



Universidade de Brasília

Instituto de Psicologia

Departamento de Processos Psicológicos Básicos

Programa de Pós-Graduação em Ciências do Comportamento

Dissertação de Mestrado

**COMPARAÇÃO NUMÉRICA SIMBÓLICA PRECOCE COMO UM
MARCADOR DE DESENVOLVIMENTO MATEMÁTICO: NORMAS
BRASILEIRAS E EVIDÊNCIAS LONGITUDINAIS DO NUMERACY
SCREENER**

Bruna Mariz Silva Egreja

Brasília, fevereiro de 2026



Universidade de Brasília

Instituto de Psicologia

Departamento de Processos Psicológicos Básicos

Programa de Pós-Graduação em Ciências do Comportamento

**COMPARAÇÃO NUMÉRICA SIMBÓLICA PRECOCE COMO UM
MARCADOR DE DESENVOLVIMENTO MATEMÁTICO: NORMAS
BRASILEIRAS E EVIDÊNCIAS LONGITUDINAIS DO NUMERACY
SCREENER**

Bruna Mariz Silva Egreja

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciências do Comportamento, do Instituto de Psicologia da Universidade de Brasília, como parte dos requisitos para obtenção do título de Mestre em Ciências do Comportamento (Área de concentração: Cognição e Neurociências do Comportamento).

Orientador: Prof. Dr. Ricardo José de Moura.

Brasília, fevereiro de 2026.

Comissão Examinadora

Prof. Dr. Ricardo José de Moura (Presidente) Programa de Pós-Graduação em Ciências
do Comportamento Universidade de Brasília - UnB

Prof. Dr.^a Vitor Geraldi Haase (Membro externo) Programa de Pós-Graduação em
Cognição e Comportamento Universidade Federal do Minas Gerais - UFMG

Prof. Dr. Rui de Moraes Junior (Membro interno) Programa de Pós-Graduação em
Ciências do Comportamento Universidade de Brasília - UnB

Prof.^a Luciano Grudtner Buratto (Membro suplente) Programa de Pós-Graduação em
Ciências do Comportamento Universidade de Brasília – UnB

Este trabalho teve apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), por meio de concessão de bolsa de Mestrado

DEDICATÓRIA

*À Dona Sônia e Dona Iza, minhas avós,
que me ensinaram, antes de qualquer teoria, sobre amor incondicional, cuidado e
coragem. Ter o colo de vocês como primeiro abrigo foi a maior sorte que eu dei na
vida. Tudo que eu escrevo é continuação do que vocês escreveram antes.*

AGRADECIMENTOS

Agradeço, primeiramente, à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo financiamento que viabilizou esta pesquisa. O apoio concedido foi fundamental para que este trabalho pudesse ser desenvolvido com a dedicação e o rigor que a ciência exige.

À Universidade de Brasília, que tem sido minha casa há sete anos, expresso profunda gratidão. Um espaço onde encontrei formação técnica, mas também identidade, pertencimento e propósito, aprendendo ciência como compromisso social.

Aos meus professores, agradeço pela generosidade intelectual, pela exigência e pelo exemplo. Cada aula, cada debate e cada orientação contribuíram para a construção do olhar que hoje sustenta esta pesquisa. Dos meus professores, o primeiro foi o meu pai, a quem devo o despertar da curiosidade sobre o mundo e o apreço pelo conhecimento científico. Agradeço imensamente à minha mãe, que é presença constante e, assim como ele, sustenta com muito amor e incentivo cada etapa do meu percurso. A ambos devo as bases que me trouxeram até aqui.

Ao professor Mauro, pela acolhida carinhosa e pelos aprendizados no Laboratório de Psicologia Evolucionista, onde aprendi, na prática, o que significa fazer ciência com seriedade e responsabilidade. Foi nesse espaço que compreendi que ser cientista é também aprender a perguntar melhor, a duvidar com método e a construir conhecimento de forma coletiva.

Ao meu orientador, Ricardo, agradeço por ter sido parceiro presente e disponível, pelo humor afiado que tornou o processo mais leve e pelo conhecimento compartilhado com generosidade. Sua orientação foi mais do que técnica; foi diálogo, construção conjunta e confiança. Em meio à energia caótica que coincidentemente nos caracteriza, encontramos método, rigor e direção. Não teria conseguido se não fosse com você.

À minha irmã, que é minha melhor amiga e me apoia incondicionalmente, participando de cada passo de todas as minhas conquistas, deixo o meu mais amoroso agradecimento. Aos amigos que dividiram a jornada e as angústias deste percurso, minha gratidão pelo apoio constante, pelas conversas que aliviaram os dias difíceis e pela celebração sincera de cada pequena conquista.

A todos os alunos de pesquisa e estágio que, de alguma forma, contribuíram com muita dedicação para que esta pesquisa se tornasse realidade, deixo meu mais sincero agradecimento. Este trabalho é resultado de uma trajetória construída coletivamente.

ÍNDICE

Resumo	10
Abstract (Resumo em inglês)	11
Apresentação	12
Introdução	14
Método	18
Resultados	21
Discussão	34
Referências	40
Apêndice	46
Material Suplementar	50

Lista de Tabelas

Tabela 1. Dados descritivos nos tempos T1 e T2	22
Tabela 2. Correlações entre idade, inteligência, NS e medidas aritméticas	24
Tabela 3. Modelos de regressão linear para desempenho em adição cronometrada e TDE-II com diferentes operacionalizações da comparação numérica simbólica	30
Tabela 4. Modelos de regressão linear para desempenho em adição cronometrada e TDE-II com diferentes operacionalizações da comparação numérica não simbólica	32

Lista de Figuras

Figura 1. Histogramas (A e B) e gráfico violino (C) da distribuição de frequência dos escores simbólicos e não simbólicos	23
Figura 2. Gráfico de dispersão da associação entre adição e comparação simbólica (A) e não simbólica (B)	28

