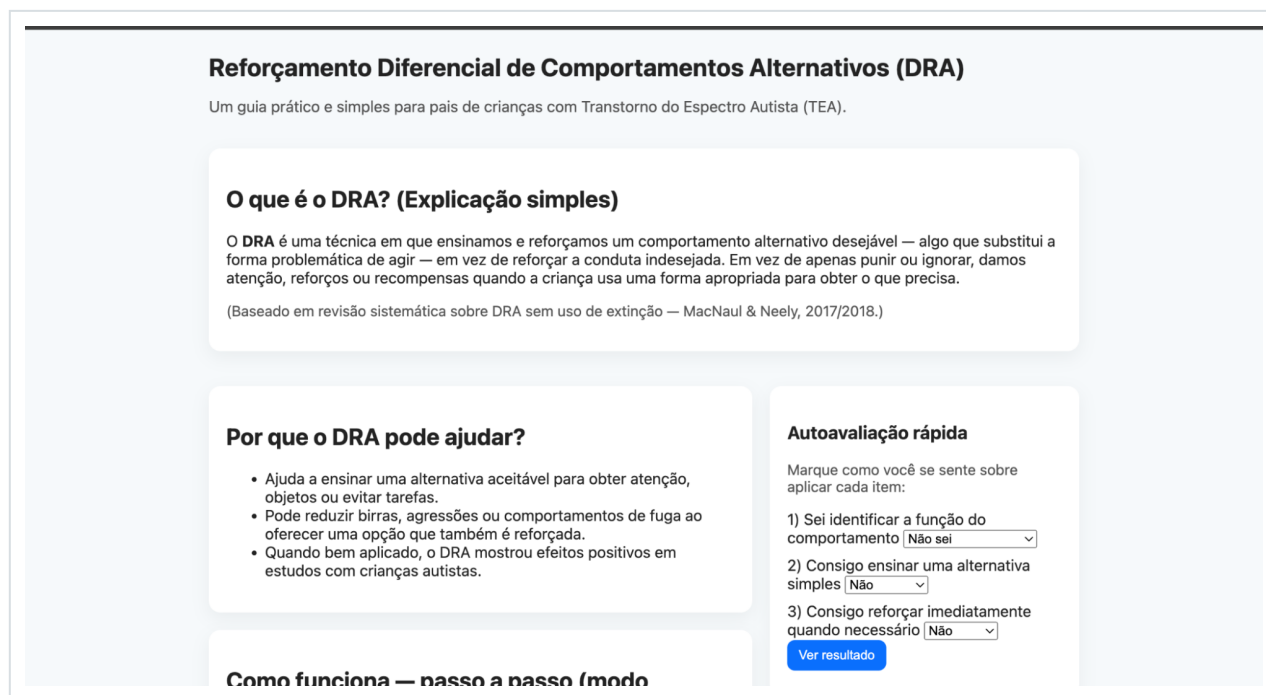


Figura 3. Tela do aplicativo-tutorial sobre DRA com painel de autoavaliação

Nota. Captura de tela do tutorial gerado por IA, com explicação do DRA baseada em revisão sistemática (MacNaul & Neely, 2017/2018) e um painel lateral de “Autoavaliação rápida” em que o usuário marca se sabe identificar a função do comportamento, ensinar uma alternativa e reforçar quando necessário.

Avaliação do Conteúdo

A análise dos dados foi realizada por meio da aplicação da rubrica de avaliação desenvolvida para este estudo, contemplando seis critérios essenciais para construção do aplicativo tutorial de DRA para apoiar pais de crianças com TEA na prevenção e no manejo de comportamentos desafiadores. Os critérios englobam aspectos de corretude, clareza e concisão, sendo pontuados em uma escala de um a quatro, em que 4 corresponderá a “completamente correto, claro ou conciso”; 3 a “quase correto, claro ou conciso”; 2 a “parcialmente correto, claro ou conciso”; e 1 a “completamente incorreto, pouco claro ou pouco conciso”. A Tabela 1 apresenta os aplicativos gerados a partir de cada prompt, juntamente com os critérios de avaliação e suas respectivas pontuações. Cada aplicativo corresponde a um prompt específico (1, 2 ou 3).

A Tabela 1 mostra as pontuações atribuídas aos três aplicativos de acordo com os critérios de corretude científica, clareza e compreensão e concisão e objetividade. O Aplicativo 1 apresentou ótima precisão conceitual (nota 4), com conteúdo coerente com os princípios da ABA e do DRA, porém obteve nota 2 em clareza, devido à linguagem técnica e pouco acessível ao público leigo, e nota 3 em concisão, por conter explicações longas e algumas repetições. O Aplicativo 2 destacou-se como o mais equilibrado, recebendo nota 4 em corretude científica, 3 em clareza e 4 em concisão, por combinar conteúdo preciso, exemplos práticos e linguagem mais fluida, mantendo o foco na aplicação. Já o Aplicativo 3 obteve nota 4 em corretude científica, 2 em clareza e 3 em concisão, demonstrando excelente embasamento teórico, mas linguagem excessivamente acadêmica e trechos extensos, o que compromete parcialmente a

acessibilidade. De modo geral, os três aplicativos mostram alta qualidade conceitual, mas o Aplicativo 2 se sobressai por alcançar melhor equilíbrio entre precisão científica e compreensão prática.

Tabela 1. Critérios e pontuações da avaliação do conteúdo dos aplicativos desenvolvidos a partir dos prompts

Critério	Aplicativo 1	Aplicativo 2	Aplicativo 3
Corretude	4	4	4
Clareza	2	3	2
Concisão	3	4	3
Total	9/12	11/12	9/12

Nota. Critérios e pontuações da avaliação do conteúdo dos três aplicativos tutoriais gerados por IA, lado a lado.

Avaliação do Aplicativo

A Tabela 2 apresenta os critérios utilizados para a avaliação dos três aplicativos desenvolvidos, considerando organização, clareza, usabilidade, acessibilidade, pertinência ao público-alvo e aplicabilidade. Cada aplicativo recebeu pontuações de 1 a 5, conforme o desempenho em cada aspecto.

O Aplicativo 1 obteve nota 4 em organização e estrutura, 3 em clareza, 3 em usabilidade, 3 em acessibilidade visual, 4 em pertinência ao público-alvo e 5 em aplicabilidade prática e alinhamento com a ABA. O conteúdo é conceitualmente correto e coerente, mas a linguagem técnica e o design simples limitaram a compreensão e a atratividade.

O Aplicativo 2 recebeu 4 em organização, 4 em clareza, 4 em usabilidade, 4 em acessibilidade, 5 em pertinência e 5 em aplicabilidade, destacando-se pelo equilíbrio entre teoria e prática, com linguagem acessível, exemplos contextualizados e boa experiência de navegação, embora ainda possa aprimorar o layout visual.

O Aplicativo 3 obteve 5 em organização e estrutura, 3 em clareza, 4 em usabilidade, 4 em acessibilidade, 4 em pertinência e 5 em aplicabilidade prática e alinhamento com a ABA. Apresenta a estrutura mais completa e embasamento científico sólido, mas a linguagem excessivamente acadêmica e uma questão do teste desalinhada ao público-alvo reduziram sua acessibilidade.

Assim, a Tabela 2 evidencia que todos os aplicativos apresentam boa coerência conceitual e aplicabilidade prática, variando principalmente no nível de clareza e adequação linguística, sendo o Aplicativo 2 o que mais se aproxima das necessidades e do perfil do público-alvo.

Tabela 2. Critérios e pontuações da avaliação do conteúdo dos aplicativos desenvolvidos a partir dos prompts

Critério	Aplicativo 1	Aplicativo 2	Aplicativo 3
Organização	4	4	4
Clareza	3	4	3
Usabilidade	3	4	4
Acessibilidade	3	4	4
Pertinência	4	5	4
Aplicabilidade	5	5	5
Total	22/30	26/30	24/30

Nota. Critérios e pontuações da avaliação do aplicativo (usabilidade e adequação) para os três tutoriais gerados por IA.

Avaliação das Referências

Nesta etapa foi feita uma verificação detalhada das referências utilizadas, com o objetivo de confirmar sua existência, autenticidade e validade científica.

As obras de Cooper, Heron e Heward (2007) e Miltenberger (2011,2016) foram confirmadas como livros reais, amplamente reconhecidos na área da Análise do Comportamento Aplicada (ABA) e com forte base científica. A obra de Hodgdon (1995) também existe e é utilizada em contextos de intervenção com pessoas com TEA, sendo considerada válida dentro do enfoque prático de estratégias visuais.

O livro de Leaf e McEachin (1999) foi igualmente confirmado e reconhecido por sua aplicação em programas de intervenção comportamental intensiva. Já as Diretrizes do Ministério da Saúde (2014) são documentos oficiais vigentes e reforçam a validade nacional das referências utilizadas.

Por outro lado, a citação de Winborn-Kemmerer está incompleta, pois não corresponde a uma obra específica, mas a uma autora com artigos publicados. Recomenda-se substituir pela referência de um de seus trabalhos reais. O mesmo vale para as menções genéricas às associações internacionais (ABAI e Autism Speaks), que devem ser ajustadas com a indicação de documentos concretos.

De forma geral, as principais referências são autênticas, atuais e cientificamente válidas, sendo necessário apenas ajustar as que estão incompletas ou muito genéricas.

O Aplicativo 3 utilizou como referência somente o artigo anexado juntamente ao prompt.

Referências aplicativo 1:

1. Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied Behavior Analysis.

2. Miltenberger, R. G. (2011). Behavior Modification: Principles and Procedures. (Capítulos práticos sobre técnicas de reforçamento).
3. Winborn-Kemmerer, M. (Revisão em intervenção comportamental aplicada a crianças com TEA – consulte guias locais de intervenção em autismo).
4. Diretrizes e manuais de sociedades de autismo (procure por documentos atualizados do Ministério da Saúde do seu país ou associações locais para recomendações contextualizadas).

Referências aplicativo 2:

1. Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). Applied Behavior Analysis (2ª ed.). Pearson. (Livro-base sobre técnicas comportamentais como DRA).
2. Miltenberger, R. G. (2016). Behavior Modification: Principles and Procedures. Cengage Learning.
3. Hodgdon, L. A. (1995). Visual Strategies for Improving Communication. QuirkRoberts Publishing. (Estratégias visuais e apoio em rotinas).
4. Leaf, R., & McEachin, J. (1999). Programas de intervenção baseados em ABA para crianças com TEA – artigos e guias práticos.
5. Artigos e diretrizes de associações de ABA e Autismo (ex.: Association for Behavior Analysis International; Autism Speaks) sobre reforçamento e ensino de habilidades alternativas.

Avaliação Geral

Os três aplicativos desenvolvidos apresentaram uma estrutura semelhante, iniciando com uma introdução conceitual sobre o Reforçamento Diferencial de Comportamentos Alternativos (DRA), abordando sua definição, objetivos e funcionamento. Além disso, o segundo e o terceiro aplicativos incluíram, ao final, um pequeno teste avaliativo, com o intuito de verificar a compreensão do usuário sobre o conteúdo apresentado. A seguir, foi realizada uma descrição detalhada e análise individual de cada aplicativo, acompanhada de exemplos ilustrativos.

Análise Individual dos Aplicativos

Aplicativo 1 – Introdução ao DRA

O aplicativo é composto por oito seções principais, cada uma representada por um “card” informativo, o que facilita a leitura e a navegação:

1. O que é DRA – Introduz o conceito de reforçamento diferencial, explicando o princípio de reforçar comportamentos alternativos em vez dos comportamentos problema.
2. Princípios básicos (3 passos fáceis) – Apresenta o passo a passo do método: identificação do comportamento problema, escolha da alternativa adequada e reforço imediato.
3. Como usar em situações do dia a dia – Fornece exemplos práticos aplicados a birras, agressividade e dificuldades na rotina, com orientações passo a passo para cada caso.

4. Métodos de ensino – Explica as estratégias de instrução, modelação e modelagem, incentivando o uso combinado dessas técnicas para acelerar o aprendizado.
5. Roteiro prático – Oferece fichas rápidas de aplicação, destacando a importância da imediatividade do reforço.
6. Simulador rápido – Traz um recurso interativo que permite ao usuário selecionar uma situação e visualizar um roteiro simulado de resposta parental.
7. Checklist rápido para pais – Resume as etapas essenciais para a implementação do DRA.
8. Segurança e limites – Orienta sobre quando buscar acompanhamento profissional em casos de risco.

Aspectos Técnicos e de Design

- Linguagem: Apresenta boa clareza conceitual, porém utiliza termos técnicos que podem dificultar a compreensão por parte de pais e cuidadores.
- Design: a interface é simples e funcional, mas carece de apelo visual e de melhor organização dos elementos. O layout prioriza o conteúdo textual, resultando em uma apresentação que parece um pouco “carregada” e sem hierarquia visual definida.
- Estrutura semântica: uso adequado de cabeçalhos e hierarquia de conteúdo, favorecendo a acessibilidade.
- Interatividade: o simulador prático utiliza JavaScript para permitir a visualização dinâmica de roteiros comportamentais, favorecendo o aprendizado ativo.

Aplicativo 2 – Tutorial DRA para Pais: Crianças com TEA

O conteúdo busca apresentar o conceito de forma simplificada, com exemplos cotidianos, atividades interativas e um quiz para verificação de aprendizagem. O formato é dinâmico e mais detalhado em comparação ao primeiro aplicativo, incorporando seções voltadas à prática em família e ao treino por simulação.

Estrutura e Conteúdo: O aplicativo está organizado em nove seções principais, apresentadas em cartões informativos e interativos:

1. O que é o DRA? – Explica o conceito de forma introdutória, com exemplos de aplicação imediata.
2. Por que usar DRA? – Aponta benefícios e razões práticas para adoção do método, destacando segurança e redução de comportamentos desafiadores.
3. Passos simples para aplicar o DRA – Apresenta um guia com seis etapas que incluem identificação da função do comportamento, ensino da alternativa e reforço imediato.
4. Métodos de ensino – Descreve instrução, modelação e modelagem com feedback imediato, explicando como aplicá-las.
5. Exemplos práticos – Mostra situações do cotidiano (birras, agressividade e rotina) com roteiros passo a passo.
6. Atividade prática (role-play) – Encoraja os cuidadores a praticar os roteiros, com exemplos de fala e comportamento.

7. Dicas práticas e de segurança – Reforça a importância da consistência, segurança e colaboração entre cuidadores.
8. Perguntas de verificação (quiz) – Apresenta questões de múltipla escolha com feedback automático.
9. Referências bibliográficas – Lista autores clássicos e fontes relevantes sobre ABA e DRA.

Aspectos Técnicos e de Design

- Linguagem: apresenta avanços em relação ao primeiro aplicativo, utilizando explicações mais diretas e exemplos concretos. Entretanto, ainda há certa densidade técnica, o que pode dificultar a leitura de pais com menor familiaridade com termos da Análise do Comportamento.
- Design: o layout é limpo e agradável, com bom uso de cores suaves e tipografia legível, mas a organização visual ainda se mostra simples e um pouco fragmentada, com blocos extensos de texto. Faltam elementos que criem hierarquia visual mais clara entre etapas e exemplos.
- Interatividade: o quiz e a atividade de role-play são pontos positivos, estimulando a prática ativa e o aprendizado aplicado.
- Usabilidade: boa responsividade e fluidez, embora com pouca diferenciação visual entre seções.

Análise do Teste (Quiz de Verificação)

O aplicativo inclui um teste interativo composto por quatro questões de múltipla escolha, com feedback automático imediato após cada resposta.

O quiz tem como principal objetivo avaliar a compreensão conceitual dos pais sobre o DRA e sua aplicação prática, reforçando o aprendizado ativo.

As perguntas abordam:

- O conceito de DRA e sua função (reforçar o comportamento alternativo).
- A escolha de respostas adequadas a situações cotidianas (ex.: birras e agressividade).
- A importância do reforço imediato e da modelagem durante o ensino.
- A postura adequada do cuidador diante de comportamentos desafiadores.

O recurso utiliza cores e mensagens automáticas (“Correto”, “Incorreto”, “Dica”) para guiar o usuário, tornando o aprendizado mais envolvente.

Apesar disso, o formato é visualmente simples e pouco atrativo, sem elementos visuais que indiquem progresso, pontuação ou estímulo de continuidade.

Perguntas de verificação (múltipla escolha)

Responda para verificar sua compreensão. Receba feedback imediato.

1. O que é DRA?

A) Punir o comportamento indesejado até ele desaparecer.

B) Reforçar um comportamento alternativo que atende à mesma função.

C) Ignorar todas as formas de comportamento.

Correto — boa aplicação!

2. Em uma birra para conseguir um brinquedo, qual alternativa é adequada?

A) Ensinar a criança a pedir o brinquedo com uma palavra ou gesto e reforçar quando pedir.

B) Dar o brinquedo imediatamente para evitar a birra.

C) Levar a criança embora sem explicar nada.

Resposta incorreta. Dica: pense na função do comportamento e escolha a alternativa que a substitui.

3. Qual é a melhor prática ao ensinar uma alternativa?

A) Esperar semanas para reforçar — reforço deve ser lento.

B) Reforçar apenas com críticas para corrigir.

C) Modelar, pedir imitação e reforçar imediatamente os esforços.

Resposta incorreta. Dica: pense na função do comportamento e escolha a alternativa que a substitui.

4. Se a criança bater quando está frustrada, o que fazer?

A) Recompensar por bater para reduzir a frustração.

B) Ensinar e reforçar uma alternativa (ex.: pedir ajuda) e dar atenção quando a alternativa for usada.

C) Brigar com a criança para mostrar que não é aceitável.

Correto — boa aplicação!

Você acertou 2 de 4 perguntas.

Figura 4. Tela de perguntas de verificação do aplicativo-tutorial

Nota. Captura de tela da seção “Perguntas de verificação (múltipla escolha)” do aplicativo, com feedback imediato após cada resposta (“Correto — boa aplicação!” ou “Resposta incorreta”) e o placar final “Você acertou 2 de 4 perguntas.”.

Aplicativo 3 — Tutorial DRA para Pais (Versão com Base Científica MacNaul & Neely, 2017)

O objetivo é orientar os pais sobre como aplicar o DRA de maneira segura e estruturada, oferecendo explicações, exemplos de situações reais, um espaço de autoavaliação, e um teste automatizado de múltipla escolha para reforçar o aprendizado.

Estrutura e Conteúdo

O aplicativo está organizado em dez seções principais, divididas entre o corpo principal e uma área lateral interativa:

1. O que é o DRA? – Explica o conceito de reforçamento diferencial, destacando a importância de reforçar comportamentos alternativos em vez de punir ou ignorar.
2. Por que o DRA pode ajudar? – Lista os benefícios práticos do método e apresenta dados de pesquisa que sustentam sua eficácia.
3. Como funciona — passo a passo – Apresenta cinco etapas práticas para identificar a função do comportamento, ensinar a alternativa e aplicar o reforço de forma imediata e ajustada.
4. Exemplos práticos – Mostra situações cotidianas como birras, agressividade e transições difíceis, com descrições claras de como ensinar e reforçar a alternativa.

5. Dicas práticas para pais – Reúne orientações objetivas sobre reforço rápido, calma emocional e segurança durante a aplicação.
6. Espaço de treino (roteiro) – Guia o usuário para praticar o DRA em casa, com etapas graduais e instruções diretas.
7. Autoavaliação rápida – Ferramenta interativa em formato de escala (“não sei”, “com ajuda”, “sozinho(a)”), que gera automaticamente um feedback orientador com base nas respostas do usuário.
8. Teste (múltipla escolha) – Avaliação automática de quatro perguntas sobre conceitos centrais e práticas do DRA.
9. Perguntas conceituais e situacionais – Complementam o teste principal, reforçando o entendimento sobre o conteúdo aplicado.
10. Referências – Apresenta a citação da revisão sistemática que fundamenta o aplicativo, com DOI e acesso direto.

Aspectos Técnicos e de Design

- Linguagem: mais técnica e próxima de textos científicos, com citações diretas de estudos e terminologia comportamental precisa. Apesar disso, o texto mantém explicações adicionais em linguagem simples, o que cria um bom equilíbrio entre rigor e acessibilidade.
- Design: o layout é mais organizado que nas versões anteriores, com estrutura em duas colunas (conteúdo e barra lateral), o que favorece a navegação e a leitura. Ainda assim, a interface permanece visualmente básica, com foco no conteúdo textual e poucos elementos gráficos.
- Interatividade: o destaque vai para o teste automático e a autoavaliação, que oferecem respostas personalizadas e reforçam o engajamento do usuário.
- Usabilidade: boa legibilidade e responsividade, com estrutura intuitiva e feedback imediato ao usuário após cada teste.

Análise do Teste e Autoavaliação

O aplicativo incorpora dois tipos de instrumentos avaliativos:

1. Autoavaliação rápida – Permite que o usuário reflita sobre sua confiança em aplicar o DRA. As respostas geram um retorno automático com orientações práticas, indicando se o cuidador deve buscar mais apoio profissional ou continuar praticando com segurança.
2. Teste de múltipla escolha (4 questões) – Avalia a compreensão dos princípios fundamentais do DRA, como:
 - Conceito de reforçamento alternativo.
 - Resultados da revisão de MacNaul & Neely (2017).
 - Estratégias adequadas de ensino e reforço.
 - Procedimentos em situações de risco (autoagressão, agressividade).

Cada resposta é corrigida automaticamente, e o sistema fornece o gabarito e explicações detalhadas após o envio, reforçando o aprendizado.

Teste (múltipla escolha)

1. O que é DRA?

- Reforçar apenas o comportamento-problema até ele desaparecer.
- Ensinar e reforçar um comportamento alternativo que substitua o comportamento-problema.
- Punir a criança sempre que o comportamento ocorrer.

2. Quando a revisão de MacNaul & Neely encontrou resultados positivos em estudos de DRA sem extinção?

- Em nenhum estudo.
- Em 9 dos 10 estudos incluídos (maioria mostraram efeitos positivos).
- Em todos os estudos, sem exceção.

3. Qual a melhor prática ao ensinar a alternativa?

- Reforçar rapidamente a alternativa e treinar em calma antes da crise.
- Ignorar a criança até que ela mude sozinha.
- Trocar punições por recompensas apenas após um mês.

4. Se o comportamento for perigoso

Figura 5. Seção do teste automatizado (múltipla escolha com gabarito e explicações)

Nota. Tela da seção de teste (múltipla escolha) do aplicativo-tutorial sobre DRA. Cada resposta é corrigida automaticamente e o sistema fornece o gabarito, com itens sobre o conceito de DRA e a revisão de MacNaul & Neely.

4 DISCUSSÃO

O presente estudo investigou a precisão e a utilidade de respostas geradas por um chatbot de inteligência artificial (ChatGPT-5) para apoiar pais de crianças com Transtorno do Espectro Autista (TEA) na prevenção e no manejo de comportamentos desafiadores, utilizando o Reforçamento Diferencial de Comportamentos Alternativos (DRA). Para isso, foram aplicados três prompts diferentes, que variaram quanto à complexidade das instruções e às demandas de produção (explicação conceitual, exemplos práticos, atividades, testes avaliativos e referências). As respostas foram avaliadas segundo critérios de corretude científica, clareza, concisão, organização, acessibilidade e aplicabilidade prática.

De forma geral, os resultados indicaram que os três aplicativos produzidos apresentaram alta precisão conceitual, com alinhamento consistente aos princípios da Análise do Comportamento e às evidências sobre o DRA. No entanto, houve diferenças importantes entre eles: o Aplicativo 1 mostrou-se conceitualmente correto, mas com linguagem excessivamente técnica; o Aplicativo 3 apresentou a estrutura mais completa e o maior rigor científico, porém manteve alto nível de complexidade textual; e o Aplicativo 2 destacou-se por combinar precisão, clareza e aplicabilidade, obtendo o melhor equilíbrio entre linguagem acessível, exemplos contextualizados e organização pedagógica. Essas diferenças confirmam que a formulação do prompt exerce influência direta sobre a qualidade e a utilidade do conteúdo gerado pela IA.

Além disso, a avaliação funcional dos aplicativos HTML demonstrou que todos apresentaram boa responsividade, navegação intuitiva e coerência interna. Entretanto, o Aplicativo 2 e o Aplicativo 3 incorporaram elementos interativos (quiz, roteiros e autoavaliação) que ampliaram seu potencial pedagógico e facilitaram o aprendizado ativo por parte dos pais.

Assim, o estudo mostra que ferramentas de IA generativa podem produzir materiais educativos de alta qualidade para o apoio parental no TEA, desde que bem direcionadas por prompts adequados. Ao mesmo tempo, evidencia que a linguagem e a estrutura do comando influenciam diretamente a clareza e a aplicabilidade das respostas, destacando a importância do prompt engineering na construção de tecnologias assistivas mais eficazes. Nesse sentido, o uso de tecnologias baseadas em inteligência artificial, como chatbots, tem sido apontado como uma alternativa promissora para ampliar o acesso à orientação parental, especialmente diante de barreiras como tempo, custo e disponibilidade de profissionais (Entenberg et al., 2021).

Comparação com a Literatura

Os resultados deste estudo corroboram, expandem e, em alguns pontos, tensionam aspectos discutidos na literatura sobre o uso de IA generativa em contextos clínicos, educativos e parentais. Primeiro, os achados corroboram Sampaio et al. (2024), que afirmam que a qualidade da resposta da IA depende significativamente da forma como o comando é estruturado.

A relevância de ferramentas que auxiliem pais no manejo de comportamentos desafiadores também é evidenciada por pesquisas sobre capacitação parental. Estudos indicam que intervenções voltadas ao ensino de estratégias comportamentais para cuidadores podem reduzir comportamentos disruptivos em indivíduos com TEA e melhorar indicadores emocionais dos próprios cuidadores. Nesse sentido, recursos digitais que ampliem o acesso a orientações baseadas em evidências podem funcionar como complemento a programas tradicionais de capacitação parental (Bagaiolo et al., 2018).

No presente estudo, o Aplicativo 2, derivado do prompt mais completo, claro e detalhado, obteve as melhores pontuações em clareza, concisão, usabilidade e aplicabilidade, confirmando diretamente que prompts robustos geram respostas superiores. Isso reforça a tese de que a IA não produz conhecimento original, mas reorganiza informações conforme a direção recebida, e que a engenharia de prompts é central para resultados confiáveis

Os achados também corroboram McFayden et al. (2024), que identificaram alta correteza conceitual e consistência geral nas respostas da IA sobre autismo, porém com variações importantes na clareza e adequação ao público. Da mesma forma, neste estudo, todos os aplicativos apresentaram alta precisão conceitual (notas 4/4 em correteza), mas houve grandes diferenças na clareza: enquanto o Aplicativo 2 foi acessível, o Aplicativo 1 e o 3 apresentaram linguagem mais técnica, confirmando a oscilação prevista pelos autores. Além disso, os resultados expandem a literatura ao demonstrar, de forma aplicada e empírica, como diferentes prompts impactam diretamente a qualidade de materiais educativos completos em HTML, e não apenas textos isolados — um campo ainda pouco explorado. Estudos anteriores, como Kim et al. (2025), testaram respostas em linguagem natural, mas não avaliaram a construção de aplicativos funcionais, com quiz, autoavaliação e navegação real. Assim, este estudo amplia o entendimento sobre a utilidade prática da IA como produtora de ferramentas pedagógicas complexas para pais de crianças com TEA.

O estudo também expande a literatura ao aplicar critérios de avaliação estruturados (3 Cs e rubrica de 6 aspectos) de forma sistemática. Enquanto pesquisas anteriores avaliaram clareza e completude de forma qualitativa, esta investigação introduz um método replicável com escalas claras, contribuindo para avaliações mais objetivas em estudos futuros sobre IA. Por outro lado, os resultados tensionam parcialmente a literatura. Kim et al. (2025) afirmam que as respostas da IA tendem a apresentar baixa contextualização para situações reais. Neste estudo, porém, os prompts mais robustos geraram aplicativos com exemplos aplicados, atividades de role-play, autoavaliação e testes automáticos/mostrando que tais limitações podem ser superadas com instruções específicas. Assim, embora a limitação exista, ela não é absoluta; depende do tipo de comando utilizado. Outro ponto de tensão é que parte da literatura (Sampaio et al., 2024) enfatiza o risco de respostas persuasivas porém incorretas (alucinações). Contudo, na presente pesquisa, não foram identificadas alucinações conceituais relevantes: o modelo demonstrou alinhamento consistente à ABA e ao DRA. Isso não contradiz a literatura, mas oferece nuances importantes: em contextos bem delineados, com instruções precisas, o risco de erro diminui significativamente.

Contribuições metodológicas

Este estudo apresenta contribuições metodológicas relevantes, especialmente ao demonstrar como diferentes formulações de prompts influenciam a qualidade das respostas geradas pela IA. A aplicação de três prompts distintos permitiu identificar variações claras na precisão conceitual, clareza e aplicabilidade prática dos conteúdos, reforçando empiricamente o papel do prompt engineering como variável determinante no desempenho dos modelos generativos.

Outra inovação importante foi a adaptação de conteúdos conceituais, textos instrucionais e questões avaliativas para o formato de aplicativos HTML. Diferentemente de estudos que analisam apenas respostas textuais, esta pesquisa avaliou a capacidade da IA de transformar conhecimento científico em ferramentas digitais completas — incluindo exemplos práticos, roteiros, autoavaliação e quizzes. Essa abordagem amplia o entendimento sobre o uso da IA não apenas como produtora de texto, mas como construtora de materiais psicoeducativos aplicáveis.

O estudo também utilizou duas rubricas complementares — uma para avaliar o conteúdo (corretude, clareza e concisão) e outra para avaliar os aplicativos (organização, linguagem, usabilidade, acessibilidade, pertinência e aplicabilidade). Esse sistema de avaliação estruturado aumenta a objetividade e a reprodutibilidade do método.

Além disso, a verificação da autenticidade das referências citadas pela IA contribuiu para discutir a confiabilidade das fontes geradas pelo modelo, identificando quando a IA apresenta citações corretas ou incompletas.

Assim, o estudo oferece um protocolo metodológico conciso e replicável, útil para futuras pesquisas sobre o uso da IA em intervenções digitais e materiais educativos para famílias de crianças com TEA.

Contribuições Aplicadas

Os resultados deste estudo mostram que o aplicativo gerado pela IA tem potencial real como ferramenta de apoio para pais de crianças com TEA. Além de apresentar conteúdo conceitualmente correto sobre o DRA, o aplicativo organiza essas informações em um formato digital acessível, interativo e de fácil navegação. Recursos como exemplos práticos, checklists, roteiros de atuação, simuladores e quizzes tornam o material mais didático e favorecem o aprendizado imediato, permitindo que os cuidadores consultem e pratiquem as estratégias diretamente em situações do cotidiano.

O aplicativo também pode funcionar como complemento às intervenções clínicas, reforçando orientações já trabalhadas em terapia e ampliando a autonomia dos cuidadores. Sua disponibilidade instantânea e responsiva facilita o uso em momentos de necessidade, tornando-o uma alternativa prática e escalável para ampliar a disseminação de estratégias baseadas em evidências. Dessa forma, o estudo demonstra que a IA pode contribuir para a criação de recursos digitais úteis e aplicáveis ao manejo preventivo de comportamentos desafiadores.

Limitações e pesquisas futuras

Este estudo apresenta algumas limitações que precisam ser consideradas. A análise concentrou-se em um único modelo de IA (ChatGPT-5), impossibilitando comparações com outras tecnologias ou versões que poderiam produzir resultados diferentes. Além disso, embora os aplicativos tenham sido avaliados quanto à clareza, precisão e organização, não houve participação direta de pais de crianças com TEA, o que limita a compreensão sobre a eficácia real dos materiais no cotidiano familiar. A pesquisa também não investigou de forma aprofundada a ocorrência de eventuais “alucinações” do modelo, apesar da verificação de referências, o que significa que respostas incorretas podem emergir em outros contextos. Outro ponto é que a avaliação foi conduzida por um pesquisador com formação específica, podendo divergir da percepção de cuidadores leigos sobre clareza e aplicabilidade. Finalmente, os aplicativos foram testados apenas em ambiente de simulação, sem acompanhamento prolongado de uso.

Diante dessas limitações, pesquisas futuras podem ampliar a investigação incluindo pais, cuidadores e profissionais como avaliadores diretos dos aplicativos, possibilitando analisar impacto real, aplicabilidade prática e compreensão do conteúdo. Estudos comparativos entre diferentes modelos de IA também seriam relevantes para identificar variações de consistência, clareza e confiabilidade. Investigações longitudinais poderiam examinar o efeito do uso contínuo desses aplicativos na aprendizagem dos cuidadores e na aplicação correta do DRA em situações reais. Além disso, pesquisas futuras podem explorar recursos de personalização automática, permitindo que a IA ajuste exemplos, instruções e níveis de detalhe ao perfil da família. Por fim, recomenda-se o desenvolvimento de métodos mais robustos para detectar e mitigar possíveis erros conceituais ou referências imprecisas geradas pela IA, aumentando a segurança no uso educacional e clínico dessas ferramentas.

5 CONCLUSÃO

O presente estudo investigou a precisão e a utilidade de materiais educativos gerados por um chatbot de inteligência artificial para apoiar pais de crianças com Transtorno do Espectro Autista (TEA) na aplicação do Reforçamento Diferencial de Comportamentos Alternativos (DRA). Os resultados demonstraram que os três aplicativos produzidos apresentaram alta precisão conceitual e alinhamento com os princípios da Análise do Comportamento Aplicada. No entanto, foram observadas diferenças relevantes quanto à clareza e à acessibilidade, destacando-se o Aplicativo 2 como o mais equilibrado entre rigor teórico e linguagem acessível, enquanto o Aplicativo 3 apresentou maior profundidade, porém com maior complexidade textual.

Esses achados reforçam a influência direta da formulação do prompt na qualidade das respostas geradas pela inteligência artificial, evidenciando que instruções mais bem estruturadas resultam em conteúdos mais claros e aplicáveis ao contexto parental. Além disso, o estudo demonstra que a IA pode produzir não apenas textos, mas também materiais educativos

interativos, com potencial para apoiar o ensino de estratégias comportamentais de forma acessível. Apesar disso, limitações como a ausência de validação com usuários reais e a restrição a um único modelo de IA indicam a necessidade de investigações futuras.

De modo geral, conclui-se que a inteligência artificial, quando utilizada de forma orientada, pode atuar como uma ferramenta complementar no apoio a pais de crianças com TEA, ampliando o acesso a informações baseadas em evidências e favorecendo a aplicação de práticas preventivas no cotidiano, sem substituir a atuação de profissionais especializados.

REFERÊNCIAS BIBLIOGRÁFICAS

- American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). American Psychiatric Publishing.
- Bagaiolo, L. F., Pacífico, C. R., Moya, A. C. C., Mizael, L. F., Jesus, F. S., Zavitoski, M., Sasaki, T., & Asevedo, G. R. C. (2018). Capacitação parental para comunicação funcional e manejo de comportamentos disruptivos em indivíduos com transtorno do espectro autista. *Cadernos de Pós-Graduação em Distúrbios do Desenvolvimento*, 18(2), 46–64. <https://doi.org/10.5935/cadernosdisturbios.v18n2p46-64>
- Carvalho, A. D. de, & Moreira, M. B. (2022). Reforçamento diferencial de comportamentos alternativos como método de prevenção e manejo de comportamentos desafiadores no autismo. *Walden4*.
- Entenberg, G., Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2021). Conversational agents (chatbots) in mental health and parenting interventions: A systematic review. *Social Sciences*, 10(11), 426. <https://doi.org/10.3390/socsci10110426>
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Pearson/Merrill-Prentice Hall.
- Goulart, P., & Assis, G. J. A. de. (2002). Estudos sobre autismo em análise do comportamento: Aspectos metodológicos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 4(2), 165–176.
- Guimarães, M., Baptista, J., & Seabra-Santos, M. J. (2018). Parent training on the management of disruptive behavior problems in children with autism spectrum disorder. *Análise Psicológica*, 36(1), 75–87. <https://doi.org/10.14417/ap.1374>
- Hodgdon, L. A. (1995). *Visual strategies for improving communication*. QuirkRoberts Publishing.
- Kim, J. (2025). Testing the capability of generative artificial intelligence for parent and caregiver support. *Family Relations*, 74(1), 45–62.
- Leaf, R., & McEachin, J. (1999). *A work in progress: Behavior management strategies and a curriculum for intensive behavioral treatment of autism*. DRL Books.
- MacNaul, H. L., & Neely, L. C. (2018). Systematic review of differential reinforcement of alternative behavior without extinction for individuals with autism. *Behavior Modification*, 42(3), 398–421. <https://doi.org/10.1177/0145445517740321>
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135–148. <https://doi.org/10.1089/cyber.2023.0202>
- Miltenberger, R. G. (2011). *Behavior modification: Principles and procedures* (5th ed.). Wadsworth Cengage Learning.
- Miltenberger, R. G. (2016). *Behavior modification: Principles and procedures*. Cengage Learning.

Ministério da Saúde. (2014). Diretrizes de atenção à reabilitação da pessoa com transtorno do espectro do autismo (TEA). Autor.

Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: Reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432e008>

Vollmer, T. R., Peters, K. P., Kronfli, F. R., Lloveras, L. A., & Ibañez, V. F. (2020). On the definition of differential reinforcement of alternative behavior. *Journal of Applied Behavior Analysis*, 53(3), 1299–1303.

Wang, J., Liu, Z., Zhao, L., Wu, Z., Ma, C., Yu, S., Dai, H., Yang, Q., Liu, Y., Zhang, S., Shi, E., Pan, Y., Zhang, T., Zhu, D., Li, X., Jiang, X., Ge, B., Yuan, Y., Shen, D., & Zhang, S. (2024). Review of large vision models and visual prompt engineering. *Information Fusion*, 104, 102187. <https://doi.org/10.1016/j.inffus.2024.102187>

EDITORA WALDEN4 · LIVROS

Editora Walden4 — sobre o ombro de gigantes



Publicações especializadas em Análise do Comportamento e práticas baseadas em evidências científicas, conectando a ciência psicológica ao público.

- 13 anos de experiência · 62 livros publicados
- Da sua ideia à publicação, com curadoria editorial
- Disponível no Google Play e em PDF

walden4.com.br/livros

CAPÍTULO

05

Efeito de diferentes níveis de prompt na adequação de aplicativos educacionais de inteligência artificial (IA) para o ensino de alfabetização a crianças autistas na educação básica



Ana Luiza Cardoso Cerqueira

Márcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo: Este presente estudo tem como objetivo pesquisar o efeito de diferentes níveis de prompt sobre a adequação de um aplicativo educacional desenvolvido por inteligência artificial para o ensino de rudimentos de leitura a crianças autistas na educação básica. Também busca-se a compreensão de como a variação no detalhamento e clareza das ordens estabelecidas nas inteligências artificiais influencia a qualidade das respostas na pertinência pedagógica e a eficácia do conteúdo fornecido. O estudo ainda pretende analisar se o conteúdo produzido atende adequadamente às demandas requeridas ao público alvo, considerando as individualidades cognitivas e socioemocionais.

Método: O estudo é qualitativo, onde se observa a qualidade das respostas geradas por uma inteligência artificial pelo aplicativo Chatgpt a partir de três níveis de prompt (genérico, intermediário e detalhado). Os conteúdos gerados pela IA foram avaliados por meio de quatro categorias: pertinência pedagógica, clareza linguística, coerência das atividades e interatividade/engajamento.

Resultados: Os resultados prompts genéricos geraram conteúdos mais simples, com conteúdos mais rasos e superficiais, os prompts intermediários apresentaram uma qualidade mediana com um pouco mais de variedade e os prompts detalhados geraram resultados mais satisfatórios onde os aplicativos são mais completos, detalhados e adequados ao público autista infantil.

Conclusão. As inteligências artificiais têm avançado de forma significativa no nosso cotidiano e na forma como interagimos com tecnologias e também pode contribuir na criação de atividades mais inclusivas e adaptadas, porém isso vai depender do nível de detalhamento que uma pessoa trazer e das instruções. Prompts bem aplicados são capazes de oferecer respostas mais satisfatórias, reforçando a importância da melhor comunicação no uso da IA para o melhor desenvolvimento de crianças autistas.

Palavras-chave: Aplicativos educacionais; Inteligência Artificial; Ensino de alfabetização; Transtorno do Espectro Autista.

1 INTRODUÇÃO

Contextualização do Tema

O Transtorno do Espectro Autista (TEA) é um termo "guarda-chuva" na qual é utilizado para descrever as dificuldades e comprometimentos no neurodesenvolvimento, especialmente na socialização, a comunicação e manifestação de comportamentos estereotipados (Hassona et

al., 2024). Indivíduos com TEA também podem apresentar dificuldades em como expressar suas necessidades, emoções e realizar outras atividades cotidianas. A intensidade desses fatores dependerá do nível de suporte que um determinado indivíduo pode se enquadrar tanto no primeiro nível de suporte, segundo ou terceiro nível. De acordo com as dificuldades citadas anteriormente muitos pais têm procurado diversas estratégias e tratamentos eficazes para suas crianças para se desenvolverem da melhor forma e uma dessas estratégias incluí o uso de inteligências artificiais como o "Chatgpt" para auxiliar por exemplo na desenvoltura da fala para crianças autistas não verbais ou que apresentam dificuldade na fala de forma significativa.

Diante do cenário atual podemos ver que o uso das inteligências artificiais tem sido cada vez mais rápido e acessível podendo ser acessados por diversos indivíduos inclusive pais e responsáveis que buscam compreender melhor como podem prestar o melhor suporte para os seus filhos e colaborar para o seu pleno desenvolvimento. Nesse contexto de avanço e uso crescente de tecnologias, estudos como os realizados pelo (Kim et al., 2025) analisou a qualidade das respostas estabelecidas como o Chatgpt, e colocou perguntas comuns de pais e cuidadores sobre o desenvolvimento infantil. Com o avanço do uso de tecnologias para fornecer informações sobre parentalidade e com isso é fundamental avaliar sua eficácia e confiabilidade dos conteúdos fornecidos, Os autores aplicaram 11 testes abordando cinco tópicos frequentes relacionados ao cuidado de crianças pequenas, utilizando tanto fontes controladas, como boletins informativos profissionais, quanto fontes abertas, como modelos de linguagem de IA. A análise considerou critérios como: corretude, clareza e conexão.