Resumo

Este estudo estabeleceu dados normativos brasileiros e examinou evidências de validade para o Numeracy Screener (NS) como um instrumento breve para avaliar o julgamento precoce de magnitude numérica na educação infantil. Um grupo de crianças foi avaliado na pré-escola e acompanhado um ano depois no ensino fundamental, possibilitando análises longitudinais das habilidades numéricas iniciais e do desempenho matemático subsequente. As distribuições de escores apresentaram variabilidade adequada para fins de rastreamento. As análises correlacionais indicaram associações consistentes entre o desempenho no NS e desfechos em aritmética, particularmente para a tarefa de comparação simbólica. As análises da curva ROC demonstraram boa capacidade discriminativa para identificar crianças com baixo desempenho em aritmética. Modelos de regressão hierárquica mostraram que a comparação numérica simbólica explicou variância única no desempenho aritmético posterior, mesmo após o controle de idade e inteligência, ao passo que a comparação não simbólica não acrescentou poder explicativo significativo. Análises adicionais indicaram que medidas contínuas de operacionalização do desempenho na tarefa – Escore Total e Índice de Eficiência – foram preditores mais robustos do desempenho matemático posterior do que medidas dicotômicas – performar acima do Nível de Chance e do ponto de corte da curva ROC. Esses achados destacam o papel do processamento numérico simbólico precoce como um marcador cognitivo da competência aritmética posterior. A versão brasileira do NS demonstrou propriedades psicométricas satisfatórias e mostrou-se um instrumento de rastreamento livre e gratuito para a identificação precoce de crianças em risco para dificuldades matemáticas, com implicações importantes para o monitoramento educacional e a intervenção precoce.

Palavras-chave: Numeracia Infantil; Comparação Numérica; Aritmética; Educação Infantil; Estudo Longitudinal; Psicometria.

Abstract

This study established Brazilian normative data and examined validity evidence for the Numeracy Screener (NS) as a brief instrument to assess early numerical magnitude judgment in early childhood education. A group of children was assessed in preschool and followed up one year later in elementary school, enabling longitudinal analyses of early numerical skills and subsequent mathematical performance. Score distributions presented adequate variability for screening purposes. Correlational analyses indicated consistent associations between NS performance and arithmetic outcomes, particularly for the symbolic comparison task. ROC curve analyses demonstrated good discriminative capacity to identify children with low arithmetic performance. Hierarchical regression models showed that symbolic numerical comparison explained unique variance in later arithmetic performance, even after controlling for age and intelligence, whereas non-symbolic comparison did not add significant explanatory power. Additional analyses indicated that continuous measures operationalizing task performance—Total Score and Efficiency Index—were more robust predictors of later mathematical performance than dichotomous measures—performing above Chance Level and the ROC curve cut-off point. These findings highlight the role of early symbolic numerical processing as a cognitive marker of later arithmetic competence. The Brazilian version of the NS demonstrated satisfactory psychometric properties and proved to be a free and open screening instrument for the early identification of children at risk for mathematical difficulties, with important implications for educational monitoring and early intervention.

Key words: Children Numeracy; Number Comparison; Arithmetic; Preschool; Longitudinal Study; Psychometrics.

Apresentação

O manuscrito apresenta um estudo longitudinal, submetido na revista *Early Childhood Research Quarterly*, que teve como objetivo estabelecer dados normativos brasileiros e investigar evidências de validade do Numeracy Screener (NS) como instrumento breve para avaliação do julgamento de magnitude numérica na educação infantil. A pesquisa foi conduzida com uma coorte de crianças avaliadas na pré-escola e acompanhadas um ano depois, no início do ensino fundamental, permitindo a investigação prospectiva da relação entre habilidades numéricas iniciais e desempenho matemático subsequente. O estudo contemplou análises psicométricas voltadas à obtenção de evidências de validade baseadas na estrutura interna, nas relações com variáveis externas e na capacidade discriminativa do instrumento, bem como a elaboração de parâmetros normativos para a população brasileira.

O material suplementar desta dissertação consiste no link para acesso e download do livreto oficial do Numeracy Screener, disponibilizado no site oficial do instrumento. As normas da APA (7ª edição) foram utilizadas para a elaboração do manuscrito.

Comparação simbólica precoce de números como marcador do desenvolvimento matemático: normas brasileiras e evidência longitudinal a partir do Numeracy Screener

Resumo

Este estudo estabeleceu dados normativos brasileiros e examinou evidências de validade para o Numeracy Screener (NS) como um instrumento breve para avaliar o julgamento precoce de magnitude numérica na educação infantil. Um grupo de crianças foi avaliado na pré-escola e acompanhado um ano depois no ensino fundamental, possibilitando análises longitudinais das habilidades numéricas iniciais e do desempenho matemático subsequente. As distribuições de escores apresentaram variabilidade adequada para fins de rastreamento. As análises correlacionais indicaram associações consistentes entre o desempenho no NS e desfechos em aritmética, particularmente para a tarefa de comparação simbólica. As análises da curva ROC demonstraram boa capacidade discriminativa para identificar crianças com baixo desempenho em aritmética. Modelos de regressão hierárquica mostraram que a comparação numérica simbólica explicou variância única no desempenho aritmético posterior, mesmo após o controle de idade e inteligência, ao passo que a comparação não-simbólica não acrescentou poder explicativo significativo. Análises adicionais indicaram que medidas contínuas de operacionalização do desempenho na tarefa – Escore Total e Índice de Eficiência – foram preditores mais robustos do desempenho matemático posterior do que medidas dicotômicas – performar acima do Nível de Chance e do ponto de corte da curva ROC. Esses achados destacam o papel do processamento numérico simbólico precoce como um marcador cognitivo da competência aritmética posterior. A versão brasileira do NS

demonstrou propriedades psicométricas satisfatórias e mostrou-se um instrumento de rastreamento livre e gratuito para a identificação precoce de crianças em risco para dificuldades matemáticas, com implicações importantes para o monitoramento educacional e a intervenção precoce.

Palavras-chave: Numeracia Infantil; Comparação Numérica; Aritmética; Educação Infantil; Estudo Longitudinal; Psicometria.