Problema de Pesquisa

O uso das IAs também podem carregar uma série de consequências e prejuízos tanto na educação quanto nas informações que enviam para terceiros que utilizam de IAs para sanar seus problemas e dúvidas (Sampaio et al., 2024) alguns dos problemas apresentados são: Vieses dos desenvolvedores e das bases de treinamentos, possibilidades de compartilhar artigos e pesquisas sem sequer existirem, falta de transparência de algoritmos, impossibilidade de replicação de certas respostas. O avanço das inteligências artificiais generativas tem ampliado o acesso a informações em diferentes áreas, incluindo a educação e o suporte a pessoas portadoras de necessidades especiais. No entanto, surge um problema que tem se discutido atualmente: até que ponto é possível confiar na qualidade, precisão e adequação das respostas fornecidas por essas ferramentas? A confiança nos resultados gerados pela IA é abalada por fatores como a ausência de referências confiáveis, a tendência a produzir respostas genéricas e, em alguns casos, a possibilidade de apresentar informações incorretas como se fossem verdadeiras.

Na educação inclusiva, esse problema torna-se ainda mais severo. No caso de crianças autistas, o uso de aplicativos baseados em IA requer cautela, pois informações imprecisas ou referências incorretas podem comprometer o processo de aprendizagem e reforçar práticas inadequadas. Assim, a confiança não se restringe à qualidade técnica das respostas, mas envolve também a verificação da veracidade das referências, bem como a segurança pedagógica e ética no uso dessas tecnologias.

Pergunta de Pesquisa

Considerando que o avanço das inteligências artificiais generativas tem proporcionado novas possibilidades para a educação, inclusive no âmbito da educação inclusiva; que essas ferramentas podem apoiar práticas pedagógicas ao oferecer recursos personalizados de ensino; e que, ao mesmo tempo, existem desafios relacionados à confiabilidade das respostas, à coerência das referências e à adequação pedagógica de seus conteúdos, especialmente quando aplicados a crianças autistas na fase de alfabetização, torna-se relevante investigar: “Como diferentes níveis de prompt influenciam a adequação de um aplicativo educacional criado por inteligência artificial para o ensino de rudimentos de leitura a crianças autistas na educação básica?”

Relevância da Pesquisa

A relevância desta presente pesquisa está na crescente utilização de inteligências artificiais como ferramenta para o desenvolvimento de recursos e materiais educacionais, especialmente voltados para públicos com necessidades especiais específicas, como crianças portadoras do TEA no ensino regular. Investigar como diferentes níveis de prompt influenciam a adequação pedagógica de aplicativos educacionais contribui para a compreensão das potencialidades e limitações da IA no contexto escolar, oferecendo suporte e auxílio para educadores, desenvolvedores e profissionais da psicologia educacional. Além disso, a pesquisa pode auxiliar na criação de ferramentas com os objetivos de serem mais eficazes, seguras e inclusivas, promovendo práticas pedagógicas fundamentadas no conhecimento científico sobre o autismo e garantindo que os recursos digitais também possam complementar, de forma responsável, o processo de aprendizagem.

Revisão da Literatura

O uso de inteligência artificial generativa, como o ChatGPT, tem se consolidado como uma ferramenta mais rápida e acessível em diversos contextos, incluindo a educação e a pesquisa científica. (Sampaio et al., 2024) destacam que plataformas baseadas em IA estão transformando a forma como se produz e se acessa o conhecimento, oferecendo respostas rápidas e personalizadas, mas também trazem um importante aviso para a necessidade de avaliação rigorosa quanto à precisão e adequação do conteúdo gerado. O êxito da utilização de inteligência artificial para o público com TEA está diretamente relacionado a relevância e clareza do conteúdo gerado conforme a discussão trazida anteriormente por um dos autores, a adequação desses materiais feitas por IA está diretamente relacionado à forma de instrução do sistema, o que justifica a análise da variabilidade de prompts que determina a qualidade das respostas e sua aplicabilidade prática na educação básica.

No contexto de informações voltadas a cuidadores e familiares, (Kim et al., 2024) estudaram a capacidade da IA generativa de fornecer orientações no contexto parental. A pesquisa mostrou que apesar das respostas possam ser úteis como apoio inicial, elas apresentam limita-

ções significativas quanto à personalização e profundidade, sendo necessárias intervenções humanas para garantir que a informação seja contextualizada e adequada à realidade particular de cada criança.

Outros estudos como o de (Hassona et al., 2024), demonstram a aplicação prática da IA em conversas sobre o autismo, mostrando que ferramentas como o Chatgpt podem oferecer informações e sugestões de estratégias educativas. Contudo, a pesquisa evidencia que a interação com a IA depende de uma comunicação precisa e estruturada: respostas genéricas ou mal direcionadas podem gerar interpretações inadequadas ou pouco úteis para famílias e profissionais. Outros estudos apresentados, como o de (McFayden et al., 2024), demonstraram a aplicação prática da IA pelo aplicativo “ChatGPT” em conversas sobre o TEA, investigando a capacidade de uma inteligência artificial em fornecer informações e sugestões de estratégias educativas, inclusive com a geração de múltiplas respostas e referências, ressaltando a importância da precisão na comunicação.

Além de perguntas e respostas as IAs generativas também são capazes produzir aplicativos de acordo com os prompts solicitados, essa possibilidade pode facilitar o trabalho de pais e profissionais transformando o conhecimento científico em tecnologia comportamental. No entanto, ainda não há pesquisas avaliando a validação desse uso pois o uso de inteligências artificiais de forma global ainda é algo recente. Dentre as diversas áreas de atuação no TEA temos os problemas relacionados à comunicação apresentando características como por exemplo ecolalia e apraxia. Uma área de pesquisa e intervenção é o ensino de leitura e escrita para crianças que se enquadram nesse público. De acordo com o presente contexto educacional, essas evidências indicam que os aplicativos voltados ao ensino de crianças autistas podem se beneficiar do uso de inteligência artificial, desde que haja atenção ao detalhamento dos prompts e à adequação pedagógica do conteúdo gerado. Assim, a literatura aponta tanto a capacidade quanto às limitações dessas tecnologias, reforçando a importância de pesquisas que investiguem de forma sistemática como diferentes níveis de prompts influenciam a eficácia e a relevância dos recursos educacionais baseados em IA.

Por exemplo, de acordo com as autoras (Alves et al., 2016), sobre o ensino para aprendizes com autismo em relação a sílabas simples, leitura combinatória e leitura e leitura com compreensão (2016), elas tinham como objetivo avaliar a eficiência da aplicação da combinação de leitura oral e leitura com compreensão em crianças diagnosticadas com TEA visando desenvolver com precisão a habilidade de ler palavras formadas por sílabas simples e interpretá-las corretamente. Participaram da pesquisa: Três crianças do sexo masculino com autismo leve, não alfabetizados, com idades entre 5 anos e 9 meses e 9 anos e 9 meses. Todos obtinham comunicação verbal e eram estudantes de escolas comuns. Durante o procedimento o ensino foi realizado em sessões realizadas apenas uma vez por semana com aproximadamente 50 minutos de duração, totalizando entre 15 e 26 sessões por candidato, foi estruturado em seis conjuntos de ensino onde o primeiro envolvia o ensino de nomeação primeiro de sílabas e depois de palavras, as sílabas simples do tipo consoante-vogal foram organizadas em seis conjuntos, cada um com grupos silábicos específicos e então a pesquisadora escrevia as sílabas e palavras com canetinhas coloridas (consoantes em preto, vogais em vermelho e bordas em azul) e ensinava a nomeação em voz alta, depois dessa etapa. As crianças aprendiam a associar palavras impressas em cores com palavras impressas em preto, para generalizar a lei-

tura oral e por fim. As crianças eram ensinadas a nomear oralmente figuras relacionadas às palavras aprendidas, antes e após cada etapa eram realizados testes avaliando a forma como as crianças absorveram os conteúdos aprendidos nas etapas e sondas de leitura recombinativa avaliando a capacidade dos participantes de ler palavras novas formadas por sílabas recombinadas das palavras ensinadas. Como conclusão, o procedimento foi eficaz para ensinar leitura combinatória com compreensão, com poucas sessões de ensino e baixo número de erros e também apresentaram melhora significativa na leitura oral e na compreensão de palavras simples.

Tabela 1. Organização do conteúdo pedagógico do programa de leitura, por conjuntos de palavras

Conjuntos	Grupos	Palavras: caderno e fichário	Figuras e palavras: testes	Recombinação
T, L, M	T	tatu, teto, ti, tutu, tio, iate		
	L	luta, lei, tela, leite, lua, lata	lua, mola, tatu, tela, tomate, mato	
	M	mola, mato, tomate, mala, meia, muleta		moto, muito, lima, remo, batata, fita, tabela, fubá, rua, abalou
F, B, R	F	fila, fome, mofo, folia, fatia, família		
	B	bola, baú, bife, bateu, bela, baleia	fila, bola, bife, baleia, rato, rio	
	R	rato, rio, rima, rifa, rabo, rolo		
P, N, V	P	pato, mapa, apito, pipa, palito, pia		
	N	nome, pena, menina, panela, neto, aluno	pato, aluno, pena, menina, nove, vela	
	V	nove, luva, novela, vovó, vela, violeta		pulo, tijolo, nova, roda, sabiá, pifou, semana, viola e feijoadada
S, D, J	S	sapo, sofá, sala, sapato, sete, sabonete		
	D	dado, roda, dia, sala, idade, data	sapo, sapato, dados, roda, janela, jiló	
	J	sujo, juba, jiló, pajé, beijo, janela		

Nota. Estrutura do programa de ensino de leitura. Para cada conjunto de consoantes (T, L, M; F, B, R; P, N, V; S, D, J) são indicados os grupos, as palavras de treino (caderno e fichário), as palavras e figuras de teste e as palavras de recombinação. Transcrição fiel da figura original.

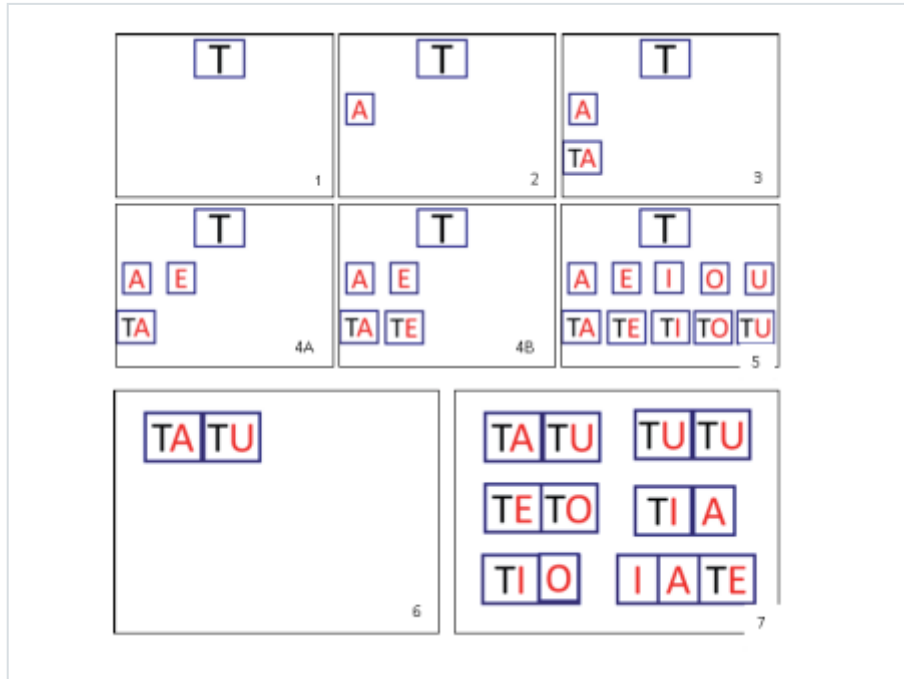


Figura 1. Esquema de combinação de letras em sílabas e palavras

Nota. Diagrama do material de alfabetização: consoantes (em destaque) e vogais são combinadas progressivamente para formar sílabas (TA, TE, TI, TO, TU) e, em seguida, palavras como TATU e TIA, ilustrando a lógica de recombinação.

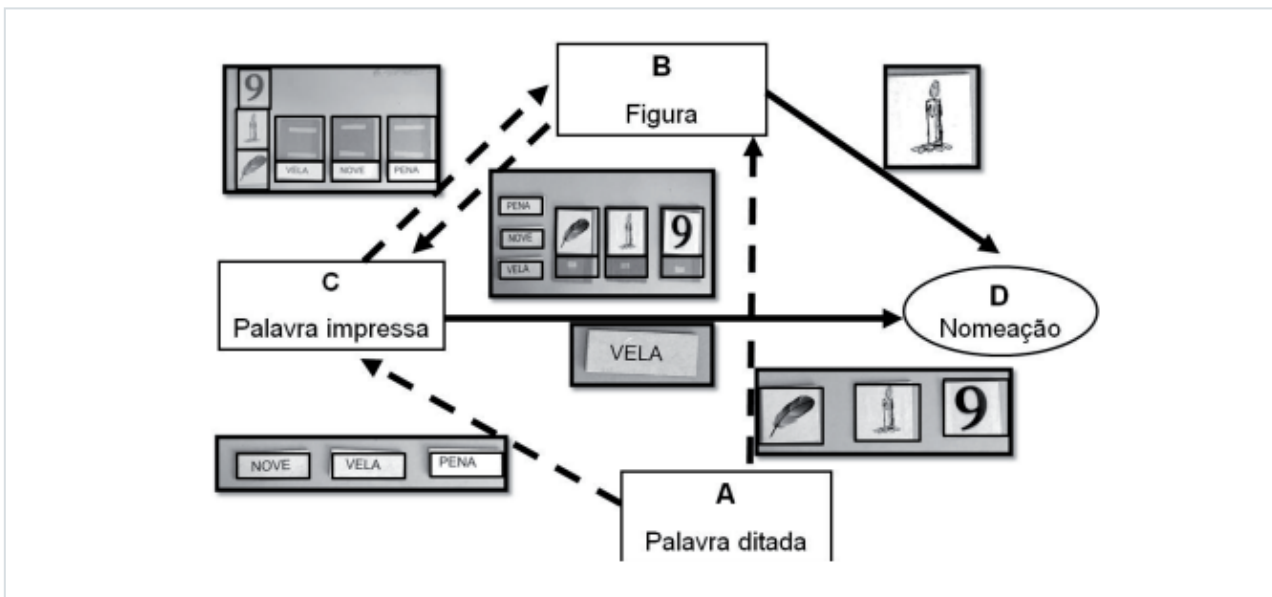


Figura 2.

Nota. Diagrama das relações condicionais do programa de leitura: entre a palavra ditada (A), a figura (B), a palavra impressa (C) e a nomeação (D). As setas cheias indicam as relações ensinadas e as tracejadas, as relações de recombinação testadas (emergência da leitura).

Figura 2. Mostra testes e figuras mostrando a relação entre a palavra escrita e a figura correspondente, essencial para garantir que a criança não estava apenas "decodificando" sons, mas compreendendo o significado do que lia. Extraída da pesquisa por (Alves et al., 2016).

Objetivo da Pesquisa

Esta pesquisa tem como objetivo analisar e estudar o efeito de diferentes níveis de prompt sobre a adequação de um aplicativo educacional desenvolvido por inteligência artificial para o ensino de rudimentos de leitura a crianças autistas na educação básica. Procura-se entender o processo de variação no detalhamento e na clareza das instruções fornecidas à inteligência artificial influenciando a qualidade, eficácia e relevância do conteúdo gerado, sua pertinência pedagógica e a eficácia das respostas. Além disso, a pesquisa pretende avaliar se o material e o conteúdo fornecido atende de forma adequada às necessidades do público-alvo.

2 MÉTODO

Delineamento de Pesquisa

O presente estudo é caracterizado como uma pesquisa exploratória-descritiva de abordagem qualitativa e quantitativa, com o objetivo de compreender como diferentes níveis de prompt influenciam a adequação de um aplicativo educacional desenvolvido por inteligência artificial para o ensino de rudimentos de leitura a crianças autistas na educação básica.

Será necessário então, de acordo com a verificação, gerar conteúdos educacionais pela IA com prompts em três níveis distintos: genérico, intermediário e detalhado. Para isso, serão gerados conteúdos educacionais pela IA com prompts em três níveis distintos: genérico, intermediário e detalhado. Esses conteúdos serão analisados por meio do método dos “3C’s” (Clareza, Coerência e Compreensibilidade), além de outras medidas, a fim de verificar sua adequação pedagógica e seu potencial de uso com o público-alvo. As observações serão submetidas à análise qualitativa, permitindo identificar categorias, padrões e nuances que evidenciam limitações, potencialidades e oportunidades de melhoria do aplicativo. Esse delineamento qualitativo possibilita investigar de forma aprofundada como a IA pode ser utilizada para gerar recursos educacionais e inclusivos.

Modelos Utilizados

Para a realização desta pesquisa, foi utilizada uma inteligência artificial do tipo modelo de linguagem de larga escala (Large Language Model – LLM), desenvolvida pela empresa OpenAI. O modelo escolhido foi o Chat GPT, baseado na arquitetura GPT-5, que utiliza redes neurais profundas treinadas com grandes volumes de dados textuais provenientes de diversas fontes, como artigos científicos, textos educacionais e materiais disponíveis publicamente na internet. O Chat GPT é capaz de compreender e gerar textos coerentes em linguagem natural a partir de comandos escritos, conhecidos como prompts. Essa característica o torna adequado para fins educacionais, permitindo a criação de conteúdos personalizados conforme as instruções fornecidas pelo usuário.

Durante a pesquisa, o modelo foi utilizado para gerar sugestões de atividades e recursos pedagógicos voltados ao ensino de rudimentos de leitura a crianças autistas. Os prompts foram elaborados em diferentes níveis de detalhamento, com o objetivo de analisar como a comple-

xidade das instruções influencia a qualidade, adequação e pertinência pedagógica das respostas. Nesta pesquisa será usado uma conta nova no Aplicativo “Chatgpt” v. 5 e serão aplicadas no contexto dos efeitos de diferentes níveis de prompt sobre a adequação de um aplicativo educacional criado por inteligência artificial para ensinar formas de melhorar a leitura para crianças autistas na educação básica.

Prompts Utilizados

Prompt Genérico (nível 1)

Características: instrução simples, pouco detalhada, sem especificações sobre idade, necessidades especiais ou estilo pedagógico. **Objetivo:** ensinar a ler um conjunto de 6 palavras (nomes de objetos e animais). **Descrição:** Esse prompt foi elaborado com base no modelo de pesquisa Kim et al. (2024) Que analisa o desempenho das inteligências artificiais diante de situações hipotéticas diante de instruções amplas. Esse tipo de prompt conta com uma linguagem mais básica permitindo avaliar a capacidade da IA gerar conteúdos educativos sem receber instruções sobre o público alvo ou objetivos pedagógicos.

Prompt: "Crie um aplicativo html educacional para ensinar leitura a crianças."

Prompt Intermediário (nível 2)

Características: instrução mais detalhada, incluindo faixa etária e objetivo de ensino, mas sem muitas orientações sobre o perfil de aprendizagem. **Objetivo:** ensinar a criança a ler letras, sílabas e palavras simples.

Prompt: "Crie um aplicativo html com atividades de leitura para crianças de 6 a 8 anos, incluindo exercícios de letras, sílabas e palavras simples." O procedimento da pesquisa foi estruturado em duas etapas principais: fase de treino e fase de teste, com o objetivo de observar o comportamento da inteligência artificial diante de diferentes níveis de prompt e avaliar a adequação do conteúdo gerado para o ensino de leitura a crianças autistas."

Na fase de treino, também foram elaborados três tipos de prompts (nível 1 – simples, nível 2 – intermediário e nível 3 – detalhado), solicitando que a inteligência artificial criasse atividades de leitura voltadas a crianças de 6 a 8 anos. Cada nível de prompt foi aplicado em múltiplas tentativas, permitindo que o modelo recebesse feedback qualitativo sobre a clareza, coerência e adequação pedagógica das respostas produzidas. Esse processo teve como objetivo melhorar a interação entre o pesquisador e o modelo, simulando um processo de aprendizado iterativo por parte da IA.

Na fase de teste, os prompts utilizados foram novamente aplicados, sem fornecimento de feedback, para verificar se as respostas apresentaram maior estabilidade e adequação após o período de treino. As respostas foram registradas e organizadas em categorias analíticas, de acordo com critérios de avaliação pedagógica, como: pertinência ao público-alvo, clareza das instruções, estrutura das atividades e coerência entre objetivo e conteúdo proposto.

Ao final, os dados gerados nas diferentes fases foram comparados qualitativamente, buscando identificar padrões, limitações e variações nas produções da IA em função do nível de detalhamento do prompt. Todo o processo foi conduzido com base em princípios metodológi-

cos de estudos exploratórios qualitativos sobre o uso de inteligências artificiais em contextos educacionais, conforme modelos descritos na literatura recente (Kim et al., 2024; Sampaio et al., 2023).

Prompt Detalhado (nível 3)

Características: instrução detalhada, incluindo idade, perfil de aprendizagem (crianças autistas), objetivos pedagógicos, estratégias inclusivas, exemplos de atividades e instruções claras de interação. **Objetivo:** Promover atividades educativas altamente personalizadas que possam desenvolver o desenvolvimento da leitura em crianças autistas de 6 até 8 anos com progressão pedagógica estruturada com adaptações para o contexto de cada criança visando a autonomia, atenção e engajamento sócio emocional, onde o aplicador também é capaz de acompanhar o progresso e fazer as adaptações necessárias.

Prompt: "Crie um aplicativo html educacional para ensinar leitura a crianças autistas de 6 a 8 anos, incorporando atividades que usem estímulos visuais, auditivos. Desenvolva exercícios progressivos de reconhecimento de letras, formação de sílabas e leitura de palavras simples, incluindo reforços visuais e auditivos. As instruções devem ser curtas, claras e interativas. Sugira mecanismos para monitorar o progresso individual da criança e permitir ajustes personalizados no ritmo de aprendizagem. Inclua estratégias inclusivas para manter a atenção, estimular autonomia e promover engajamento sócio emocional."

Avaliação do Conteúdo Gerado

Esta análise terá como referência os critérios de qualidade utilizados em estudos passados sobre a utilização das IA'S onde é possível encontrar nessas duas pesquisas objetivos com o intuito de explorar a qualidade dos prompts utilizados nas respostas de acordo com os questionamentos dos pais sobre a parentalidade e em outra pesquisa observar as respostas voltadas para pais que tinham seus questionamentos sobre o autismo e como lidar com o transtorno em seus filhos (Kim et al., 2025) e (McFayden et al., 2024). Essas pesquisas apresentam a relevância das categorias como clareza, correção, concisão, confiabilidade e legitimidade na análise das respostas feitas pelas inteligências artificiais que durante esse estudo, de acordo com esses conceitos serão aplicadas no contexto da criação de aplicativos para auxiliar educadores e pais em relação a crianças autistas.

Dessa forma, para que os objetivos deste estudo sejam alcançados, é necessário evidenciar esses parâmetros de avaliação para o desenvolvimento de ferramentas educacionais. A ponte entre a literatura citada e a presente análise reside na aplicação rigorosa dessas categorias no contexto da criação de aplicativos voltados a educadores e pais de crianças autistas.

Para avaliar a eficácia dos prompts e das respostas geradas, serão utilizados os seguintes critérios fundamentais:

Concisão: analisar se as respostas elaboradas pelas IA responde às demandas de forma objetiva, evitando repetições.

Clareza: verificar se há compreensão e coerência de acordo com as solicitações enviadas por educadores e responsáveis. **Correção:** analisar as respostas elaboradas pelas IA de acordo com o contexto descrito pelas solicitações e demandas pedagógicas.

Confiabilidade: observar se as informações trazidas pelas IA foram baseadas em pesquisas e fontes científicas validadas. Legibilidade: examinar a compreensão e a acessibilidade da linguagem descrita.

Além das categorias supracitadas, esta pesquisa empregará uma rubrica de avaliação expandida. Este crivo metodológico incluirá itens adicionais que permitem mensurar a adaptabilidade do conteúdo e o nível de personalização pedagógica, indo além das métricas padrão de processamento de linguagem natural.

Os resultados obtidos serão agrupados em tabelas comparativas, permitindo observar as variações de qualidade entre os aplicativos gerados a partir de diferentes tipos de prompts. Por fim, os dados serão interpretados à luz das pesquisas de referência, buscando identificar padrões de desempenho, bem como as limitações da IA na geração de recursos educacionais para contextos inclusivos.

Os resultados obtidos também serão agrupados em tabelas comparativas permitindo observar as diferenças de qualidade entre os aplicativos gerados a partir de cada tipo de prompt e os dados serão interpretados à luz das pesquisas de referência, buscando identificar padrões de desempenho da IA e possíveis limitações em relação à geração de aplicativos educacionais para contextos inclusivos.

Tabela 2. Apresenta uma análise qualitativa dos três aplicativos educacionais gerados pela inteligência artificial ChatGPT, com base em diferentes níveis de prompts: genérico (nível 1), intermediário (nível 2) e detalhado (nível 3). A avaliação foi realizada considerando quatro critérios principais: pertinência pedagógica, clareza linguística, coerência das atividades e interatividade/engajamento. Cada critério foi classificado em três níveis: Excelente (3), Bom (2) e Insatisfatório (1)

CRITÉRIO DE AVALIAÇÃO	EXCELENTE (3)	BOM (2)	INSATISFATÓRIO (1)
PERTINÊNCIA PEDAGÓGICA	Conteúdo totalmente alinhado aos objetivos de ensino; atividades coerentes e adequadas para crianças autistas.	Conteúdo parcialmente alinhado aos objetivos; algumas atividades relevantes.	Conteúdo pouco ou nada alinhado aos objetivos; atividades inadequadas.
CLAREZA LINGUÍSTICA	Texto claro, fácil de compreender e bem estruturado; instruções objetivas.	Texto compreensível, mas com pequenas ambiguidades ou instruções pouco claras.	Texto confuso, mal estruturado ou difícil de compreender.
ADEQUAÇÃO AO PÚBLICO-ALVO	Conteúdo adaptado às necessidades cognitivas, sócio emocionais e sensoriais de crianças autistas.	Conteúdo parcialmente adaptado; algumas necessidades do público-alvo não são contempladas.	Conteúdo não adaptado às necessidades específicas do público-alvo.
COERÊNCIA DAS ATIVIDADES	Atividades organizadas e sequenciais; há relação clara entre objetivo e exercício.	Algumas atividades coerentes; outras sem relação direta com os objetivos.	Atividades desorganizadas; sem relação aparente entre objetivo e exercício.
INTERATIVIDADE/ENGAJAMENTO	Conteúdo estimulante e interativo, com estratégias que promovem participação ativa da criança.	Conteúdo pouco interativo; engajamento limitado.	Conteúdo estático ou sem estratégias de engajamento.

Nota. Análise qualitativa dos três aplicativos de alfabetização segundo rubrica de três níveis (Excelente = 3, Bom = 2, Insatisfatório = 1) por critério.

Procedimento

A pesquisa foi elaborada em três etapas principais: elaboração dos prompts, geração e análise dos conteúdos e avaliação da adequação pedagógica desenvolvida.

Na primeira etapa, foram elaborados diferentes modelos de prompts, variando em nível de detalhamento e clareza, para solicitar à inteligência artificial (ChatGPT) a criação de atividades voltadas ao ensino de estratégias de leitura para crianças autistas na educação básica. Serão utilizados três níveis: prompts simples (gerais e curtos), prompts intermediários (com instruções moderadas) e prompts detalhados (com instruções completas, especificando o perfil do público-alvo, objetivos pedagógicos e exemplos de atividades).

Durante o procedimento proposto na segunda etapa, foram enviados prompts na IA para gerar aplicativos educacionais. Esses materiais serão coletados e organizados para análise comparativa, buscando identificar variações na qualidade, adequação e coerência das respostas em relação ao público-alvo. Foram aplicados com três níveis distintos de prompt: nível 1 (genérico), nível 2 (intermediário) e nível 3 (detalhado) com o intuito de observar as variações na qualidade e adequação pedagógica das respostas. Cada prompt foi inserido separadamente, em momentos distintos, garantindo que o modelo não mantivesse histórico de conversas anteriores, com o objetivo de evitar interferências nas respostas. A criação de cada aplicativo foi realizada em ambiente controlado, seguindo o mesmo procedimento: inserção do comando, geração do código HTML e registro das respostas produzidas.

Na terceira etapa, será realizada uma análise qualitativa dos conteúdos gerados, com base em critérios de pertinência pedagógica, clareza linguística e adequação às necessidades cognitivas e socioemocionais de crianças com autismo. Os resultados serão descritos de maneira interpretativa, dando destaque a padrões, limitações e potencialidades observadas nas produções da inteligência artificial em função dos diferentes níveis de prompt utilizados.

3 RESULTADOS

Coleta de Dados

A coleta de dados foi programada alternando entre os dias 8, 10 e 12 de outubro de 2025, utilizando a plataforma ChatGPT para a geração dos aplicativos educacionais em formato HTML. Foram aplicados com três níveis distintos de prompt: nível 1 (genérico), nível 2 (intermediário) e nível 3 (detalhado) com o intuito de observar as variações na qualidade e adequação pedagógica das respostas.

Cada prompt foi inserido separadamente, em momentos distintos, garantindo que o modelo não mantivesse histórico de conversas anteriores, com o objetivo de evitar interferências nas respostas. A criação de cada aplicativo foi realizada em ambiente controlado, seguindo o mesmo procedimento: inserção do comando, geração do código HTML e registro das respostas produzidas.

Durante o processo, não houve erros técnicos significativos, apenas pequenas variações na estrutura do código e no formato das atividades apresentadas. Essas diferenças foram consideradas parte do comportamento natural da IA diante de diferentes níveis de detalhamento das instruções.

Os aplicativos podem ser acessados pelos links:

<https://chatgpt.com/canvas/shared/69df0b9968508191899333c1ac207181>

<https://chatgpt.com/canvas/shared/6937ab8d0bf081919da42f7fa1c3f705>

<https://chatgpt.com/canvas/shared/6937ac67ceb48191a2e17171e0acde0f>

Exemplos de Telas

Podemos observar que o aplicativo “Aprender a Ler - Versão Interativa” foi desenvolvido com o objetivo de auxiliar crianças em fase de alfabetização por meio de uma abordagem lúdica e acessível. A interface apresenta um design organizado e amigável, com cores suaves, elementos visuais simples e uma estrutura que facilita a navegação, mesmo para usuários iniciantes.

Na parte superior da tela, encontra-se o título do aplicativo, acompanhado de um controle deslizante que permite ajustar a velocidade da narração. Esse recurso possibilita adaptar o ritmo do áudio às necessidades da criança, tornando o aprendizado mais confortável e personalizado.

A área principal do aplicativo é composta por um cartão central que exibe uma imagem ilustrativa (representada por emoji) correspondente à palavra que deve ser formada. Logo abaixo da imagem, há um espaço com caixas vazias onde o usuário deve montar a palavra correta. Para isso, a criança pode utilizar o conjunto de letras disponíveis no “pool de letras”, localizado logo abaixo, arrastando-as ou clicando nelas para preencher os espaços.

O aplicativo oferece botões interativos que auxiliam no processo de aprendizagem. O botão “Ouvir palavra” permite que a criança escute a pronúncia completa da palavra, facilitando a associação entre som e escrita. Já o botão “Mostrar sílabas” apresenta a divisão silábica, contribuindo para a compreensão da estrutura das palavras. Além disso, o botão “Limpar” possibilita reiniciar a tentativa, enquanto o botão “Próxima” leva o usuário a uma nova palavra e imagem, promovendo a continuidade da atividade de forma dinâmica.

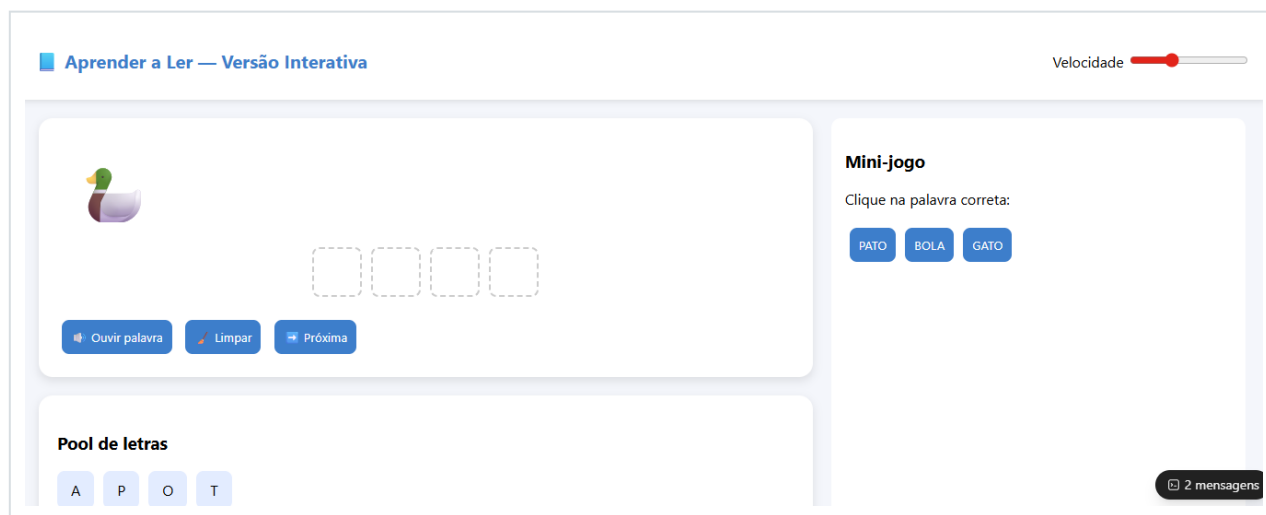


Figura 3. Tela de instrução do aplicativo “Aprender a Ler — Versão Interativa” (App 1)

Nota. Captura de tela do aplicativo gerado por IA. Apresenta uma figura (pato), espaços para montar a palavra com um “Pool de letras” (A, P, O, T), controles “Ouvir palavra”, “Limpar” e “Próxima”, e um “Mini-jogo” para clicar na palavra correta (PATO, BOLA, GATO), com ajuste de velocidade.

Na tela de instrução do aplicativo “Leitura Fácil: Atividades 6–8 anos”, é apresentada um guia simples e acolhedor para que a criança compreenda como interagir com o jogo educativo. Nessa tela, aparecem informações explicando que o aplicativo é dividido em três partes: le-

tras, sílabas e palavras onde o objetivo é aprender brincando. O texto orienta que, no modo Letras, a criança deve tocar nas letras para ouvir o som e descobrir qual delas inicia a palavra mostrada. No modo Sílabas, ela deve montar a palavra correta juntando as sílabas correspondentes à figura que aparece na tela. Já no modo Palavras, a criança precisa formar palavras completas usando as letras disponíveis, podendo clicar ou arrastar cada letra para o espaço correto.

Também são explicados os botões que aparecem no canto superior e lateral da tela: o controle de velocidade e volume da voz, o botão de resetar o progresso, e o indicador de pontuação que mostra quantos acertos a criança já obteve. Além disso, há uma pequena dica para o professor ou responsável, sugerindo que sejam feitas perguntas abertas, como “Qual som essa letra faz?”, para reforçar a aprendizagem de forma lúdica e interativa.

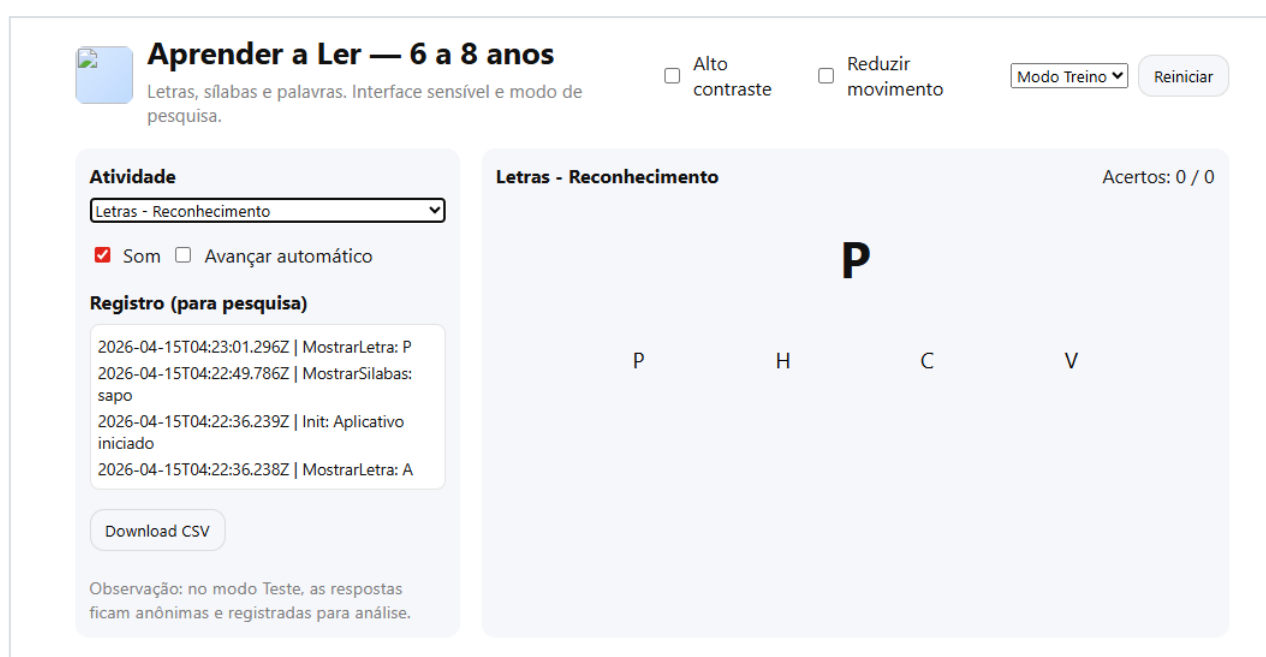


Figura 4. Exemplo de Tela de Instrução do App 2

Nota. Captura de tela do aplicativo “Aprender a Ler” (6 a 8 anos) em Modo Treino, na atividade de reconhecimento de letras (alvo “P”, com alternativas P, H, C, V). Há opções de acessibilidade (alto contraste, reduzir movimento), reforço sonoro e um registro de eventos com horário (“Registro para pesquisa”) exportável em CSV.

O aplicativo criado: “Leitura Amiga Versão Inclusiva para Crianças Autistas” foi criado com o objetivo de ensinar leitura de forma simples, de fácil acesso e personalizada a crianças de 6 a 8 anos com Transtorno do Espectro Autista (TEA). Ele utiliza estímulos visuais e auditivos, combinando entre imagens, sons, cores e reforços positivos para favorecer o engajamento, a atenção e o aprendizado gradual das crianças.

É visto na tela inicial que o aplicativo apresenta instruções curtas e claras, acompanhadas por ícones e narração em voz suave, para que a criança entenda sem complicações o que deve fazer. O ambiente é colorido, calmo, com elementos visuais limpos voltados para o público infantil e botões grandes, promovendo conforto sensorial e evitando sobrecarga de estímulos.

O conteúdo está dividido em três níveis progressivos de aprendizagem: Nível fácil; Reconhecimento de Letras: a criança ouve o som de uma letra e escolhe a opção correta entre três, observando uma imagem relacionada e nível intermediário; Formação de Sílabas: o jogo apresenta figuras e pede que a criança selecione a sílaba inicial ou final correta, estimulando a associação entre som e forma e por fim o Nível avançado com Leitura de Palavras Simples: a criança escuta a palavra e precisa escolher entre três opções escritas, reforçando a leitura global e o reconhecimento visual.

O aplicativo conta também com um sistema simples de progresso individual, que registra quantas atividades foram concluídas com sucesso. Esse recurso permite que o educador, cuidador ou terapeuta monitore o avanço e ajuste o ritmo de aprendizagem conforme a necessidade da criança.