Introdução

O desenvolvimento de habilidades numéricas básicas na infância é amplamente reconhecido como um pilar do desenvolvimento cognitivo e educacional das crianças, sendo considerado central para a aquisição futura de competências aritméticas mais complexas (Mazzocco et al., 2011; Dehaene, 2011; Purpura & Lonigan, 2013; Geary et al., 2018). Assim, a detecção precoce de dificuldades na aquisição dessas habilidades é crucial, uma vez que o fracasso em matemática tende a ser cumulativo e progressivo; a ausência de domínio de um conceito fundamental frequentemente compromete a aquisição bem-sucedida de habilidades mais avançadas (Caron, 2007; Geary, 2011). Evidências sugerem que as habilidades matemáticas iniciais constituem os preditores mais robustos do sucesso acadêmico subsequente, superando fatores como leitura e atenção (Duncan et al., 2007). Além disso, tais habilidades se associam a diversos desfechos relevantes ao longo da vida, incluindo inserção no mercado de trabalho, saúde mental e física, empregabilidade e produtividade na vida adulta (Parsons & Bynner, 2005; Romano et al., 2010; OECD, 2013).

Entre as habilidades numéricas básicas observadas no início da escolarização, destacam-se as representações de magnitude numérica. Essas incluem a capacidade de comparar conjuntos, contar e enumerar elementos, estimar quantidades, bem como aprender a sequência numérica simbólica e compreender que cada símbolo representa um valor (Piazza, 2010). Estudos como os de Sarnecka e Carey (2008) e LeFevre et al. (2010) identificam a habilidade de estimar e comparar magnitudes — tanto em formatos simbólicos (utilizando dígitos da lista numérica) quanto não-simbólicos (utilizando conjuntos de objetos ou pontos) — como um componente central do desenvolvimento numérico inicial. Essa habilidade de comparação, particularmente em sua forma simbólica, também é destacada como um importante preditor transversal e longitudinal do sucesso aritmético (Bartelet et al., 2014; Xenidou-Dervou et al., 2017), uma vez que envolve a compreensão do significado de cada símbolo numérico em termos de sua magnitude.

A capacidade de representar magnitudes emerge precocemente no desenvolvimento humano e está ancorada principalmente em dois sistemas cognitivos: o Sistema Numérico Aproximado (Approximate Number System, ANS) e o Sistema de Rastreamento de Objetos (Object Tracking System, OTS). O ANS permite a representação e comparação aproximada e contínua de quantidades, enquanto o OTS possibilita o rastreamento preciso de pequenos conjuntos, operando com representações exatas, porém limitadas numericamente. Ambos os sistemas são observados em bebês e em animais não humanos (Feigenson et al., 2004; Cantlon, 2012).

Isso sugere que a capacidade de estimar e comparar magnitudes constituiu uma característica essencial ao longo da evolução da espécie humana, tendo sido selecionada por possibilitar a sobrevivência diferencial de nossos ancestrais, por exemplo, ao

estimar qual árvore possuía mais frutos ou qual grupo apresentava maior número de membros (Dehaene, 2011).

Estudos demonstram que a precisão das representações numéricas não-simbólicas, frequentemente medida por tarefas de comparação de conjuntos, está consistentemente associada ao desempenho aritmético, mesmo após o controle de variáveis cognitivas de domínio geral, como memória de trabalho e inteligência fluida (Halberda et al., 2008; Mazzocco et al., 2011). Por outro lado, alguns achados sugerem que a capacidade de controle inibitório em tarefas dessa natureza constitui um preditor mais forte da habilidade aritmética do que o próprio componente de comparação não-simbólica (Gilmore et al., 2013). Além disso, evidências apontam para a importância da comparação não-simbólica como ponto de partida para a aquisição das habilidades de comparação simbólica (Feigenson et al., 2013).

O desempenho em tarefas de comparação de magnitude simbólica (por exemplo, julgar qual de dois numerais arábicos é maior), por sua vez, parece constituir um preditor ainda mais crítico do sucesso aritmético, explicando uma parcela significativa da variância no desempenho em matemática entre crianças do ensino fundamental (Holloway & Ansari, 2009; De Smedt et al., 2009). Isso evidencia que o progresso no mapeamento entre símbolos numéricos e suas quantidades correspondentes é uma habilidade essencial para a aprendizagem formal da matemática (Rousselle & Noël, 2007; Lyons et al., 2014). Dessa forma, a identificação de possíveis déficits na aquisição dessas habilidades básicas nos primeiros anos da educação formal, com vistas à orientação de intervenções, é fundamental para o pleno desenvolvimento de competências matemáticas mais complexas ao longo do tempo.

Neste contexto, Nosworthy et al. (2013) desenvolveram o Numeracy Screener (NS), um instrumento breve e de baixo custo, adequado para contextos escolares ou

clínicos, destinado à mensuração das habilidades de comparação de magnitude simbólica e não-simbólica, inicialmente aplicado a crianças entre 6 e 9 anos de idade. O instrumento, em formato papel e lápis e com tempo controlado, demonstrou sensibilidade a variações desenvolvimentais e capacidade preditiva do desempenho aritmético. Os resultados indicaram que a acurácia na comparação de números, tanto simbólicos quanto não simbólicos, correlacionou-se com diferenças individuais no desempenho aritmético. Essa relação mostrou-se ainda mais forte para a comparação de numerais arábicos, que explicou variância única no desempenho mesmo após o controle de variáveis como leitura, QI e memória de trabalho. Assim, a tarefa configura-se como uma ferramenta promissora e acessível para triagem universal do desempenho matemático em idades precoces.

O NS avalia o julgamento de magnitude numérica em duas partes: uma simbólica e outra não-simbólica. Em cada parte, as crianças devem selecionar, dentre pares apresentados, o número ou conjunto de pontos de maior magnitude dentro de um tempo determinado. O desempenho é mensurado pelo número de respostas corretas. Esse instrumento tem sido aplicado em estudos com estudantes de diferentes nacionalidades, e um trabalho de meta-análise (Díaz-Simón et al., resultados não publicados) evidenciou correlações significativas com o desempenho em matemática em diferentes faixas etárias e contextos culturais. O NS surge em um contexto no qual há escassez de instrumentos padronizados e validados na literatura brasileira para a avaliação de habilidades numéricas básicas, especialmente em crianças mais jovens, como aquelas na educação infantil ou nos anos iniciais do ensino fundamental. Essa lacuna se insere em um cenário de desigualdades educacionais e altas taxas de evasão e repetência associadas a dificuldades precoces em matemática (Silva & Ciasca, 2022; Rodrigues et al., 2023).

Consequentemente, esforços recentes têm buscado desenvolver e validar instrumentos capazes de avaliar competências numéricas iniciais de maneira sensível e ecologicamente válida. Entre esses instrumentos destacam-se o PRONUMERO (Gomides et al., 2022), o subteste de Aritmética do Teste de Desempenho Escolar – TDE II (Stein et al., 2019) e a Bateria de Avaliação de Competências Matemáticas – BAC-MAT (Rodrigues & Ciasca, 2020). No entanto, alguns desses instrumentos requerem treinamento técnico especializado ou recursos nem sempre disponíveis em contextos escolares com restrições, o que frequentemente limita seu uso a contextos clínicos. Nesse sentido, a disponibilidade de instrumentos breves, confiáveis, de fácil aplicação e com normas locais representa uma contribuição significativa para a triagem em larga escala do desempenho matemático precoce.

Considerando esse cenário, o NS apresenta-se como alinhado a critérios de viabilidade e custo-efetividade, podendo ser utilizado como ferramenta de triagem tanto por professores quanto por profissionais da educação e psicólogos. Diante desse panorama, o presente estudo teve como objetivo estabelecer normas brasileiras para o NS e reunir evidências de validade a partir de uma amostra de crianças matriculadas na rede pública de ensino regular, inicialmente no segundo ano da educação infantil. O desempenho das crianças em ambas as partes do NS foi correlacionado com diferentes medidas de desempenho aritmético obtidas no mesmo ano e no ano subsequente (primeiro ano do ensino fundamental). Os dados obtidos visam contribuir para a ampliação do repertório de instrumentos disponíveis para a avaliação e o monitoramento do desenvolvimento numérico infantil no Brasil.

Método

Procedimentos de coleta de dados

A coleta de dados foi conduzida de forma longitudinal, em duas ondas. A primeira onda ocorreu em 2024 e incluiu uma amostra de 159 crianças, com idade média de 5 anos e 9 meses, matriculadas no último ano da educação infantil em quatro escolas da rede pública de ensino de Brasília, Brasil. A segunda onda de coleta ocorreu no ano seguinte. Nesse momento, todas as crianças participantes haviam transitado de suas escolas de educação infantil para diferentes escolas de ensino fundamental, tanto públicas quanto privadas, o que contribuiu para a perda amostral (tratada posteriormente nas análises). A equipe de pesquisa entrou em contato com os responsáveis por todas as crianças que participaram da primeira onda. Alguns responsáveis não responderam, mudaram de cidade ou optaram por não continuar no estudo, resultando em uma amostra longitudinal final de 61 crianças.

Na primeira onda, toda a coleta de dados foi realizada nas próprias escolas em que as crianças estavam matriculadas. Já na segunda onda, um pequeno número de crianças ($n = 4$) foi avaliado no laboratório de pesquisa da universidade, devido à impossibilidade de acesso às escolas. Em ambos os contextos, as avaliações foram conduzidas individualmente. O estudo incluiu exclusivamente crianças sem histórico relatado de diagnóstico de transtornos do neurodesenvolvimento ou de aprendizagem, conforme informações fornecidas pelos responsáveis.

Instrumentos

Numeracy Screener (NS; onda 1): O Numeracy Screener (NS) consiste em um livreto de fácil impressão contendo uma tarefa de comparação numérica dividida em duas partes: simbólica e não-simbólica. Na parte simbólica, são apresentados 56 pares de numerais indo-arábicos variando de 1 a 9. Na parte não-simbólica, os pares consistem em conjuntos de pontos (também variando de 1 a 9), com diferentes

tamanhos. Em ambas as partes, a criança é instruída a riscar o número ou o conjunto de pontos de maior magnitude em cada par apresentado. As duas partes apresentam estímulos distintos e a criança dispõe de um tempo previamente definido para completar cada uma delas. Para crianças da educação infantil, o limite de tempo é de 2 minutos.

Antes de cada parte, a criança recebe instruções do examinador e realiza uma página de treino, composta por 9 itens de prática, com feedback fornecido após cada resposta, a fim de esclarecer dúvidas. Uma vez confirmado que a criança compreendeu a tarefa, ela passa para os itens de teste na página seguinte, momento em que o cronômetro é iniciado. O escore é calculado pela soma dos itens corretamente respondidos dentro do limite de tempo.

Tarefa de adição (onda 1): O desempenho aritmético das crianças na primeira onda foi avaliado por meio de uma tarefa de adição com números de um dígito. A tarefa foi iniciada com um item de exemplo ($1 + 1$), durante o qual o examinador verificou se a criança compreendia o conceito de adição, oferecendo explicações com o uso de dedos e objetos em caso de erro. Em seguida, foram apresentados 17 itens de teste: $3 + 1$; $1 + 2$; $2 + 3$; $2 + 2$; $4 + 1$; $4 + 2$; $3 + 4$; $3 + 3$; $6 + 1$; $4 + 4$; $7 + 2$; $2 + 6$; $4 + 5$; $7 + 4$; $2 + 8$; $6 + 6$; $8 + 5$. Os itens foram apresentados oralmente pelo examinador, nessa ordem. As crianças não podiam utilizar papel e lápis, mas eram informadas de que poderiam usar os dedos para contar, caso desejassem. O escore corresponde ao número total de respostas corretas. Não houve limite de tempo, e nenhum feedback de desempenho foi fornecido durante a aplicação.

Matrizes Progressivas Coloridas de Raven (CPM) (onda 1): Como medida de inteligência na primeira onda, foi utilizada a versão brasileira do teste neuropsicológico de inteligência fluida Matrizes Progressivas Coloridas de Raven (CPM) (Angelini et al., 1999), aplicada de acordo com as normas descritas no manual. A adaptação brasileira do

instrumento apresenta consistência interna satisfatória (coeficiente de divisão ao meio = .85) (Paula et al., 2018). O teste gera um escore de inteligência com base na capacidade de raciocínio por matrizes e é composto por três partes, cada uma contendo 12 itens de dificuldade crescente. A criança deve apontar para a alternativa visual (entre seis opções) que melhor completa um padrão apresentado no topo da página. Apenas uma criança da amostra apresentou desempenho abaixo do percentil 10 em relação à amostra normativa do teste. No entanto, esse não foi considerado critério de exclusão, uma vez que a inteligência foi utilizada neste estudo apenas como variável de controle.