Além disso, há estratégias inclusivas incorporadas à interface e ao funcionamento, como pausas curtas entre as tarefas, reforço multimodal (visual e auditivo), repetição controlada e possibilidade de ouvir novamente as instruções. Essas características ajudam a manter a atenção, estimular a autonomia e promover o engajamento sócio emocional

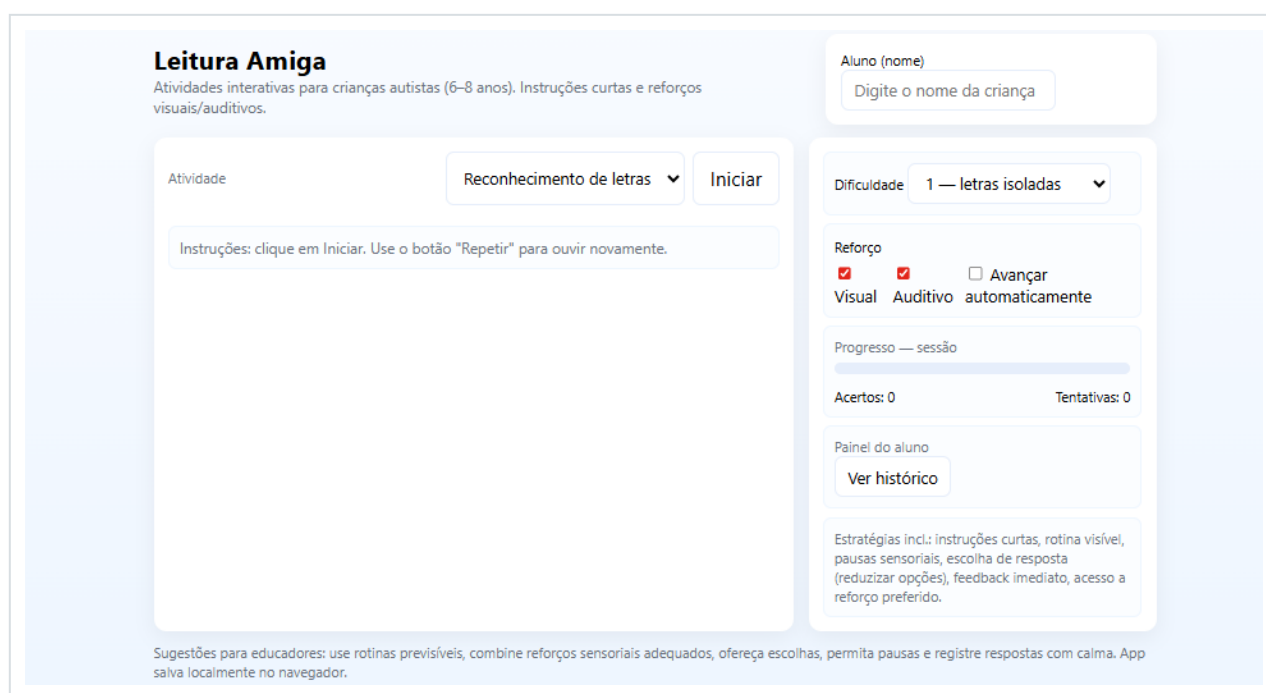


Figura 5. Tela de instrução do aplicativo “Leitura Amiga” (App 3)

Nota. Captura de tela do aplicativo gerado por IA, voltado a crianças autistas de 6–8 anos. Inclui seleção de atividade (“Reconhecimento de letras”), reforço visual e auditivo, controle de dificuldade, acompanhamento de acertos/tentativas e estratégias de instrução curta e pausas sensoriais.

Análise do Conteúdo

A análise dos dados dessa pesquisa foi conduzida de forma qualitativa, levando em conta a comparação entre as respostas geradas pela inteligência artificial em cada nível de prompt (genérico, intermediário e detalhado). O objetivo foi identificar os padrões, variações e possíveis aprimoramentos no desempenho da IA diante de diferentes graus de especificidade nas instruções.

Os dados foram organizados com base nos critérios definidos como: Pertinência pedagógica, clareza linguística, coerência nas atividades, e interatividade/engajamento assim dando uma abertura para uma análise comparativa entre os os prompts gerados.

De modo geral, observou-se que os prompts mais detalhados (nível 3) resultaram em respostas mais completas e coerentes, com atividades que apresentavam maior adequação pedagógica e sensibilidade às necessidades de crianças autistas trazendo também fontes científicas utilizadas. Já os prompts intermediários (nível 2) mostraram equilíbrio entre detalhamento e objetividade, produzindo resultados com certa satisfação, embora com menor profundidade em estratégias inclusivas. Por outro lado, os prompts genéricos (nível 1) geraram conteúdos mais superficiais, com instruções pouco contextualizadas e menor variedade de atividades.

A Tabela 2 mostra uma análise qualitativa dos três aplicativos gerados pelo Chatgpt, a partir dos diferentes níveis de prompt (genérico, intermediário e detalhado), revelou padrões evidentes sobre o impacto do detalhamento das instruções na qualidade pedagógica do material. A avaliação foi conduzida com base em quatro critérios centrais: pertinência pedagógica, clareza linguística, coerência das atividades e interatividade/engajamento.

De maneira geral, observou-se que o aplicativo criado a partir do prompt genérico (nível 1) apresentou desempenho menor em quase todos os critérios apresentados. Na pertinência pedagógica, as atividades não estabeleceram com objetivos claros de ensino de leitura, tampouco demonstravam preocupação com o perfil de crianças autistas. A clareza linguística foi apenas moderada, pois embora o texto fosse simples, faltavam instruções completas e direcionadas. A coerência das atividades também foi avaliada como insatisfatória, uma vez que os exercícios eram desconexos e não seguiam uma progressão lógica. Por fim, o nível de interatividade foi baixo, com escassez de elementos motivacionais ou recursos que pudessem favorecer o engajamento da criança.

Em contraste, o aplicativo derivado do prompt intermediário (nível 2) apresentou avanços importantes. Houve maior correspondência entre o conteúdo gerado e o objetivo de ensino, refletindo uma pertinência pedagógica mais aceitável. A linguagem utilizada demonstrou maior clareza, com instruções compreensíveis e melhor organizadas. As atividades, apesar de coerentes em suas etapas, ainda careciam de transições mais naturais entre os exercícios. Em termos de interatividade, o aplicativo apresentou recursos simples, como cliques e pequenos feedbacks, que contribuíram para um nível moderado de engajamento.

Por outro lado, o aplicativo gerado por meio do prompt detalhado (nível 3) alcançou o desempenho mais satisfatório em todos os critérios da rubrica. A pertinência pedagógica foi considerada excelente, visto que o conteúdo apresentou objetivos claramente delineados e estratégias adequadas ao ensino de leitura para crianças autistas. A clareza linguística também se destacou, com descrições completas e linguagem alinhada ao contexto educacional. A coerência das atividades foi mostrada como superior, apresentando progressão lógica e organi-

zação consistente. Além disso, o nível de interatividade e engajamento foi significativamente maior, com propostas que motivaram a participação ativa da criança e incorporaram elementos de estímulo visual e verbal.

Tabela 3. Tabela de Análise de Conteúdo pela IA

Critério	App 1	App 2	App 3
Pertinência pedagógica	Insatisfatório (1):	Bom (2):	Bom (3):
Clareza linguística	Bom (2):	Excelente (3):	Excelente (3):
Coerência das atividades	Insatisfatório (1)	Bom (2)	Excelente (3)
Interatividade/engajamento	Insatisfatório (1)	Bom (2)	Excelente (3)

Nota. Síntese da análise de conteúdo dos três aplicativos de alfabetização (App 1, App 2 e App 3) realizada pela IA.

Diante disso, a análise qualitativa dos dados sugere que o uso de prompts mais estruturados e contextualizados amplia a confiabilidade e aplicabilidade dos recursos educacionais criados por inteligências artificiais, sobretudo quando voltados ao ensino de leitura para crianças autistas.

4 DISCUSSÃO

Resumo

O objetivo do presente trabalho foi analisar o efeito de diferentes níveis de prompt na qualidade e adequação pedagógica de aplicativos educacionais voltados para crianças com autismo gerados pelo aplicativo de inteligência artificial intitulado “Chatgpt” para o ensino de rudimentos de leitura na educação básica. Para tanto, adotou-se um método qualitativo comparativo, no qual três níveis distintos de prompt genérico, intermediário e detalhado foram aplicados ao modelo de IA com o propósito de observar como o grau de especificidade das instruções influenciava o conteúdo produzido.

No geral, os resultados demonstraram que o detalhamento do prompt exerce influência significativa sobre a qualidade pedagógica e a estrutura das atividades criadas pela IA. Os aplicativos gerados a partir de prompts pouco específicos apresentaram limitações claras quanto à pertinência pedagógica, à coerência das atividades e ao engajamento proposto às crianças. Esse padrão indica que, quando a IA recebe instruções vagas, tende a produzir materiais igualmente superficiais, pouco contextualizados e insuficientes para atender às demandas de crianças autistas, que requerem maior precisão e adaptação.

Por outro lado, prompts intermediários mostraram melhora considerável no desempenho da IA, sugerindo que mesmo um aumento moderado na especificidade das instruções já produz efeitos positivos na qualidade das respostas. No entanto, a ausência de informações aprofundadas sobre o perfil da criança, estratégias inclusivas ou objetivos pedagógicos complexos resultou em materiais ainda limitados em profundidade e personalização.

O desempenho superior foi observado quando a IA recebeu prompts detalhados. Nessa condição, os aplicativos gerados demonstraram maior alinhamento pedagógico, melhor coerência na organização das atividades e maior potencial de engajamento, com elementos adaptados às necessidades cognitivas e socioemocionais de crianças autistas. Esses achados reforçam a importância do papel do usuário humano no processo: quanto mais preciso o comando, maior a probabilidade de a IA produzir conteúdos relevantes, apropriados e consistentes.

Comparação com a Literatura

Os resultados da seguinte pesquisa corroboram com os estudos encontrados em literaturas específicas sobre o uso de inteligências artificiais em contextos educacionais e para apoiar o grupo de pais e responsáveis. Estudos como os de (Kim et al., 2025) e de (Hassona et al., 2024) apontam que as IA's podem gerar conteúdos e materiais relevantes, entretanto a qualidade das respostas apresentadas são diretamente influenciadas pela forma como os prompts são aplicados. Da mesma forma, Sampaio et al. menciona que a IA destaca que ela não possui total compreensão contextual ou pedagógica por si só, necessitando de instruções mais diretas para fornecer conteúdos mais adequados. Portanto, os dados desta pesquisa reforçam que o detalhamento do prompt é essencial para que a inteligência artificial produza respostas coerentes, pedagógicas e sensíveis às necessidades do público voltado para os pais e responsáveis que requerem melhor auxílio para suas crianças autistas.

Contribuições metodológicas

Este estudo contribui metodologicamente ao demonstrar uma forma de avaliação de categoria qualitativa para os conteúdos educacionais gerados por IA, estruturado a partir da variação sistemática de níveis de prompt nos níveis básico, moderado e avançado, analisando as diferenças de conteúdo que a inteligência artificial gera através de diversos prompts usados. A aplicação do processo de teste, aliado à análise segundo critérios pedagógicos, forneceu uma abordagem replicável para futuras pesquisas. O estudo também avança ao propor uma rubrica específica para avaliar pertinência pedagógica, clareza linguística, coerência das atividades e interatividade/engajamento critérios que até hoje são pouco explorados de modo integrado em pesquisas com IA na educação inclusiva.

Contribuições Aplicadas

Cada app gerado pelas IAs serviu como teste para se analisar um suposto avanço e uso das IAs para contribuir de forma significativa em um bom desenvolvimento das crianças com autismo já que provavelmente é uma das maiores preocupações que um pai ou responsável pode adquirir sabendo que possui um filho com esse transtorno, este estudo contribui metodologicamente ao demonstrar um modelo de avaliação qualitativa para materiais educacionais gerados por IA e uma suposta contribuição futura para pesquisas posteriores com temas semelhantes e também foi uma pesquisa estruturada a partir da variação sistemática de níveis de prompt. A aplicação do processo de teste, aliado à análise segundo critérios pedagógicos. O estudo também avança ao propor uma rubrica específica para avaliar pertinência pe-

dagógica, clareza linguística, coerência das atividades e interatividade/engajamento critérios pouco explorados hoje em dia de modo integrado em pesquisas com IA na educação inclusiva.

Limitações e pesquisas futuras

Entre as limitações do estudo, destaca-se o fato de que apenas um modelo de inteligência artificial foi utilizado, além da análise ter se concentrado exclusivamente em atividades voltadas ao ensino inicial de leitura. Foi notado durante a pesquisa uma dificuldade significativa de gerar aplicativos com IA e de acessar pois diversas vezes os links apareciam com erro gerando inacessibilidade sendo necessário pedir a inteligência artificial refazer o mesmo link diversas vezes, assim dificultando de forma significativa o fácil acesso dos aplicativos de IA principalmente para aqueles que não são familiarizados com tecnologias e também não houve participação direta de crianças, educadores ou familiares na avaliação prática dos aplicativos, o que restringe a análise ao nível de conteúdo. Pesquisas futuras podem incluir a testagem dos aplicativos em ambiente escolar, avaliar indicadores de aprendizagem real e ampliar os prompts para outras áreas de desenvolvimento, como habilidades socioemocionais e comunicação alternativa. Sugere-se também que diferentes modelos de IA sejam comparados, investigando potenciais riscos, vieses e responsabilidade no uso dessas tecnologias na educação especial.

5 CONCLUSÃO

Este trabalho teve como objetivo analisar a qualidade dos aplicativos com atividades pedagógicas por uma inteligência artificial a partir de diferentes níveis de especificidade nos prompts, considerando o contexto onde pais e responsáveis buscam ferramentas que possam auxiliar no desenvolvimento dos seus filhos autistas. Para tanto, realizou-se uma pesquisa qualitativa, fundamentada na comparação das respostas da IA conforme os critérios de pertinência pedagógica, clareza linguística, coerência das atividades e interatividade/engajamento, mostrando que os prompts mostraram respostas mais satisfatórias e alinhadas às necessidades que as crianças precisavam enquanto os prompts genéricos resultaram em materiais superficiais e com conteúdo mais escasso.

Conclui-se, portanto, que a inteligência artificial apresenta potencial para contribuir com a elaboração de recursos pedagógicos direcionados aos pais e responsáveis que buscam ferramentas que possam auxiliar seus filhos autistas. Contudo, sua eficácia depende do nível de precisão das orientações prestadas e do acompanhamento crítico de um profissional da educação ou da psicologia. Assim, esta pesquisa evidencia que a IA pode ser uma ferramenta complementar relevante no futuro, desde que utilizada de modo consciente, ético e supervisionado.

REFERÊNCIAS BIBLIOGRÁFICAS

- Carvalho, P. F., & Moreira, M. B. (2024). Intervenções baseadas em brincadeiras para crianças autistas. *Walden4*.
- Hassona, R. (2024). "My kid has autism": An interesting conversation with ChatGPT. *Journal of Autism and Developmental Disorders*, 54(3), 1184–1187. <https://doi.org/10.1007/s10803-023-06015-7>
- Kim, Y., Vilches, S. L., Shapiro, S., & Clarkson, A. (2025). Testing the capability of generative artificial intelligence for parent and caregiver information seeking. *Family Relations*, 74(3), 1266–1284. <https://doi.org/10.1111/fare.13167>
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135–148. <https://doi.org/10.1089/cyber.2023.0202>
- Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: Reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, Artigo e008. <https://doi.org/10.1590/1678-98732432e008>

Avaliação Neuropsicológica no TEA



Mapeamento completo do perfil cognitivo de crianças com Transtorno do Espectro Autista, com instrumentos validados e adaptados.

- Atenção, memória, linguagem e funções executivas
- Relatório com análise quantitativa e qualitativa
- Devolutiva presencial e articulação com escola e equipe

walden4.com.br/autismo/avaliacao

CAPÍTULO

06

Efeito do Nível de Estruturação de Prompts na Qualidade de Tutoriais Gerados por IA sobre Intervenções em Autismo



João Vítor Mansilha Silva

Márcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. O estudo teve como objetivo investigar como diferentes níveis de estruturação de prompts influenciam a qualidade de tutoriais produzidos pela Inteligência Artificial generativa DeepSeek sobre intervenções baseadas em Reforçamento Não Contingente (NCR) para comportamentos autolesivos em crianças com Transtorno do Espectro Autista. Buscou-se verificar tanto a precisão conceitual quanto à clareza, concisão, qualidade das referências, estrutura do aplicativo gerado e a consistência após reformulações.

Método. Foi conduzido um delineamento experimental de sujeito único (intrassujeito), no qual o modelo DeepSeek-V3 foi tratado como participante experimental. Foram elaborados três prompts com diferentes níveis de estruturação (baixo, médio e alto), cada um solicitando a criação de um tutorial em HTML sobre intervenção com NCR. Os tutoriais foram avaliados por meio de três rubricas: (1) Qualidade do conteúdo (corretude, clareza, concisão, referências), (2) Estrutura do aplicativo (responsividade, usabilidade, clareza instrucional), e (3) Consistência após reformulação (manutenção da estrutura após solicitação de alteração de cor).

Resultados. Os três tutoriais foram gerados com precisão conceitual elevada, atingindo pontuação máxima em corretude. Prompts mais estruturados produziram materiais mais complexos, com casos práticos, gráficos interativos, sistemas de registro e resumos finais. Contudo, maior detalhamento não resultou necessariamente em maior pontuação geral: o nível 1 apresentou média 3,71; nível 2 também 3,71; e o nível 3 obteve 3,57, principalmente devido a pequenas inconsistências em referências e usabilidade (botão de PDF inoperante). A análise de consistência mostrou que, no nível 1, a IA modificou exclusivamente a cor solicitada, preservando 100% da estrutura original.

Conclusão. O nível de estruturação dos prompts influencia a densidade, abrangência e complexidade dos tutoriais gerados, mas não necessariamente melhora a pontuação geral de qualidade. A IA demonstrou estabilidade, precisão conceitual e boa organização instrucional em todos os níveis, indicando potencial significativo como ferramenta de apoio para elaboração de materiais clínicos baseados em evidências. Destaca-se que, apesar dos resultados positivos, pesquisas futuras devem testar prompts realmente pouco estruturados (sem mencionar a intervenção) e comparar o desempenho entre diferentes IAs.

Palavras-chave: Inteligência Artificial generativa; DeepSeek; prompts; intervenção comportamental; Reforçamento Não Contingente; autismo; comportamentos autolesivos; ABA

1 INTRODUÇÃO

Contextualização do Tema

Atualmente, os grandes sistemas de Inteligência Artificial (IA) generativa como o CHAT GPT e outros modelos de linguagem de grande escala (LLMs) aplicam modelos generativos, os quais são utilizados para entender e interpretar a linguagem humana, baseando-se em textos para realizar previsões de respostas que se assemelham ao uso humano (Sampaio et al., 2024). Dessa forma, surgem diversas possibilidades de uso para essas plataformas, que incluem o de aprendizado e o de aperfeiçoamento profissional de forma adaptada às necessidades de cada indivíduo, o que pode viabilizar o uso dessas Inteligências Artificiais generativas no aprimoramento da atuação profissional direcionada ao tratamento de indivíduos com Transtorno do Espectro Autista (TEA).

Nesse sentido, seria possível suprir uma lacuna observada na prática profissional relacionada ao autismo, em que há uma escassez de profissionais qualificados com uma formação teórico-técnica necessária para tratamento do TEA, que é uma condição complexa do neurodesenvolvimento (Holanda & Maia, 2024). Além disso, muitos profissionais têm o acesso dificultado à literatura científica atualizada acerca do autismo e suas intervenções, devido a fatores como a barreira linguística, já que a grande parte dessas evidências está reportada em língua inglesa, e a linguagem científica (Cruz & Moreira, 2021). Portanto, a introdução ao uso de IAs generativas no âmbito profissional e da saúde permite os profissionais a terem acesso a evidências científicas atualizadas no campo do TEA, diminuindo limitações linguísticas e de compreensão dos conteúdos (Sampaio et al., 2024).

Problema de Pesquisa

Entretanto, é imprescindível considerar que as IAs generativas são passíveis de erros, dentre eles a possibilidade de alucinações e da presença de vieses em suas respostas. Alucinações ocorrem quando a resposta produzida pela máquina é sintaticamente possível, mas incompatível com a realidade em termos de fatos e dados. Já os vieses advêm da reprodução de premissas e escolhas realizadas durante o processo de desenvolvimento e criação dessas IAs, em que as visões dos desenvolvedores e trabalhadores que fazem o treinamento e rotulação dos dados e bancos de dados escolhidos para o treinamento são incorporadas nas respostas produzidas pela máquina (Sampaio et al., 2024). Diante o exposto, é dada a necessidade de confirmar se, em meio às diversas formas e produtos da Inteligência Artificial generativa, o uso mais direto para a elaboração de tutoriais de intervenções para profissionais da área da psicologia é confiável.

Pergunta de Pesquisa

Considerando então, a possibilidade das IAs serem passíveis de erros e o problema apresentado, surge a seguinte pergunta de pesquisa: qual o efeito de diferentes níveis de estruturação dos prompts em um modelo de Inteligência Artificial generativa sobre a qualidade da elaboração de um tutorial para profissionais da psicologia sobre autismo?

Relevância da Pesquisa

A escassez de profissionais devidamente capacitados para atuar com TEA é um fenômeno presente na realidade brasileira (Holanda & Maia, 2024). Fatores como a dificuldade em acessar literatura científica atualizada em razão de barreiras linguísticas, ou ainda a necessidade de compreender achados complexos em estratégias práticas de intervenção, são obstáculos recorrentes na prática clínica. Desse modo, compreender de que maneira a Inteligência Artificial pode suprir parte dessas lacunas, garantindo qualidade e aplicabilidade nos conteúdos produzidos, torna-se altamente relevante.

O estudo nos ajuda a aprofundar o entendimento sobre como diferentes formas de estruturação de prompts em sistemas de Inteligência Artificial generativa podem atuar na qualidade de materiais produzidos para fins profissionais, especialmente tutoriais de intervenção psicológica voltados ao tratamento do Transtorno do Espectro Autista (TEA). As descobertas podem ter aplicações práticas importantes, como: ampliar o acesso de psicólogos a conteúdos técnicos baseados em evidências, facilitar a atualização profissional contínua e contribuir para a formação de práticas mais eficazes e acessíveis no atendimento a crianças autistas. Além disso, o estudo pode introduzir novas metodologias para avaliar a confiabilidade do uso de IA em contextos clínicos, estimulando colaborações entre psicologia, ciência da computação e educação.

Revisão da Literatura

As Inteligências Artificiais generativas caracterizam-se por utilizarem Modelos de Linguagem de Grande Escala (LLMs), os quais buscam entender e interpretar a linguagem humana, aplicando modelos generativos que empregam regras para prever o uso humano com base em textos e imitá-lo da forma mais fiel possível. Dessa forma, o funcionamento das IAs pode ser explicado por meio dos conceitos de aprendizado de máquina e do método Transformers para processamento de textos.

O aprendizado de máquina é um campo da IA responsável pelo desenvolvimento de algoritmos (regras) capazes de efetuar ações sem programação humana direta, agindo de forma autônoma com base em conjuntos de dados para a realização de previsões que se assemelham aos padrões humanos. O modelo Transformers é um método de análise de texto que se fundamenta no conceito de atenção, no qual relaciona elementos do texto independentemente de sua distância e os classifica em termos de importância, auxiliando na construção do texto, já que a previsão de qual palavra será selecionada se dá com base no contexto do texto como um todo e nos elementos mais relevantes para o que é solicitado (Sampaio et al., 2024).

Esses conceitos aplicam-se a todas as IAs generativas, as quais se apresentam em diferentes aplicações, entre elas a DeepSeek, que será de especial interesse para este trabalho. A DeepSeek é uma LLM que se destacou em relação a outros modelos já estabelecidos no mercado devido ao seu baixo custo, alta precisão e confiabilidade de dados. Ela demonstrou superar IAs como Gemini, GPT e Llama no quesito precisão, somente sendo superada pelo Claude. Entretanto, apresentou custo consideravelmente menor com relação a essa última LLM e uma maior lentidão quando comparada às outras IAs citadas (Gao et al., 2025).

Por outro lado, indicou melhor organização de informações e recuperação de dados ao ser avaliada em relação às IAs Grok, Chat GPT e Gemini. Com isso, garante uma maior confiabilidade de dados estruturados, fundamentados e verificáveis, ocupando uma posição de destaque para pesquisa científica e recuperação de informação técnica (Souza & Weigang, 2025). Nesse contexto, será feito o uso da DeepSeek neste trabalho em virtude de seu baixo custo para profissionais e alta precisão e confiabilidade nos conteúdos, tendo em vista a qualidade do tutorial elaborado sobre a intervenção com reforçamento não contingente (NCR) para comportamentos autolesivos em crianças autistas.

O TEA é caracterizado pelo Manual diagnóstico estatístico de transtornos mentais V (DSM-5) como um transtorno do neurodesenvolvimento, que apresenta critérios diagnósticos relacionados à limitações persistentes na comunicação e interação social, nas habilidades relacionais, além da presença de padrões excessivos de comportamentos restritos e repetitivos (Associação Psiquiátrica Americana, 2013). Em uma visão analítico-comportamental, esses comportamentos que compõem os critérios diagnósticos do DSM-5 estão sujeitos à mudança por meio de interações com o ambiente, de modo que são provocados por certos eventos e mantidos ao longo do tempo por suas consequências. Portanto, é necessário considerá-las como variáveis controladoras do comportamento para a realização de uma intervenção individualizada e eficaz (Goulart & Assis, 2002).

Com base nesses conceitos, em intervenções baseadas na Análise do Comportamento Aplicada (ABA) são realizados procedimentos como o de reforçamento não contingente (noncontingent reinforcement, NCR), que consiste na disponibilização de um reforçador de uma maneira não contingente, ou seja, sem que haja a necessidade de uma resposta específica para a ocorrência da consequência de tempos em tempos (Moreira & Medeiros, 2019). Nessa perspectiva, a aplicação do NCR em manipulações experimentais demonstrou-se muito eficaz para o tratamento de comportamentos-problema, dentre eles, os comportamentos autolesivos (Ceppi & Benvenuti, 2011), ou SIB (do inglês, self-injurious behaviors), que podem ser descritos como comportamentos que geram alguma forma de dano físico à própria pessoa que os emite (Cruz & Moreira, 2021).

Um dos motivos da aplicação do NCR para diminuição do SIB é que, diferentemente de outros procedimentos como reforçamento diferencial de comportamentos alternativos (DRA) e reforçamento diferencial de outros comportamentos (DRO), não realiza o processo de extinção. A extinção ocorre quando não há a apresentação de uma consequência para o comportamento emitido, sendo um modo eficaz de diminuição da frequência de uma resposta. Entre-

tanto, para fins práticos, um dos primeiros efeitos da extinção é o aumento abrupto de emissão do comportamento (Ceppi & Benvenuti, 2011), o que pode prejudicar fisicamente o indivíduo que está passando pela intervenção.

Com isso, Vollmer et al. (1993) investigaram o efeito da atenção no tratamento do SIB utilizando NCR e DRO em seu procedimento. Os efeitos do tratamento foram comparados em um delineamento multielemento intrassujeito e um delineamento de linha de base múltipla entre sujeitos. Para o reforço não contingente (NCR) era fornecida atenção às participantes sob um esquema de tempo fixo, em que as respostas da participante não influenciavam na frequência em que o reforço era fornecido. Nesse procedimento foi realizado um desvanecimento (fading), de uma taxa inicial de 6 por minuto (atenção contínua) para uma taxa final de 0,2 por minuto (uma entrega a cada 5 minutos), em que a frequência da atenção era reduzida gradualmente quando a taxa de SIB caía para um nível mais baixo. Foi demonstrado que o NCR suprimiu o SIB de forma mais eficaz que o DRO e que atenuou os efeitos de aumento das respostas induzidos pela extinção.

Embora a literatura científica ofereça evidências robustas sobre como intervir em comportamentos-problema em crianças com TEA, persiste uma lacuna significativa entre o conhecimento produzido academicamente e sua apropriação por pais, cuidadores e profissionais sem formação especializada. Iniciativas como o manual de Cruz & Moreira (2021) representam esforços voltados à redução dessa distância, ao traduzir procedimentos analítico-comportamentais em linguagem mais acessível. Contudo, a mera disponibilização de materiais informativos mostra-se insuficiente sem a mediação de processos estruturados de treinamento.

Nesse cenário, torna-se relevante considerar estratégias de expansão das intervenções baseadas em ABA para além do contexto clínico especializado. Ferreira et al. (2016) verificaram o efeito de um pacote de treino sobre a precisão de implementação de tentativas discretas por cinco cuidadores de crianças diagnosticadas com TEA, utilizando um delineamento intrassujeito com comparação de medidas pré e pós-tratamento. O pacote foi composto por quatro componentes aplicados de forma sequencial: instruções, nas quais os experimentadores apresentavam verbalmente e por escrito o conteúdo dos programas de ensino, esclarecendo dúvidas e descrevendo os comportamentos esperados em cada etapa; modelação, em que os próprios experimentadores demonstravam ao vivo a execução correta das tentativas discretas enquanto o cuidador observava; role-play com feedback imediato, no qual o cuidador praticava a aplicação dos programas com um confederado que simulava o comportamento da criança, recebendo correções e elogios em tempo real a cada tentativa executada; e vídeo-feedback, em que o cuidador assistia à gravação de seu próprio desempenho na fase de linha de base enquanto o experimentador sinalizava os acertos e erros de implementação. Os resultados demonstraram que todos os participantes alcançaram 100% de precisão de desempenho no pós-teste, com manutenção das habilidades adquiridas na aplicação de programas subsequentes, indicando aquisição e generalização do repertório ensinado.

Entretanto, conforme apontado por Baer et al. (1968), para que uma prática científica cumpra sua função social, ela precisa ser não apenas eficaz, mas também tecnológica, ou seja, descrita com precisão suficiente para ser replicada de forma independente por qualquer profissio-

nal minimamente treinado. Sob essa perspectiva, mesmo diante de evidências que demonstram a viabilidade do treinamento de cuidadores (Ferreira et al., 2016), o conhecimento produzido frequentemente não se converte em tecnologia disseminável em larga escala, permanecendo restrito ao contexto de pesquisa. Estudos mais recentes, como o de McFayden et al. (2024), voltam-se justamente para esse problema ao investigar o uso de ferramentas como o ChatGPT-4 na orientação de pais e cuidadores, ampliando a discussão para além dos procedimentos clínicos tradicionais. Nesse contexto, modelos de Inteligência Artificial generativa apresentam-se como uma alternativa promissora para transformar conhecimento científico em tecnologia aplicável de forma imediata e acessível.

McFayden et al. (2024) investigaram o uso do ChatGPT-4 como recurso de informação para pais e cuidadores de crianças com TEA, analisando aspectos como correção, clareza e concisão das respostas, além da linguagem empregada e do nível de compreensão e aplicabilidade prática pelo instrumento PEMAT-P. De forma geral, o modelo apresentou conteúdos corretos e bem estruturados, mas com pouca utilidade prática, sobretudo quando se tratava de orientar ações concretas. Outro ponto de destaque foi a inconsistência nas referências fornecidas, muitas vezes desatualizadas ou incorretas, o que comprometeu a confiabilidade do material. A simples regeneração das respostas não trouxe avanços significativos, reforçando a ideia de que a qualidade está muito mais relacionada à forma como o prompt é elaborado do que ao número de tentativas de geração. Também foi observado que o modelo alternava entre uma linguagem mais médica e outra alinhada a uma visão afirmativa da neurodiversidade, revelando a influência de vieses em sua produção textual.

Em replicação ao estudo de McFayden et al. (2024), Kim et al. (2025) ampliaram a investigação sobre a qualidade das respostas de modelos de IA generativa para o contexto mais amplo da educação parental, comparando o desempenho do ChatGPT-3.5 e do ChatGPT-4 com o de um boletim informativo profissional de parentalidade, o Just in Time Parenting (JITP.info), a partir de cinco tópicos temáticos frequentes entre cuidadores de crianças pequenas: promoção do sono, estabelecimento de limites, autocuidado parental, manejo de birras e demonstração de afeto. Para a avaliação dos produtos gerados, os autores adaptaram o sistema de três critérios proposto por McFayden et al. (2024), substituindo a concisão pela conexão relacional, considerada mais relevante para o domínio da parentalidade, e acrescentaram avaliações de reprodutibilidade das citações e de legibilidade pelo índice Flesch-Kincaid. Os resultados indicaram que o ChatGPT produziu respostas predominantemente corretas e claras, porém com limitações expressivas quanto à dimensão relacional entre cuidador e criança, com citações frequentemente ausentes ou imprecisas, e com índices de legibilidade situados entre os graus 10 e 14 da escala Flesch-Kincaid, valores substancialmente superiores aos do boletim de referência, o que pode representar uma barreira para públicos com menor escolaridade. Os autores concluíram que a qualidade das respostas variou conforme o tipo de questão apresentada e que a forma de elaboração do prompt exerceu papel determinante sobre a qualidade do material gerado, convergindo com os achados de McFayden et al. (2024) ao reforçar que limitações quanto à aplicabilidade prática, à consistência das referências e à adequação ao público-alvo permanecem como desafios centrais dos modelos de IA generativa em contextos de orientação e cuidado.

Esses resultados reforçam a necessidade de articular o rigor conceitual dos estudos experimentais com a aplicabilidade prática das informações geradas por modelos de IA, o que direciona o foco da presente pesquisa para a análise da qualidade, clareza e utilidade dos materiais produzidos. Ao propor a análise da elaboração de tutoriais sobre intervenções baseadas no reforçamento não contingente (NCR) para comportamentos autolesivos em crianças com TEA, torna-se claro que não basta avaliar apenas a precisão conceitual do conteúdo. É igualmente necessário verificar se o material é aplicável à prática clínica, se apresenta instruções claras e seguras, e se as referências fornecidas são confiáveis e acessíveis.

Objetivo da Pesquisa

Portanto, o objetivo do presente estudo é: Investigar o efeito de diferentes níveis de estruturação de prompts na qualidade de um aplicativo tutorial elaborado pelo modelo de IA generativa DeepSeek sobre intervenções com reforçamento não contingente (NCR) para comportamentos autolesivos em crianças autistas.

2 MÉTODO

Delineamento de Pesquisa

Para fins deste trabalho iremos considerar a IA como um participante de pesquisa, pois há possibilidade de considerar a Inteligência Artificial como uma unidade experimental única, de modo semelhante ao que ocorre em pesquisas com participantes humanos. Em ambos os casos, parte-se da obtenção de um desempenho inicial estável em linha de base para, em seguida, manipular sistematicamente a variável independente e observar seus efeitos sobre o comportamento. Assim como humanos respondem a instruções e condições ambientais específicas, a IA responde às instruções fornecidas por meio dos prompts, estabelecendo uma relação direta entre a condição de entrada e a qualidade da resposta.

Dessa forma, realizamos um estudo de delineamento experimental de caso único (intra-sujeito), a partir de um método quantitativo de pesquisa utilizando o modelo DeepSeek, abordando três níveis de estruturação de perguntas (prompts) sendo eles: 1) menos estruturado 2) médio e 3) mais estruturado.

O delineamento experimental de sujeito único é uma abordagem de pesquisa na qual cada sujeito é seu próprio controle, sendo exposto a todas as condições do experimento (controle e experimental) enquanto suas respostas são medidas repetidamente para verificar se existe uma relação ordenada entre as condições manipuladas (variáveis independentes) e as alterações observadas (variáveis dependentes) (Sampaio et al., 2008).

Um exemplo semelhante seria de realizar um estudo com professores, nas quais se poderia solicitar a elaboração de uma prova a partir de diferentes níveis de instrução. Em uma condição inicial, o professor receberia uma instrução simples, como “elabore uma prova para sua disciplina”, resultando em um produto com menor direcionamento. Em uma segunda condição, a instrução seria de nível intermediário, por exemplo: “elabore uma prova para sua disciplina contendo questões objetivas e dissertativas, abrangendo conteúdos do último bimes-

tre”. Por fim, em uma terceira condição, o professor receberia uma instrução altamente estruturada, especificando não apenas o formato das questões, mas também a quantidade, os critérios de dificuldade e a distribuição dos conteúdos a serem cobrados. A qualidade das provas elaboradas em cada condição seria, então, avaliada a partir de critérios previamente definidos, permitindo observar de que maneira o grau de estruturação da instrução influencia o desempenho do professor. De forma análoga, no presente estudo a Inteligência Artificial DeepSeek foi considerada como participante, sendo exposta a diferentes níveis de estruturação de prompts e tendo suas respostas analisadas qualitativamente com base em critérios de avaliação de conteúdo.

Participante (Modelos Utilizados)

O presente estudo teve como participante o modelo de inteligência artificial generativa DeepSeek-V3, versão gratuita, utilizado sob condições experimentais controladas. Trata-se de um Large Language Model (LLM) desenvolvido pela DeepSeek, distinguindo-se por seu baixo custo operacional, alta precisão e confiabilidade na recuperação de informações estruturadas e fundamentadas. O modelo foi considerado experimentalmente ingênuo, uma vez que foi utilizada uma conta recém-criada, sem histórico de interações anteriores, garantindo que nenhum viés ou aprendizado prévio influenciasse as respostas geradas durante a coleta de dados. Sua arquitetura baseia-se no método Transformers, permitindo processamento contextualizado de linguagem natural e geração de textos coerentes e informativos, adequados para aplicações em pesquisa e recuperação técnica.

Prompts Utilizados

Foram testados diferentes prompts em uma conta já estabelecida com a finalidade de identificar fatores que auxiliariam na produção de prompts mais eficazes em cada nível de estruturação.

Os prompts testados previamente foram:

- “Suponha que sou um profissional de psicologia recém formado e tenho um paciente autista nível 2 com comportamento autolesivo de bater com o punho no queixo. Com isso, crie um tutorial em html para me ensinar uma intervenção para diminuir esse comportamento”
- “Crie um tutorial em html me ensinando uma intervenção para comportamento autolesivo em TEA, sou psicólogo”
- “Crie um tutorial HTML interativo e responsivo sobre intervenção com Reforçamento não contingente (NCR) para autolesão em TEA, destinado a psicólogos, que incluua:
 - Fundamentação científica baseada em ABA
 - Explicação passo a passo da implementação do DRA
 - Casos clínicos interativos com sistema de pontuação
 - Gráficos para registro e acompanhamento da frequência de autolesão
 - Sistema de feedback e recomendações personalizadas

- Quadro de resumo final com opção de download em PDF
- Design estético e experiência de usuário gamificada
- O tutorial deve ser clinicamente preciso, visualmente atraente e oferecer uma experiência de aprendizagem envolvente através de elementos interativos e feedback imediato.”
- “Com base nesses critérios crie um tutorial em html me ensinando uma intervenção utilizando reforçamento diferencial de comportamento alternativo para comportamento autolesivo em TEA baseado na análise do comportamento, sou psicólogo, faça um modelo em html que seja responsivo, com usabilidade, claro e com validade científica, além de ser interativo e estético, adicione também gráficos interativos (por exemplo, para acompanhar frequência de autolesão ao longo do tempo) configure o gráfico para permitir entrada manual de novos dados pelo usuário (ex.: registrar ocorrências semanais), expanda esses casos para incluir planos de intervenção sugeridos passo a passo baseados na resposta do usuário, acrescente um sistema de pontuação/feedback acumulado (ex.: progresso do psicólogo ao longo dos casos) para deixar o tutorial ainda mais gamificado, adicione também um quadro de resumo final (com pontos, acertos, recomendações e próximos passos sugeridos) ao concluir todos os casos, adicione também a opção de baixar o resumo final em PDF separado (com pontos, feedback e recomendações)”

Com base nos resultados e em sugestões de prompt do próprio Deep Seek, as versões finais são:

Prompt nível 1: Sou analista do comportamento e atendo uma criança autista que apresenta comportamentos autolesivos. Gere um app html com tutorial para me ensinar a conduzir uma intervenção baseada em reforçamento não contingente para reduzir os comportamentos autolesivos. Liste as referências.

Prompt Nível 2: Sou analista do comportamento e atendo uma criança autista que apresenta comportamentos autolesivos. Gere um app html com tutorial para me ensinar a conduzir uma intervenção baseada em reforçamento não contingente para reduzir os comportamentos autolesivos. Liste as referências. Utilize na elaboração do tutorial as estratégias analítico-comportamentais de instrução, modelação e modelagem. Crie um tutorial completo em HTML sobre intervenção com Reforçamento não contingente (NCR) para comportamento autolesivo em Transtorno do Espectro Autista (TEA), baseado nos princípios da Análise do Comportamento Aplicada (ABA). O tutorial deve ser destinado a psicólogos e possuir validade científica, com referências a estudos e práticas baseadas em evidências.

Desenvolva uma interface responsiva e esteticamente atraente, com navegação intuitiva por abas que cubram: introdução ao NCR, fundamentos teóricos, implementação passo a passo, casos práticos interativos, sistema de registro de dados e resumo final.

Inclua casos clínicos simulados onde o usuário deve selecionar intervenções apropriadas, com sistema de pontuação que recompense decisões clinicamente corretas. Adicione gráficos interativos (usando Chart.js) para registro da frequência de autolesão ao longo do tempo, permitindo entrada manual de dados semanais e cálculo automático do progresso da intervenção.

Implemente um sistema gamificado com feedback imediato sobre as escolhas do usuário, barra de progresso visual e nível de proficiência baseado no desempenho. Ao final, apresente um quadro de resumo com pontuação total, análise de pontos fortes, áreas para melhoria e recomendações personalizadas de próximos passos.

Inclua a funcionalidade de gerar e baixar um PDF deste resumo final, contendo todos os dados de desempenho e recomendações. Utilize um design moderno com esquema de cores profissional, ícones intuitivos e layout responsivo que funcione em dispositivos móveis e desktop.

Todo o conteúdo deve ser clinicamente preciso, com explicações claras sobre análise funcional, seleção de comportamentos alternativos, esquemas de reforçamento e estratégias de generalização, sempre fundamentadas na literatura científica da área.

Prompt Nível 3: O prompt 3 foi idêntico ao prompt 2, entretanto foi anexado o artigo de Vollmer et al. (1993) a partir do seguinte comando adicionado ao prompt 2: utilize o artigo de pesquisa anexo como base para a elaboração do tutorial.

Avaliação do Produto Gerado

A Rubrica 1 apresenta a rubrica de avaliação dos 3Cs (Correctness, Clarity e Conciseness), elaborada com base no estudo de McFayden et al. (2024). Cada critério foi avaliado em quatro categorias de desempenho, que variam de 1 (insuficiente) a 4 (excelente).

No critério Corretude (Correctness), a categoria 4 corresponde a respostas que apresentam informações cientificamente precisas, atualizadas e sem erros factuais, com afirmações verificáveis em fontes confiáveis. A categoria 3 indica respostas majoritariamente corretas, mas com pequenas imprecisões ou omissões não essenciais. A categoria 2 aplica-se a respostas com vários erros menores ou interpretações ambíguas dos fatos, enquanto a categoria 1 representa respostas com informações incorretas, desatualizadas ou enganosas.

O critério Clareza (Clarity) avalia o grau de compreensão e organização textual. Na categoria 4, a resposta é totalmente compreensível, utiliza linguagem acessível e apresenta estrutura lógica, permitindo fácil entendimento da ideia central. A categoria 3 descreve respostas claras na maior parte, mas que contêm trechos densos, termos técnicos ou frases ambíguas. A categoria 2 refere-se a respostas cuja estrutura ou vocabulário dificultam a compreensão em vários pontos, enquanto a categoria 1 caracteriza respostas confusas, desorganizadas ou com linguagem inadequada ao público-alvo.