Desempenho aritmético (onda 2): Na segunda onda de coleta, as crianças, então matriculadas no primeiro ano do ensino fundamental, responderam ao subteste de Aritmética do Teste de Desempenho Escolar – Segunda Edição (TDE II) (Stein, 2019) e à bateria PRONUMERO (Gomides et al., 2022), como medidas de habilidade matemática. Ambos os instrumentos foram aplicados de acordo com as normas descritas em seus respectivos manuais. O subteste de Aritmética do TDE II avalia o desempenho da criança em cálculos aritméticos básicos, incluindo adição, subtração, multiplicação e divisão, além de questões mais conceituais, como contagem, comparação numérica verbal e representação em linha numérica, apresentadas em ordem crescente de complexidade. A tarefa requer a aplicação de procedimentos formais de cálculo e conhecimentos matemáticos adquiridos no contexto escolar.

Adição cronometrada (onda 2): Essa tarefa foi utilizada para complementar a avaliação da habilidade aritmética, em razão da diferença de natureza entre as medidas: enquanto uma enfatiza aspectos mais conceituais, a outra avalia a eficiência de processamento sob restrição de tempo. Foi utilizada a Bateria de Avaliação do Processamento Numérico e Cálculo (PRONUMERO), especificamente sua primeira parte, aplicada conforme as diretrizes do manual. A parte administrada consiste em 12

operações de adição. Para as análises, foi calculado um escore baseado no número de respostas corretas, uma vez que não há normas disponíveis correspondentes ao ano escolar das crianças participantes deste estudo. Além das tarefas mencionadas, os protocolos de avaliação incluíram outras tarefas que não foram consideradas no escopo do presente estudo.

Procedimentos éticos

A pesquisa foi aprovada pelo Comitê de Ética em Pesquisa (CEP-UnB) da Universidade de Brasília (protocolo 02345418.6.0000.5540) e contou com autorização da Secretaria de Educação e da direção de cada uma das escolas participantes. Termos de Consentimento Livre e Esclarecido foram fornecidos aos responsáveis pelas crianças nas quatro escolas, informando que se tratava de uma coleta de dados longitudinal. As crianças cujos responsáveis devolveram o termo assinado foram chamadas nominalmente durante o horário de aula e realizaram individualmente o protocolo de avaliação. A ordem de apresentação das tarefas, bem como a ordem das partes simbólica e não-simbólica do NS, foi randomizada entre as crianças, a fim de minimizar possíveis efeitos de ordem.

Resultados

A Tabela 1 apresenta estatísticas descritivas da amostra nas duas ondas do presente estudo. Inicialmente, são reportadas análises transversais (onda 1), seguidas pelas análises longitudinais (desfechos da onda 2).

Tabela 1

Dados descritivos das ondas 1 (O1) e 2 (O2).

	Média	DP	Mínimo	Máximo
O1 - Idade (meses)	68.97	3.82	60	78
O1 - CPM	18.94	4.64	7	30
O1 - NS - Simbólico	34.21	11.82	7	56
O1 - NS – Não-simbólico	39.03	7.23	22	54
O1 – Tarefa de Adição	8.70	5.25	0	17
O2 – Idade (meses)	69.70	3.82	62	78
O2 - PRONUMERO (Adição cronometrada)	5.33	3.49	0	12
O2 - TDE II (Aritmética)	12.87	3.37	6	20

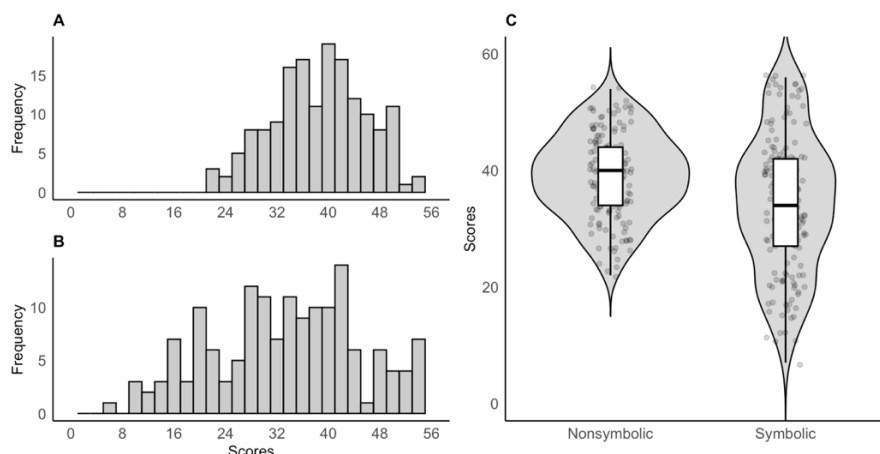
O1: Primeira onda de coleta (segundo ano da Educação Infantil).

O2: Segunda onda de coleta (Primeiro ano do Ensino Fundamental).

O desempenho no NS diferiu sistematicamente entre suas duas partes. As crianças obtiveram escores significativamente mais elevados na tarefa de comparação não-simbólica em comparação à tarefa de comparação simbólica, conforme indicado por um teste t para amostras pareadas, $t(158) = -6,38$, $p < .001$, com tamanho de efeito moderado ($d = -0,46$, IC 95% $[-0,62, -0,31]$). A inspeção das distribuições de escores (Apêndice 1; Figura 1) indicou que as distribuições eram monotônicas e não apresentavam anomalias evidentes.

Figura 1

Histogramas (A and B) e gráfico de violino (C) das distribuições de escores das partes simbólica e não-simbólica.



Considerando que crianças em idade pré-escolar variam quanto ao reconhecimento e à compreensão de números indo-arábicos, dois escores adicionais foram derivados do NS. Primeiramente, investigou-se se o desempenho das crianças excedia o nível de acaso. O desempenho ao nível do acaso foi estimado por meio de um modelo binomial, assumindo respostas dicotômicas (correta/incorreta) e uma probabilidade de acerto ao acaso de .50 para cada item. Como a tarefa era limitada por tempo (2 minutos), as crianças diferiram quanto ao número de itens respondidos. Assim, o limiar de desempenho acima do acaso foi calculado individualmente para cada criança, com base no número de itens respondidos (n_i), utilizando a distribuição binomial $B(n_i, .50)$. O desempenho acima do acaso foi definido como o menor número de respostas corretas que excedia o percentil 95 dessa distribuição. Esse procedimento assegurou que a probabilidade de observar tal desempenho por acaso fosse inferior a 5% para cada criança, permitindo identificar desempenhos estatisticamente incompatíveis com respostas aleatórias, ao mesmo tempo em que controlava o erro do tipo I em nível individual.

Na parte simbólica do NS, 24 crianças (14,9%) apresentaram desempenho abaixo do nível de acaso. Para fins de comparação, o mesmo procedimento foi aplicado à parte

não-simbólica, e nenhuma criança apresentou desempenho abaixo do acaso. Em segundo lugar, para avaliar a fluência com que as crianças executavam a tarefa, foi calculado o escore de eficiência inversa (*Inverse Efficiency Score*, IES), obtido pela divisão do número de respostas corretas pelo tempo necessário para completar a tarefa.

Associações transversais entre o desempenho em adição na pré-escola e o Numeracy Screener

Conforme apresentado na Tabela 2 e na Figura 2, o desempenho na tarefa de adição aplicada na educação infantil apresentou correlações de magnitude fraca a moderada com as partes simbólica e não-simbólica do NS.

Tabela 2

Correlações entre os escores nas variáveis idade, inteligência, Numeracy Screener e medidas de desempenho matemático.

	1	2	3	4	5	6	7	8	9	10
1. Idade	1									
2. CPM Raven	0.19*	1								
3. NS não-simbólico (total)	0.19*	0.22*	1							
4. NS não-simbólico (IES)	0.05	0.16*	-0.13	1						
5. NS simbólico (total)	0.30*	0.36*	0.55*	0.06	1					
6. NS simbólico (IES)	0.21*	0.32*	0.17*	0.39*	0.68*	1				
7. NS simbólico (chance)	-0.21	-0.27	-0.15	-0.40	-0.60	-0.89	1			
7. T1 Adição	0.36*	0.46*	0.22*	0.25*	0.51*	0.45*	-0.33*	1		
8. T2 Adição cronometrada	-0.02	0.32*	0.27*	0.07	0.40*	0.34*	-0.24	0.49*	1	

9. T2 TDEII	0.16	0.39*	0.39*	0.18	0.45*	0.41*	-0.23	0.61*	0.73*	1
		*	*		**	*		**	**	

A acurácia diagnóstica das duas partes do NS na identificação de crianças com baixo desempenho em adição foi avaliada por meio de análises de curvas ROC. No que se refere à tarefa de adição, o critério adotado para identificar desempenhos inferiores foi o percentil 25, consistente com práticas normativas em avaliação psicológica e educacional, nas quais escores situados no primeiro quartil da distribuição são considerados estatisticamente inferiores ao padrão de desempenho da amostra de referência (Anastasi & Urbina, 1997; AERA, APA & NCME, 2014). Esse critério não implica interpretação diagnóstica, mas funciona como marcador estatístico de desempenho discrepante em relação ao centro da distribuição. No presente estudo, a aplicação desse ponto de corte baseou-se em percentis derivados da distribuição da amostra (Apêndice 1), de modo que escores abaixo do percentil 25 foram classificados como indicativos de possível prejuízo no desempenho numérico para fins descritivos e normativos.

Para a parte simbólica, a área sob a curva indicou boa capacidade discriminativa na amostra total ($AUC = .83$). O ponto de corte ótimo estimado pelo índice de Youden foi de 29,5 acertos, resultando em sensibilidade de .77 e especificidade de .79. A aplicação desse ponto de corte resultou em elevado valor preditivo negativo ($NPV = .91$) e baixa razão de verossimilhança negativa ($LR- = .29$), indicando forte capacidade da parte simbólica em descartar crianças sem risco matemático em contextos de triagem. A análise estratificada revelou desempenho consistente ao longo do ano letivo, com $AUC = .82$ no primeiro semestre e $AUC = .80$ no segundo semestre. No primeiro semestre, o mesmo ponto de corte (29,5) esteve associado a sensibilidade de .84 e especificidade de .76 ($LR- = .21$). No segundo semestre, o ponto de corte ótimo deslocou-se para 33,5 acertos, com

sensibilidade de .86 e especificidade de .72 ($LR- = .20$), mantendo perfil favorável para triagem.

Em contraste, a parte não-simbólica apresentou desempenho discriminativo mais modesto. Na amostra total, $AUC = .65$, indicando capacidade discriminativa limitada. O ponto de corte ótimo foi de 34,5 acertos, com sensibilidade de .46 e especificidade de .79. Esse perfil esteve associado a valores preditivos positivos moderados ($PPV = .42$), valor preditivo negativo de .82 e uma razão de verossimilhança negativa menos informativa ($LR- = .68$), refletindo menor capacidade de exclusão de risco em comparação à parte simbólica. Nas análises estratificadas, o desempenho permaneceu modesto, com $AUC = .67$ no primeiro semestre e $AUC = .62$ no segundo semestre. No primeiro semestre, o ponto de corte (34,5) produziu sensibilidade e especificidade relativamente equilibradas (.50 e .77, respectivamente). Já no segundo semestre, o ponto de corte (44,5) maximizou a sensibilidade (1,00), mas à custa de baixa especificidade (.26), indicando maior taxa de falsos positivos e instabilidade no desempenho discriminativo. Em conjunto, os resultados indicam que a parte simbólica do NS apresenta desempenho discriminativo robusto e estável ao longo do ano letivo para identificação de risco matemático, enquanto a parte não-simbólica apresenta desempenho substancialmente inferior e menos consistente.