O critério Concisão (Conciseness) analisa o equilíbrio entre objetividade e completude. A categoria 4 é atribuída a respostas que apresentam exatamente o necessário para responder à questão, sem redundâncias nem omissões. A categoria 3 corresponde a respostas ligeiramente longas ou breves, mas que mantêm a essência do conteúdo. A categoria 2 aplica-se a respostas com excesso de detalhes irrelevantes ou falta de informações importantes, enquanto a categoria 1 caracteriza respostas extremamente prolixas ou tão curtas que comprometem a compreensão.

Por sua vez, o critério Referências Bibliográficas, analisa a conformidade, verificabilidade e relevância das fontes citadas. As referências avaliadas como excelentes estão completas, corretamente formatadas conforme a APA e associadas a publicações verificáveis, preferencialmente de periódicos indexados no Qualis A ou B. As categorias inferiores refletem inconsistências formais, fontes não rastreáveis ou inexistentes, e, no pior caso, ausência de referências válidas.

A pontuação final da avaliação é obtida pela média aritmética das três dimensões, resultando em um escore global entre 1 e 4, que representa a qualidade geral da resposta analisada (Adaptado de McFayden et al., 2024).

A Rubrica 1 apresenta os critérios e categorias de avaliação de respostas geradas por Inteligência Artificial, considerando os aspectos de corretude, clareza, concisão e referências bibliográficas. Cada critério é descrito a seguir, com suas respectivas categorias de desempenho.

1. Corretude (Correctness)

- Categoria 4 – Excelente: a resposta apresenta informações cientificamente precisas, atualizadas e sem erros factuais. Todas as afirmações podem ser verificadas em fontes confiáveis.
- Categoria 3 – Boa: a resposta é majoritariamente correta, apresentando apenas pequenas imprecisões ou omissões não essenciais.
- Categoria 2 – Regular: há vários erros menores ou interpretações ambíguas dos fatos.
- Categoria 1 – Insuficiente: a resposta contém informações incorretas, desatualizadas ou enganosas.

2. Clareza (Clarity)

- Categoria 4 – Excelente: a explicação é completamente compreensível, utiliza linguagem acessível e possui estrutura lógica que facilita a apreensão da ideia central.
- Categoria 3 – Boa: a resposta é clara em sua maior parte, embora contenha trechos densos, termos técnicos excessivos ou frases ambíguas.
- Categoria 2 – Regular: a estrutura ou o vocabulário dificultam a compreensão em diversos pontos.
- Categoria 1 – Insuficiente: a resposta é confusa, desorganizada ou utiliza linguagem inadequada ao público-alvo.

3. Concisão (Conciseness)

- Categoria 4 – Excelente: a resposta fornece exatamente o necessário para responder à questão, sem redundâncias ou omissões.
- Categoria 3 – Boa: a resposta é ligeiramente longa ou breve demais, mas mantém a essência do conteúdo.
- Categoria 2 – Regular: apresenta excesso de detalhes irrelevantes ou carece de informações importantes.

- Categoria 1 – Insuficiente: é extremamente prolixa ou tão breve que compromete a compreensão.

4. Referências Bibliográficas

- Categoria 4 – Excelente: as referências estão completas e corretas conforme as normas da APA, correspondem a artigos ou fontes verificáveis e provêm de periódicos indexados reconhecidos (preferencialmente Qualis A ou B).
- Categoria 3 – Boa: as referências são majoritariamente corretas e verificáveis, com pequenas inconsistências formais ou uso de periódicos de menor classificação.
- Categoria 2 – Regular: há erros de formatação ou inclusão de fontes não verificáveis, mas sem comprometer totalmente a credibilidade.
- Categoria 1 – Insuficiente: as referências estão incorretas, incompletas ou referem-se a fontes inexistentes ou não rastreáveis.

A Rubrica 2 apresenta a rubrica utilizada para avaliar a estrutura do aplicativo, composta pelos critérios de Responsividade, Usabilidade e Clareza Instrucional. Cada critério foi avaliado em quatro categorias de desempenho, variando de 1 (insuficiente) a 4 (excelente), com base na adequação técnica e na experiência do usuário.

O critério Responsividade refere-se à capacidade do aplicativo de adaptar-se adequadamente a diferentes tamanhos de tela e dispositivos, como celulares, tablets e computadores. A categoria 4 foi atribuída quando o aplicativo manteve a legibilidade, a proporção dos elementos e a funcionalidade completa em todos os dispositivos. A categoria 3 indicou boa adaptação, com pequenas distorções visuais ou ajustes mínimos necessários. A categoria 2 corresponde a situações em que o aplicativo funcionava em mais de um dispositivo, mas apresentava perdas perceptíveis de layout ou desempenho. Já a categoria 1 foi utilizada quando o aplicativo demonstrou falhas significativas de adaptação, comprometendo o uso em diferentes telas.

O critério Usabilidade avaliou o grau de facilidade de uso, coerência e funcionamento do aplicativo. A categoria 4 foi atribuída a aplicativos intuitivos, de fácil navegação, com todas as funções operando corretamente e fluxo de interação eficiente. A categoria 3 descreveu aplicativos de navegação simples, com funcionamento adequado das funções principais, embora com pequenas inconsistências. A categoria 2 representou aplicativos que exigiam maior esforço cognitivo do usuário, apresentando menus confusos ou funções parcialmente inoperantes. Por fim, a categoria 1 caracterizou aplicativos de difícil uso, com falhas frequentes, botões inativos ou estrutura de navegação confusa.

O critério Clareza Instrucional avaliou a presença e a qualidade das orientações fornecidas ao usuário sobre como utilizar o aplicativo. A categoria 4 foi atribuída quando o aplicativo apresentou instruções claras, objetivas e acessíveis, visíveis e compreensíveis para qualquer usuário. A categoria 3 indicou que as instruções eram compreensíveis, porém poderiam ser mais visíveis ou detalhadas. A categoria 2 referiu-se a instruções vagas, incompletas ou pouco acessíveis, que exigiam que o usuário deduzisse o funcionamento. Por fim, a categoria 1 representou ausência total de instruções ou explicações, tornando a utilização confusa ou inviável para novos usuários.

A Rubrica 2 apresenta os critérios e categorias de avaliação de aplicativos, considerando os aspectos de responsividade, usabilidade e clareza instrucional. Cada critério é descrito a seguir, com suas respectivas categorias de desempenho.

1. Responsividade

- Categoria 4 – Excelente: o aplicativo adapta-se perfeitamente a diferentes tamanhos de tela (celular, tablet e computador), mantendo legibilidade, proporção de elementos e funcionalidade completa em todos os dispositivos.
- Categoria 3 – Boa: o aplicativo apresenta adaptação adequada à maioria dos dispositivos, com pequenas distorções visuais ou ajustes mínimos necessários.
- Categoria 2 – Regular: o aplicativo funciona em mais de um tipo de dispositivo, mas com perda perceptível de layout, legibilidade ou desempenho.
- Categoria 1 – Insuficiente: o aplicativo apresenta falhas graves de adaptação, comprometendo a visualização e o uso em diferentes dispositivos.

2. Usabilidade

- Categoria 4 – Excelente: o aplicativo é intuitivo, de fácil navegação e todas as funções operam corretamente sem erros. O fluxo de interação é coerente e eficiente.
- Categoria 3 – Boa: a navegação é simples e as funções principais operam corretamente, embora pequenas falhas ou inconsistências possam ocorrer ocasionalmente.
- Categoria 2 – Regular: o aplicativo é utilizável, mas exige esforço cognitivo elevado, apresenta menus confusos ou funções parcialmente inoperantes.
- Categoria 1 – Insuficiente: o aplicativo é de difícil uso, com falhas frequentes, botões inativos ou navegação confusa.

3. Clareza Instrucional

- Categoria 4 – Excelente: o aplicativo apresenta instruções claras, objetivas e acessíveis sobre como utilizá-lo, localizadas de forma visível e compreensível para qualquer usuário.
- Categoria 3 – Boa: as instruções estão disponíveis e são compreensíveis, mas poderiam ser mais visíveis ou detalhadas.
- Categoria 2 – Regular: as instruções são vagas, incompletas ou pouco acessíveis, exigindo que o usuário deduza o funcionamento.
- Categoria 1 – Insuficiente: o aplicativo não fornece instruções ou explicações de uso, tornando sua utilização confusa ou inviável para novos usuários.

A Rubrica 3 apresenta a rubrica utilizada para avaliar a consistência na reformulação do tutorial, desenvolvida para mensurar o grau de fidelidade do material após a solicitação de alteração de cor, conforme descrito na seção Procedimentos. Essa avaliação tem como objetivo verificar se a inteligência artificial é capaz de alterar exclusivamente o elemento requisitado (a cor), mantendo inalteradas as demais partes do tutorial original.

A rubrica é composta por cinco subcritérios: Precisão ao reformular, Conteúdo, Estrutura (sequência), Layout e Outros aspectos. Cada subcritério foi avaliado em quatro categorias de desempenho, variando de 1 (insuficiente) a 4 (excelente).

O subcritério Precisão ao reformular avalia se a IA modificou apenas o aspecto solicitado (a cor) preservando a lógica, a sequência e o conteúdo original do tutorial. A categoria 4 indica reformulação precisa, enquanto categorias inferiores refletem alterações indevidas que comprometem a consistência do material.

O Conteúdo examina a coerência interna e a integridade das informações e instruções após a modificação. A Estrutura (sequência) avalia a manutenção da ordem e da progressão lógica das etapas. O Layout considera a clareza visual, a padronização e a legibilidade do tutorial reformulado. Por fim, Outros aspectos englobam a compatibilidade com o material original, a preservação de exemplos, imagens ou elementos instrucionais complementares.

A pontuação final é obtida pela média aritmética dos cinco subcritérios, resultando em um índice de consistência entre 1 e 4.

Adicionalmente, será realizada uma comparação entre os tutoriais reformulados pela IA analisada e pela Google AI, com o objetivo de identificar diferenças estruturais, de clareza e de precisão nas modificações realizadas. O relatório comparativo descreve detalhadamente essas variações, destacando o grau de consistência e fidelidade mantido por cada sistema de IA.

A Rubrica 3 apresenta os subcritérios e categorias de avaliação aplicados à reformulação de um tutorial, considerando aspectos de precisão ao reformular, conteúdo, estrutura (sequência), layout e outros aspectos. Cada subcritério é descrito a seguir, com suas respectivas categorias de desempenho.

1. Precisão ao Reformular

- Categoria 4 – Excelente: após a solicitação de alteração de cor, o tutorial mantém integralmente o conteúdo, a clareza e a estrutura didática originais, modificando apenas o aspecto solicitado.
- Categoria 3 – Boa: o tutorial preserva a maior parte das informações e da lógica original, apresentando pequenas alterações não intencionais que não comprometem a compreensão.
- Categoria 2 – Regular: o tutorial apresenta mudanças perceptíveis em diferentes partes, incluindo elementos não relacionados à cor, afetando parcialmente sua coerência.
- Categoria 1 – Insuficiente: o tutorial reformulado altera significativamente o conteúdo, a estrutura ou a finalidade, descaracterizando o material original.

2. Conteúdo

- Categoria 4 – Excelente: todas as informações permanecem corretas, coerentes e completas após a modificação de cor. As instruções são claras e correspondem ao propósito inicial do tutorial.

- Categoria 3 – Boa: o conteúdo mantém coerência geral, mas contém pequenas omissões, redundâncias ou reformulações desnecessárias.
- Categoria 2 – Regular: o tutorial apresenta inconsistências informativas, perda parcial de etapas ou contradições entre instruções.
- Categoria 1 – Insuficiente: o conteúdo é incoerente, com instruções incorretas, desorganizadas ou que comprometem o entendimento da tarefa.

3. Estrutura (Sequência)

- Categoria 4 – Excelente: a sequência lógica e hierárquica das etapas é mantida, preservando a ordem e o encadeamento instrucional após a modificação de cor.
- Categoria 3 – Boa: a estrutura geral é preservada, com pequenas variações que não afetam a compreensão.
- Categoria 2 – Regular: há reordenação indevida ou mudanças na hierarquia das etapas que dificultam a continuidade.
- Categoria 1 – Insuficiente: há reordenação indevida ou mudanças significativas na hierarquia das etapas que comprometem a lógica e a compreensão do tutorial.

4. Layout

- Categoria 4 – Excelente: o formato e o estilo do tutorial permanecem consistentes, com boa organização visual, padronização e clareza na apresentação das etapas.
- Categoria 3 – Boa: o layout é legível e coerente, apresentando apenas pequenas inconsistências visuais ou de formatação.
- Categoria 2 – Regular: o tutorial é funcional, mas apresenta layout irregular ou visualmente desorganizado.
- Categoria 1 – Insuficiente: o layout é confuso, com má formatação ou inconsistência visual que prejudica a compreensão do conteúdo.

5. Outros Aspectos

- Categoria 4 – Excelente: o tutorial mantém compatibilidade total com a proposta original, preservando todos os elementos instrucionais, exemplos e recursos complementares.
- Categoria 3 – Boa: há pequenas diferenças na compatibilidade ou substituições não intencionais em elementos secundários.
- Categoria 2 – Regular: o tutorial apresenta incompatibilidades ou perda de partes complementares, como exemplos ou imagens.
- Categoria 1 – Insuficiente: o tutorial torna-se inconsistente com a proposta inicial, apresentando falhas, erros ou omissões críticas.

Procedimento

Foi criada uma nova conta do aplicativo DeepSeek sem qualquer atividade anterior. Em seguida foi aberto novo chat e adicionado à caixa de diálogo o prompt de nível 1, em seguida foi solicitado a mudança da cor de fundo do tutorial sem demais alterações por meio do

prompt: “Altere a cor de fundo do tutorial mantendo exatamente todos os demais elementos presentes nele, sem alterá-los”. Além disso, foi pressionado o botão “continuar” para que o código fosse finalizado após a inserção dos prompts na caixa de diálogo. Cabe ressaltar que os botões “DeepThink” e “Search não foram utilizados durante a coleta

Em sequência, foi aberto um novo chat e os passos citados acima foram aplicados para os prompts de nível 2 e 3 de forma idêntica ao procedimento realizado para o prompt nível 1, com exceção do passo de alterar cor de fundo do tutorial.

3 RESULTADOS

Dados Gerais

A coleta de dados foi realizada dia 27 de outubro de 2025 e foram registrados os tempos necessários para a elaboração de cada uma das respostas.

Prompt nível 1: 02 minutos e 12 segundos

- Trocar de cor: 02 minutos e 10 segundos
- Link: <https://chat.deepseek.com/share/6uba1n5fnrpql86psu>

Prompt nível 2: 07 minutos e 21 segundos

- Link: <https://chat.deepseek.com/share/xbq2z2uc4v89lz194j>

Prompt nível 3: 05 minutos e 20 segundos

- Link: <https://chat.deepseek.com/share/uy6idl6jwd7ua6018y>

As respostas foram coletadas, analisadas qualitativamente e posteriormente foram classificadas a partir dos critérios descritos na Avaliação do Produto Gerado.

De maneira geral, a Inteligência Artificial gerou todos os tutoriais conforme solicitado, sem apresentar interrupções ou problemas durante o processo de produção. Todas as respostas foram estruturadas de forma contínua e coerente, mantendo o fluxo da geração do código e do conteúdo instrucional. Apenas no prompt de nível 2 foi necessário acionar o botão “continuar” para que a IA finalizasse a produção do código HTML, o que não interferiu na qualidade ou completude do material gerado. Nos demais níveis, a elaboração ocorreu integralmente de forma automática, evidenciando estabilidade no funcionamento da ferramenta ao longo da coleta.

Para ilustrar visualmente a configuração estrutural dos tutoriais produzidos nos três níveis de prompts, foram elaboradas figuras correspondentes às principais seções presentes em cada material. As imagens permitem observar como o modelo organizou elementos instrucionais, apresentou definições, estruturou passos de implementação e inseriu componentes interativos e de registro de dados, evidenciando as diferenças de complexidade entre os tutoriais 1, 2 e 3. A seguir, descrevemos cada figura, destacando o que cada uma representa no contexto dos produtos gerados pela IA.

A Figura 1 apresenta um exemplo da seção de definição utilizada no Tutorial 1, evidenciando sua estrutura mais simples, predominantemente textual, voltada para introduzir o conceito de Reforçamento Não Contingente (NCR) de forma direta. Na Figura 2, observa-se o formato adotado para o passo a passo desse mesmo tutorial, com instruções sequenciais e linguagem objetiva. A Figura 3 demonstra um exemplo de implementação prática contido no Tutorial 1, composto por um cenário clínico resumido e a aplicação do procedimento NCR.

A Figura 4 ilustra a introdução ao NCR no Tutorial 2, já com um layout mais elaborado, presença de navegação lateral e indicadores de progresso. A Figura 5 mostra a seção de fundamentos teóricos, que no nível 2 inclui explicações conceituais detalhadas e referências explícitas ao embasamento em ABA. A Figura 6 apresenta a estrutura do passo a passo nesse segundo tutorial, composta por etapas enumeradas e recursos visuais que conferem maior clareza instrucional. A Figura 7 demonstra um dos casos práticos interativos, característicos do nível 2, nos quais o usuário seleciona respostas clínicas. A Figura 8 mostra o sistema de registro de dados, que nesse nível inclui campos organizados e orientação para acompanhamento de frequência. A Figura 9 apresenta o resumo final do Tutorial 2, integrando síntese dos conteúdos vistos e fechamento do percurso instrucional.

A Figura 10 apresenta a introdução ao NCR no Tutorial 3, que mantém estrutura semelhante ao nível 2, porém aprimorada pelo uso da literatura fornecida (Vollmer et al., 1993). A Figura 11 demonstra os fundamentos teóricos desse terceiro tutorial, com explicações alinhadas diretamente ao artigo-base utilizado na elaboração. A Figura 12 apresenta o passo a passo no nível 3, organizado com maior detalhamento e precisão conceitual. A Figura 13 ilustra os casos práticos interativos desse nível, que seguem o mesmo formato do nível 2, porém com conteúdo mais robusto. A Figura 14 demonstra a seção de registro de dados no Tutorial 3, com campos orientados e integração de elementos vinculados ao estudo referenciado. Por fim, a Figura 15 apresenta o resumo final do tutorial, que sintetiza o conteúdo do nível 3 com base nas instruções e no artigo anexo, reforçando a profundidade e completude característica desse nível de estruturação.

A íntegra dos textos gerados pode ser vista nos ANEXOs I, II e III.



Figura 1. Exemplo Definição no Tutorial Referente ao Tutorial 1

Nota. Tela inicial do Tutorial 1 sobre Reforçamento Não Contingente (NCR), com a definição da estratégia — apresentação de reforçadores em intervalos fixos ou variáveis, independentemente do comportamento — e uma “Dica” sobre sua eficácia para comportamentos mantidos por reforçamento automático.

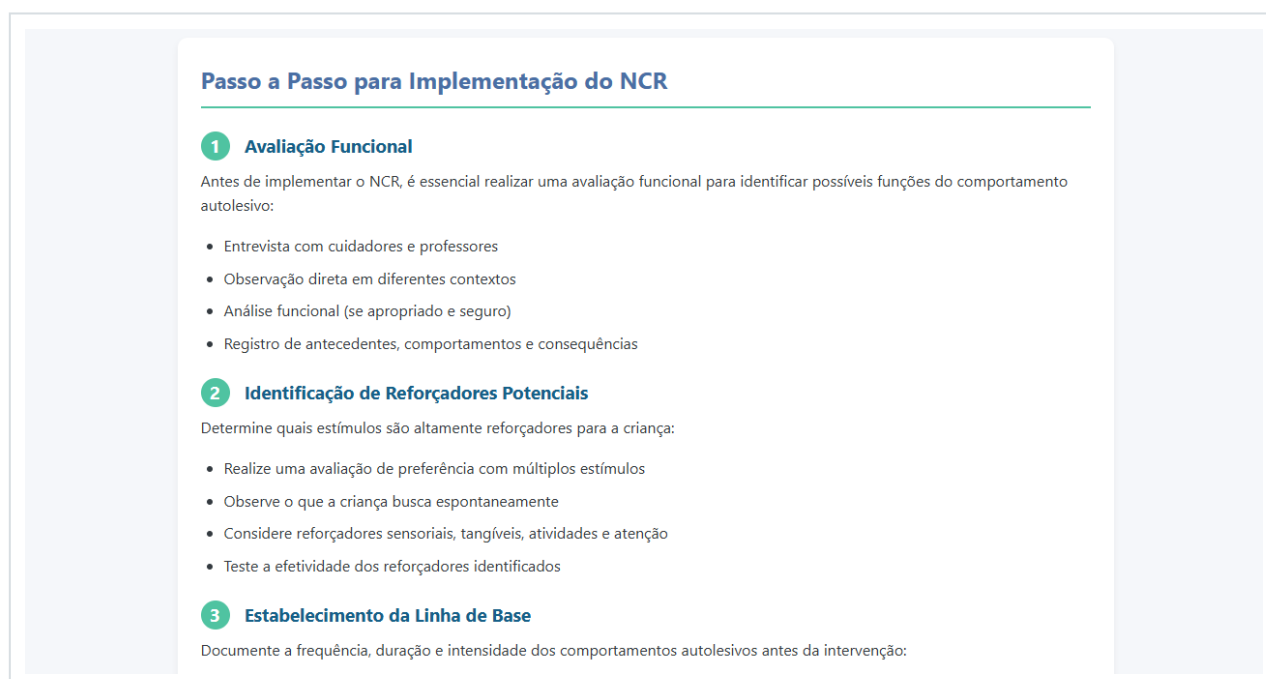


Figura 2. Exemplo de Passo a Passo Referente ao Tutorial 1

Nota. Tela da seção “Passo a Passo” do Tutorial 1, que detalha a sequência de implementação do NCR.

Exemplo de Implementação Prática

Cenário:
Criança de 8 anos com autismo que apresenta bater a cabeça contra superfícies duras, mantido por reforçamento automático (estimulação vestibular/proprioceptiva).

Procedimento NCR:

1. **Avaliação de preferência:** Identificou-se que balançar em uma rede e compressão articular são altamente preferidos.
2. **Linha de base:** Média de 15 episódios de bater a cabeça por sessão de 15 minutos.
3. **Esquema inicial:** FT 30s (fornecimento de balanço na rede a cada 30 segundos).
4. **Implementação:** Sessões de 10 minutos, 3 vezes ao dia.
5. **Resultados:** Após 5 sessões, redução para 3 episódios por sessão.
6. **Diluição:** Aumento gradual para FT 1min, depois FT 2min, mantendo a redução.

Figura 3. Exemplo de Implementação Prática Referente ao Tutorial 1

Nota. Exemplo de implementação prática do conteúdo do Tutorial 1, gerado por IA, ilustrando como a orientação do tutorial se traduz em uma atividade aplicada.

Tutorial NCR - Reforçamento Não Contingente
Progresso: 16% Iniciante

[Introdução ao NCR](#) [Fundamentos Teóricos](#) [Implementação Passo a Passo](#) [Casos Práticos](#) [Registro de Dados](#) [Resumo Final](#)

Introdução ao Reforçamento Não Contingente (NCR)

O que é Reforçamento Não Contingente?

O Reforçamento Não Contingente (NCR) é uma intervenção baseada em Análise do Comportamento Aplicada (ABA) que envolve a apresentação de um estímulo reforçador em intervalos de tempo fixos ou variáveis, independentemente do comportamento do indivíduo.

No contexto de comportamentos autolesivos em TEA, o NCR visa reduzir esses comportamentos ao fornecer acesso livre ao reforçador que mantém o comportamento problemático, diminuindo assim sua motivação.

Quando Utilizar o NCR?

O NCR é especialmente eficaz quando:

- O comportamento autolesivo é mantido por reforçamento automático (sensorial)
- Há dificuldade em identificar a função específica do comportamento

Figura 4. Exemplo de Introdução ao NCR Referente ao Tutorial 2

Nota. Aba “Introdução ao NCR” do tutorial, apresentando o NCR como intervenção baseada em ABA e os critérios de quando utilizá-lo em comportamentos autolesivos no TEA. Exibe a barra de progresso e as abas do tutorial.

Tutorial NCR - Reforçamento Não Contingente Progresso: 32% **Iniciante**

Introdução ao NCR **Fundamentos Teóricos** Implementação Passo a Passo Casos Práticos Registro de Dados Resumo Final

Fundamentos Teóricos do NCR

Base Conceitual na Análise do Comportamento

O NCR fundamenta-se no princípio da saciação, onde o acesso não contingente ao reforçador diminui o valor motivacional desse estímulo, reduzindo consequentemente a probabilidade de comportamentos mantidos por esse reforçador.

Esta intervenção altera a operação estabelecadora (EO) que evoca o comportamento problema, tornando o reforçador menos potente.

Análise Funcional do Comportamento

Antes de implementar o NCR, é essencial realizar uma análise funcional para identificar os possíveis reforçadores mantenedores do comportamento autolesivo:

- **Reforço Social Positivo:** Atenção, acesso a itens ou atividades
- **Reforço Social Negativo:** Fuga/esquiva de demandas

Figura 5. Exemplo de Fundamentos Teóricos Referente ao Tutorial 2

Nota. Tela da seção “Fundamentos Teóricos” do Tutorial 1, com a base conceitual do Reforçamento Não Contingente.

Tutorial NCR - Reforçamento Não Contingente Progresso: 48% **Iniciante**

Introdução ao NCR Fundamentos Teóricos **Implementação Passo a Passo** Casos Práticos Registro de Dados Resumo Final

Implementação Passo a Passo do NCR

- 1 Realizar Análise Funcional**

Identifique a função do comportamento autolesivo através de:

 - Entrevista com cuidadores (FAI - Functional Assessment Interview)
 - Observação direta (ABC recording)
 - Análise funcional experimental (condições padrão ou brief)
- 2 Selecionar o Reforçador**

Com base na análise funcional, selecione o reforçador identificado como mantenedor do comportamento:

 - Para comportamentos mantidos por atenção: ofereça atenção não contingente
 - Para comportamentos mantidos por fuga: ofereça pausas não contingentes
 - Para comportamentos mantidos por acesso a itens: ofereça acesso livre ao item
 - Para comportamentos mantidos automaticamente: ofereça estimulação alternativa

Figura 6. Exemplo de Implementação Passo a Passo Referente ao Tutorial 2

Nota. Tela da seção “Implementação Passo a Passo” do Tutorial 1, com as etapas práticas de aplicação do NCR.

Tutorial NCR - Reforçamento Não Contingente Progresso: 64% **Iniciante**

Introdução ao NCR Fundamentos Teóricos Implementação Passo a Passo **Casos Práticos** Registro de Dados Resumo Final

Casos Práticos Interativos

Caso 1: Lucas, 8 anos

Lucas apresenta comportamentos de bater a cabeça contra superfícies duras. A análise funcional indicou que o comportamento é mantido por reforço automático (estimulação vestibular). A frequência basal é de 15 episódios por hora.

Qual seria a intervenção NCR mais apropriada?

- Fornecer atenção a cada 2 minutos, independentemente do comportamento
- Oferecer acesso a balanço terapêutico a cada 30 segundos
- Permitir fuga de demandas a cada 5 minutos
- Fornecer comida preferida a cada minuto

Figura 7. Exemplo de Casos Práticos Referente ao Tutorial 2

Nota. Tela da seção “Casos Práticos” do Tutorial 2, com exemplos de aplicação do NCR.

Tutorial NCR - Reforçamento Não Contingente Progresso: 80% **Iniciante**

Introdução ao NCR Fundamentos Teóricos Implementação Passo a Passo Casos Práticos **Registro de Dados** Resumo Final

Registro de Dados e Progresso

Registro da Frequência de Comportamento Autolesivo

Adicione dados semanais da frequência de comportamentos autolesivos para visualizar o progresso da intervenção:

Semana: Frequência (pc): **Adicionar Dados**

Frequência (por hora)

1,0
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2

— Frequência de Comportamento Autolesivo (por hora)

Figura 8. Exemplo de Registro de Dados Referente ao Tutorial 2

Nota. Tela da seção “Registro de Dados” do Tutorial 2: o usuário insere a frequência semanal do comportamento autolesivo e visualiza o gráfico de progresso da intervenção.

The screenshot shows the 'Resumo Final e Recomendações' page of the 'Tutorial NCR - Reforçamento Não Contingente'. The page is titled 'Resumo Final e Recomendações' and features a progress bar at the top right indicating 'Progresso: 100%' and a 'Intermediário' level. The navigation menu includes 'Introdução ao NCR', 'Fundamentos Teóricos', 'Implementação Passo a Passo', 'Casos Práticos', 'Registro de Dados', and 'Resumo Final'. The main content area is titled 'Seu Desempenho no Tutorial' and is divided into three columns:

- Pontos Fortes:**
 - Compreensão básica dos princípios do NCR
 - Capacidade de identificar algumas funções do comportamento
- Áreas para Melhoria:**
 - Aprofundar conhecimento sobre análise funcional experimental
 - Praticar identificação de funções do comportamento em casos complexos
 - Desenvolver habilidades para implementação prática do NCR
- Próximos Passos Recomendados:**
 - Revisar fundamentos da análise do comportamento aplicada
 - Praticar com mais casos simulados antes da implementação clínica
 - Participar de supervisão com analista do comportamento experiente
 - Manter registro sistemático de dados durante implementações

At the bottom of the page, there are two buttons: 'Gerar Relatório em PDF' and 'Reiniciar Tutorial'.

Figura 9. Exemplo de Resumo Final Referente ao Tutorial 2

Nota. Tela da seção “Resumo Final” do Tutorial 2, que sintetiza os principais pontos do tutorial sobre NCR.

The screenshot shows the 'Introdução ao Reforçamento Não Contingente (NCR)' page of the 'Reforçamento Não Contingente (NCR) para Autolesão em TEA'. The page is titled 'Reforçamento Não Contingente (NCR) para Autolesão em TEA' and features a subtitle 'Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)'. The navigation menu includes 'Introdução', 'Fundamentos Teóricos', 'Implementação', 'Casos Práticos', 'Registro de Dados', and 'Resumo'. The main content area is titled 'Introdução ao Reforçamento Não Contingente (NCR)' and contains the following text:

O Reforçamento Não Contingente (NCR) é uma intervenção baseada em Análise do Comportamento Aplicada (ABA) que envolve a apresentação de um estímulo reforçador independentemente do comportamento do indivíduo. Diferente de procedimentos de reforçamento diferencial, no NCR o reforço não está vinculado a nenhuma resposta específica.

Por que usar NCR para comportamento autolesivo?

De acordo com Vollmer et al. (1993), o NCR pode ser uma alternativa eficaz ao Reforçamento Diferencial de Outros Comportamentos (DRO) para reduzir comportamentos autolesivos mantidos por reforçamento social positivo (atenção).

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. Journal of Applied Behavior Analysis, 26(1), 9-21.

Figura 10. Exemplo de Introdução ao NCR Referente ao Tutorial 3

Nota. Aba “Introdução ao NCR” do Tutorial 3, com a definição e as indicações de uso do Reforçamento Não Contingente.

Reforçamento Não Contingente (NCR) para Autolesão em TEA
Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução **Fundamentos Teóricos** Implementação Casos Práticos Registro de Dados Resumo

Fundamentos Teóricos do NCR

Mecanismos de Ação

O NCR reduz comportamentos problema através de dois mecanismos principais:

- 1. Extinção**
Ao fornecer o reforço funcional independentemente do comportamento-alvo, elimina-se a relação contingente entre o comportamento e sua consequência reforçadora. Isso enfraquece a resposta ao longo do tempo.
- 2. Saciação/Operação Estabelecadora Abolitiva**
O fornecimento frequente e não contingente do reforço pode reduzir seu valor motivacional, diminuindo a probabilidade de comportamentos mantidos por esse reforço.

Figura 11. Exemplo de Fundamentos Teóricos Referentes ao Tutorial 3

Nota. Tela da seção “Fundamentos Teóricos” do Tutorial 3.

Reforçamento Não Contingente (NCR) para Autolesão em TEA
Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução Fundamentos Teóricos **Implementação** Casos Práticos Registro de Dados Resumo

Implementação do NCR: Passo a Passo

Passo 1: Realizar Análise Funcional

Identifique a função do comportamento autolesivo através de uma análise funcional. Condições típicas incluem:

- **Sozinho:** Avalia se o comportamento é automaticamente reforçado
- **Atenção:** Avalia se o comportamento é mantido por reforçamento social positivo
- **Demanda:** Avalia se o comportamento é mantido por fuga/esquiva
- **Brincar:** Condição controle com ambiente enriquecido

Passo 2: Selecionar o Reforçador

Com base na análise funcional, selecione o reforçador identificado como mantenedor do comportamento. No caso de comportamento mantido por atenção, use atenção como reforçador no NCR.

Figura 12. Exemplo de Implementação Passo a Passo Referente ao Tutorial 3

Nota. Tela da seção “Implementação Passo a Passo” do Tutorial 3.

Casos Práticos Interativos

Teste seus conhecimentos com estes cenários baseados em casos reais. Selecione a intervenção mais apropriada para cada situação.

Seu Progresso

Nível: Iniciante

Pontuação: 0/0

Caso 1: Lucas, 8 anos

Lucas é uma criança de 8 anos com TEA que se engaja em bater a cabeça contra a parede. Uma análise funcional revelou que esse comportamento é mantido por atenção. Durante a linha de base, a frequência é de aproximadamente 6 episódios por hora. Qual seria a abordagem inicial mais apropriada para implementar NCR?

Iniciar com atenção contínua (esquema denso) e depois realizar fading gradual

Implementar imediatamente um esquema de atenção a cada 30 minutos

Usar um esquema de tempo variável desde o início

Implementar DRO em vez de NCR

Figura 13. Exemplo de Casos Práticos Referentes ao Tutorial 3

Nota. Tela da seção “Casos Práticos” do Tutorial 3, com exemplos de aplicação do NCR.

Registro de Dados e Progresso

Utilize esta ferramenta para registrar a frequência de comportamentos autolesivos ao longo do tempo e monitorar o progresso da intervenção com NCR.

Semana (ex: 1) Frequência (episódios/hora) Adicionar Dados

Frequência de Autolesão (episódios/hora)

Frequência (episódios/hora)

Semanas de Intervenção

Redefinir Dados

Interpretação do Gráfico

O gráfico acima mostra a frequência do comportamento autolesivo ao longo do tempo. Uma tendência decrescente indica que a intervenção está sendo eficaz. Se a frequência aumentar ou se estabilizar em níveis altos, considere ajustar o esquema de NCR (ex: retornar temporariamente a um esquema mais denso).

Dica Baseada em Evidências

No estudo de Vollmer et al. (1993), o NCR resultou em reduções significativas no comportamento autolesivo já nas primeiras sessões de tratamento. Se não observar redução após 2-3 semanas com um esquema adequado, reavalie a análise funcional.

Figura 14. Exemplo Registro de Dados Referente ao Tutorial 3

Nota. Tela da seção “Registro de Dados” do Tutorial 3, para acompanhamento da frequência do comportamento autolesivo ao longo da intervenção.

Resumo e Recomendações

Seu Desempenho no Tutorial
Pontuação Total: 0/0

Nível de Proficiência: Iniciante

Pontos Fortes

Áreas para Melhoria

- Aplicação de estratégias de fading
- Adaptação do NCR para diferentes funções comportamentais

Próximos Passos Recomendados
Recomenda-se revisar os fundamentos teóricos antes de implementar NCR na prática clínica.
Consulte as referências científicas para embasar suas intervenções.

Principais Conclusões sobre NCR

- NCR é uma intervenção baseada em evidências para reduzir comportamentos autolesivos mantidos por reforçamento social
- Deve ser precedida por análise funcional para identificar o reforçador mantenedor
- É mais eficaz quando iniciado com esquema denso seguido de fading gradual
- Pode atenuar efeitos colaterais da extinção e é mais fácil de implementar que DRO
- Pode ser combinado com reforçamento de comportamentos alternativos

Figura 15. Exemplo de Resumo Referente ao Tutorial 3

Nota. Tela da seção “Resumo” do Tutorial 3, com a síntese final do conteúdo sobre NCR.

Comparação Rubrica

A Tabela 1 apresenta a síntese das pontuações atribuídas aos três tutoriais gerados pelos diferentes níveis de estruturação dos prompts, reunindo em um único quadro comparativo todos os critérios avaliados como corretude, clareza, concisão, referências bibliográficas, responsividade, usabilidade, clareza instrucional, média geral e consistência após reformulação. Essa

apresentação permite visualizar de forma clara e imediata o desempenho de cada tutorial em relação aos parâmetros estabelecidos, facilitando a comparação entre os níveis e a identificação de diferenças decorrentes do grau de detalhamento do prompt.

Tabela 1. Pontuações atribuídas aos três tutoriais

Critério Avaliado	Tutorial 1	Tutorial 2	Tutorial 3
Corretude	4 (Excelente)	4 (Excelente)	4 (Excelente)
Clareza	3 (Boa)	3 (Boa)	3 (Boa)
Concisão	3 (Boa)	4 (Excelente)	4 (Excelente)
Referências Bibliográficas	4 (Excelente)	4 (Excelente)	3 (Boa)
Responsividade	4 (Excelente)	4 (Excelente)	4 (Excelente)
Usabilidade	4 (Excelente)	3 (Boa)	3 (Boa)
Clareza Instrucional	4 (Excelente)	4 (Excelente)	4 (Excelente)
Média Geral	3,71	3,71	3,57
Consistência após Reformulação	4 (Excelente)	—	—

Nota. Pontuações atribuídas aos três tutoriais gerados por IA em cada critério avaliado, conforme a rubrica do estudo.

No Tutorial 1, a resposta apresentou corretude classificada como excelente (4), indicando precisão conceitual no conteúdo produzido. A clareza foi avaliada como boa (3), porém observou-se o uso de termos técnicos ao longo do tutorial. A concisão foi considerada boa (3), em razão do texto ser mais breve, ainda que suficiente para cumprir sua finalidade. As referências bibliográficas foram avaliadas como excelentes (4). Em relação à estrutura do aplicativo, foram atribuídas pontuações excelentes (4) em responsividade, usabilidade e clareza instrucional, indicando coerência na navegação e compreensão do material. Quanto à avaliação após a solicitação de alteração de cor de fundo, foram atribuídas pontuações excelentes (4) em todos os subcritérios: precisão ao reformular, conteúdo, estrutura (sequência), layout e outros aspectos, mantendo integralmente o tutorial original. Dessa forma, o Tutorial 1 apresentou pontuação média de 3,71/4, e pontuação 4 na reformulação. Vale ressaltar que o tutorial gerado nesse nível possuía menor quantidade de páginas, botões e funcionalidades, o que pode ter facilitado a manutenção da consistência após a modificação.

No Tutorial 2, observou-se corretude excelente (4) e clareza boa (3), novamente com presença de termos técnicos. A concisão foi classificada como excelente (4), e as referências bibliográficas também receberam pontuação excelente (4). Sobre a estrutura do aplicativo, a responsividade foi avaliada como excelente (4), enquanto a clareza instrucional também obteve pontuação excelente (4). Entretanto, a usabilidade foi classificada como boa (3), pois na página de casos práticos, ao selecionar uma resposta, não era possível alterá-la após a escolha,

além de o botão de gerar PDF não estar funcional. Com isso, o Tutorial 2 apresentou pontuação média de 3,71/4, mantendo o mesmo desempenho geral do nível 1, porém com maior complexidade estrutural.

Por fim, no Tutorial 3, observou-se corretude excelente (4) e clareza boa (3), novamente com recorrência de termos técnicos. A concisão foi classificada como excelente (4). Já as referências bibliográficas foram avaliadas como boas (3), devido a uma citação de periódico apresentada de forma incorreta. Em relação à estrutura do tutorial, a responsividade apresentou pontuação excelente (4) e a clareza instrucional também excelente (4). A usabilidade recebeu pontuação boa (3) em razão de o botão de gerar PDF não estar funcional. Assim, o Tutorial 3 apresentou pontuação média de 3,57/4, sendo a menor entre os três níveis avaliados, embora mantendo coerência geral no material produzido.

Comparação de Conteúdo

A Tabela 2 apresenta uma síntese comparativa dos conteúdos incluídos nos tutoriais produzidos a partir dos três níveis de estruturação dos prompts, permitindo observar de maneira clara como cada versão incorporou os diferentes componentes instrucionais previstos. Nota-se que os três tutoriais contemplaram elementos essenciais, como a introdução ao NCR e o passo a passo de implementação, garantindo consistência mínima entre as versões. Contudo, conteúdos de maior complexidade, como casos práticos, sistemas de registro de dados e resumo final, aparecem apenas nos tutoriais gerados pelos prompts dos níveis 2 e 3, evidenciando que níveis mais altos de estruturação resultam em materiais mais completos e alinhados às demandas de um aplicativo instrucional voltado à prática profissional. Dessa forma, a tabela demonstra que a abrangência e a profundidade do conteúdo produzido pela IA variam diretamente em função do detalhamento presente no prompt.

Tabela 2. Comparação de Conteúdo Presente nos Tutoriais Produzidos

Conteúdo	Tutorial 1	Tutorial 2	Tutorial 3
Introdução ao NCR	X	X	X
Fundamentos teóricos		X	X
Passo a passo implementação	X	X	X
Exemplos de implementação	X	X	X
Casos práticos (Exercícios)		X	X
Registro de Dados		X	X
Resumo		X	X

Nota. Comparação do conteúdo presente em cada um dos três tutoriais, indicando quais elementos foram contemplados em cada versão.

Comparação Depois da Alteração

Foi solicitada à IA para alterar a cor do fundo para checar se as demais propriedades do Tutorial seriam mantidas após a reformulação e se somente a cor de fundo seria alterada. Em seguida, foi solicitado ao Google AI Studio que realizasse uma comparação detalhada entre o código original e o código gerado após a alteração da cor de fundo. O objetivo dessa análise foi verificar se a modificação solicitada afetou outros elementos estruturais, instrucionais ou visuais do material. A ferramenta forneceu um relatório técnico comparativo, destacando as diferenças encontradas linha a linha e avaliando o grau de fidelidade entre as versões.