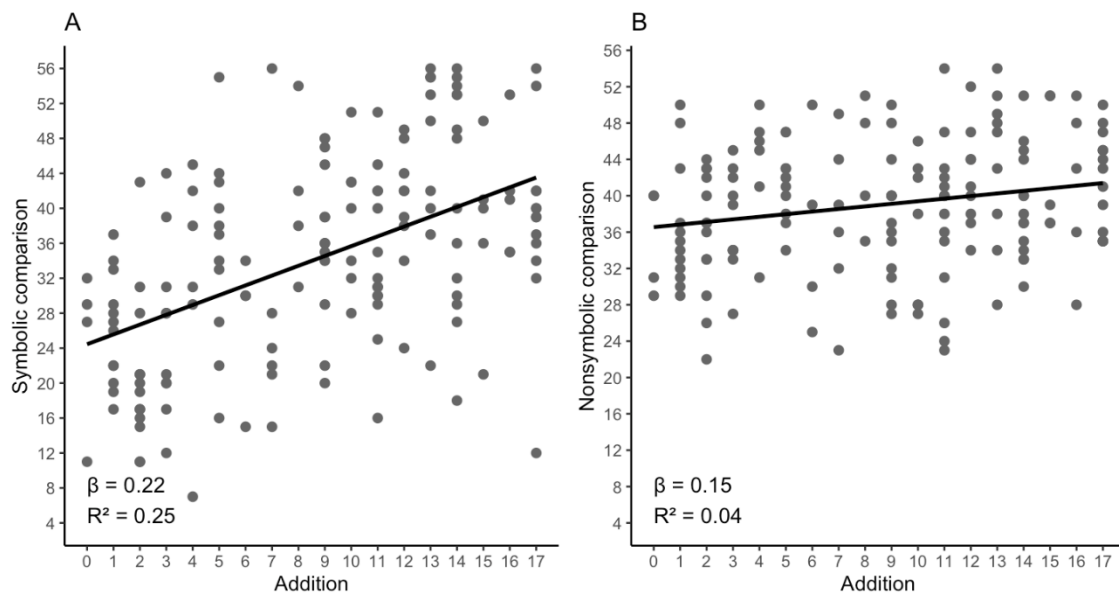
Para examinar de forma mais aprofundada a associação entre habilidades numéricas iniciais e o desempenho aritmético na educação infantil, testamos a relação entre a tarefa de adição e o *Numeracy Screener* utilizando escores contínuos. Essas análises foram conduzidas separadamente para os componentes de comparação simbólica e não simbólica, a fim de avaliar suas contribuições únicas e compartilhadas para o desempenho aritmético. Inicialmente, foi realizada uma regressão múltipla hierárquica para examinar se o desempenho em comparação simbólica explica variância única no

desempenho em adição, controlando por idade, semestre escolar e inteligência geral (CPM). No primeiro bloco, idade ($\beta = 0,27$, $p < 0,001$) e Raven ($\beta = 0,40$, $p < 0,001$) explicaram uma proporção significativa da variância em adição, $R^2 = .29$ (R^2 ajustado = .28), $F(3, 155) = 21,52$, $p < .001$. No segundo bloco, a inclusão do desempenho em comparação simbólica ($\beta = 0,34$, $p < .001$) resultou em melhora significativa no ajuste do modelo, $\Delta R^2 = .09$, $F(1, 154) = 22,43$, $p < .001$, elevando o R^2 final do modelo para .37 (R^2 ajustado = .36), indicando que a associação entre comparação simbólica e desempenho matemático permaneceu robusta mesmo após o controle de fatores desenvolvimentais (idade, semestre) e cognitivos (CPM).

O mesmo modelo de regressão múltipla hierárquica foi testado para o desempenho na comparação não simbólica. O primeiro bloco é o mesmo reportado no parágrafo anterior. No segundo bloco, a inclusão da comparação não simbólica ($\beta = 0,08$, $p = 0,228$) não resultou em melhora significativa no ajuste do modelo, $\Delta R^2 = .006$, $F(1, 154) = 1,27$, $p = .26$, elevando o R^2 final do modelo apenas marginalmente para .29 (R^2 ajustado = .27). Portanto, após o controle de fatores desenvolvimentais (idade, semestre) e cognitivos (CPM), não houve evidência de associação independente entre a comparação não simbólica e o desempenho matemático.

Figura 2

Diagrama de dispersão mostrando a associação entre o desempenho em adição e os escores de comparação simbólica (A) e não simbólica (B).



Associações longitudinais entre o desempenho no Numeracy Screener na pré-escola e o desempenho posterior em adição

A influência longitudinal do NS sobre o desempenho matemático das crianças foi analisada por meio do subteste de Aritmética do TDE-II e da tarefa de adição cronometrada da bateria PRONUMERO. Essas duas tarefas foram selecionadas porque a primeira avalia aspectos mais conceituais do conhecimento matemático, enquanto a segunda avalia eficiência, por se tratar de uma tarefa cronometrada. Foram observadas associações consistentes entre o NS e medidas subsequentes de desempenho matemático, particularmente para a comparação simbólica ($r = .45$ com o TDE II; $r = .40$ com a adição cronometrada), sugerindo continuidade entre habilidades numéricas iniciais avaliadas na pré-escola e o desempenho matemático formal observado no primeiro ano do ensino fundamental.

É importante destacar que a bateria PRONUMERO não possui normas referentes ao ano escolar das crianças da amostra; portanto, o desempenho não foi comparado com a amostra normativa do instrumento. No que diz respeito ao subteste de Aritmética do

TDE II, após a conversão dos escores totais das crianças em percentis de acordo com a amostra normativa do teste para o primeiro ano, nenhuma das crianças do estudo apresentou desempenho abaixo do percentil 25, o qual é indicado no manual como um sinal de dificuldade matemática.

O poder preditivo do NS administrado na educação infantil sobre o desempenho matemático no ensino fundamental foi investigado por meio de uma série de regressões hierárquicas que incluíram idade e inteligência no primeiro nível do modelo e o escore de aritmética no segundo. Diferentes formas de operacionalização do desempenho no NS foram testadas: escore total de acurácia, escore de eficiência inversa (IES), classificação com base no nível de acaso (apenas para a parte simbólica) e classificação com base no ponto de corte derivado da curva ROC apresentada anteriormente. A amostra na T2 incluiu 9 crianças (14%) que apresentaram desempenho abaixo do nível de acaso na parte simbólica do NS na T1, mantendo, assim, proporção semelhante.