O resultado dessa verificação está apresentado a seguir, com base no relatório gerado automaticamente pelo sistema: Os dois códigos HTML-CSS apresentados são praticamente idênticos, com uma única e sutil diferença no estilo CSS. A divergência reside especificamente na cor de fundo (`background-color`) definida para o elemento `body`. No primeiro código, a cor de fundo é `#f5f7fa`, um tom de cinza muito claro, quase branco. Já no segundo código, essa propriedade foi alterada para `#e8f4f8`, que representa um azul-claro bastante suave. Fora essa modificação na paleta de cores, não há outras distinções detectáveis no CSS ou em toda a estrutura e conteúdo HTML dos dois documentos.

4 DISCUSSÃO

Resumo

O objetivo do presente trabalho foi investigar como diferentes níveis de estruturação de prompts influenciam a qualidade de tutoriais produzidos pela Inteligência Artificial DeepSeek sobre intervenções baseadas em reforçamento não contingente para comportamentos autolesivos em crianças com TEA. Para isso, conduziu-se um delineamento experimental de sujeito único, no qual o modelo de IA foi exposto a três níveis distintos de prompts, variando do menos ao mais estruturado. Cada resposta gerada foi avaliada a partir de rubricas específicas que analisaram aspectos de conteúdo, clareza, concisão, uso de referências, além de critérios ligados à estrutura do aplicativo produzido. Também foi examinada a capacidade da IA de manter consistência após uma modificação pontual no código, permitindo comparar de forma sistemática o impacto do grau de detalhamento dos prompts sobre a qualidade final dos tutoriais.

De modo geral, os resultados demonstraram que o modelo DeepSeek apresentou alto nível de consistência e precisão em todas as etapas do experimento. A IA foi capaz de gerar tutoriais completos, tecnicamente corretos e visualmente organizados, apresentando desempenho estável nos diferentes níveis de estruturação dos prompts. Observou-se que, mesmo diante de solicitações pontuais de modificação como a alteração da cor de fundo, o modelo manteve a integridade do código original, preservando o conteúdo instrucional, a sequência lógica das etapas e a coerência visual do layout. Essa estabilidade reforça a capacidade do sistema de compreender comandos específicos sem comprometer a estrutura geral do material, o que representa um indicador relevante de confiabilidade técnica e funcional.

Além disso, a análise comparativa conduzida pelo Google AI Studio confirmou a manutenção da fidelidade estrutural nas versões modificadas, corroborando os achados obtidos pelas rubricas de avaliação. As alterações detectadas limitaram-se exclusivamente ao parâmetro estético solicitado, não sendo identificadas perdas de conteúdo, mudanças indevidas de elementos instrucionais ou inconsistências de formatação. Assim, tanto os dados quantitativos quanto as observações qualitativas indicam que o modelo DeepSeek apresentou desempenho satisfatório na reformulação de tutoriais, alcançando pontuações elevadas em precisão, coerência e clareza, e demonstrando potencial de aplicação confiável em contextos educacionais e clínicos mediados por Inteligência Artificial.

Verificou-se ainda que, nos níveis de estruturação mais altos dos prompts, os tutoriais tendem a ser mais completos e detalhados, incorporando elementos adicionais como fundamentação teórica ampliada, gráficos interativos, casos práticos e sistemas de feedback automatizado. Essa tendência sugere que a qualidade e a profundidade do material gerado pela IA estão diretamente relacionadas ao grau de especificidade e clareza das instruções fornecidas. Em outras palavras, quanto mais estruturado o prompt, maior a capacidade da IA de organizar o conteúdo de forma lógica, integrando aspectos técnicos, visuais e instrucionais de maneira coesa e alinhada às demandas profissionais da área da Psicologia.

Comparação com a Literatura

Os achados deste estudo dialogam diretamente com pesquisas recentes que investigam o uso de Inteligências Artificiais generativas na produção e transmissão de conhecimento técnico, especialmente no campo do autismo. Estudos como McFayden et al. (2024) demonstram que modelos generativos tendem a apresentar boa precisão conceitual, mas podem falhar na aplicabilidade prática e na consistência das referências, um padrão parcialmente reproduzido no presente trabalho, sobretudo nos tutoriais dos níveis 2 e 3, cujos botões de geração de PDF apresentaram funcionamento inoperante e onde foram identificadas imprecisões pontuais nas citações bibliográficas. Entretanto, diferentemente do relatado por esses autores, o modelo DeepSeek apresentou alto nível de estabilidade estrutural mesmo quando solicitado a realizar modificações pontuais, revelando maior controle sobre a manutenção do conteúdo original.

De modo convergente, Kim et al. (2025), em replicação do estudo de McFayden et al. (2024) conduzida com um público mais amplo de cuidadores de crianças em desenvolvimento típico, também observaram que a qualidade das respostas geradas por modelos de IA varia significativamente conforme o tipo e o grau de especificidade da pergunta apresentada, e que as respostas tendem a apresentar limitações quanto à dimensão relacional e à adequação ao contexto do público-alvo. Esse padrão encontra correspondência nos presentes resultados: prompts mais estruturados resultaram em tutoriais mais completos, detalhados e funcionalmente ricos, ao passo que o nível menos estruturado produziu material mais enxuto, porém igualmente preciso do ponto de vista conceitual. Ademais, assim como verificado por Kim et al. (2025), a forma de elaboração do prompt mostrou-se um fator determinante para a qualidade do produto gerado, indicando que a especificidade instrucional opera como uma variável de controle sobre o comportamento verbal do modelo. De forma geral, o presente estudo amplia o corpo de evidências ao associar o nível de estruturação do prompt à qualidade ins-

trucional, algo ainda pouco explorado de maneira sistemática na literatura, e ao fazê-lo com um modelo distinto do ChatGPT, o que contribui para a generalização dos achados a outros sistemas de IA generativa.

Cabe destacar ainda que os três tutoriais obtiveram pontuação máxima no critério de correção, o que indica que o modelo DeepSeek foi capaz de reproduzir com precisão os princípios da Análise do Comportamento Aplicada subjacentes ao NCR, incluindo a necessidade de avaliação funcional prévia, a seleção adequada do reforçador e o procedimento de desvanecimento gradual do esquema de reforçamento. Esse resultado é particularmente relevante em um contexto clínico, pois confirma que o material gerado pela IA não apenas organiza a informação de forma estruturada, mas também a representa de maneira tecnicamente fidedigna, atendendo ao critério de tecnologia descrito por Baer et al. (1968), segundo o qual uma prática científica deve ser descrita com precisão suficiente para ser replicada de forma independente. A estabilidade conceitual observada nos três níveis de estruturação sugere, portanto, que o modelo possui um repertório técnico consolidado sobre intervenções comportamentais para TEA, independentemente da complexidade do prompt fornecido.

Por outro lado, a recorrência de termos técnicos nos três tutoriais, identificada nos critérios de clareza avaliados como bons em vez de excelentes, aponta para uma limitação relevante no que diz respeito à acessibilidade do material para públicos com menor familiaridade com a terminologia analítico-comportamental. Esse achado converge com a preocupação levantada por Kim et al. (2025) acerca dos altos índices de legibilidade das respostas geradas por modelos de IA, que podem comprometer a utilidade prática dos materiais para profissionais em formação inicial ou para cuidadores sem formação especializada. Nesse sentido, pesquisas futuras poderiam investigar o efeito da inserção de instruções explícitas sobre o nível de linguagem desejado nos prompts, verificando se tal ajuste é capaz de tornar os tutoriais mais acessíveis sem comprometer sua precisão conceitual.

Contribuições teóricas

Um dos aspectos mais inovadores do presente estudo reside na proposta metodológica de tratar a Inteligência Artificial como participante de pesquisa em um delineamento experimental de sujeito único. Embora estudos envolvendo IA como objeto de análise sejam crescentes na literatura, a maioria os posiciona como ferramenta ou produto a ser avaliado, sem conferir ao modelo o estatuto epistemológico de unidade experimental. A decisão de adotar essa perspectiva não é meramente terminológica: ela carrega implicações conceituais e metodológicas que merecem apreciação crítica.

Do ponto de vista da Análise do Comportamento, o delineamento intrassujeito é historicamente associado ao estudo do comportamento de organismos individuais sob condições controladas, com ênfase na variabilidade intraindividual e na demonstração de relações funcionais entre variáveis independentes e dependentes (Sidman, 1960; Sampaio et al., 2008). A lógica subjacente a esse delineamento pressupõe que o mesmo organismo serve como seu próprio controle, eliminando a variabilidade interindividual que caracteriza os desenhos de grupo. Ao aplicar esse raciocínio à IA, o presente estudo explora uma analogia funcional relevante: assim como um participante humano é exposto sequencialmente a diferentes condi-

ções experimentais tendo seu comportamento medido em cada uma delas, o modelo Deep-Seek foi submetido a três níveis crescentes de estruturação de prompts, e suas respostas foram analisadas sistematicamente com base em critérios previamente definidos.

Essa analogia encontra respaldo na tradição de pesquisas em controle instrucional dentro da Análise do Comportamento. Estudos clássicos sobre o efeito de regras sobre o comportamento humano demonstraram que a precisão, a especificidade e o grau de estruturação das instruções verbais influenciam diretamente a qualidade e a consistência das respostas dos participantes (Zettle & Hayes, 1982; Hayes et al., 1986). Quando um professor recebe uma instrução vaga, como "elabore uma prova", versus uma instrução altamente especificada, que detalha formato, quantidade de questões e critérios de dificuldade, o produto resultante difere substancialmente. O presente estudo replica essa lógica no contexto da interação humano-IA: o prompt funciona operacionalmente como uma regra verbal que especifica contingências e direciona o comportamento verbal do modelo, de modo que o grau de sua estruturação atua como variável independente sobre a qualidade da resposta gerada.

Essa reconceptualização metodológica apresenta tanto potencial heurístico quanto limitações que precisam ser reconhecidas. Do lado das contribuições, ela permite a aplicação de um arcabouço experimental consolidado, o delineamento de sujeito único, a um novo tipo de "participante", abrindo espaço para investigações sistemáticas e replicáveis sobre o comportamento de modelos de linguagem. Ademais, ao controlar o histórico de interações por meio do uso de uma conta recém-criada, o estudo operacionaliza o conceito de ingenuidade experimental, análogo ao procedimento de controle de história de aprendizagem em estudos com animais ou humanos, o que representa um cuidado metodológico relevante e pouco discutido na literatura sobre avaliação de sistemas de IA.

Por outro lado, a analogia entre participante humano e modelo de linguagem exige cautela conceitual. O comportamento humano sob controle instrucional envolve história de reforçamento, processos de aprendizagem, motivação e variabilidade genuína ao longo do tempo. Os LLMs, por sua vez, não aprendem durante a sessão experimental, não acumulam história comportamental entre interações isoladas e não possuem estados motivacionais. Suas respostas são determinadas por parâmetros estocásticos fixos e pela arquitetura do modelo, o que significa que a "variabilidade" observada entre condições reflete essencialmente a sensibilidade do modelo ao input, e não um processo de aprendizagem ou adaptação em curso. Nesse sentido, a IA se assemelha menos a um organismo em experimentação e mais a um sistema de processamento cujos outputs são função direta da estrutura do estímulo de entrada.

Ainda assim, é precisamente essa previsibilidade que confere validade à proposta metodológica: se os outputs de um LLM são diretamente determinados pelos inputs, então a manipulação sistemática da estrutura do prompt constitui, de fato, uma manipulação experimental legítima de uma variável independente sobre uma variável dependente mensurável. Isso torna o modelo de IA um participante analiticamente útil, especialmente em pesquisas cujo objetivo é investigar as propriedades funcionais do próprio prompt, e não as capacidades do modelo em sentido amplo.

Em síntese, a contribuição teórica central deste estudo consiste em propor e operacionalizar um modelo metodológico no qual a IA é tratada como unidade experimental em um delineamento intrassujeito, fundamentando essa escolha na tradição analítico-comportamental de controle instrucional. Essa abordagem amplia o repertório metodológico disponível para pesquisadores que investigam a interação humano-IA e abre caminho para estudos futuros que explorem, com maior rigor experimental, os efeitos de diferentes variáveis de prompt sobre a qualidade, a precisão e a utilidade clínica das respostas geradas por modelos de linguagem em contextos aplicados, como o da psicologia e da educação especial.

Contribuições metodológicas

Metodologicamente, o estudo contribui ao demonstrar a viabilidade do uso de delineamentos experimentais de sujeito único aplicados a sistemas de IA, permitindo análises rigorosas e comparações diretas entre condições experimentais. A aplicação de rubricas detalhadas, inspiradas em instrumentos validados na literatura, fortalece a precisão da avaliação e cria um modelo replicável para futuras pesquisas que busquem mensurar a qualidade das respostas de IA.

Análise de consistência pós-reformulação na avaliação de sistemas generativos

Outro avanço metodológico foi a inclusão da análise de consistência após reformulação, que se mostrou uma forma efetiva de testar a estabilidade do modelo frente a modificações pontuais, revelando sua capacidade de alterar apenas o elemento solicitado sem comprometer a integridade do restante do material produzido. Essa abordagem expande o escopo metodológico disponível para investigações sobre segurança, confiabilidade e precisão de sistemas generativos.

Uso de IA para avaliar diferenças nos códigos

Adicionalmente, a utilização de uma ferramenta externa, o Google AI Studio, para a verificação comparativa entre versões do código representa uma estratégia de triangulação metodológica que fortalece a credibilidade dos dados obtidos pelas rubricas, permitindo validar de forma objetiva e independente as conclusões derivadas da avaliação qualitativa.

Contribuições Aplicadas

Do ponto de vista aplicado, o estudo demonstra que modelos de IA generativa podem atuar como ferramentas potentes para auxiliar profissionais da Psicologia na produção de materiais instrucionais sobre intervenções fundamentadas em ABA. Os tutoriais gerados nos níveis mais altos de estruturação apresentam potenciais usos em formação profissional, revisão de conteúdo e apoio clínico, integrando explicações conceituais, casos práticos e sistemas de registro. Além disso, a alta responsividade e clareza instrucional observadas nos três níveis indicam que aplicações desse tipo podem facilitar a disseminação de práticas baseadas em evidências, especialmente em contextos com escassez de materiais acessíveis e atualizados, como o cenário brasileiro de atendimento ao TEA descrito por Holanda & Maia (2024). Tais achados sugerem que, quando bem orientada por meio de prompts estruturados, a IA pode

colaborar significativamente na qualificação da prática clínica envolvendo comportamentos autolesivos em TEA, contribuindo para reduzir a distância entre o conhecimento produzido academicamente e sua apropriação por profissionais em contextos de recursos limitados.

Limitações e pesquisas futuras

Algumas limitações devem ser consideradas. Embora o prompt de nível 1 tenha sido definido como o menos estruturado, ele ainda continha uma indicação explícita sobre qual intervenção deveria ser utilizada, direcionando a IA ao uso do NCR. Isso significa que o modelo não precisou selecionar autonomamente a melhor intervenção com base no comportamento-alvo e no contexto apresentado. Assim, uma etapa importante para pesquisas futuras seria testar um prompt verdadeiramente pouco estruturado, sem qualquer referência à intervenção desejada, a fim de verificar se a IA seria capaz de identificar de forma independente uma estratégia comportamental adequada e, especialmente, se escolheria o NCR ou alguma outra intervenção baseada em evidências. Essa investigação permitiria avaliar a capacidade do modelo de raciocinar sobre casos clínicos, selecionar procedimentos compatíveis com a análise funcional implícita e construir tutoriais coerentes sem direcionamento prévio.

Outra limitação refere-se à ausência de avaliação por juízes independentes com expertise na área da Análise do Comportamento Aplicada, o que poderia conferir maior validade externa às pontuações atribuídas pelas rubricas. A inclusão de profissionais especializados como avaliadores em estudos futuros permitiria verificar o grau de concordância entre a avaliação realizada pelos pesquisadores e o julgamento clínico qualificado, fortalecendo a credibilidade dos critérios de qualidade adotados. Além disso, recomenda-se comparar os resultados com a performance de outras IAs generativas, possibilitando identificar diferenças intermodelo na sensibilidade às instruções e na precisão das escolhas clínicas. Por fim, seria relevante explorar a produção de tutoriais voltados a pais e cuidadores, examinando como diferentes níveis de detalhamento do prompt influenciam acessibilidade, clareza e aplicabilidade prática para esse público, bem como investigar a replicação deste procedimento com participantes humanos, a fim de verificar se os efeitos observados se mantêm em contextos aplicados ou em condições mais próximas da prática clínica.

5 CONCLUSÃO

O presente estudo investigou o efeito de diferentes níveis de estruturação de prompts sobre a qualidade de tutoriais produzidos pelo modelo de Inteligência Artificial generativa DeepSeek acerca de intervenções com reforçamento não contingente para comportamentos autolesivos em crianças com Transtorno do Espectro Autista. Os resultados obtidos permitem afirmar que o nível de estruturação dos prompts influencia diretamente a densidade, a abrangência e a complexidade dos materiais gerados, sem que isso implique, necessariamente, melhora proporcional na pontuação geral de qualidade. Os três tutoriais atingiram pontuação máxima em corretude e apresentaram desempenho consistente nos demais critérios ava-

liados, o que indica que o modelo possui repertório técnico consolidado sobre os princípios analítico-comportamentais subjacentes ao NCR, independentemente do grau de detalhamento da instrução fornecida.

Do ponto de vista do conteúdo instrucional, prompts mais estruturados resultaram em materiais mais completos, incorporando elementos como fundamentos teóricos ampliados, casos práticos interativos, sistemas de registro de dados e resumos com feedback automatizado. Essa relação entre especificidade instrucional e complexidade do produto gerado converge com os achados de McFayden et al. (2024) e Kim et al. (2025), que identificaram o grau de elaboração do prompt como variável determinante para a qualidade das respostas de modelos generativos. Ao mesmo tempo, o modelo demonstrou estabilidade estrutural frente a solicitações pontuais de modificação, preservando integralmente o conteúdo instrucional, a sequência lógica das etapas e a coerência visual do layout após a alteração da cor de fundo, o que representa um indicador relevante de confiabilidade técnica.

A recorrência de termos técnicos nos três tutoriais, refletida nas pontuações de clareza classificadas como boas em vez de excelentes, aponta para uma limitação relacionada à acessibilidade do material para públicos com menor familiaridade com a terminologia analítico-comportamental. Esse dado reforça a necessidade de que pesquisas futuras investiguem o efeito da inserção de instruções explícitas sobre o nível de linguagem desejado nos prompts, verificando se tal ajuste é capaz de ampliar a utilidade prática dos tutoriais sem comprometer sua precisão conceitual. Também se recomenda que estudos subsequentes testem prompts sem qualquer indicação prévia da intervenção a ser utilizada, a fim de avaliar a capacidade do modelo de selecionar autonomamente estratégias comportamentais adequadas a partir de uma descrição do comportamento-alvo e do contexto clínico.

De modo geral, os achados do presente trabalho indicam que modelos de IA generativa, quando orientados por prompts adequadamente estruturados, apresentam potencial significativo como ferramentas de apoio à elaboração de materiais instrucionais baseados em evidências para profissionais da Psicologia que atuam com TEA. A combinação entre precisão conceitual, organização instrucional e estabilidade estrutural observada nos tutoriais gerados sugere que essa tecnologia pode contribuir para reduzir a distância entre o conhecimento científico produzido academicamente e sua apropriação por profissionais em contextos com acesso limitado a recursos atualizados, como o cenário brasileiro de atendimento ao autismo descrito por Holanda & Maia (2024). Tais contribuições, contudo, devem ser interpretadas com cautela, considerando as limitações metodológicas do presente estudo e a necessidade de replicação com avaliadores independentes e com outros modelos de IA generativa.

REFERÊNCIAS BIBLIOGRÁFICAS

- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. American Psychiatric Publishing.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, 1(1), 91–97.

- Ceppi, B., & Benvenuti, M. (2011). Análise funcional do comportamento autolesivo. *Revista de Psiquiatria Clínica*, 38(6), 247–253.
- Cruz, A. E. A. Q., & Moreira, M. B. (2021). *Autismo: estratégias científicas para lidar com comportamentos desafiadores* (1ª ed.). Instituto Walden4.
- Ferreira, L. A., Silva, Á. J. M., & Barros, R. S. (2016). Ensino de aplicação de tentativas discretas a cuidadores de crianças diagnosticadas com autismo. *Perspectivas em Análise do Comportamento*, 7(1), 101–113.
- Gao, T., Jin, J., Ke, Z. T., & Moryoussef, G. (2025). A comparison of DeepSeek and other LLMs. arXiv. <https://arxiv.org/abs/2502.03688>
- Goulart, P., & de Assis, G. J. A. (2002). Estudos sobre autismo em análise do comportamento: aspectos metodológicos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 4(2), 151–165.
- Hayes, S. C., Brownstein, A. J., Zettle, R. D., Rosenfarb, I., & Korn, Z. (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of the Experimental Analysis of Behavior*, 45(3), 237–256. <https://doi.org/10.1901/jeab.1986.45-237>
- Holanda Júnior, F. W. N., & Maia, S. N. dos S. (2024). Prática profissional de psicólogos que atuam com intervenção para o transtorno do espectro autista. *Interação Em Psicologia*, 28(2). <https://doi.org/10.5380/riep.v28i2.86890>
- Kim, Y. B., Vilches, S. L., Shapiro, S., & Clarkson, A. (2025). Testing the capability of generative artificial intelligence for parent and caregiver information seeking. *Family Relations*, 74(3), 1266–1284. <https://doi.org/10.1111/fare.13167>
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135–148. <https://doi.org/10.1089/cyber.2023.0202>
- Moreira, M. B., & Medeiros, C. A. (2019). *Princípios básicos de análise do comportamento* (2ª ed.). Artmed.
- Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., de Freitas, C. S., Telles, M., Teixeira, J. S., da Escóssia, F., & dos Santos, L. C. (2024). ChatGPT and other artificial intelligences will transform scientific research: Reflections on their uses. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432E008>
- Sampaio, A. A. S., Azevedo, F. H. B., Cardoso, L. R. D., Lima, C., Pereira, M. B. R., & Andery, M. A. P. A. (2008). Uma introdução aos delineamentos experimentais de sujeito único. *Interação em Psicologia*, 12(1), 151–164. <https://doi.org/10.5380/psi.v12i1.9537>
- Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. Basic Books.
- Souza, M. E. de C., & Weigang, L. (2025). Grok, Gemini, ChatGPT and DeepSeek: Comparison and applications in conversational artificial intelligence. Zenodo. <https://doi.org/10.5281/zenodo.14885243>
- Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9–21.
- Zettle, R. D., & Hayes, S. C. (1982). Rule-governed behavior: A potential theoretical framework for cognitive-behavioral therapy. In P. C. Kendall (Ed.), *Advances in cognitive-behavioral research and therapy* (Vol. 1, pp. 73–118). Academic Press.

ANEXOS

Anexo I - Transcrição do produto pelo prompt nível 1

Reforçamento Não Contingente (NCR)

Tutorial para intervenção em comportamentos autolesivos em crianças autistas

O que é Reforçamento Não Contingente (NCR)?

O Reforçamento Não Contingente (NCR) é uma estratégia comportamental que envolve a apresentação de reforçadores de alta qualidade em intervalos fixos ou variáveis de tempo, independentemente do comportamento do indivíduo. Esta abordagem visa reduzir comportamentos problemáticos ao diminuir a motivação para emití-los.

Dica: O NCR é particularmente eficaz para comportamentos mantidos por reforçamento automático (como autoestimulação) ou quando a função do comportamento é difícil de determinar.

Passo a Passo para Implementação do NCR

1 Avaliação Funcional

Antes de implementar o NCR, é essencial realizar uma avaliação funcional para identificar possíveis funções do comportamento autolesivo:

Entrevista com cuidadores e professores

Observação direta em diferentes contextos

Análise funcional (se apropriado e seguro)

Registro de antecedentes, comportamentos e consequências

2 Identificação de Reforçadores Potenciais

Determine quais estímulos são altamente reforçadores para a criança:

Realize uma avaliação de preferência com múltiplos estímulos

Observe o que a criança busca espontaneamente

Considere reforçadores sensoriais, tangíveis, atividades e atenção

Teste a efetividade dos reforçadores identificados

3 Estabelecimento da Linha de Base

Documente a frequência, duração e intensidade dos comportamentos autolesivos antes da intervenção:

Registre por pelo menos 3-5 sessões

Use medidas objetivas e quantificáveis

Estabeleça critérios claros para o que constitui o comportamento autolesivo

4 Determinação do Esquema de Reforçamento Inicial

Selecione um esquema de reforçamento que forneça o reforçador com frequência suficiente para competir com o comportamento autolesivo:

Comece com intervalos curtos (ex: 30 segundos a 2 minutos)

Considere esquemas de tempo fixo (FT) ou variável (VT)

O objetivo é saturar o ambiente com reforçadores

5 Implementação do Procedimento NCR

Execute o procedimento de acordo com o esquema estabelecido:

Forneça o reforçador nos intervalos programados, independentemente do comportamento

Não forneça atenção ou reação aos comportamentos autolesivos (quando seguro)

Mantenha sessões curtas inicialmente (5-15 minutos)

Gradualmente aumente a duração das sessões

6 Coleta de Dados Contínua

Monitore continuamente a ocorrência de comportamentos autolesivos:

Registre a frequência em cada sessão

Compare com a linha de base

Ajuste o procedimento conforme necessário

7 Diluição do Esquema de Reforçamento

Quando os comportamentos autolesivos diminuírem significativamente, comece a aumentar gradualmente os intervalos entre reforçadores:

Aumente os intervalos em pequenos incrementos

Mantenha a redução dos comportamentos problema

Se os comportamentos aumentarem, retorne ao esquema anterior

8 Generalização e Manutenção

Implemente estratégias para garantir que os ganhos se mantenham em diferentes contextos:

Treine outros cuidadores para implementar o procedimento

Aplique em diferentes ambientes

Incorpore o NCR em rotinas diárias

Atenção: Sempre priorize a segurança da criança. Em casos de comportamentos autolesivos severos, considere o uso de equipamentos de proteção e consulte sempre uma equipe multidisciplinar.

Exemplo de Implementação Prática

Cenário:

Criança de 8 anos com autismo que apresenta bater a cabeça contra superfícies duras, mantido por reforçamento automático (estimulação vestibular/proprioceptiva).

Procedimento NCR:

Avaliação de preferência: Identificou-se que balançar em uma rede e compressão articular são altamente preferidos.

Linha de base: Média de 15 episódios de bater a cabeça por sessão de 15 minutos.

Esquema inicial: FT 30s (fornecimento de balanço na rede a cada 30 segundos).

Implementação: Sessões de 10 minutos, 3 vezes ao dia.

Resultados: Após 5 sessões, redução para 3 episódios por sessão.

Diluição: Aumento gradual para FT 1min, depois FT 2min, mantendo a redução.

Referências Bibliográficas

Carr, J. E., Coriaty, S., & Dozier, C. L. (2000). Current issues in the function-based treatment of aberrant behavior in individuals with developmental disabilities. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 91-112). Context Press.

Hagopian, L. P., Dozier, C. L., Rooker, G. W., & Jones, B. A. (2013). Assessment and treatment of severe problem behavior. In G. J. Madden (Ed.), *APA handbook of behavior analysis* (Vol. 2, pp. 353-386). American Psychological Association.

Lomas, J. E., Fisher, W. W., & Kelley, M. E. (2010). The effects of variable-time delivery of food items and praise on problem behavior reinforced by escape. *Journal of Applied Behavior Analysis*, 43(3), 425-435.

Tucker, M., Sigafos, J., & Bushell, H. (1998). Use of noncontingent reinforcement in the treatment of challenging behavior: A review and clinical guide. *Behavior Modification*, 22(4), 529-547.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9-21.

Este material é destinado a profissionais qualificados. Consulte sempre a literatura científica atualizada e supervisores especializados.

© 2023 - Tutorial NCR para Comportamentos Autolesivos

Anexo II - Transcrição do produto pelo prompt nível 2

Página 1

 Tutorial NCR - Reforçamento Não Contingente

Progresso: 16%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Introdução ao Reforçamento Não Contingente (NCR)

O que é Reforçamento Não Contingente?

O Reforçamento Não Contingente (NCR) é uma intervenção baseada em Análise do Comportamento Aplicada (ABA) que envolve a apresentação de um estímulo reforçador em intervalos de tempo fixos ou variáveis, independentemente do comportamento do indivíduo.

No contexto de comportamentos autolesivos em TEA, o NCR visa reduzir esses comportamentos ao fornecer acesso livre ao reforçador que mantém o comportamento problemático, diminuindo assim sua motivação.

Quando Utilizar o NCR?

O NCR é especialmente eficaz quando:

O comportamento autolesivo é mantido por reforçamento automático (sensorial)

Há dificuldade em identificar a função específica do comportamento

Intervenções contingentes apresentam limitações práticas ou éticas

Busca-se uma intervenção de baixa demanda para o cliente

Vantagens do NCR

Redução rápida de comportamentos problemáticos

Fácil implementação após análise funcional

Baixa taxa de comportamentos colaterais indesejados

Pode ser combinado com outras intervenções baseadas em ABA

Referência: Carr, J. E., Coriaty, S., & Dozier, C. L. (2000). Current issues in the function-based treatment of aberrant behavior in individuals with developmental disabilities. In J. Austin & J. E. Carr (Eds.), Handbook of applied behavior analysis (pp. 91-112). Context Press.

Próximo: Fundamentos Teóricos

Página 2



Tutorial NCR - Reforçamento Não Contingente

Progresso: 32%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Fundamentos Teóricos do NCR

Base Conceitual na Análise do Comportamento

O NCR fundamenta-se no princípio da saciação, onde o acesso não contingente ao reforçador diminui o valor motivacional desse estímulo, reduzindo conseqüentemente a probabilidade de comportamentos mantidos por esse reforçador.

Esta intervenção altera a operação estabelecadora (EO) que evoca o comportamento problema, tornando o reforçador menos potente.

Análise Funcional do Comportamento

Antes de implementar o NCR, é essencial realizar uma análise funcional para identificar os possíveis reforçadores mantenedores do comportamento autolesivo:

Reforço Social Positivo: Atenção, acesso a itens ou atividades

Reforço Social Negativo: Fuga/esquiva de demandas

Reforço Automático: Estimulação sensorial interna

O NCR é particularmente eficaz para comportamentos mantidos por reforço automático.

Evidências Científicas

Estudos demonstram a eficácia do NCR na redução de comportamentos autolesivos:

Vollmer, Iwata, Zarcone, Smith, & Mazaleski (1993) - Redução de 80-90% em comportamentos autolesivos

Hagopian, Fisher, & Legacy (1994) - Eficácia em casos de automutilação grave

Tucker, Sigafos, & Bushell (1998) - Manutenção dos efeitos em follow-up de 6 meses

Referência: Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Non-contingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9-21.

Anterior: Introdução

Próximo: Implementação

Tutorial NCR - Reforçamento Não Contingente para Comportamento Autolesivo em TEA

Desenvolvido para psicólogos com base em práticas baseadas em evidências da Análise do Comportamento Aplicada (ABA)

Página 3



Tutorial NCR - Reforçamento Não Contingente

Progresso: 48%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Implementação Passo a Passo do NCR

1

Realizar Análise Funcional

Identifique a função do comportamento autolesivo através de:

Entrevista com cuidadores (FAI - Functional Assessment Interview)

Observação direta (ABC recording)

Análise funcional experimental (condições padrão ou brief)

2

Selecionar o Reforçador

Com base na análise funcional, selecione o reforçador identificado como mantenedor do comportamento:

Para comportamentos mantidos por atenção: ofereça atenção não contingente

Para comportamentos mantidos por fuga: ofereça pausas não contingentes

Para comportamentos mantidos por acesso a itens: ofereça acesso livre ao item

Para comportamentos mantidos automaticamente: ofereça estimulação alternativa

3

Determinar o Esquema de Entrega

Estabeleça um esquema de entrega do reforçador:

Inicie com um esquema de alta densidade (ex: FR1 ou VR1)

Use intervalos de tempo fixos (FT) ou variáveis (VT)

O intervalo inicial deve ser menor que a taxa basal do comportamento

4

Implementar o NCR

Procedimentos de implementação:

Apresente o reforçador de acordo com o esquema estabelecido

Não exija nenhum comportamento do cliente para receber o reforçador

Ignore comportamentos autolesivos (quando seguro)

Monitore continuamente a frequência do comportamento

5

Desvanecimento do Esquema

Após redução significativa do comportamento, inicie o desvanecimento:

Aumente gradualmente o intervalo entre entregas do reforçador

Mude para esquemas de reforçamento mais naturais

Introduza gradualmente demandas durante as sessões

Referência: Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of noncontingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis*, 27(2), 317-325.

Anterior: Fundamentos

Próximo: Casos Práticos

Tutorial NCR - Reforçamento Não Contingente para Comportamento Autolesivo em TEA

Desenvolvido para psicólogos com base em práticas baseadas em evidências da Análise do Comporta

Página 4

 Tutorial NCR - Reforçamento Não Contingente

Progresso: 64%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Casos Práticos Interativos

Caso 1: Lucas, 8 anos

Lucas apresenta comportamentos de bater a cabeça contra superfícies duras. A análise funcional indicou que o comportamento é mantido por reforço automático (estimulação vestibular). A frequência basal é de 15 episódios por hora.

Qual seria a intervenção NCR mais apropriada?

Fornecer atenção a cada 2 minutos, independentemente do comportamento

Oferecer acesso a balanço terapêutico a cada 30 segundos

Permitir fuga de demandas a cada 5 minutos

Fornecer comida preferida a cada minuto

Próximo Caso

Próximo: Registro de Dados

Tutorial NCR - Reforçamento Não Contingente para Comportamento Autolesivo em TEA

Desenvolvido para psicólogos com base em práticas baseadas em evidências da Análise do Com

Página 5

 Tutorial NCR - Reforçamento Não Contingente

Progresso: 80%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Registro de Dados e Progresso

Registro da Frequência de Comportamento Autolesivo

Adicione dados semanais da frequência de comportamentos autolesivos para visualizar o progresso da intervenção:

Semana

Frequência (por hora)

Adicionar Dados

Análise do Progresso

Com base nos dados registrados, o sistema calculará automaticamente:

Taxa de redução do comportamento

Eficácia da intervenção

Necessidade de ajustes no protocolo

Adicione dados para ver a análise do progresso.

Anterior: Casos Práticos

Próximo: Resumo Final

Tutorial NCR - Reforçamento Não Contingente para Comportamento Autolesivo em TEA

Desenvolvido para psicólogos com base em práticas baseadas em evidências da Análise do Comportamento Aplicada (ABA)

Página 6

 Tutorial NCR - Reforçamento Não Contingente

Progresso: 100%

Iniciante

Introdução ao NCR

Fundamentos Teóricos

Implementação Passo a Passo

Casos Práticos

Registro de Dados

Resumo Final

Resumo Final e Recomendações

Seu Desempenho no Tutorial

Pontos Fortes

Áreas para Melhoria

Próximos Passos Recomendados

Gerar Relatório em PDF

Reiniciar Tutorial

Referências Científicas

Carr, J. E., Coriaty, S., & Dozier, C. L. (2000). Current issues in the function-based treatment of aberrant behavior in individuals with developmental disabilities. In J. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 91-112). Context Press.

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9-21.

Hagopian, L. P., Fisher, W. W., & Legacy, S. M. (1994). Schedule effects of noncontingent reinforcement on attention-maintained destructive behavior in identical quadruplets. *Journal of Applied Behavior Analysis*, 27(2), 317-325.

Tucker, M., Sigafos, J., & Bushell, H. (1998). Use of noncontingent reinforcement in the treatment of challenging behavior: A review and clinical guide. *Behavior Modification*, 22(4), 529-547.

Tutorial NCR - Reforçamento Não Contingente para Comportamento Autolesivo em TEA

Desenvolvido para psicólogos com base em práticas baseadas em evidências da Análise do Comportamento Aplicada (ABA)

Anexo III - Transcrição do produto pelo prompt nível 3

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Introdução ao Reforçamento Não Contingente (NCR)

O Reforçamento Não Contingente (NCR) é uma intervenção baseada em Análise do Comportamento Aplicada (ABA) que envolve a apresentação de um estímulo reforçador independentemente do comportamento do indivíduo. Diferente de procedimentos de reforçamento diferencial, no NCR o reforço não está vinculado a nenhuma resposta específica.

Por que usar NCR para comportamento autolesivo?

De acordo com Vollmer et al. (1993), o NCR pode ser uma alternativa eficaz ao Reforçamento Diferencial de Outros Comportamentos (DRO) para reduzir comportamentos autolesivos mantidos por reforçamento social positivo (atenção).

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9-21.

Vantagens do NCR sobre DRO

Reduz comportamentos induzidos por extinção (como agressão e respostas emocionais)

Taxa de reforço mais alta e previsível

Mais fácil de implementar por cuidadores

Menos propenso a criar privação de interação social

Quando considerar NCR

O NCR é particularmente indicado quando:

Uma análise funcional identificou que o comportamento é mantido por reforçamento social positivo

O indivíduo tem um repertório comportamental limitado

Há necessidade de uma intervenção rápida para reduzir comportamentos de alto risco

Cuidadores têm dificuldade em implementar procedimentos complexos de monitoramento

Próximo: Fundamentos Teóricos

Página 2

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Fundamentos Teóricos do NCR

Mecanismos de Ação

O NCR reduz comportamentos problema através de dois mecanismos principais:

1. Extinção

Ao fornecer o reforço funcional independentemente do comportamento-alvo, elimina-se a relação contingente entre o comportamento e sua consequência reforçadora. Isso enfraquece a resposta ao longo do tempo.

2. Saciação/Operação Estabelecadora Abolitiva

O fornecimento frequente e não contingente do reforço pode reduzir seu valor motivacional, diminuindo a probabilidade de comportamentos mantidos por esse reforço.

Evidências Científicas

O estudo de Vollmer et al. (1993) demonstrou que:

NCR e DRO foram igualmente eficazes na redução de comportamentos autolesivos mantidos por atenção

NCR atenuou comportamentos induzidos por extinção em 2 de 3 participantes

NCR resultou em taxas de reforço mais altas do que DRO

NCR foi mais fácil de implementar do que DRO

Análise Funcional como Pré-requisito

Antes de implementar NCR, é essencial realizar uma análise funcional para identificar a função do comportamento autolesivo. O NCR é mais eficaz quando utiliza o reforçador identificado como mantenedor do comportamento.

No estudo de Vollmer et al. (1993), todos os participantes tiveram autolesão mantida por atenção, confirmada através de análise funcional baseada no modelo de Iwata et al. (1982).

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities*, 2, 1-20.

Próximo: Implementação

Página 3

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Implementação do NCR: Passo a Passo

Passo 1: Realizar Análise Funcional

Identifique a função do comportamento autolesivo através de uma análise funcional. Condições típicas incluem:

Sozinho: Avalia se o comportamento é automaticamente reforçado

Atenção: Avalia se o comportamento é mantido por reforçamento social positivo

Demanda: Avalia se o comportamento é mantido por fuga/esquiva

Brincar: Condição controle com ambiente enriquecido

Passo 2: Selecionar o Reforçador

Com base na análise funcional, selecione o reforçador identificado como mantenedor do comportamento. No caso de comportamento mantido por atenção, use atenção como reforçador no NCR.

Passo 3: Determinar o Esquema Inicial

Inicie com um esquema denso de reforçamento. Vollmer et al. (1993) começaram com atenção contínua (6 interações por minuto) e depois realizaram fading gradual.

Passo 4: Implementar o NCR

Forneça o reforçador em intervalos de tempo fixos (esquema de tempo fixo - FT), independentemente do comportamento do indivíduo.

Exemplo de Implementação

Para um esquema FT de 10 segundos:

A cada 10 segundos, forneça atenção por 10 segundos

Não espere nenhum comportamento específico

Ignore o comportamento autolesivo (extinção)

Passo 5: Realizar Fading do Esquema

Gradualmente, aumente o intervalo entre as entregas de reforço conforme o comportamento autolesivo diminui. Vollmer et al. (1993) reduziram gradualmente de 6/min para 0,2/min (uma entrega a cada 5 minutos).

Passo 6: Monitorar e Ajustar

Registre a frequência do comportamento autolesivo e ajuste o esquema conforme necessário. Se o comportamento aumentar, retorne temporariamente a um esquema mais denso.

Precauções e Considerações

Evite reforço acidental do comportamento autolesivo (não forneça atenção imediatamente após o comportamento)

Combine com reforçamento de comportamentos alternativos quando possível

Considere características individuais e contexto ao implementar

Próximo: Casos Práticos

Página 4

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Casos Práticos Interativos

Teste seus conhecimentos com estes cenários baseados em casos reais. Selecione a intervenção mais apropriada para cada situação.

Seu Progresso

Nível: Intermediário

Pontuação: 1/2

Caso 1: Lucas, 8 anos

Lucas é uma criança de 8 anos com TEA que se engaja em bater a cabeça contra a parede. Uma análise funcional revelou que esse comportamento é mantido por atenção. Durante a linha de base, a frequência é de aproximadamente 6 episódios por hora. Qual seria a abordagem inicial mais apropriada para implementar NCR?

Iniciar com atenção contínua (esquema denso) e depois realizar fading gradual

Implementar imediatamente um esquema de atenção a cada 30 minutos

Usar um esquema de tempo variável desde o início

Implementar DRO em vez de NCR

Correto! Vollmer et al. (1993) recomendaram iniciar com um esquema denso (atenção contínua) e depois realizar fading gradual conforme o comportamento diminui.

Próximo Caso

Página 5

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Registro de Dados e Progresso

Utilize esta ferramenta para registrar a frequência de comportamentos autolesivos ao longo do tempo e monitorar o progresso da intervenção com NCR.

Semana (ex: 1)

Frequência (episódios/hora)

Adicionar Dados

Redefinir Dados

Interpretação do Gráfico

O gráfico acima mostra a frequência do comportamento autolesivo ao longo do tempo. Uma tendência decrescente indica que a intervenção está sendo eficaz. Se a frequência aumentar ou se estabilizar em níveis altos, considere ajustar o esquema de NCR (ex: retornar temporariamente a um esquema mais denso).

Dica Baseada em Evidências

No estudo de Vollmer et al. (1993), o NCR resultou em reduções significativas no comportamento autolesivo já nas primeiras sessões de tratamento. Se não observar redução após 2-3 semanas com um esquema adequado, reavalie a análise funcional.

Próximo: Resumo

Página 6

Reforçamento Não Contingente (NCR) para Autolesão em TEA

Tutorial baseado em evidências para psicólogos - Baseado em Vollmer et al. (1993)

Introdução

Fundamentos Teóricos

Implementação

Casos Práticos

Registro de Dados

Resumo

Resumo e Recomendações

Seu Desempenho no Tutorial

Pontuação Total: 1/2

Nível de Proficiência: Intermediário

Pontos Fortes

Reconhecimento da importância da análise funcional

Áreas para Melhoria

Aplicação de estratégias de fading

Adaptação do NCR para diferentes funções comportamentais

Próximos Passos Recomendados

Bom progresso! Reveja os fundamentos teóricos e pratique com mais casos clínicos.

Consulte as referências científicas para embasar suas intervenções.

Principais Conclusões sobre NCR

NCR é uma intervenção baseada em evidências para reduzir comportamentos autolesivos mantidos por reforçamento social

Deve ser precedida por análise funcional para identificar o reforçador mantenedor

É mais eficaz quando iniciado com esquema denso seguido de fading gradual

Pode atenuar efeitos colaterais da extinção e é mais fácil de implementar que DRO

Pode ser combinado com reforçamento de comportamentos alternativos

Referências Científicas

Vollmer, T. R., Iwata, B. A., Zarcone, J. R., Smith, R. G., & Mazaleski, J. L. (1993). The role of attention in the treatment of attention-maintained self-injurious behavior: Noncontingent reinforcement and differential reinforcement of other behavior. *Journal of Applied Behavior Analysis*, 26(1), 9-21.

Iwata, B. A., Dorsey, M. F., Slifer, K. J., Bauman, K. E., & Richman, G. S. (1982). Toward a functional analysis of self-injury. *Analysis and Intervention in Developmental Disabilities*, 2, 1-20.

Mace, F. C., & Lalli, J. S. (1991). Linking descriptive and experimental analyses in the treatment of bizarre speech. *Journal of Applied Behavior Analysis*, 24, 553-562.

Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, 18, 111-126.

Gerar Relatório em PDF

Atendimento especializado em Autismo



Uma equipe completa de especialistas para apoiar crianças e adolescentes no espectro autista e suas famílias.

- Psicoterapia ABA · Grupo de Habilidades Sociais
- Fonoaudiologia · Terapia Ocupacional · Psicopedagogia
- Supervisão escolar · Nutrição

CAPÍTULO

07

Comparação da qualidade técnica e científica de aplicativo gerado por Inteligência Artificial para ensinar crianças autistas a discriminarem suas emoções de acordo com o contexto



Valéria Costa França

Márcio Moreira

Centro Universitário de Brasília

RESUMO

Objetivo do estudo. O objetivo da presente pesquisa foi avaliar a qualidade técnica, metodológica e científica de aplicativos gerados por uma inteligência artificial (Claude Sonnet 4.5) para ensinar crianças autistas a discriminarem emoções de acordo com o contexto por meio de historinhas.

Método. Para isso, foram elaborados três prompts com diferentes níveis de especificidade, e cada um deles gerou um aplicativo distinto, posteriormente avaliado em quatro dimensões: adequação aos princípios da Análise do Comportamento Aplicada (ABA), adequação às características do TEA, funcionalidade técnica e estrutura metodológica.

Resultados. De forma geral, os resultados mostraram que todos os aplicativos apresentaram algum nível de adequação às propostas, porém apenas o aplicativo gerado pelo prompt intermediário estruturou corretamente as quatro fases experimentais, apesar de ter apresentado um erro crucial relativo ao feedback. Outras limitações identificadas envolveram a ausência de salvamento adequado de dados, falhas metodológicas nas fases de pré e pós-teste e baixa variedade de estímulos.

Conclusão. De modo geral, os achados mostram que a IA é uma ferramenta promissora para a criação de recursos educacionais acessíveis, permitindo que profissionais e cuidadores produzam materiais personalizados mesmo sem conhecimentos avançados em programação. Apesar das falhas observadas, os aplicativos gerados oferecem uma base inicial que pode ser aprimorada e utilizada como ponto de partida para o desenvolvimento de tecnologias assistivas mais robustas.

Palavras-chave: autismo, TEA, transtorno do espectro autista, inteligência artificial, aplicativo.

1 INTRODUÇÃO

Contextualização do Tema

Faz parte da história do ser humano aprender a identificar problemas e encontrar soluções e um dos desafios atuais está demonstrado no Censo Demográfico 2022 que, de acordo com a Agência de notícias do IBGE(2025), identificou 2,4 milhões de pessoas com diagnóstico de transtorno do espectro autista (TEA), o que corresponde a 1,2% da população brasileira. Como se trata da primeira vez em que a pesquisa aborda esse tema, não há parâmetros ofici-

ais de comparação, mas, de acordo com Goulart & Assis (2002) é notório o aumento de interesse pelo tema, visto que houve um aumento significativo na produção literária em relação ao assunto.

Por exemplo, numa pesquisa rápida no Google Scholar, ao se pesquisar o termo “autismo”, até o ano de 2000, aparecem 6.420 documentos. Quando se utiliza o mesmo termo e o período é ajustado para 2001 até os dias atuais o número sobe para 66.400, ou seja, um aumento de mais de 10 vezes. Os dados demonstram um aumento significativo na produção de material sobre o assunto, mas somente com essa informação não é possível avaliar o quanto esses estudos têm impactado os pais e demais cuidadores das crianças com autismo, havendo a necessidade de se avaliar estratégias que possam transformar o conteúdo produzido em opções de auxílio prático para os interessados.

Uma habilidade básica para a interação social, que crianças com Transtorno do Espectro Autista frequentemente apresentam dificuldade, é o reconhecimento de expressões faciais de emoções. A ausência dessa capacidade pode prejudicar a interpretação dos estados emocionais de outras pessoas e prejudicar o ajuste de seu comportamento às demandas de diferentes contextos sociais. Essas dificuldades podem estar relacionadas, entre outros fatores, a padrões atípicos de observação das faces, como a fixação em regiões que não fornecem pistas relevantes para a identificação das emoções, bem como diferenças na ativação de áreas cerebrais envolvidas no processamento de expressões faciais. Diante dessas limitações, diversos estudos têm investigado procedimentos de ensino capazes de ampliar o repertório de reconhecimento de emoções em crianças com TEA, buscando desenvolver intervenções baseadas em evidências que favoreçam a aprendizagem dessa habilidade e contribuam para a melhoria das interações sociais desses indivíduos (Kusano et al., 2023).

Kinser et al. (2019, como citado em Kim et al., 2025) informam que as tecnologias da informação e comunicação (TICs) são onipresentes para a geração atual de pais, com 96% dos pais de crianças de até 3 anos acessando a internet e 94% possuindo um smartphone. Ao investigar as necessidades dos pais, Rubio Hernández et al. (2021, como citado em Kim et al., 2025) constataram que os pais se autoavaliaram com menor competência em relação ao funcionamento familiar e expressaram necessidade e desejo de obter mais informações sobre como atender às necessidades socioemocionais, regular as próprias emoções durante a definição de limites e práticas disciplinares, além de receber mais apoio para compartilhar os sonhos e objetivos de seus filhos.

Narciso et al. (2024, como citado em Santos et al., 2025) discutem que as Tecnologias Digitais da Informação e comunicação (TDICs) são ferramentas que oferecem possibilidades significativas para melhorar a comunicação, a socialização e as habilidades cognitivas de indivíduos autistas, podendo apresentar soluções, como jogos digitais, Comunicação Aumentativa e Alternativa de Alta Tecnologia, Robôs, Realidade Virtual, e aplicações de Inteligência Artificial (IA), sendo recursos valiosos que podem ser adaptados às necessidades individuais, proporcionando um meio potencializador para a expressão e interação social.

De acordo com Pontes et al. (2020) a Inteligência Artificial - IA permite a interação entre um programa e um mundo externo da educação com o objetivo de investigar o aprendizado, onde quer que ocorra a fim de apoiar a educação formal bem como a aprendizagem ao longo

da vida, auxiliando professores no atendimento das necessidades de todos os alunos, ajudando a diminuir as lacunas de desempenho e equipando os pais para melhorar a aprendizagem de seus filhos.

Problema de Pesquisa

Sampaio et al. (2024) alertam para as questões relacionadas à confiabilidade das informações obtidas por meio das IAs. Segundo os autores, a UNESCO lidera uma discussão relevante sobre IA e educação onde aponta a importância de promover a colaboração internacional para desenvolver e implementar soluções de IA para a educação e garantir que a IA seja usada de forma ética e acessível a todos. Afirmam, também, que já estão sendo produzidos relatórios que apontam benefícios e riscos da IA na educação, enfatizando a necessidade de cautela na sua implementação para evitar disparidades de desempenho e promover um quadro abrangente para lidar com privacidade e segurança de dados.

Diante disso surge a necessidade de se avaliar a utilidade e confiabilidade dos aplicativos gerados por inteligências artificiais para ensinar crianças autistas a discriminarem emoções de acordo com o contexto.

Pergunta de Pesquisa

Considerando que as Tecnologias Digitais da Informação e comunicação (TDICs) são ferramentas que oferecem possibilidades significativas para melhorar a comunicação, a socialização e as habilidades cognitivas de indivíduos autistas Narciso et al. (2024, como citado em Santos et al., 2025) considerando que a Inteligência Artificial - IA permite a interação entre um programa e um mundo externo da educação com o objetivo de investigar o aprendizado, onde quer que ocorra a fim de apoiar a educação formal bem como a aprendizagem ao longo da vida (Pontes et al., 2020) e considerando que, de acordo com o estudo desenvolvido por Lima (2017), o uso de histórias infantis, se mostrou eficaz no treino de reconhecimento de expressões faciais de emoções em crianças com autismo, a pergunta base da presente pesquisa é “As inteligências artificiais são úteis e confiáveis para ensinar crianças autistas a discriminarem emoções de acordo com o contexto utilizando histórias?”

Relevância da Pesquisa

Segundo Correia (2014, como citado em Felix & Moreira, 2023) as emoções exercem papel fundamental na comunicação e na interação social, sendo a capacidade de reconhecê-las essencial para o estabelecimento de relações saudáveis. Entretanto, crianças com Transtorno do Espectro Autista (TEA) frequentemente apresentam dificuldades neste domínio, o que pode comprometer a comunicação, a compreensão das normas sociais e a qualidade das interações interpessoais.

Em novembro de 2022, o ChatGPT (Chat Generative Pre-Trained Transformer) foi lançado com o propósito de estabelecer interações em linguagem natural e em formato de diálogo, apresentando capacidade de responder e criar trocas comunicativas semelhantes às humanas, além de sustentar conversações longas (Sampaio et al., 2024).

Nesse contexto, Almeida et al. (2023, como citado em Sampaio et al., 2024) destacam que os sistemas utilizados para compreender e interpretar a linguagem humana, conhecidos como modelos de linguagem, operam a partir da previsão de palavras ou símbolos subsequentes em uma sentença. Para isso, são treinados com grandes volumes de textos escritos e aplicam modelos generativos, que aprendem como as palavras são combinadas para formar novos enunciados, de modo a imitar o uso humano da linguagem. Apesar de seu potencial, esses modelos apresentam limitações, como as chamadas alucinações, que ocorrem quando produzem respostas sintaticamente plausíveis, mas incorretas em termos de fatos e dados (Hacker et al., 2023, como citado em Sampaio et al., 2024; Alkaisse & Mcfarlane, 2023, como citado em Sampaio et al., 2024).

Assim, diante das possibilidades apresentadas pelas IAs, mas, também, de suas limitações, a presente pesquisa mostra-se relevante por buscar avaliar a utilidade e a confiabilidade das IAs como ferramentas de apoio no processo de reconhecimento de emoções por crianças autistas.

Revisão da Literatura

O Transtorno do Espectro Autista (TEA) configura-se como uma condição do neurodesenvolvimento, cujas manifestações centrais envolvem prejuízos persistentes na comunicação e na interação social em múltiplos contextos, acompanhados por padrões restritos e repetitivos de comportamento, interesses ou atividades, que estavam presentes durante o período inicial do desenvolvimento Associação Psiquiátrica Americana [APA] (2022). O Transtorno do Espectro Autista (TEA) tem origem multifatorial, relacionada a fatores genéticos e ambientais, ainda que sua causa exata seja desconhecida. A condição apresenta expressiva variabilidade, impactando diferentemente as dimensões sociais, comunicativas e comportamentais (Schaeffer, 2017, como citado em Felix & Moreira, 2023).

As emoções exercem papel fundamental na comunicação e na interação social, sendo a capacidade de reconhecê-las essencial para o estabelecimento de relações saudáveis. Entretanto, crianças com Transtorno do Espectro Autista (TEA) frequentemente apresentam dificuldades neste domínio, o que pode comprometer a comunicação, a compreensão das normas sociais e a qualidade das interações interpessoais. Tais limitações impõem barreiras à inclusão social e podem resultar em respostas mal adaptativas ao ambiente (Correia, 2014, como citado em Felix & Moreira, 2023).

Diante dessas dificuldades no reconhecimento de emoções, diferentes estudos têm buscado desenvolver e avaliar estratégias de ensino que favoreçam esse repertório em crianças com TEA. Nesse sentido, destaca-se a pesquisa apresentada a seguir.

O trabalho científico desenvolvido por Lima (2017) se propôs a avaliar a eficácia do uso de histórias infantis, no treino de reconhecimento de expressões faciais de emoções em crianças com autismo. Para Skinner (1974/2004, como citado em Lima, 2017) em práticas culturais como contar histórias, relatar contos e mitos são apresentadas instruções sobre o modo de se comportar em uma determinada sociedade. Desse modo, contar histórias adquire a função de ensinar comportamentos importantes para a sobrevivência das culturas.

O teste de Lima (2017) contou com a participação de seis crianças com diagnóstico de TEA leve a moderado na faixa etária entre seis e sete anos. Os participantes foram encaminhados do Setor de atendimento ao Transtorno do Espectro do Autismo do Serviço de Atendimento em Reabilitação de uma ONG. Após a realização da entrevista com os pais e da aplicação dos testes CARS e ABLA, foram selecionados os participantes que se encaixavam nos critérios de inclusão (Lima, 2017).

Foram utilizados um notebook, folhas de registro específicas de cada fase, lápis, borracha e uma câmera filmadora para a gravação das sessões. Como estímulos do Pré e Pós-Teste foram utilizadas imagens impressas em forma de desenho e fotografia de quatro expressões faciais básicas (alegria, tristeza, raiva e medo). As imagens eram de pessoas de gêneros (masculino e feminino), faixas etárias (crianças, adultos e idosos) e raças (negra, parda e branca) diferentes. Foram utilizados 10 conjuntos de estímulos com expressões faciais de emoções variadas em cada conjunto. Como estímulos de treino foram utilizadas 20 histórias. Cinco histórias para cada uma das quatro expressões faciais utilizadas no estudo (Lima, 2017).

O procedimento foi composto de cinco fases:

- Pré-Teste: objetivo de quantificar as expressões faciais de emoções reconhecidas pelos participantes.
- Treino: objetivo de ensinar o reconhecimento de expressões faciais de emoções.
- Pós-Teste: mesmo procedimento do Pré-Teste.
- Teste de Generalização (objetivo de verificar se os participantes generalizavam o comportamento de reconhecer expressões faciais de emoções para situações de interação social apresentadas em vídeo).
- Follow-Up (objetivo verificar se a habilidade de reconhecer as expressões faciais de emoções ensinadas na Fase de Treino se mantinham no repertório dos participantes após transcorridas quatro semanas da aplicação da Fase 4).

Os resultados do Pré-teste mostraram que todos os participantes emitiram entre 70% e 90% de respostas incorretas. Estes resultados corroboram outros encontrados na literatura os quais apontam que crianças com autismo apresentam dificuldade no reconhecimento de expressões faciais emocionais quando comparadas a crianças com desenvolvimento típico. A comparação da porcentagem de respostas corretas emitidas pelos participantes no Pré-Teste e Pós-Teste mostrou que ocorreu um aumento no número de emissões de respostas corretas no Pós-Teste, tendo todos os participantes atingido o critério de desempenho (Lima, 2017).

De acordo com Lima (2017) os dados mostrando que todos os participantes atingiram o critério de desempenho requerido durante a Fase de Treino e apresentaram desempenho generalizado durante o Teste de Generalização, indicam que as histórias infantis utilizadas no estudo foram um recurso lúdico-didático eficaz para o refinamento da habilidade do reconhecimento de expressões faciais de emoções.

Esses dados, segundo Lima (2017), diferem dos encontrados em outros estudos, os quais relatam a eficácia dos procedimentos utilizados em suas respectivas pesquisas para o ensino de reconhecimento de expressões faciais, mas relatam também a não ocorrência de respostas

generalizadas de reconhecimento de expressões emocionais em outros contextos, principalmente em contextos sociais. Lima (2017) ressalta a efetividade do procedimento de ensino utilizado em gerar comportamentos generalizados, apontando a probabilidade de que os repertórios adquiridos se generalizem também para contextos da vida cotidiana.

No Follow-Up, de acordo com Lima (2017), os resultados sugerem que algumas crianças podem não ter mantido o repertório de reconhecimento de expressões emocionais por não apresentarem linguagem vocal. Molini (2001, como citado em Lima, 2017) observou que crianças com autismo sem comprometimento da linguagem verbal, tendem a apresentar melhores desempenhos sócio-cognitivos. Considerando os resultados apresentados por Lima (2017), especialmente no que se refere à eficácia do ensino e à generalização do repertório, torna-se relevante analisar outros estudos que investigam diferentes procedimentos de ensino dessa mesma habilidade, possibilitando a comparação entre estratégias.

Em complemento a essa discussão, o estudo de Araújo et al. (2024) investigaram o reconhecimento e a nomeação de emoções em crianças e jovens adultos com Transtorno do Espectro Autista (TEA) e Síndrome de Down, por meio de tarefas aplicadas pela família e mediadas por recursos tecnológicos. Participaram do estudo dez indivíduos com desenvolvimento atípico, com idades entre três e 20 anos, cujas atividades foram conduzidas em ambiente domiciliar com orientação dos pesquisadores.

Araújo et al. (2024) utilizaram um delineamento experimental de múltiplas sondagens, no qual foram avaliadas tarefas de identidade, reconhecimento, nomeação e imitação de emoções, a partir de diferentes tipos de estímulos, como faces humanas (femininas e masculinas), emojis e imagens de familiares. O procedimento foi organizado nas seguintes fases:

- Linha de base: avaliação inicial do repertório dos participantes, sem fornecimento de dicas ou consequências para acertos e erros;
- Ensino: treino de reconhecimento e nomeação de emoções com uso de dicas, reforçamento e repetição até o critério de aprendizagem;
- Generalização: verificação da transferência das habilidades aprendidas para funções inversas e novos estímulos;
- Linha de base final: reavaliação do desempenho dos participantes após o ensino, para comparação com os dados iniciais.

Os resultados de Araújo et al. (2024) indicaram que, na linha de base, os participantes apresentavam desempenhos variáveis, com dificuldades influenciadas pela valência emocional, pelo tipo de estímulo e pelo gênero das faces. Após a implementação do procedimento de ensino, observou-se aumento significativo no número de acertos, além da generalização das habilidades de reconhecimento e nomeação para diferentes estímulos, independentemente de sua natureza.

Destaca-se, ainda, que o estudo de Araújo et al. (2024) evidenciou a viabilidade da participação ativa dos familiares na aplicação das intervenções, mesmo em contexto remoto, contribuindo para o desenvolvimento do repertório emocional dos participantes. Esses achados re-

forçam a eficácia de procedimentos estruturados de ensino, bem como a importância do envolvimento familiar e do uso de tecnologias como estratégias complementares no desenvolvimento de habilidades socioemocionais em indivíduos com TEA.

Nessa mesma direção, ao ampliar a análise sobre diferentes estratégias de ensino do reconhecimento de emoções, destaca-se a pesquisa realizada por Kusano et al. (2023), que teve como objetivo comparar a eficácia de dois procedimentos de ensino para o reconhecimento de expressões faciais de emoções. Especificamente, os autores investigaram os efeitos do procedimento de emparelhamento ao modelo (Matching to Sample – MTS) e do mesmo procedimento associado à estimulação vestibular (MTS/ES), buscando verificar se a estimulação sensorial poderia favorecer a aprendizagem dessa habilidade.

O estudo de Kusano et al. (2023) foi conduzido com um menino de três anos e oito meses diagnosticado com TEA, que frequentava uma instituição especializada em atendimento a crianças com autismo. Foi utilizado um delineamento de múltiplas sondagens entre comportamentos, no qual diferentes pares de emoções foram ensinados em condições alternadas de ensino. O procedimento de emparelhamento ao modelo consistia na apresentação de um estímulo modelo (palavra referente à emoção) e na escolha, pelo participante, entre estímulos de comparação representados por imagens de expressões faciais. Na condição MTS/ES, o mesmo procedimento era realizado após a criança receber estimulação vestibular, obtida por meio do deslocamento da criança sentada em um skate antes das tentativas de ensino.

Kusano et al. (2023) organizaram a pesquisa em diferentes fases. Inicialmente, foram realizadas sondagens de linha de base para avaliar o repertório prévio do participante em relação ao reconhecimento de expressões faciais. Em seguida, ocorreram as fases de ensino, nas quais diferentes pares de emoções foram treinados por meio dos procedimentos experimentais. Durante essas fases, eram realizadas sondagens periódicas para verificar a aquisição das relações entre palavras e expressões faciais. Além disso, foram conduzidas sondagens de generalização, nas quais o desempenho do participante foi avaliado utilizando fotografias de faces reais diferentes das utilizadas durante o ensino.

Os resultados de Kusano et al. (2023) indicaram que o participante apresentou melhora no desempenho ao longo das sessões de ensino, aprendendo a relacionar as palavras às expressões faciais nas duas condições experimentais. Entretanto, não foram observadas diferenças consistentes que indicassem maior eficácia do procedimento associado à estimulação vestibular em comparação ao MTS tradicional. Além disso, verificou-se que a aprendizagem obtida com desenhos de expressões faciais não se generalizou automaticamente para fotografias de faces reais, uma vez que o desempenho do participante permaneceu baixo nas sondagens realizadas com esse tipo de estímulo.

Na discussão dos resultados, os autores destacam que alguns aspectos do arranjo experimental podem ter influenciado o desempenho observado no primeiro estudo, especialmente a organização das tentativas de comparação, que poderia favorecer estratégias de resposta baseadas na rejeição de estímulos incorretos ou na exclusão. Ainda assim, para Kusano et al. (2023), os resultados evidenciam que o procedimento de emparelhamento ao modelo pode ser uma estratégia eficaz para o ensino de reconhecimento de emoções em crianças com

TEA. Por outro lado, a ausência de generalização entre diferentes tipos de estímulos sugere a necessidade de incluir maior variedade de exemplos durante o ensino, a fim de favorecer a transferência do aprendizado para contextos mais naturais.

A partir dessas evidências, observa-se que, embora diferentes procedimentos de ensino apresentem eficácia, ainda há desafios relacionados à generalização e à adaptação das estratégias. Nesse cenário, novas abordagens, especialmente aquelas baseadas em tecnologia, passam a ser investigadas como alternativas complementares. Em novembro de 2022 o ChatGPT (Chat Generative Pre-Trained Transformer) foi construído para estabelecer com os usuários interações em linguagem natural e formato de diálogo com capacidade de responder e criar diálogos similares aos de humanos, mantendo uma conversação longa (Sampaio et al., 2024). De acordo com Sampaio et al. (2024), as capacidades do ChatGPT e sua aplicabilidade em diferentes áreas do conhecimento, como marketing, direito, jornalismo, programação, ciência de dados e afins, impulsionaram um debate que vem sendo travado há alguns anos: a automação do trabalho e a substituição de profissionais humanos por tecnologias.

Sampaio et al. (2024) explicam que as LLMs estão embasadas em três conceitos fundamentais : aprendizado de máquina, o método Transformers para processamento de texto e a importância dos dados massivos.

Aprendizado de máquina é um campo da inteligência artificial que desenvolve algoritmos capazes de tomar decisões sem programação humana direta, aprendendo padrões e agindo de forma independente. Baseia-se em conjuntos de dados para treinar modelos que prevejam corretamente os estados desses dados, criando hipóteses com base em dados passados para fazer previsões futuras (Jung, 2022, como citado em Sampaio et al., 2024). Existem métodos de aprendizado de máquina com diferentes níveis de interferência humana, mas o objetivo é que o algoritmo aprenda a realizar a ação de forma autônoma, com revisão humana.

O método Transformers é uma abordagem de análise textual fundamentada no conceito de atenção, no qual cada elemento do texto é representado como um vetor em relação a todos os outros elementos do texto. A técnica de transformer permite que termos sejam relacionados entre si independentemente da sua distância no texto e classifica os elementos em termos de importância, tudo isso com grande capacidade de paralelismo (Vaswani et al., 2017, como citado em Sampaio et al., 2024). De acordo com Sampaio et al. (2024), o modelo Transformers permitiu a explosão do tamanho dos modelos de linguagem natural para se tornarem LLMs, pois fez com que eles se tornassem generalizáveis.

De acordo com Hacker et al. (2023, como citado em Sampaio et al., 2024), o ChatGPT e outros LLMS não são muito diferentes de um papagaio que tenta repetir o que aprendeu após doses massivas de treinamento. O modelo elabora palavra por palavra, baseado na probabilidade de um humano ter utilizado aquela sequência de palavras. Exatamente por isso, tais modelos nunca repetem exatamente as respostas e, com alguma frequência, erram ao responder, causando o que tem sido chamado de alucinações. Diante desses vieses que as IAs podem apresentar vários estudiosos começaram a avaliar a qualidade das respostas obtidas por meio desses serviços.

Kim et al. (2025) buscaram explorar a qualidade das respostas de inteligência artificial generativa (IA) a perguntas comuns sobre parentalidade, a partir de diversas fontes de informações digitais disponíveis. Com esse objetivo, realizaram um teste dedutivo utilizando 11 condições de teste em cinco tópicos. Para ajudar a controlar a variedade sempre mutável de informações na internet aberta, os autores criaram um universo de dados a partir do conteúdo de 40 boletins de parentalidade de oito páginas cada. Isso garantiu uma fonte de informação de alta qualidade e previamente revisada.

De acordo com Kim et al. (2025) foram comparados o ChatGPT-3.5 e o ChatGPT-4 com a ferramenta de busca do site do boletim de parentalidade (Testes 1 e 2), assim como com o universo conhecido do conteúdo online do boletim JITP.info e a internet aberta. Cinco perguntas abertas sobre tópicos foram geradas, com base em uma análise de assuntos que o Google Analytics demonstrou serem frequentemente pesquisados no site do boletim JITP.info. Esses cinco tópicos representam desafios comuns da parentalidade:

- Incentivar o sono;
- Estabelecer limites e disciplina;
- Autocuidado parental;
- Lidar com birras;
- Receber segurança de que seu filho sabe que você se importa

De acordo com os testes realizados por Kim et al. (2025), o ChatGPT forneceu, em grande parte, respostas corretas e claras, embora as citações estivessem frequentemente ausentes ou imprecisas. As respostas dos LLMs muitas vezes careciam de ênfase na conexão entre pais e filhos e no contexto do desenvolvimento infantil, além de apresentarem um aumento acentuado na dificuldade de leitura.

Kim et al. (2025) propuseram duas hipóteses específicas e partiram do princípio de que as respostas de um LLM a perguntas básicas de parentalidade seriam mais sofisticadas quando testadas em combinação com uma newsletter digital nacional de longa duração, curada por especialistas em pesquisa. Pensaram que a inclusão da newsletter levaria a melhorias mensuráveis na qualidade das respostas do LLM.

De acordo com Kim et al. (2025), essa suposição não foi apoiada por todos os testes, mas a avaliação também depende de como a sofisticação é medida. Se for utilizado o critério de práticas apropriadas ao desenvolvimento e sensibilidade ao contexto, incluindo legibilidade e conexão interpessoal, então a sofisticação tende a diminuir. Segundo os autores, isso foi mais evidente nas respostas ao tópico sensível ao desenvolvimento de birras, onde os LLMs forneceram respostas inapropriadas para a idade, mas menos relevante no tópico de estabelecimento de limites, onde os três Cs (correção, clareza e conexão) receberam pontuações altas.

No entanto, Kim et al. (2025) perceberam que a legibilidade apresentou um padrão inverso: as respostas sobre birras eram mais acessíveis, enquanto as respostas sobre limites atingiram níveis de leitura até o 14º ano. Os autores também notaram que os diferentes públicos

utilizam as IAs de diferentes maneiras sendo necessário levar essa questão em conta no que se refere a apoiar usuários de difícil alcance ou consumidores passivos da internet para melhorar os resultados familiares e infantis.

Em resumo, neste contexto, Kim et al. (2025) defendem que o ChatGPT parece ser suficiente para condições normativas, e não fabrica informações, no entanto, permanece necessário que especialistas apoiem pais e cuidadores para que obtenham respostas úteis às suas perguntas.

Diante dessas discussões sobre potencial e limitações da inteligência artificial, torna-se importante analisar evidências empíricas específicas sobre sua aplicação no contexto do TEA, especialmente no fornecimento de informações a pais e cuidadores. Com esse objetivo foi realizada uma pesquisa quantitativa por McFayden et al. (2024), utilizando o ChatGPT em abril de 2023, abordando questões frequentes sobre os assuntos seguintes: (a) informações básicas sobre o autismo; (b) mitos e concepções equivocadas; e (c) recursos disponíveis. De acordo com os autores, o objetivo principal do artigo foi avaliar se o ChatGPT-4 pode ser uma ferramenta viável para fornecer informações corretas, claras e úteis sobre o Transtorno do Espectro Autista (TEA) a pais e cuidadores.

Segundo McFayden et al. (2024) as respostas obtidas nos testes foram claras, corretas e concisas numa média de 3,7/4, entretanto com zero aplicabilidade devido à linguagem majoritariamente médica. Em relação aos mitos, o ChatGPT os refutou com precisão, mantendo boa clareza e compreensão, entretanto, novamente, demonstrando aplicabilidade prática inexistente. Quanto aos recursos, esse foi o único domínio em que houve maior aplicabilidade com respostas claras e compreensíveis, embora ainda limitadas quanto a recomendações práticas. Os autores constataram que menos da metade dos links funcionavam corretamente, mas quando corretos, direcionavam para órgãos confiáveis. De acordo com a análise de McFayden et al. (2024), esse problema comprometeu a utilidade das referências.

McFayden et al. (2024) defendem que o estudo indicou que o ChatGPT apresenta informações corretas, claras e concisas sobre o autismo, sendo uma ferramenta promissora para combater a desinformação e o estigma. Entretanto alertam para suas limitações que incluem a baixa aplicabilidade prática das respostas e a inconsistência das referências fornecidas. Para os autores o ChatGPT pode ser uma ferramenta útil para compreender conceitos gerais, mas ainda não substitui fontes atualizadas nem orientação profissional.

Diante do exposto, observa-se que os estudos analisados convergem ao evidenciar a relevância do ensino do reconhecimento de emoções em crianças com TEA, ainda que por diferentes abordagens. Enquanto Lima (2017) demonstra a eficácia do uso de histórias infantis, com destaque para a generalização do repertório aprendido, Kusano et al. (2023) evidenciam a efetividade de procedimentos estruturados, embora com limitações nesse mesmo aspecto. Por sua vez, as investigações sobre o uso de inteligência artificial, como apresentado por Sampaio et al. (2024) e Kim et al. (2025), apontam o potencial dessas tecnologias, mas também suas fragilidades quanto à contextualização e aplicabilidade. De modo complementar, McFayden et al. (2024) reforçam que, apesar da clareza e correção das informações forneci-

das por sistemas como o ChatGPT, ainda há limitações práticas relevantes. Assim, os estudos indicam a necessidade de integrar estratégias eficazes de ensino com recursos tecnológicos, considerando suas potencialidades e restrições.

Objetivo da Pesquisa

Segundo Correia (2014, como citado em Felix & Moreira, 2023) as emoções desempenham papel central na comunicação e na interação social, sendo o reconhecimento delas essencial para o estabelecimento de relações saudáveis. Contudo, crianças com Transtorno do Espectro Autista (TEA) apresentam dificuldades nesse domínio, o que pode afetar a comunicação, a compreensão das normas sociais e a qualidade das interações. Nesse contexto, o lançamento do ChatGPT (Chat Generative Pre-Trained Transformer), em novembro de 2022, trouxe novas possibilidades ao oferecer interações em linguagem natural, em formato de diálogo, com capacidade de sustentar trocas comunicativas semelhantes às humanas (Sampaio et al., 2024).

Almeida et al. (2023, como citado em Sampaio et al., 2024) destacam que os modelos de linguagem, utilizados para compreender e interpretar a linguagem humana, funcionam por meio da previsão de palavras ou símbolos subsequentes, sendo treinados com grandes volumes de textos escritos e aplicando modelos generativos para imitar o uso humano da linguagem. Apesar de seu potencial, esses modelos apresentam limitações, como as chamadas alucinações, que correspondem a respostas plausíveis em termos de forma, mas incorretas em relação aos fatos (Hacker et al., 2023; Alkaisse & Mcfarlane, 2023, como citado em Sampaio et al., 2024). Considerando tais possibilidades e limitações, este estudo tem como objetivo avaliar a utilidade e a confiabilidade das IAs como ferramentas de apoio no processo de discriminação de emoções por crianças autistas de acordo com o contexto com base em historinhas.

2 MÉTODO

Delineamento de Pesquisa

Este estudo adotou um delineamento comparativo com manipulação sistemática de prompts e análise de produtos digitais gerados, visando investigar o efeito de diferentes níveis de complexidade dos prompts sobre as características técnicas, pedagógicas e a adequação científica de aplicativos desenvolvidos para o ensino de discriminação de emoções em crianças autistas, fundamentados nos princípios da Análise do Comportamento Aplicada (ABA).

A variável independente consistiu no nível de detalhamento dos prompts (básico, intermediário e avançado), definidos a partir da quantidade e organização das informações fornecidas à inteligência artificial para a geração dos aplicativos. Cada nível correspondeu a uma condição distinta de geração, resultando em três versões independentes de aplicativos, que constituíram as unidades de análise do estudo ($n = 3$).

As variáveis dependentes corresponderam às características técnicas, pedagógicas e metodológicas dos aplicativos gerados, organizadas em quatro dimensões: (1) adequação aos princípios da Análise do Comportamento Aplicada (ABA), incluindo estrutura de apresentação de estímulos, sequência de esvanecimento de dicas, sistema de registro de respostas e mecanis-

mos de reforçamento; (2) adequação às características do Transtorno do Espectro Autista (TEA), contemplando interface do usuário, acessibilidade, previsibilidade e clareza de instruções; (3) funcionalidade técnica, incluindo operacionalidade do código, responsividade e sistema de salvamento de dados; e (4) estrutura metodológica experimental, envolvendo a organização das fases de pré-teste, treino, pós-teste e generalização, progressão pedagógica, variedade de estímulos e fidelidade ao delineamento proposto por Lima (2017).

A análise comparativa foi realizada por meio de um protocolo de avaliação estruturado, composto por critérios específicos derivados da literatura em ABA e educação especial. Cada critério foi avaliado em escala ordinal de 1 a 4 pontos, permitindo uma análise qualitativa descritiva complementada por quantificação do nível de adequação dos aplicativos. Os resultados foram organizados em matriz comparativa, possibilitando a identificação de vantagens, limitações e diferenças entre as versões geradas em função do nível de detalhamento dos prompts.

O processo de geração dos aplicativos foi conduzido em sessões independentes, sem histórico conversacional prévio, utilizando o mesmo modelo de linguagem e condições de acesso, de modo a garantir controle das variáveis externas e equivalência entre as condições de geração.

O delineamento adotado não incluiu a fase de follow-up descrita no estudo de Lima (2017), em função de limitações operacionais relacionadas ao ambiente digital e à natureza automatizada da geração dos aplicativos, mantendo-se, contudo, as fases centrais de pré-teste, treino, pós-teste e generalização.

O processo de desenvolvimento seguiu uma abordagem iterativa, com registro sistemático das decisões de design e programação, assegurando a rastreabilidade e a replicabilidade dos procedimentos adotados. Esse delineamento permitiu analisar de forma sistemática a relação entre o nível de detalhamento dos prompts e a qualidade técnica, pedagógica e científica dos aplicativos gerados.

Modelos Utilizados

Neste estudo, foi utilizado o Claude Sonnet 4.5, um modelo de linguagem de grande escala desenvolvido pela Anthropic, pertencente à família Claude 4. O modelo apresenta capacidades multimodais, permitindo o processamento de texto e imagens, além de uma ampla janela de contexto que possibilita o manejo de grandes volumes de informação.

Para o desenvolvimento dos aplicativos educacionais, o modelo foi utilizado por sua capacidade de gerar código em linguagens como HTML, CSS e JavaScript, possibilitando a criação de protótipos funcionais de aplicações interativas. O acesso ao modelo foi realizado por meio da interface web Claude.ai, que permite interação conversacional e geração iterativa de artefatos, viabilizando o refinamento progressivo dos aplicativos a partir de ciclos sucessivos de instrução e resposta entre pesquisador e inteligência artificial.

Prompts Utilizados

Foram elaborados três prompts com o objetivo de avaliar como diferentes níveis de detalhamento das instruções influenciam as características dos aplicativos gerados pela inteligência artificial.

1º prompt

Crie um aplicativo HTML que ensina crianças autistas a reconhecerem emoções com base em um procedimento de discriminações condicionais nos quais os modelos são vinhetas (histórias curtas) e os estímulos de comparação são fotografias de expressões faciais. Utilize feedback visual e auditivo. Salve os resultados.

2º prompt

Crie um aplicativo HTML que ensina crianças autistas a reconhecerem emoções com base em um procedimento de discriminações condicionais nos quais os modelos são vinhetas (histórias curtas) e os estímulos de comparação são fotografias de expressões faciais. Considere os parâmetros das fases listadas abaixo. Utilize feedback visual e auditivo. Salve os resultados.

Pré-Teste: objetivo de quantificar as expressões faciais de emoções reconhecidas pelos participantes.

- Treino: objetivo de ensinar o reconhecimento de expressões faciais de emoções.
- Pós-Teste: mesmo procedimento do Pré-Teste.
- Teste de Generalização (objetivo de verificar se os participantes generalizavam o comportamento de reconhecer expressões faciais de emoções para situações de interação social apresentadas em vídeo).

3º prompt

Crie um aplicativo HTML que ensina crianças autistas a reconhecerem emoções com base em um procedimento de discriminações condicionais nos quais os modelos são vinhetas (histórias curtas) e os estímulos de comparação são fotografias de expressões faciais. Considere o artigo do upload. Utilize feedback visual e auditivo. Salve os resultados.

O artigo anexado ao prompt 3 foi: Lima, A. A. (2017). Efeitos do uso de histórias infantis sobre o reconhecimento de expressões faciais de emoções em crianças com autismo (Doctoral dissertation, Dissertação de mestrado, Universidade Federal do Pará). <https://repositorio.ufpa.br/handle/2011/13848>.

Procedimento

O estudo foi conduzido em etapas sequenciais para avaliar sistematicamente o efeito de diferentes níveis de detalhamento dos prompts sobre as características e adequação científica dos aplicativos gerados. Foram elaborados três prompts com graus crescentes de especificidade: (1) Prompt Básico, solicitando a criação de um aplicativo HTML para ensinar crianças autistas a reconhecerem emoções por meio de procedimento de discriminações condicio-

nais, utilizando vinhetas (histórias curtas) como estímulos modelo e fotografias de expressões faciais como estímulos de comparação, com instrução para salvar os resultados; (2) Prompt Intermediário, mantendo as mesmas especificações do Prompt Básico e adicionando parâmetros estruturais de quatro fases experimentais sequenciais - Pré-Teste (para quantificar as expressões faciais reconhecidas pelos participantes), Treino (para ensinar o reconhecimento de expressões faciais), Pós-Teste (replicando o procedimento do Pré-Teste) e Teste de Generalização (para verificar se os participantes generalizavam o comportamento de reconhecer expressões faciais para situações de interação social apresentadas em vídeo); e (3) Prompt Avançado, mantendo a solicitação base do Prompt Básico e incluindo como referência o artigo científico de Lima (2017), intitulado "Efeitos do uso de histórias infantis sobre o reconhecimento de expressões faciais de emoções em crianças com autismo", disponibilizado por meio de upload de arquivo.

Cada prompt foi submetido ao Claude Sonnet 4.5 em sessões independentes, sem histórico conversacional prévio, garantindo que cada geração partisse de condições iniciais equivalentes e que as diferenças observadas fossem atribuíveis exclusivamente às variações no detalhamento dos prompts.

Os três aplicativos gerados foram então submetidos a uma análise sistemática utilizando critérios predefinidos organizados em quatro dimensões principais: (1) Adequação aos Princípios da Análise do Comportamento Aplicada, avaliando a presença e implementação correta de técnicas comportamentais como reforçamento diferencial, registro de desempenho, estruturação de tentativas discretas e procedimentos baseados em discriminações condicionais; (2) Adequação às Características do Transtorno do Espectro Autista, examinando adaptações para processamento sensorial atípico, previsibilidade da interface, clareza de instruções, suporte visual e considerações sobre perfis cognitivos diferenciados; (3) Funcionalidade Técnica, verificando a operacionalidade do código HTML gerado, responsividade da interface, sistema de salvamento de resultados, ausência de erros críticos e qualidade da experiência do usuário; e (4) Estrutura Metodológica, analisando a presença e adequação das fases experimentais (pré-teste, treino, pós-teste e generalização), progressão pedagógica, variedade de estímulos e fidedignidade ao delineamento de pesquisa proposto por Lima (2017).

Cada dimensão foi avaliada por meio de análise qualitativa descritiva complementada por escalas de adequação, registrando-se tanto elementos positivos quanto limitações identificadas em cada aplicativo. A análise comparativa entre os três níveis de detalhamento dos prompts permitiu identificar quais componentes informacionais exerceram maior influência sobre aspectos específicos da qualidade dos aplicativos gerados, fornecendo dados sobre a relação entre especificidade do prompt e adequação científica do produto final.

Avaliação do Conteúdo Gerado

O conteúdo gerado foi avaliado conforme a rubrica especificada abaixo, a qual permitiu uma análise sistemática e comparativa dos três aplicativos gerados pelos diferentes níveis de detalhamento dos prompts, identificando especificamente quais aspectos foram melhor ou pior implementados em função da especificidade das instruções fornecidas.

Dimensão 1: Adequação aos Princípios da Análise do Comportamento Aplicada (ABA)

1.1 Procedimento de Discriminação Condicional

- 4 - Excelente: Implementa claramente o paradigma de discriminação condicional com vinhetas como modelos e expressões faciais como estímulos de comparação. Apresenta relação funcional consistente entre modelo e comparações.
- 3 - Adequado: Implementa o procedimento básico de discriminação condicional, mas com algumas inconsistências na relação modelo-comparação.
- 2 - Parcial: Apresenta elementos do procedimento, mas a estrutura de discriminação condicional não está clara ou completa.
- 1 - Inadequado: Não implementa adequadamente o procedimento de discriminação condicional.

1.2 Sistema de Reforçamento

- 4 - Excelente: Fornece reforçamento positivo imediato, específico e consistente para respostas corretas. Inclui feedback diferenciado para acertos e erros.
- 3 - Adequado: Fornece reforçamento para respostas corretas, mas com atrasos ou inconsistências ocasionais.
- 2 - Parcial: Sistema de reforçamento presente, mas genérico ou pouco efetivo.
- 1 - Inadequado: Ausência de sistema de reforçamento adequado ou feedback inconsistente.

1.3 Registro de Desempenho

- 4 - Excelente: Registra sistematicamente todas as respostas (corretas/incorretas), tempo de resposta, fase do procedimento e permite exportação/salvamento dos dados.
- 3 - Adequado: Registra respostas corretas e incorretas com possibilidade de salvamento básico.
- 2 - Parcial: Registro limitado de dados ou salvamento inconsistente.
- 1 - Inadequado: Não possui sistema adequado de registro de dados.

1.4 Estruturação de Tentativas

- 4 - Excelente: Apresenta estrutura clara de tentativas discretas com início definido, oportunidade de resposta, consequência e intervalo entre tentativas.
- 3 - Adequado: Estrutura de tentativas presente, mas com transições pouco claras.
- 2 - Parcial: Tentativas desorganizadas ou sem estrutura consistente.
- 1 - Inadequado: Ausência de estruturação adequada de tentativas.

Dimensão 2: Adequação às Características do Transtorno do Espectro Autista

2.1 Interface Visual

- 4 - Excelente: Interface limpa, organizada, com mínimo de estímulos distratores, cores não excessivamente estimulantes, layout previsível e consistente.
- 3 - Adequado: Interface clara, mas com alguns elementos potencialmente distratores.

- 2 - Parcial: Interface confusa ou com excesso de estímulos visuais.
- 1 - Inadequado: Interface desorganizada, caótica ou sensorialmente sobrecarga.

2.2 Clareza de Instruções

- 4 - Excelente: Instruções simples, diretas, com suporte visual, apresentadas passo a passo. Linguagem acessível e concreta.
- 3 - Adequado: Instruções claras, mas poderiam ser mais simplificadas ou ter mais suporte visual.
- 2 - Parcial: Instruções vagas ou excessivamente complexas.
- 1 - Inadequado: Instruções confusas, abstratas ou ausentes.

2.3 Previsibilidade

- 4 - Excelente: Sequência de eventos altamente previsível, com indicadores claros de progresso e transições suaves entre fases.
- 3 - Adequado: Sequência previsível, mas com algumas transições abruptas.
- 2 - Parcial: Baixa previsibilidade ou mudanças inesperadas frequentes.
- 1 - Inadequado: Sequência imprevisível ou caótica.

2.4 Adaptabilidade Sensorial

- 4 - Excelente: Permite ajustes de volume, brilho, velocidade; considera sensibilidades sensoriais com opções de redução de estímulos.
- 3 - Adequado: Algumas opções de ajuste sensorial disponíveis.
- 2 - Parcial: Poucas ou nenhuma opção de adaptação sensorial.
- 1 - Inadequado: Não considera necessidades sensoriais atípicas.

Dimensão 3: Funcionalidade Técnica

3.1 Operacionalidade do Código

- 4 - Excelente: Código executa perfeitamente sem erros, bugs ou travamentos. Todos os recursos funcionam conforme esperado.
- 3 - Adequado: Código funciona adequadamente com erros menores que não comprometem a experiência.
- 2 - Parcial: Erros frequentes que afetam a usabilidade.
- 1 - Inadequado: Código não funciona ou apresenta erros críticos.

3.2 Responsividade

- 4 - Excelente: Interface adapta-se perfeitamente a diferentes tamanhos de tela e dispositivos.
- 3 - Adequado: Interface funciona bem na maioria dos dispositivos com pequenas inconsistências.
- 2 - Parcial: Problemas significativos de responsividade em alguns dispositivos.

- 1 - Inadequado: Não responsivo ou inutilizável em dispositivos móveis/tablets.

3.3 Sistema de Salvamento

- 4 - Excelente: Salva automaticamente todos os dados de forma confiável com opções de exportação em formatos úteis (CSV, JSON).
- 3 - Adequado: Salva dados básicos com funcionalidade limitada de exportação.
- 2 - Parcial: Salvamento inconsistente ou dados incompletos.
- 1 - Inadequado: Não salva dados ou sistema não funcional.

Dimensão 4: Estrutura Metodológica Experimental

4.1 Fase de Pré-Teste

- 4 - Excelente: Implementa pré-teste completo para quantificar expressões reconhecidas, com variedade adequada de estímulos e registro sistemático.
- 3 - Adequado: Pré-teste presente com algumas limitações na variedade de estímulos.
- 2 - Parcial: Pré-teste superficial ou incompleto.
- 1 - Inadequado: Ausência de pré-teste ou implementação inadequada.

4.2 Fase de Treino

- 4 - Excelente: Fase de treino estruturada com progressão clara, múltiplas oportunidades de aprendizagem, feedback consistente e variedade de vinhetas para cada emoção.
- 3 - Adequado: Treino estruturado, mas com limitações na variedade ou progressão.
- 2 - Parcial: Treino desorganizado ou com poucas oportunidades de aprendizagem.
- 1 - Inadequado: Fase de treino ausente ou inadequada.

4.3 Fase de Pós-Teste

- 4 - Excelente: Pós-teste replica fielmente o pré-teste, permitindo comparação direta de desempenho.
- 3 - Adequado: Pós-teste similar ao pré-teste com pequenas diferenças.
- 2 - Parcial: Pós-teste significativamente diferente do pré-teste.
- 1 - Inadequado: Ausência de pós-teste ou implementação inadequada.

4.4 Teste de Generalização

- 4 - Excelente: Inclui teste de generalização com estímulos novos (vídeos ou situações sociais diferentes das do treino) para verificar transferência de aprendizagem.
- 3 - Adequado: Teste de generalização presente, mas limitado em variedade ou formato.
- 2 - Parcial: Teste de generalização superficial ou incompleto.
- 1 - Inadequado: Ausência de teste de generalização.

4.5 Progressão Pedagógica

- 4 - Excelente: Progressão lógica e gradual de dificuldade, com critérios claros para avanço entre fases.

- 3 - Adequado: Progressão presente, mas com saltos de dificuldade ocasionais.
- 2 - Parcial: Progressão inconsistente ou pouco estruturada.
- 1 - Inadequado: Ausência de progressão pedagógica clara.

4.6 Variedade de Estímulos

- 4 - Excelente: Utiliza múltiplas vinhetas por emoção (mínimo 5) e expressões faciais diversificadas (diferentes gêneros, idades, etnias) conforme Lima (2017).
- 3 - Adequado: Variedade razoável de estímulos (3-4 por emoção).
- 2 - Parcial: Variedade limitada de estímulos (1-2 por emoção).
- 1 - Inadequado: Estímulos repetitivos ou insuficientes.

4.7 Fidelidade ao Delineamento de Lima (2017)

- 4 - Excelente: Implementa fielmente os elementos centrais do estudo de Lima (2017): uso de histórias, quatro emoções básicas, estrutura de fases, procedimento de discriminação condicional.
- 3 - Adequado: Implementa a maioria dos elementos do estudo com pequenas adaptações.
- 2 - Parcial: Implementa apenas alguns elementos do estudo original.
- 1 - Inadequado: Não segue o delineamento proposto por Lima (2017).

Sistema de Pontuação Geral

Pontuação total possível: 112 pontos (28 critérios × 4 pontos)

Classificação final:

- Excelente (95-112 pontos): Aplicativo altamente adequado, pronto para implementação com ajustes mínimos
- Bom (75-94 pontos): Aplicativo adequado com necessidade de refinamentos específicos
- Regular (56-74 pontos): Aplicativo requer revisões substanciais antes de implementação
- Insuficiente (28-55 pontos): Aplicativo inadequado, necessita reformulação completa

3 RESULTADOS

A Coleta de Dados

Os dados foram coletados nos dias 26 e 27 de outubro de 2025. Foi criada uma conta paga no Claude.ai e realizado o procedimento de criação de três aplicativos com base em cada prompt. Não houve intercorrências durante o processo. Os comandos foram executados de forma satisfatória.

Descrição das Telas dos Aplicativos

Aplicativo 1

O aplicativo 1 pode ser acessado clicando no link: <https://claude.ai/public/artifacts/8e97456d-e06c-44c7-a3f9-615fe6feaf65>

O aplicativo inicia com uma tela que mostra a primeira tentativa. Não é apresentada explicação inicial para o usuário. Nessa tela consta um título, uma barra de progressão, uma vinheta, quatro cards com emoji representando uma emoção conforme figura 1.



Figura 1. Exemplo da Tela Inicial do APP 1

Nota. Tela do aplicativo "Aprendendo Emoções" (APP 1). Apresenta uma breve história ("Maria estava brincando no parque...") e pergunta "Como essa pessoa está se sentindo?", com quatro alternativas representadas por emojis (Feliz, Bravo, Triste, Assustado).

Ao clicar no botão correto o aplicativo apresenta um som de celebração e em volta da imagem aparece uma borda verde, conforme Figura 2.



Figura 2. Exemplo da Tela de Teste do App 1 (Versão com Acerto)

Nota. Tela de teste do APP 1 na versão com acerto: após a escolha correta, o aplicativo exibe o retorno positivo "Muito bem! Você acertou!" e avança a barra de progresso.

Ao clicar no botão errado o aplicativo apresenta outro som e em volta da imagem aparece uma borda vermelha, conforme Figura 3.



Figura 3. Exemplo da Tela da Fase de Teste do App 1 (Versão com Erro)

Nota. Tela da fase de teste do APP 1 na versão com erro, ilustrando o retorno fornecido pelo aplicativo quando a emoção é identificada incorretamente.

Ao terminar as oito questões do teste, o aplicativo apresenta uma tela com o título, uma frase de celebração por completar toda a atividade, uma tabela com as respostas marcadas com a cor verde (acerto) ou vermelha (erro), a porcentagem de acertos, o tempo empregado para concluir cada questão, um botão com a opção de jogar novamente e outro de baixar os resultados, conforme Figura 4.



Figura 4. Exemplo da Tela Final do APP 1

Nota. Tela final do APP 1, exibida ao término da atividade de reconhecimento de emoções.

Aplicativo 2

O aplicativo 2 pode ser acessado clicando no ink: <https://claude.ai/public/artifacts/e9deadfa-1b32-4f88-9203-2d02e0170f1f>

O aplicativo inicia com uma tela que mostra os botões para cada uma das quatro fases. Não é apresentada explicação inicial para o usuário. Nessa tela consta um título e um botão para Pré-teste, Treino, Pós-teste e Generalização, conforme Figura 5.

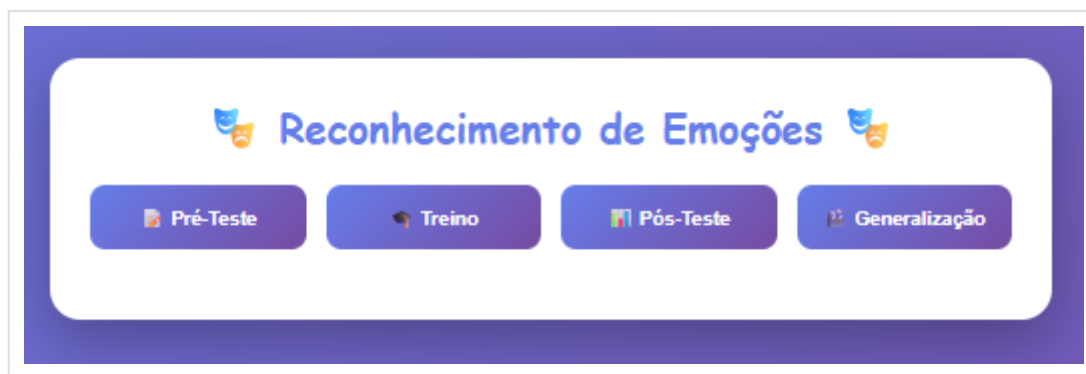


Figura 5. Exemplo da Tela Inicial do APP 2

Nota. Tela inicial do APP 2 (“Reconhecimento de Emoções”), com os botões das quatro fases do protocolo: Pré-Teste, Treino, Pós-Teste e Generalização.

Fase de Pré-teste

Ao clicar no botão de Pré-teste é apresentada uma tela onde consta um título, um botão com opção de voltar ao menu, uma barra de progressão, uma vinheta, seis cards com emoji representando uma emoção, conforme Figura 6.



Figura 6. Exemplo da Tela Inicial da Fase de Pré-teste APP 2

Nota. Tela do APP 2 (“Reconhecimento de Emoções”). A mesma captura ilustra, no texto, tanto a tela inicial da fase de Pré-teste quanto a da fase de Treino (ver nota de revisão sobre reaproveitamento da imagem).

Ao clicar no botão correto o aplicativo apresenta um som de celebração, em volta da imagem aparece uma borda verde e, no final da página, aparece uma frase de celebração e um botão para seguir para a próxima questão, conforme Figura 7.

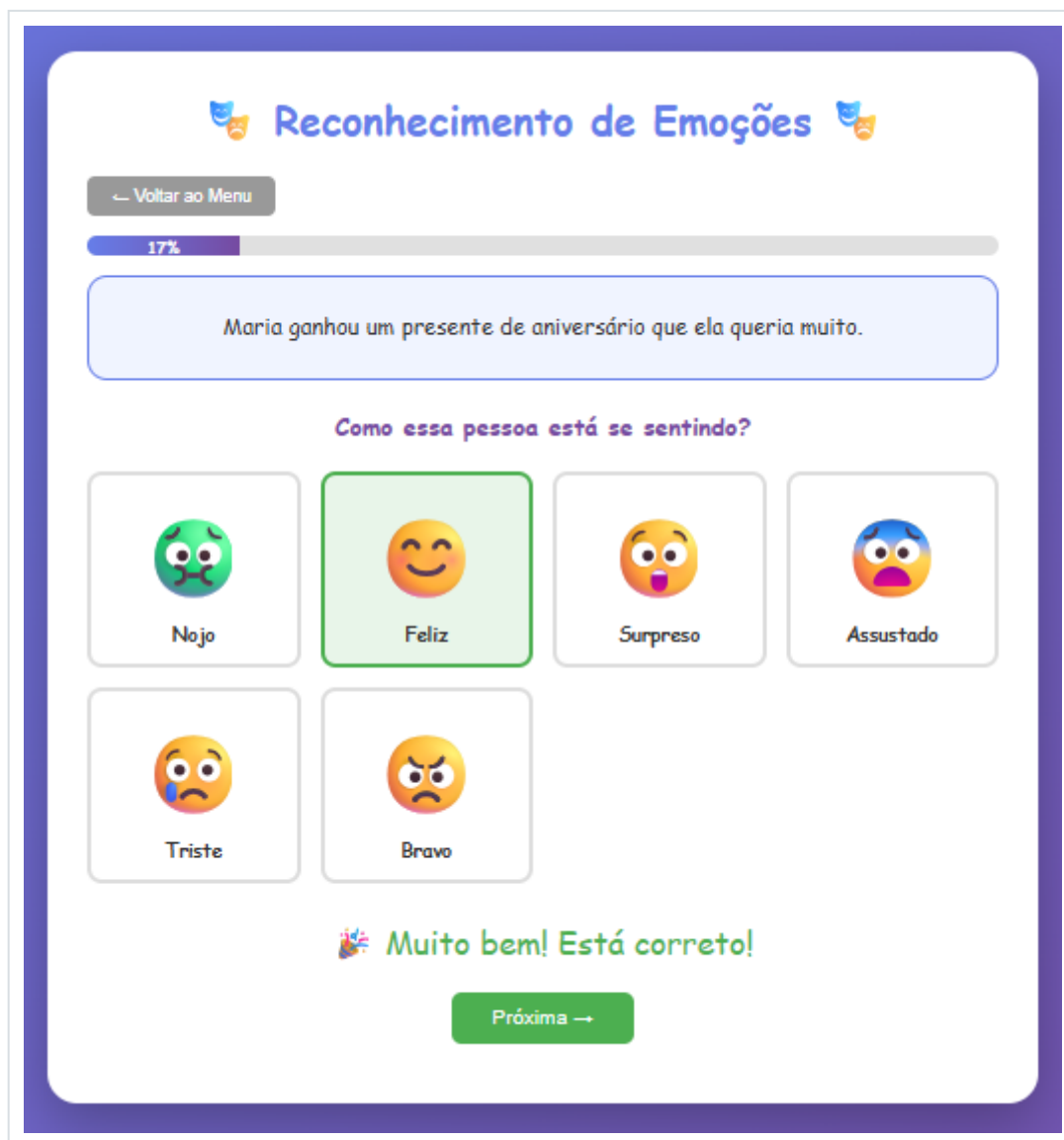


Figura 7. Exemplo da Tela da Fase de Pré-teste do APP 2 (Versão com Acerto)

Nota. Tela da fase de pré-teste do APP 2 na versão com acerto, usada para registrar o desempenho da criança antes do treino.

Ao clicar no botão errado o aplicativo apresenta outro som, em volta da imagem aparece uma borda vermelha e, no final da página, aparece uma frase orientando para tentar novamente e um botão para seguir para a próxima questão, conforme Figura 8.

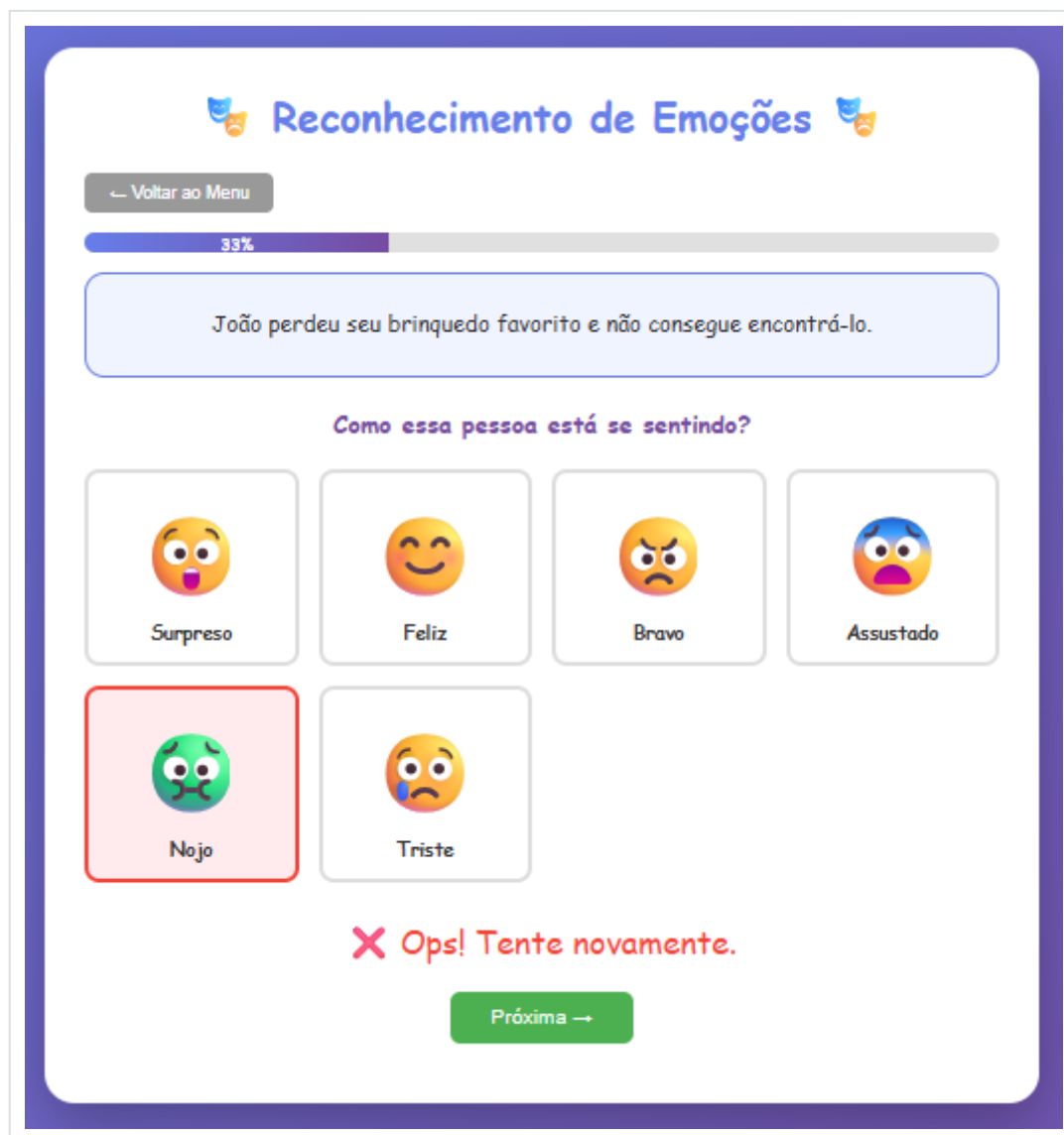


Figura 8. Exemplo da Tela da Fase de Pré-Teste do APP 2 (Versão com Erro)

Nota. Tela da fase de pré-teste do APP 2 na versão com erro, registrando uma resposta incorreta na linha de base.

Ao terminar as seis questões do teste, o aplicativo apresenta uma tela com o título, um botão com a opção de retornar ao menu, uma barra de progressão, dois quadro com os resultados em número e porcentagem, uma tabela com as respostas marcadas com a cor verde (acerto) ou vermelha (erro), o tempo empregado para concluir cada questão, um botão com a opção de baixar os resultados e outro para jogar novamente, conforme Figura 9.

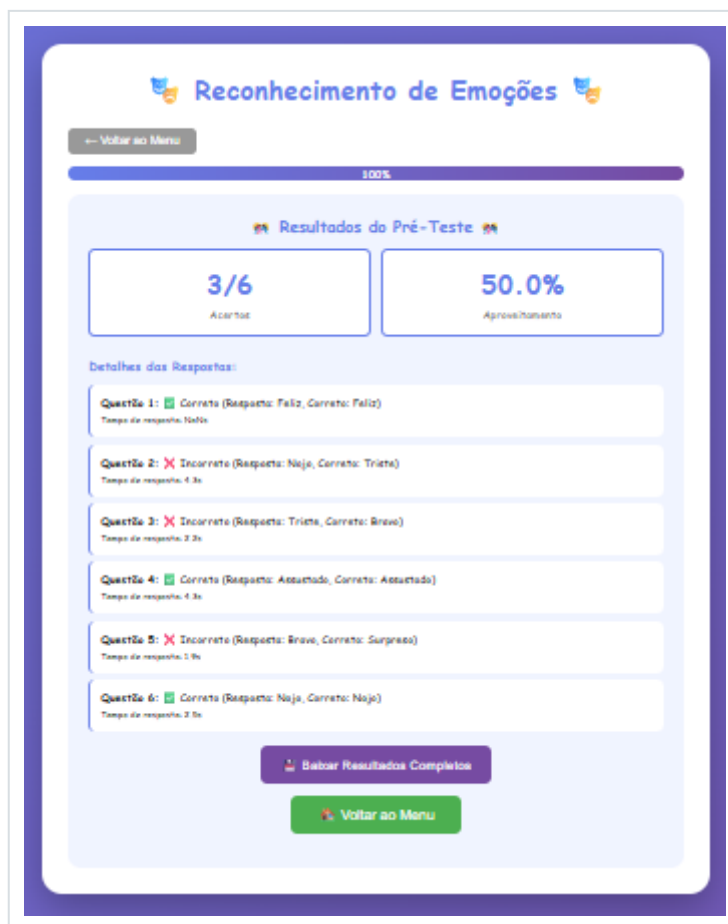


Figura 9. Exemplo da Tela Final da Fase do Pré-Teste do APP 2

Nota. Tela final da fase de pré-teste do APP 2, com o encerramento dessa etapa do protocolo.

Fase de Treino

Ao clicar no botão de Treino é apresentada uma tela (semelhante à tela inicial do Pré-teste) onde consta um título, um botão com opção de voltar ao menu, uma barra de progressão, uma vinheta, seis cards com emoji representando uma emoção, conforme Figura 10.



Figura 10. Exemplo da Tela da Fase do Treino APP 2

Nota. Tela do APP 2 (“Reconhecimento de Emoções”). A mesma captura ilustra, no texto, tanto a tela inicial da fase de Pré-teste quanto a da fase de Treino (ver nota de revisão sobre reaproveitamento da imagem).

Ao clicar no botão correto, o aplicativo apresenta um som de celebração, em volta da imagem aparece uma borda verde e, no final da página, aparece uma frase de celebração com uma explicação de como podemos ficar diante daquela situação e um botão para seguir para a próxima questão, conforme Figura 11.



Figura 11. Exemplo da Tela da Fase de Treino do APP 2 (Versão com Acerto)

Nota. Tela da fase de treino do APP 2 na versão com acerto, em que o aplicativo ensina a identificação da emoção com apoio e reforço imediato.

Ao clicar no botão errado o aplicativo apresenta outro som, em volta da imagem aparece uma borda vermelha e, no final da página, aparece a resposta correta e um botão para seguir para a próxima questão, conforme Figura 12.

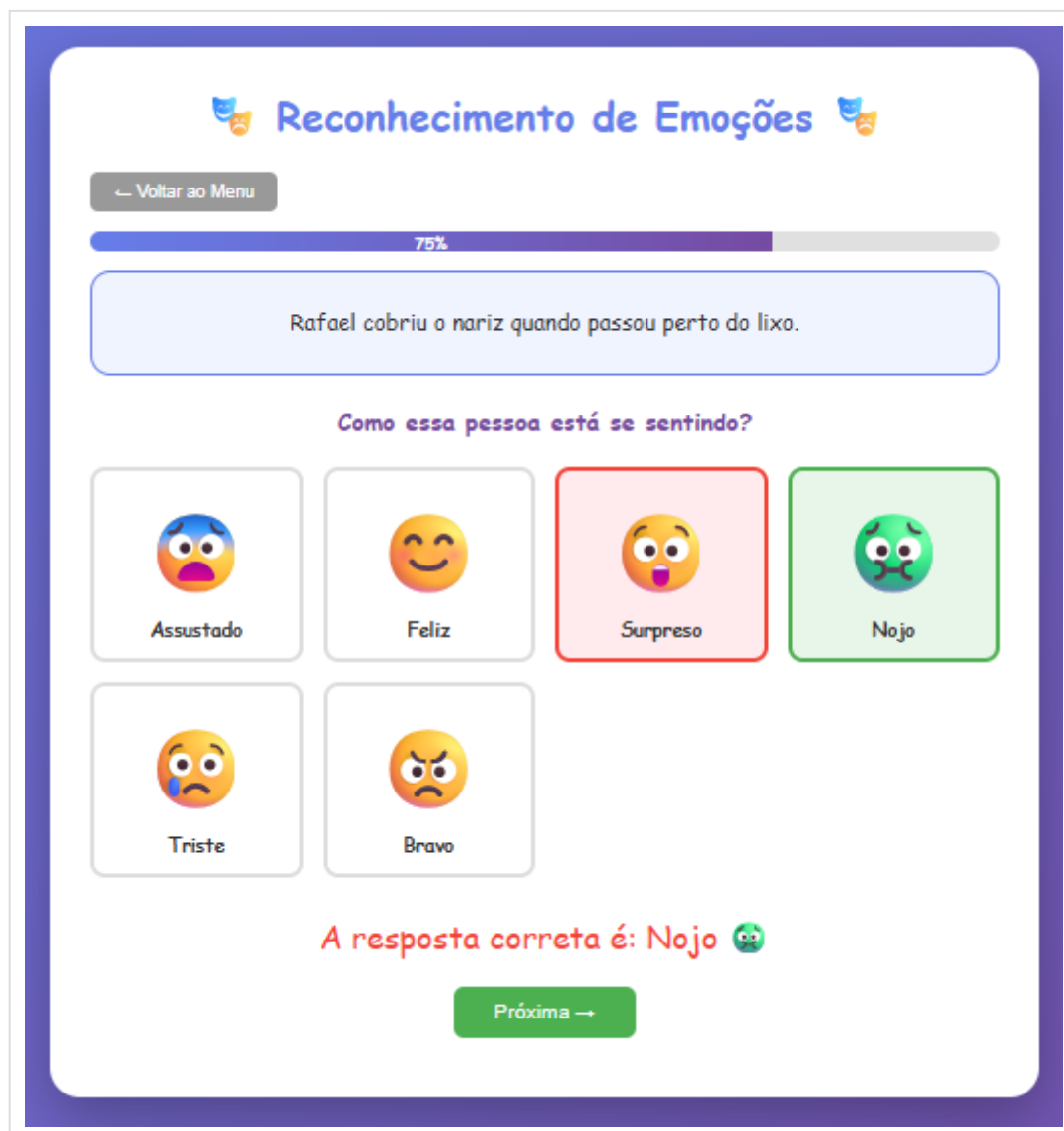


Figura 12. Exemplo da Tela da Fase de Treino do APP 2 (Versão com Erro)

Nota. Tela da fase de treino do APP 2 na versão com erro, ilustrando a correção oferecida durante o ensino.

Ao terminar as oito questões do teste, o aplicativo apresenta uma tela com o título, um botão com a opção de retornar ao menu, uma barra de progressão, dois quadro com os resultados em número e porcentagem, uma tabela com as respostas marcadas com a cor verde (acerto) ou vermelha (erro), o tempo empregado para concluir cada questão, um botão com a opção de baixar os resultados e outro para jogar novamente, conforme Figura 13.

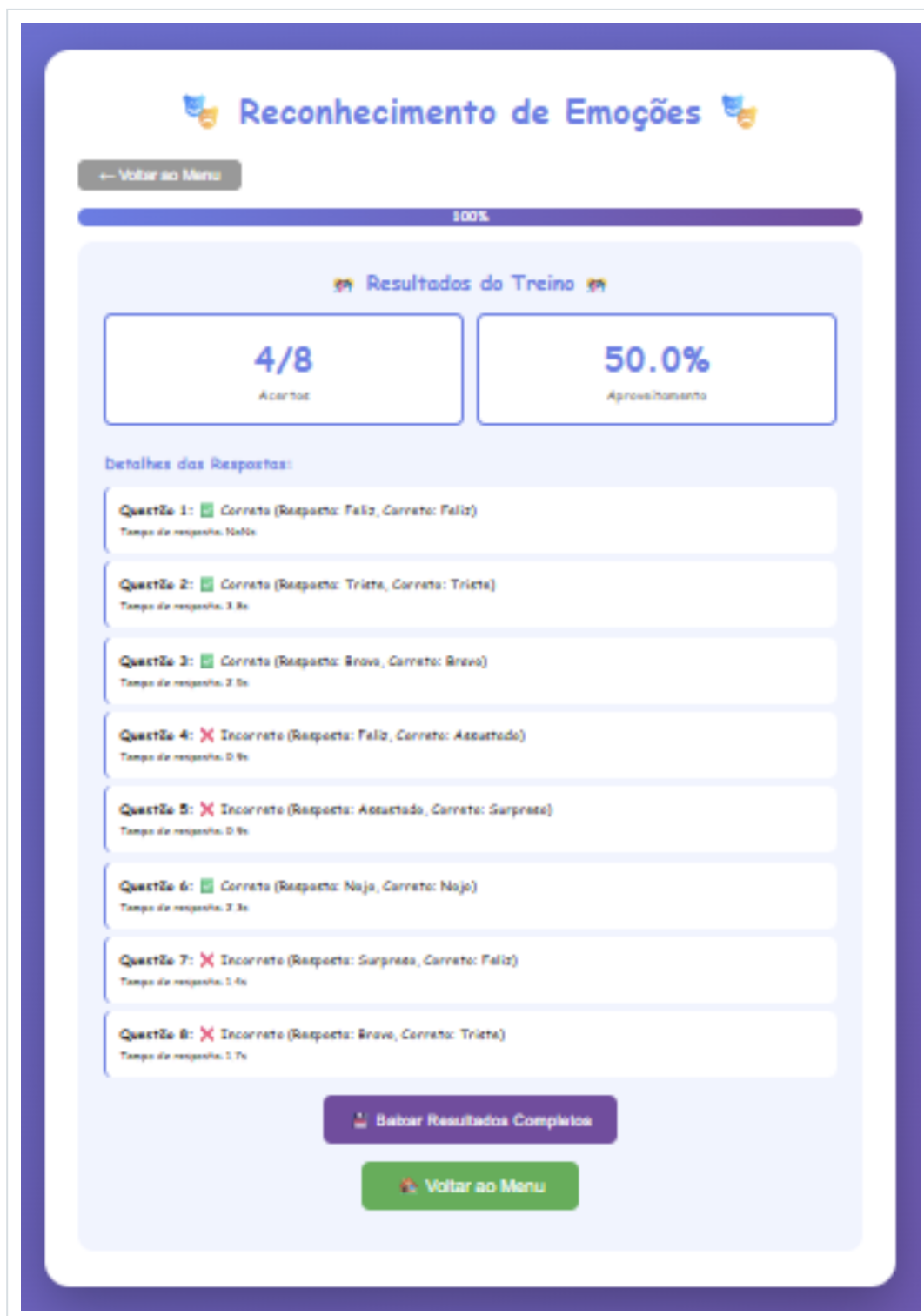


Figura 13. Exemplo da Tela da Fase Final do Treino do APP 2

Nota. Tela de resultados ao final da fase de Treino do APP 2 “Reconhecimento de Emoções”: exibe o número de acertos (4 de 8; 50%) e o detalhamento das respostas por questão (emoção respondida x correta), com opção de baixar os resultados completos.

Fase de Pós-teste

Ao clicar no botão correto o aplicativo apresenta um som de celebração, em volta da imagem aparece uma borda verde e, no final da página, aparece uma frase de celebração e um botão para seguir para a próxima questão, conforme Figura 14.



Figura 14. Exemplo da Tela da Fase de Pós-teste do APP 2 (Versão com Acerto)

Nota. Tela da fase de pós-teste do APP 2 na versão com acerto, que mede o desempenho após o treino.

Ao clicar no botão errado o aplicativo apresenta outro som, em volta da imagem aparece uma borda vermelha e, no final da página, aparece uma frase orientando para tentar novamente e um botão para seguir para a próxima questão, conforme Figura 15.

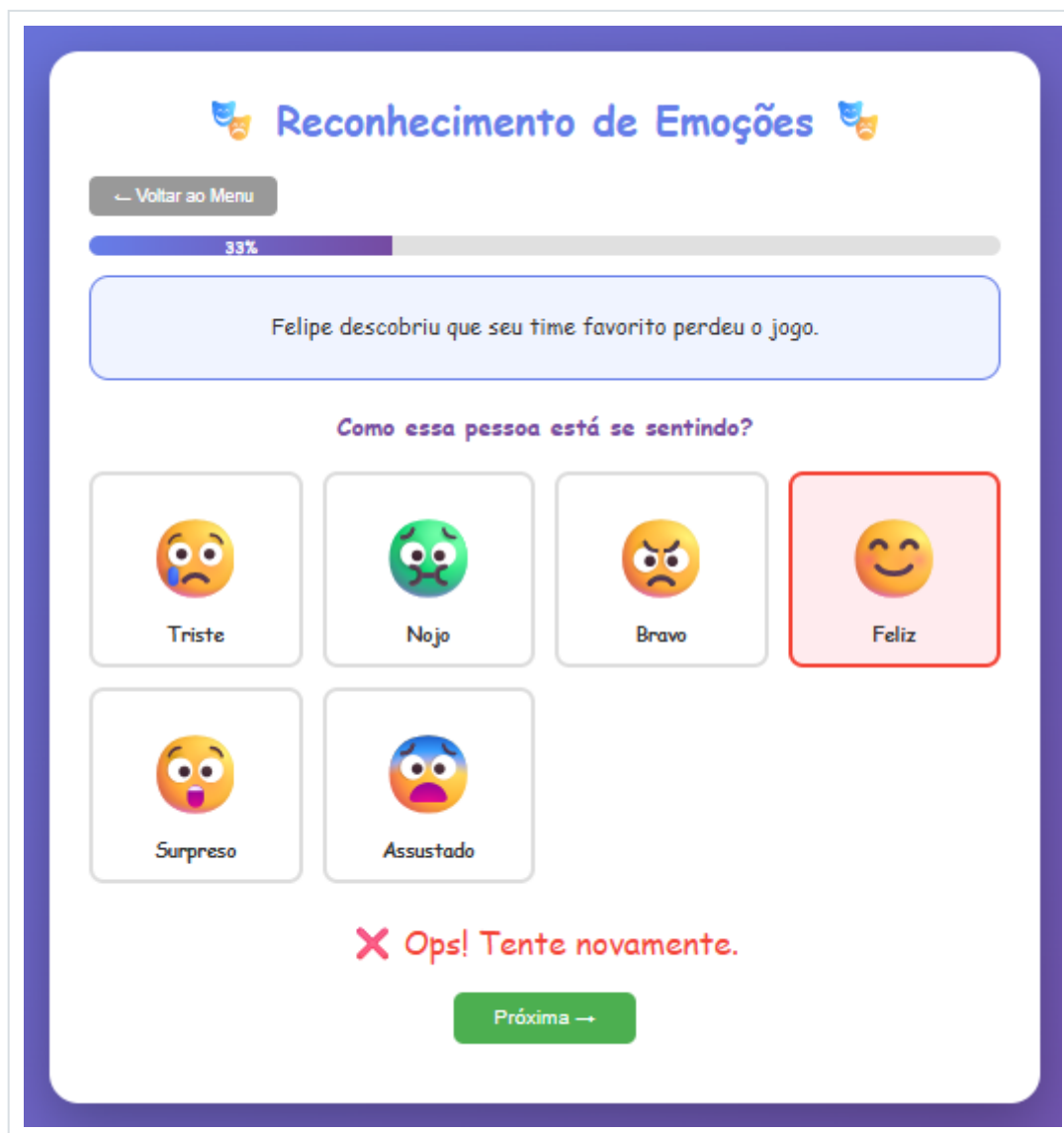


Figura 15. Exemplo da Tela da Fase de Pós-Teste do APP 2 (Versão com Erro)

Nota. Tela da fase de pós-teste do APP 2 na versão com erro.

Ao terminar as seis questões do teste, o aplicativo apresenta uma tela com o título, um botão com a opção de retornar ao menu, uma barra de progressão, dois quadro com os resultados em número e porcentagem, uma tabela com as respostas marcadas com a cor verde (acerto) ou vermelha (erro), o tempo empregado para concluir cada questão, um botão com a opção de baixar os resultados e outro para jogar novamente, conforme Figura 16.

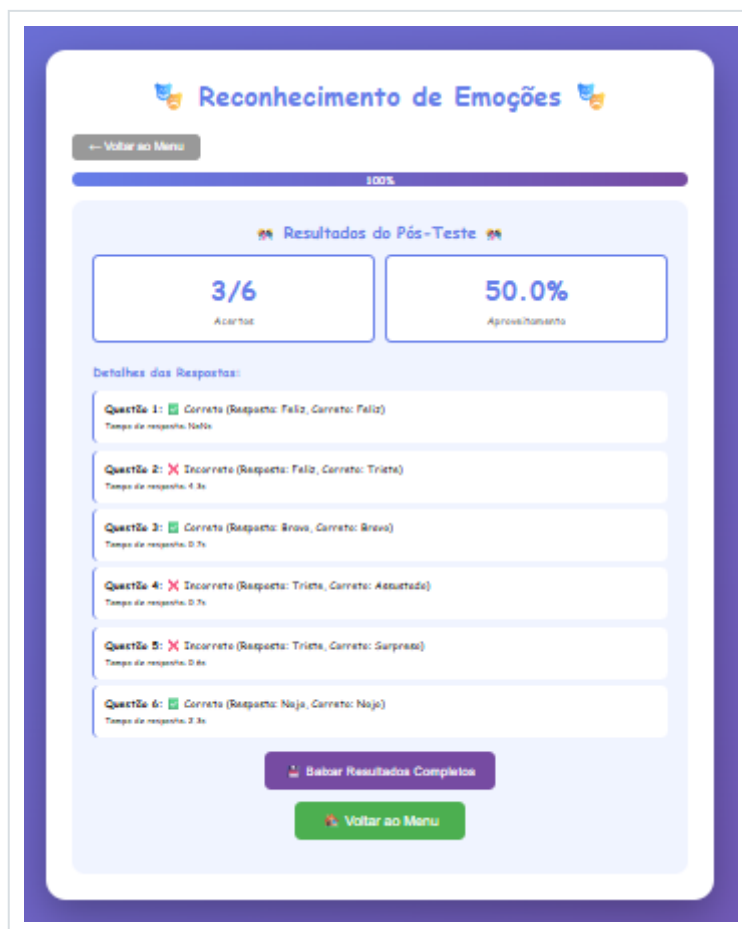


Figura 16. Exemplo da Tela da Fase Final do Pós-Teste do APP 2

Nota. Tela final da fase de pós-teste do APP 2.

Fase da Generalização

Ao clicar no botão de Generalização é apresentada uma tela onde consta um título, um botão com opção de voltar ao menu, uma barra de progressão, a pergunta “Como essa pessoa está se sentindo?”, um quadro indicando o lugar do vídeo com a sugestão do que deve constar no vídeo, seis cards com emoji representando uma emoção e um botão com seta para seguir para a próxima página, conforme Figura 17.

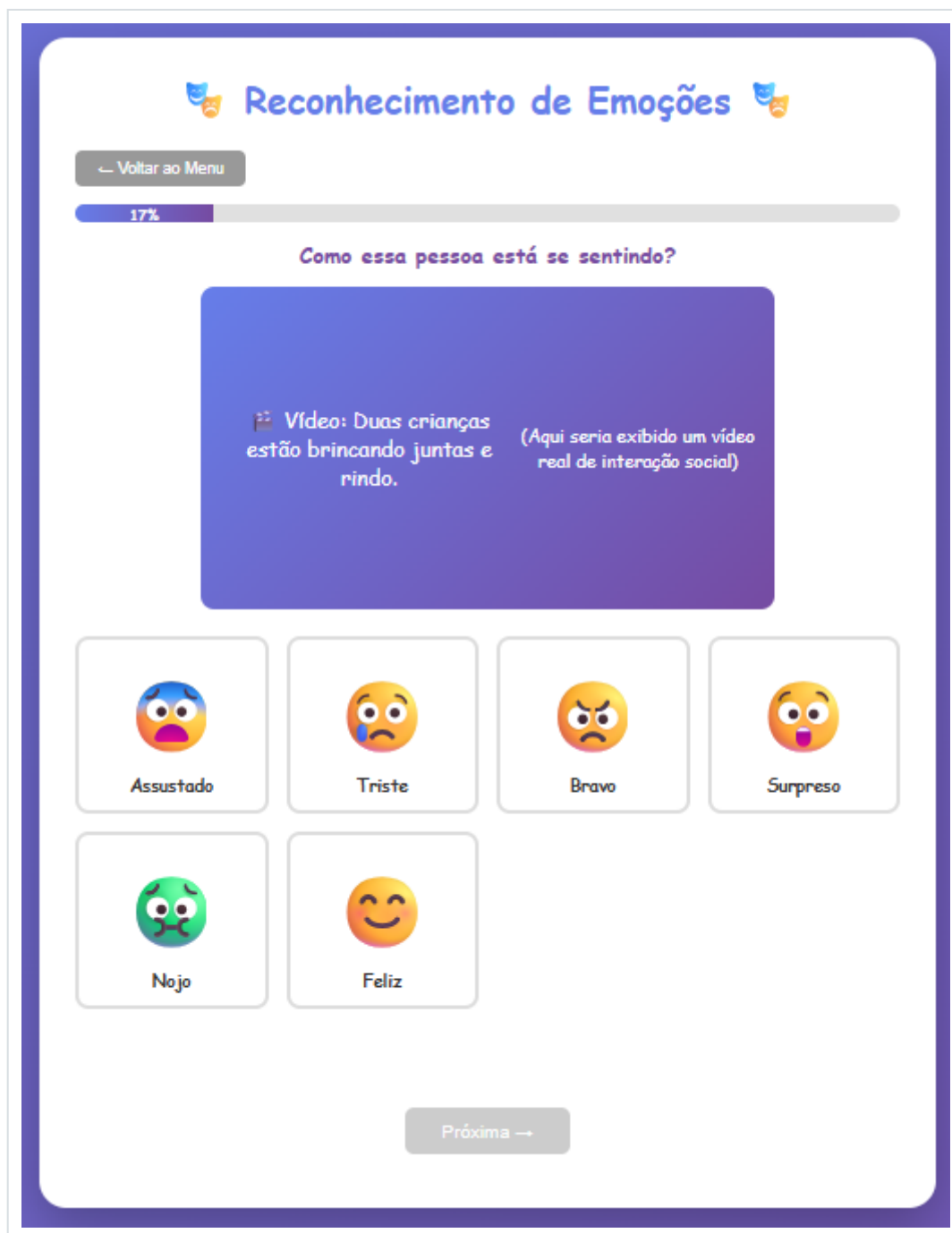


Figura 17. Exemplo da Tela da Fase de Generalização do APP 2

Nota. Tela da fase de generalização do APP 2, que avalia a transferência da habilidade aprendida para novas histórias e situações.

Ao clicar no botão correto o aplicativo apresenta um som de celebração, em volta da imagem aparece uma borda verde e, no final da página, aparece uma frase de celebração e um botão para seguir para a próxima questão, conforme Figura 18.

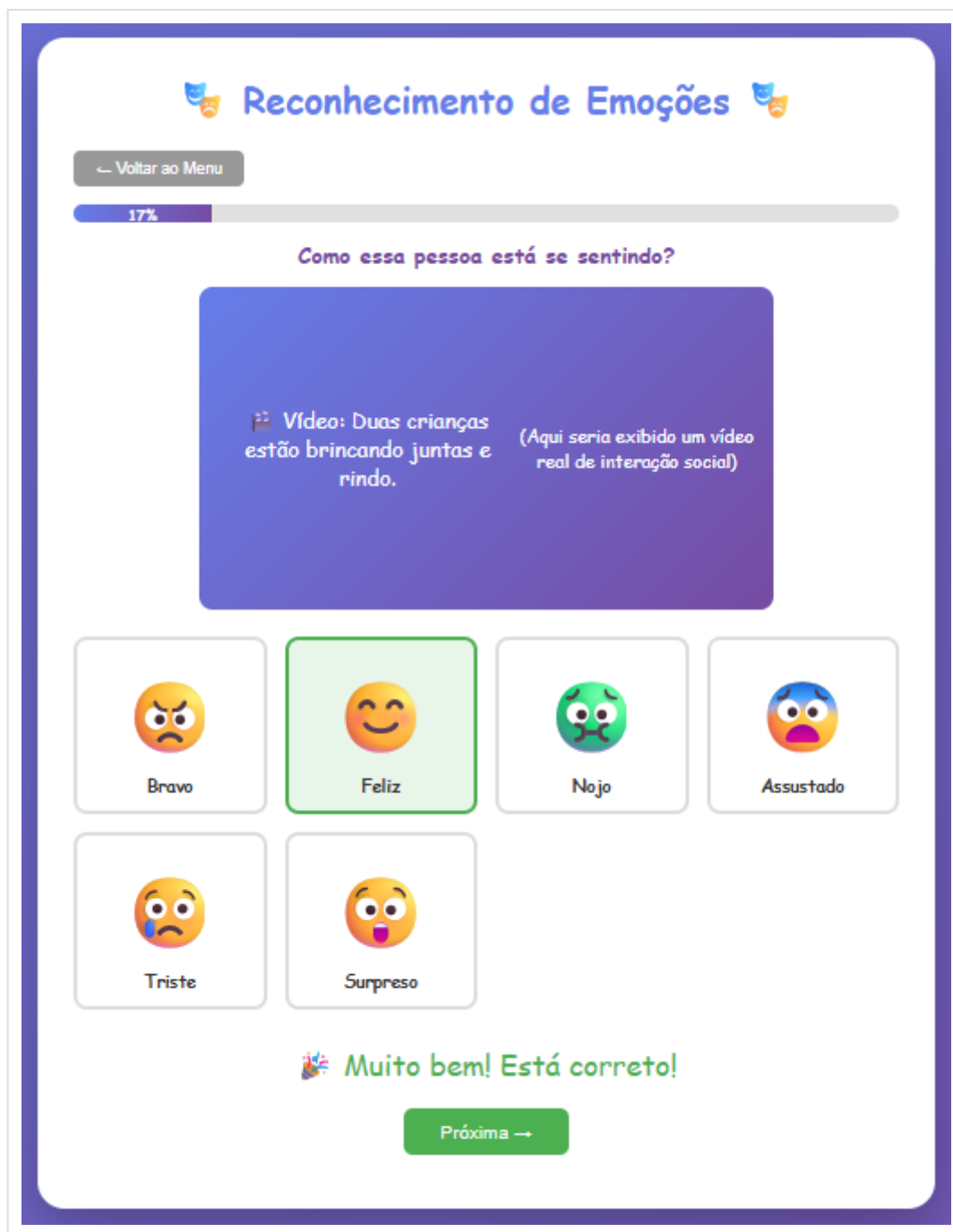


Figura 18. Exemplo da Tela da Fase de Generalização do APP 2 (Versão com Acerto)

Nota. Tela da fase de generalização do APP 2 na versão com acerto.

Ao clicar no botão errado o aplicativo apresenta um som diferente do de celebração, em volta da imagem aparece uma borda vermelha e, no final da página, aparece a frase “Ops! Tente novamente.” e um botão para seguir para a próxima questão, conforme Figura 19.

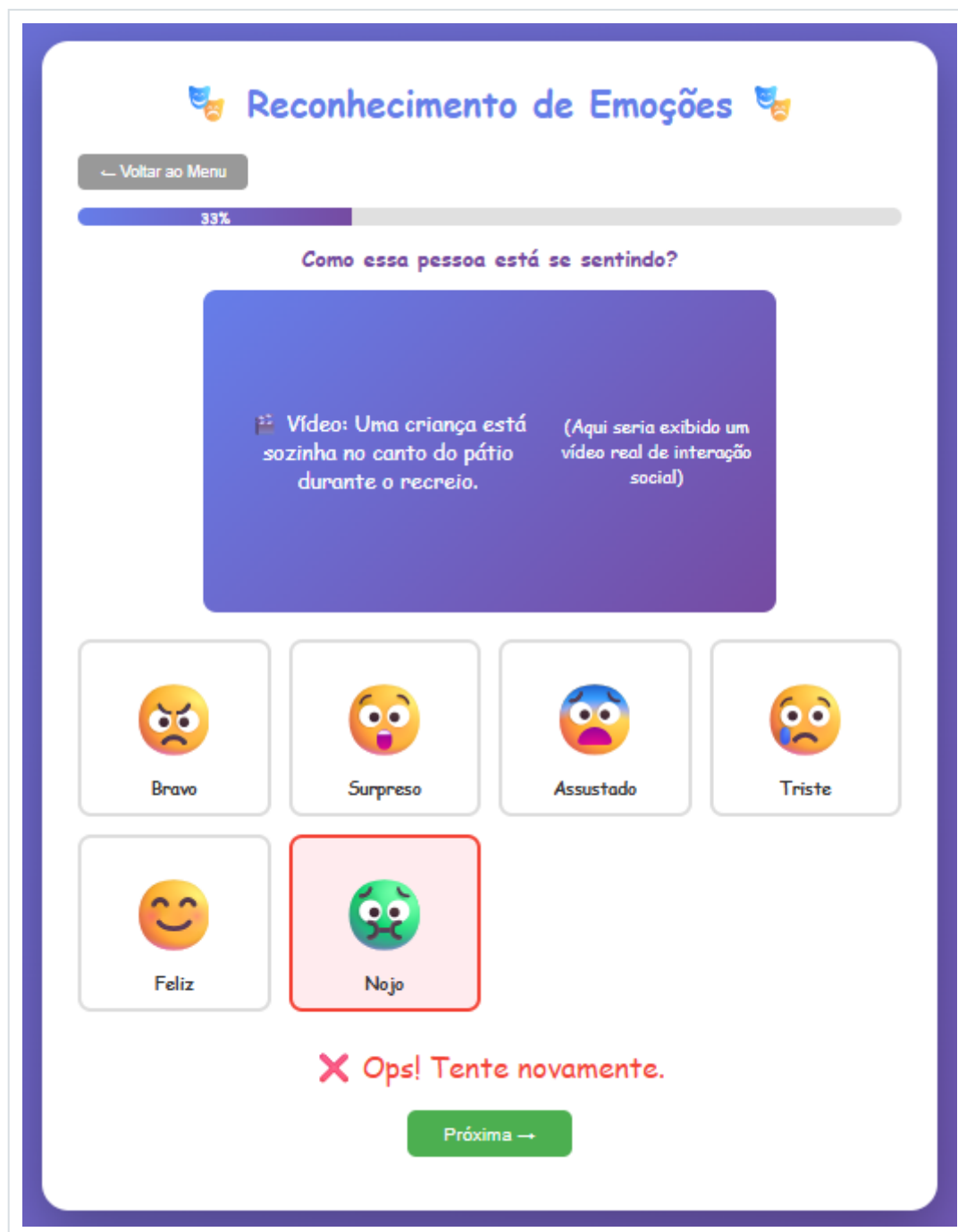


Figura 19. Exemplo da Tela da Fase de Generalização do APP 2 (Versão com Erro)

Nota. Tela da fase de generalização do APP 2 na versão com erro.

Ao terminar as seis questões do teste, o aplicativo apresenta uma tela com o título, um botão com a opção de retornar ao menu, uma barra de progressão, dois quadro com os resultados em número e porcentagem, uma tabela com as respostas marcadas com a cor verde (acerto) ou vermelha (erro), o tempo empregado para concluir cada questão, um botão com a opção de baixar os resultados e outro para voltar ao menu, conforme Figura 20.

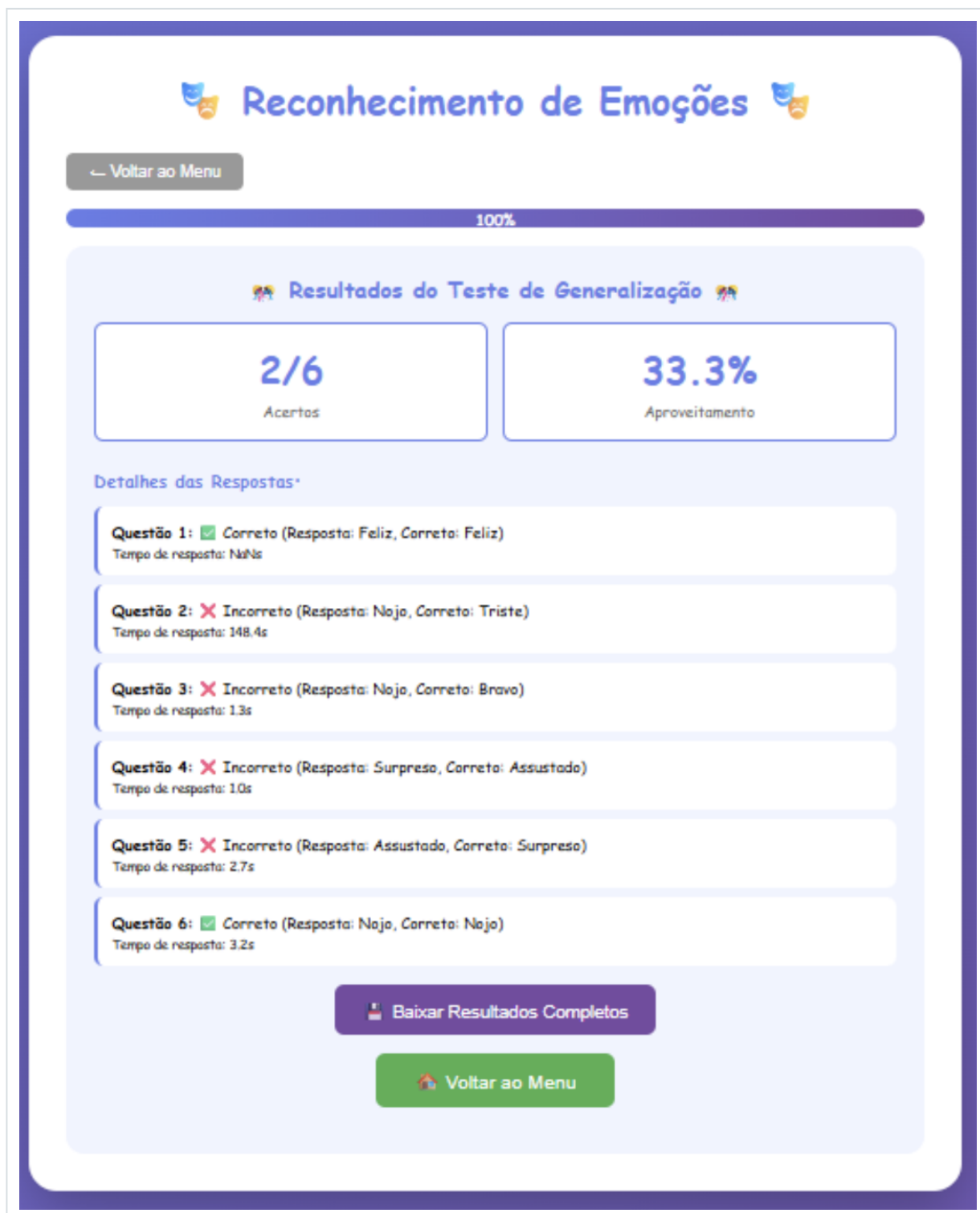


Figura 20. Exemplo da Tela Final da Fase de Generalização do APP 2

Nota. Tela de resultados ao final da fase de Generalização do APP 2: exibe o desempenho (2 de 6 acertos; 33,3%) e o detalhamento por questão, avaliando a transferência da habilidade para novas situações.

Aplicativo 3

O aplicativo 3 pode ser acessado clicando no link: <https://claude.ai/public/artifacts/3804d1e8-d561-4d8a-89f5-a2aa4a3b91ba>

O aplicativo inicia com uma tela de boas-vindas. Não é apresentada explicação inicial para o usuário. Nessa tela consta um título, um quadrado com um emoji, uma frase de boas-vindas e um convite para iniciar a atividade, conforme Figura 21.



Figura 21. Exemplo da Tela Inicial do APP 3

Nota. Tela inicial do APP 3 (“Reconhecimento de Emoções”), com a mensagem de boas-vindas e o botão “Começar Atividade”.

Ao clicar no botão “Começar a atividade”, o aplicativo inicia com uma tela que mostra a primeira tentativa. Não é apresentada explicação inicial para o usuário. Nessa tela consta um título, uma barra de progressão, uma vinheta, quatro cards com emoji representando uma emoção conforme Figura 22.

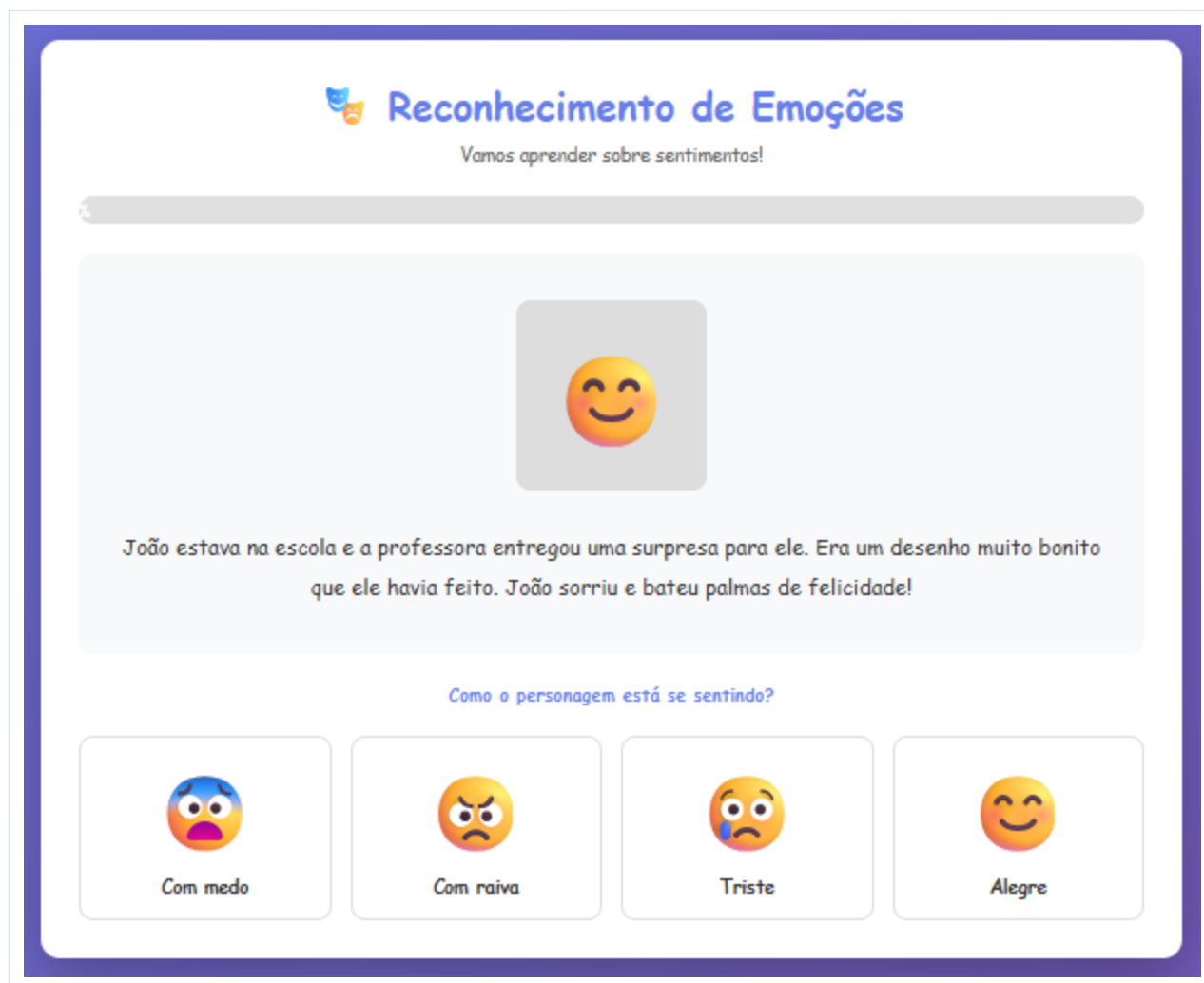


Figura 22. Exemplo da Tela de Teste do APP 3

Nota. Tela de teste do APP 3, com a apresentação de uma situação e a solicitação de identificação da emoção do personagem.

Ao clicar no botão correto o aplicativo apresenta um som, o fundo da imagem correta fica verde, aparece uma frase de celebração e um botão para a próxima história, conforme Figura 23.



Figura 23. Exemplo da Tela de Teste do APP 3 (Versão com Acerto)

Nota. Tela de teste do APP 3 na versão com acerto, exibindo o retorno positivo ao usuário.

Ao clicar no botão errado o aplicativo apresenta um som, o fundo da imagem errada fica vermelho, aparece uma frase com o feedback apresentando a opção correta e um botão para a próxima história, conforme Figura 24.



Figura 24. Exemplo da Tela de Teste do APP 3 (Versão com Erro)

Nota. Tela de teste do APP 3 na versão com erro, exibindo o retorno corretivo.

Ao terminar as 10 questões do teste, o aplicativo apresenta uma tela com o título, uma frase convidando para aprender sobre os sentimentos, um troféu, uma frase de celebração por ter completado a atividade, um quadro sobre a performance da criança (quantidade de histórias, as respostas corretas e a porcentagem de acertos), um botão com a opção recomeçar o teste e outro de salvar os resultados, conforme Figura 25.



Figura 25. Exemplo da Tela Final do APP 3

Nota. Tela final do APP 3, ao término da atividade de reconhecimento de emoções.

4 ANÁLISE DE CONTEÚDO

Na dimensão 1, conforme apresentado na Tabela 1, as respostas do pesquisador e do Google IA foram convergentes exceto quanto ao item 1.3 porque, de acordo com a avaliação do pesquisador esse item apresentou nota 2 devido ao aplicativo registrar as respostas corretas e incorretas e o tempo de execução de cada resposta, entretanto não foi possível baixar/salvar os resultados.

Tabela 1

Tabela comparativa dos resultados dos três aplicativos em relação à Dimensão 1: Adequação aos Princípios da Análise do Comportamento Aplicado (ABA). A primeira nota se refere à nota do pesquisador e a segunda à da IA.

Tabela 1. Avaliação do APP 1 — desempenho por critério (Pré-teste/Generalização)

Critério	Aplicativo 1 (P/G)	Aplicativo 2 (P/G)	Aplicativo 3 (P/G)
1.1 - Procedimento de Discriminação Condicional	4/4	4/4	4/4
1.2 - Sistema de Reforçamento	4/4	4/4	4/4
1.3 - Registro de Desempenho	2/4	2/4	2/4
1.4 - Estruturação de Tentativas	4/4	4/4	4/4

Nota. Pesquisador/Google IA (P/G)

Na dimensão 2 as respostas do pesquisador e do Google IA foram convergentes em todos os itens, merecendo atenção o ponto 2.4 devido ao aplicativo considerar algumas necessidades sensoriais atípicas como cores mais suaves e poucas distrações, entretanto não oferece a possibilidade de ajustes nos estímulos como volume, brilho e velocidade.

Tabela 2

Tabela comparativa dos resultados dos três aplicativos em relação à Dimensão 2: Adequação às Características do Transtorno do Espectro Autista. A primeira nota se refere à nota do pesquisador e a segunda à da IA.

Tabela 2. Avaliação do APP 2 — desempenho por critério (Pré-teste/Generalização)

Critério	Aplicativo 1 (P/G)	Aplicativo 2 (P/G)	Aplicativo 3 (P/G)
2.1 - Interface Visual	4/4	4/4	4/4
2.2 - Clareza de Instruções	4/4	4/4	4/4
2.3 - Previsibilidade	4/4	4/4	4/4
2.4 - Adaptabilidade Sensorial	2/2	2/2	2/2

Nota. Pesquisador/Google IA (P/G)

Na dimensão 3 as respostas do pesquisador e do Google IA foram ...

Tabela 3

Tabela comparativa dos resultados dos três aplicativos em relação à Dimensão 3: Funcionalidade Técnica. A primeira nota se refere à nota do pesquisador e a segunda à da IA.

Tabela 3. Avaliação do APP 3 — desempenho por critério (Pré-teste/Generalização)

Critério	Aplicativo 1 (P/G)	Aplicativo 2 (P/G)	Aplicativo 3 (P/G)
3.1 - Operacionalidade do Código	4/4	4/4	4/4
3.2 - Responsividade	4/4	4/4	4/4
3.3 - Sistema de salvamento	1/3	1/3	1/3

Nota. Pesquisador/Google IA (P/G)

A dimensão 4 está bem estruturada, mas na análise do item 4.1 há um erro crucial relativo ao feedback que não poderia ser dado nesta fase. A mesma coisa no pós teste. Na verdade ele fez três fases iguais. Para que o aplicativo estivesse adequado, na fase de Pré-teste (linha de base) e pós-teste não haveria feedback, apenas no treino.

Tabela 4

Tabela comparativa dos resultados dos três aplicativos em relação à Dimensão 4: Estrutura Metodológica Experimental. A primeira nota se refere à nota do pesquisador e a segunda à da IA.

Tabela 4. Comparação geral dos aplicativos — desempenho por critério (P/G)

Critério	Aplicativo 1 (P/G)	Aplicativo 2 (P/G)	Aplicativo 3 (P/G)
4.1 - Fase de pré-teste	1/1	1/1	1/1
4.2 - Fase de treino	1/3	3/3	1/3
4.3 - Fase de pós-teste	1/1	4/4	1/1
4.4 - Teste /de generalização	1/1	4/1	1/1
4.5 - Progressão pedagógica	1/2	4/2	1/2
4.6 - Variedade de estímulos	1/2	2/2	1/2
4.7 - Fidelidade ao Delineamento de Lima (2017)	1/3	4/3	1/2

Nota. Pesquisador/Google IA (P/G)

5 DISCUSSÃO

Resumo

O objetivo da presente pesquisa foi avaliar a qualidade técnica, metodológica e científica de aplicativos gerados por uma inteligência artificial (Claude Sonnet 4.5) para ensinar crianças autistas a discriminarem emoções de acordo com o contexto por meio de historinhas. Para isso, foram elaborados três prompts com diferentes níveis de detalhamento, e cada um deles gerou um aplicativo distinto, posteriormente avaliado em quatro dimensões: adequação aos princípios da Análise do Comportamento Aplicada (ABA), adequação às características do TEA, funcionalidade técnica e estrutura metodológica.

De forma geral, os resultados mostraram que todos os aplicativos apresentaram algum nível de adequação às propostas, porém apenas o aplicativo gerado pelo prompt intermediário estruturou corretamente as quatro fases experimentais, apesar de ter apresentado um erro crucial relativo ao feedback. Outras limitações identificadas envolveram a ausência de salvamento adequado de dados, falhas metodológicas nas fases de pré-teste e pós-teste e baixa variedade de estímulos.

Comparação com a Literatura

Os resultados obtidos dialogam diretamente com os achados de McFayden et al. (2024) e Kim et al. (2025), que apontam que os modelos de linguagem são capazes de gerar conteúdos corretos e coerentes, mas apresentam limitações importantes relacionadas à clareza, aplicabilidade e consistência técnica. Semelhante ao observado nesses estudos, os aplicativos desenvolvidos nesta pesquisa demonstraram precisão conceitual em vários pontos, especialmente na implementação do procedimento de discriminação condicional e no uso de feedback imediato – ambos consistentes com os princípios da ABA.

Por outro lado, também foram identificadas falhas semelhantes às relatadas na literatura, como a inconsistência no sistema de salvamento, a dificuldade em executar recomendações mais complexas (como remover feedback no pré-teste e pós-teste) e a não implementação completa de alguns parâmetros metodológicos, mesmo quando especificados. Isso reforça o argumento de Hacker et al. (2023, como citado em Sampaio et al., 2024) de que os modelos de IA tendem a produzir resultados plausíveis, mas nem sempre precisos, devido às limitações inerentes ao funcionamento probabilístico dos LLMs.

Além disso, considerando os achados de Lima (2017) sobre a importância da estrutura metodológica adequada – especialmente a diferenciação clara entre fases com e sem feedback – observa-se que apenas um dos aplicativos aproximou-se desse padrão.

Nesse contexto, o estudo de Araújo et al. (2024) contribui para ampliar a compreensão sobre o uso de tecnologias no ensino de habilidades emocionais. Os autores investigaram o reconhecimento e a nomeação de emoções em crianças e jovens adultos com TEA e Síndrome de

Down, por meio de tarefas aplicadas pela família em ambiente domiciliar e mediadas por recursos tecnológicos. Os resultados indicaram aumento significativo no número de acertos após o ensino, bem como a generalização das habilidades para diferentes estímulos.

Destaca-se, ainda, que o estudo evidenciou a viabilidade da participação ativa dos familiares na aplicação das intervenções, mesmo em contexto remoto, contribuindo para o desenvolvimento do repertório emocional dos participantes. Esses achados reforçam a eficácia de procedimentos estruturados de ensino, bem como a importância do envolvimento familiar e do uso de tecnologias como estratégias complementares no desenvolvimento de habilidades socioemocionais em indivíduos com TEA.

Assim, os resultados da presente pesquisa corroboram a literatura ao evidenciar o potencial das IAs na geração de materiais educativos úteis e expandem esse campo ao demonstrar a possibilidade de desenvolvimento de aplicativos completos. Contudo, também indicam limitações importantes, na medida em que, mesmo com instruções detalhadas, persistem falhas estruturais que demandam revisão humana especializada.

Contribuições teóricas

Esta pesquisa contribui teoricamente ao demonstrar que modelos de IA podem ser utilizados não apenas para gerar textos explicativos sobre autismo e emoções, como mostrado por estudos anteriores, mas também para produzir artefatos funcionais, como aplicativos educacionais baseados em princípios comportamentais. Assim, o estudo amplia a discussão sobre o uso de LLMs como ferramentas de criação de tecnologia assistiva (recurso, equipamento, software, serviço ou estratégia criada para aumentar, manter ou melhorar as habilidades funcionais de pessoas com deficiência ou necessidades específicas), evidenciando que sua utilidade ultrapassa a geração de respostas dialogadas e pode alcançar a produção de instrumentos de ensino.

Além disso, os resultados reforçam a importância da qualidade dos prompts — um aspecto que ainda é pouco explorado nos estudos sobre IA aplicada à educação. A comparação entre os três níveis de detalhamento mostra que prompts mais detalhados não necessariamente garantem melhor execução em todos os aspectos, o que evidencia a necessidade de pesquisas teóricas sobre engenharia de prompts no desenvolvimento de intervenções educacionais digitais.

Contribuições metodológicas

No campo metodológico, esta pesquisa dá um passo relevante ao avaliar aplicativos produzidos por IA destinados ao ensino de crianças com TEA, já que a revisão da literatura não identificou estudos com esse enfoque. Pesquisas anteriores concentravam-se principalmente na avaliação textual da precisão das respostas fornecidas pelas IAs (como McFayden et al., 2024; Kim et al., 2025) enquanto esta investigação amplia esse horizonte ao propor uma avaliação sistemática de produtos digitais completos.

Assim, ao utilizar critérios derivados da ABA, da Educação Especial e de parâmetros metodológicos comparativos, o estudo oferece uma proposta inovadora de avaliação que pode servir de referência para pesquisas futuras. Além disso, demonstra que o uso de prompts estrutura-

dos pode influenciar diretamente a qualidade pedagógica e metodológica de um aplicativo, destacando a importância da construção cuidadosa das instruções como parte essencial do processo de pesquisa.

Contribuições Aplicadas

Do ponto de vista aplicado, esta pesquisa evidencia que, apesar de limitações, os aplicativos gerados apresentam características que podem ser úteis para profissionais e cuidadores. Interfaces simples, feedback imediato, previsibilidade e ausência de estímulos distratores são elementos reconhecidamente benéficos para crianças com TEA, o que mostra que as IAs conseguem produzir materiais acessíveis e alinhados às necessidades desse público.

Além disso, o estudo demonstra que ferramentas como o Claude podem ser utilizadas por profissionais sem formação em programação para criar recursos educacionais adaptados, democratizando a produção de tecnologias assistivas. Isso representa uma contribuição prática expressiva, especialmente para contextos clínicos, escolares e familiares com recursos limitados.

Embora ainda não se trate de um produto final pronto para uso clínico, o estudo mostra que a IA já é capaz de gerar protótipos funcionais, que podem ser refinados por analistas do comportamento, professores e desenvolvedores, ampliando significativamente o acesso a materiais baseados em evidências.

Limitações e pesquisas futuras

A principal limitação deste estudo é que os aplicativos foram avaliados enquanto produtos técnicos, sem aplicação direta com o público-alvo — crianças autistas. Dessa forma, não foi possível verificar se a interface, os estímulos, o tempo de apresentação e o formato das histórias são de fato eficazes em situações reais de ensino. Outra limitação observada diz respeito ao sistema de salvamento de dados, que apresentou falhas em todas as versões, comprometendo uma das funcionalidades essenciais para pesquisas comparativas e intervenções clínicas.

Além disso, embora os prompts tenham sido elaborados de forma progressiva, a IA nem sempre conseguiu implementar instruções metodologicamente fundamentais, como a ausência de feedback nas fases de pré-teste e pós-teste, o que evidencia limitações estruturais do modelo.

Embora não tenha sido o objetivo deste trabalho comparar o aplicativo gratuito com a versão paga, como os testes iniciais foram realizados na versão gratuita, foi possível observar diferenças entre o aplicativo gerado na versão paga do modelo de IA e aquele obtido na versão gratuita. A versão paga produziu elementos visuais mais elaborados, como celebrações mais coloridas e caixas de diálogo mais evidentes nos acertos, além de apresentar uma segunda tentativa com outra história quando o usuário errava. No entanto, essas diferenças, embora esteticamente positivas, não resolveram problemas estruturais importantes. Tanto na versão paga quanto na gratuita, ao se tentar baixar os resultados em PDF ou copiar os dados em CSV,

ocorreu o mesmo erro técnico (“Uncaught InvalidCharacterError: Failed to execute 'atob' on 'Window'...”). Assim, em nenhuma das versões foi possível salvar os resultados, configurando uma limitação técnica significativa.

Para pesquisas futuras, sugere-se que novos estudos realizem:

- Testes empíricos com crianças autistas, avaliando eficácia, engajamento, compreensão e usabilidade;
- Refinamento dos prompts, a fim de investigar quais estruturas produzem maior fidelidade metodológica e adequação científica;
- Comparações entre diferentes modelos de IA, verificando qual tecnologia se mostra mais adequada para gerar recursos baseados na ABA;
- Implementações assistidas, combinando IA com desenvolvedores humanos e analistas do comportamento para produzir aplicativos completos, estáveis e validados empiricamente;
- Revisões específicas sobre funcionalidade técnica, buscando soluções que garantam sistemas confiáveis de salvamento e exportação de dados, essenciais tanto para pesquisa quanto para uso clínico e educacional.

6 CONCLUSÃO

A presente pesquisa avaliou a utilidade e a confiabilidade de aplicativos gerados por inteligência artificial para ensinar crianças autistas a discriminarem emoções por meio de histórias. Os resultados indicam que, embora os aplicativos apresentem limitações técnicas e metodológicas — especialmente no registro e salvamento de dados, na fidelidade às fases experimentais e na implementação adequada das instruções —, eles também demonstram potencial ao incorporar elementos da Análise do Comportamento Aplicada (ABA) e ao utilizar interfaces simples, previsíveis e adequadas ao perfil sensorial de crianças com TEA.

De modo geral, os achados mostram que a IA é uma ferramenta promissora para a criação de recursos educacionais acessíveis, permitindo que profissionais e cuidadores produzam materiais personalizados mesmo sem conhecimentos avançados em programação. Apesar das falhas observadas, os aplicativos gerados oferecem uma base inicial que pode ser aprimorada e utilizada como ponto de partida para o desenvolvimento de tecnologias assistivas mais robustas. Assim, com refinamento técnico e validação empírica, a colaboração entre IA e especialistas têm potencial para resultar em soluções inovadoras e eficazes para o ensino de habilidades socioemocionais a crianças autistas, contribuindo significativamente para a inclusão, autonomia e qualidade de vida dessas crianças.

Referências

- Araújo, P. P. D., Gois, J. P., Domeniconi, C., Benitez, P., Souza, C. B. A. D., Freitas, M. C. D., & Trevisan, D. (2024). Avaliação e ensino de emoções com crianças e jovens adultos com autismo ou síndrome de down mediado pela família. *Psicologia: ciência e profissão*, 44, e258093.

- Associação Psiquiátrica Americana. (2022). Manual diagnóstico e estatístico de transtornos mentais (5ª ed., revisão do texto). <https://integrada.minhabiblioteca.com.br/reader/books/9786558822851>
- Felix, L. M. de L., & Moreira, M. (2023). Autismo: Estratégias científicas para ensino de reconhecimento de emoções. Instituto Walden4.
- Goulart, P., & de Assis, G. J. A. (2002). Estudos sobre autismo em análise do comportamento: aspectos metodológicos. *Revista Brasileira de Terapia Comportamental e Cognitiva*, 4(2), 151-165.
- Kim, Y., Vilches, S. L., Shapiro, S., & Clarkson, A. (2025). Testing the capability of generative artificial intelligence for parent and caregiver information seeking. *Family Relations*, 74(3), 1266-1284.
- Kusano, M. E., Domeniconi, C., & Schmidt, A. (2023). Comparação de procedimentos para ensino de emoções faciais para criança com TEA: Um estudo exploratório. *Acta Comportamentalia*, 31(2).
- Lima, A. A. (2017). Efeitos do uso de histórias infantis sobre o reconhecimento de expressões faciais de emoções em crianças com autismo (Dissertação de mestrado, Universidade Federal do Pará). <https://repositorio.ufpa.br/handle/2011/13848>
- McFayden, T. C., Bristol, S., Putnam, O., & Harrop, C. (2024). ChatGPT: Artificial intelligence as a potential tool for parents seeking information about autism. *Cyberpsychology, Behavior, and Social Networking*, 27(2), 135-148. <https://doi.org/10.1089/cyber.2023.0202>
- Pontes, A. N., Cerqueira, T. M. G., Lima, N. M. L., de Brum, E. H. M., & Brunoni, D. (2020). Triagem do Transtorno do Espectro do Autismo em escolares: uso da inteligência artificial. *MEMNON EDIÇÕES CIENTÍFICAS LTDA*, (pp 51-145).
- Sampaio, R. C., Nicolás, M. A., Junquilha, T. A., Silva, L. R. L., Freitas, C. S. D., Telles, M., & Santos, L. C. D. (2024). ChatGPT e outras IAs transformarão a pesquisa científica: Reflexões sobre seus usos. *Revista de Sociologia e Política*, 32, e008. <https://doi.org/10.1590/1678-98732432e008>
- Santos, A. B., Santos, R. E. D. S., & da Silva, C. M. R. (2025). TEAChat: Promovendo o Suporte Educacional ao Autismo com Inteligência Artificial Conversacional. In *Workshop sobre Educação em Computação (WEI)* (pp. 1369-1379). SBC
- Siqueira, B. (2025). Censo 2022 identifica 2,4 milhões de pessoas diagnosticadas com autismo no Brasil. Agência IBGE de Notícias. <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/43464-censo-2022-identifica-2-4-milhoes-de-pessoas-diagnosticadas-com-autismo-no-brasi>

Grupo de Habilidades Sociais



Um espaço estruturado para crianças e adolescentes autistas desenvolverem habilidades sociais e emocionais — reconhecer emoções, conversar, cooperar e lidar com frustrações — com apoio baseado em evidências.

- Reconhecimento e regulação de emoções
- Interação, conversação e cooperação em grupo
- Acompanhamento orientado pela Análise do Comportamento