Os resultados indicam que o desempenho na parte simbólica do *Numeracy Screener* esteve consistentemente associado ao desempenho matemático subsequente, mensurado por duas tarefas independentes. Nos modelos de regressão, medidas contínuas de desempenho — especificamente o escore total e o escore de eficiência inversa (IES) — foram preditores significativos tanto da tarefa de adição com tempo controlado quanto do desempenho no TDE (ver Tabela 3). Especificamente, o escore simbólico total previu significativamente o desempenho em adição com tempo controlado ($\beta = .38$, $p = .006$) e no TDE ($\beta = .35$, $p = .011$). De forma semelhante, o IES também previu significativamente o desempenho em adição com tempo controlado ($\beta = .32$, $p = .019$) e os escores no TDE ($\beta = .31$, $p = .017$). De modo geral, esses resultados indicam que ambas as medidas apresentaram poder preditivo comparável para o desempenho matemático posterior.

Em contraste, classificações dicotômicas de desempenho apresentaram valor preditivo mais fraco e menos consistente. A classificação baseada no ponto de corte derivado da curva ROC alcançou significância para a tarefa de adição com tempo controlado ($\beta = -.27$, $p = .044$) e foi marginalmente significativa para o TDE ($\beta = -.25$, $p = .050$). No entanto, a classificação baseada no desempenho em relação ao nível de acaso não previu significativamente nem a adição com tempo controlado ($p = .110$) nem os escores no TDE ($p = .301$).

As comparações entre modelos utilizando o Critério de Informação de Akaike (AIC) e o Critério de Informação Bayesiano (BIC) foram consistentes com esse padrão. Para ambas as variáveis dependentes, modelos que incluíam medidas contínuas de desempenho (escore total ou IES) apresentaram, de modo geral, melhores índices de ajuste do que modelos baseados em classificações dicotômicas. Em conjunto, esses achados sugerem que indicadores contínuos que capturam o nível e a eficiência do processamento numérico simbólico fornecem uma caracterização mais informativa da competência numérica inicial do que classificações baseadas em pontos de corte.

Tabela 3

Modelos de Regressão Linear da performance em Adição Cronometrada e no TDE-II usando formas alternativas de operacionalizar o desempenho na parte simbólica.

Variável	B	SE B	β	t	p
1) Variável dependente: Adição cronometrada					
Modelo 1.1: Escore total ($R^2 = .22$, R^2 ajust. = .18, AIC = 319.47, BIC = 330.03)					
(Intercepto)	11.09	7.58	—	1.46	0.149
NS simb. (total)	0.12	0.04	0.38	2.84	0.006**
Idade	-0.18	0.11	-0.2	-1.62	0.111
COM	0.15	0.1	0.19	1.47	0.147
Modelo 1.2: Ponto de corte ROC ($R^2 = .17$, R^2 ajust. = .13, AIC = 323.20, BIC = 333.76)					
(Intercepto)	11.08	7.91	—	1.4	0.167

NS simb. (ROC)	-2.15	1.04	-0.27	-2.06	0.044*
Idade	-0.13	0.11	-0.14	-1.13	0.266
COM	0.19	0.1	0.25	1.93	0.06
Modelo 1.3: IES ($R^2 = .19$, R^2 ajust. = .15, AIC = 321.63, BIC = 332.18)					
(Intercepto)	6.8	7.59	—	0.9	0.374
NS simb. (IES)	7.76	3.21	0.32	2.41	0.019*
Idade	-0.17	0.12	-0.19	-1.51	0.138
COM	0.19	0.1	0.25	1.96	0.055
Modelo 1.4: Ponto de corte por acaso ($R^2 = .14$, R^2 ajust. = .10, AIC = 324.80, BIC = 335.36)					
(Intercepto)	11.54	8.23	—	1.4	0.166
NS simb. (acaso)	-2.07	1.27	-0.21	-1.62	0.110
Idade	-0.15	0.12	-0.16	-1.25	0.217
COM	0.22	0.1	0.3	2.33	0.024*

2) Variável dependente: TDE II

Modelo 2.1: Escore total ($R^2 = .25$, R^2 ajust. = .21, AIC = 313.04, BIC = 323.59)

(Intercepto)	6.44	7.19	—	0.9	0.374
NS simb. (total)	0.1	0.04	0.35	2.65	0.011*
idade	-0.01	0.11	-0.01	-0.1	0.924
CPM	0.17	0.09	0.24	1.84	0.07

Modelo 2.2: Ponto de corte ROC ($R^2 = .21$, R^2 ajust. = .17, AIC = 315.95, BIC = 326.50)

(Intercepto)	6.57	7.46	—	0.88	0.382
NS simb. (ROC)	-1.97	0.98	-0.25	-2	0.050*
idade	0.04	0.11	0.04	0.35	0.725
CPM	0.21	0.09	0.29	2.26	0.028*

Modelo 2.3: IES ($R^2 = .24$, R^2 ajust. = .20, AIC = 310.54, BIC = 321.09)

(Intercepto)	2.63	7.13	—	0.37	0.714
NS simb. (IES)	7.37	3.01	0.31	2.44	0.017*
idade	-0.01	0.109	0.01	-0.06	0.950
CPM	0.21	0.09	0.28	2.27	0.027*

Modelo 2.4: Ponto de corte por acaso ($R^2 = .17$, R^2 ajust. = .13, AIC = 310.54, BIC = 321.09)

(Intercepto)	5.56	7.84	—	0.72	0.474
NS simb. (acaso)	-1.25	1.22	-0.13	-1.03	0.301
idade	0.04	0.11	0.04	0.31	0.756
CPM	0.25	0.09	0.35	2.73	0.009**

* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$

Em contraste, o desempenho na parte não simbólica não se mostrou um preditor longitudinal consistente do desempenho matemático (Tabela 4). As associações observadas foram mais fracas do que aquelas encontradas para a parte simbólica e não se mantiveram robustas nos modelos de regressão. O único efeito marginalmente

significativo observado nessa parte da tarefa foi uma associação com o TDE II ($B = 0,27$, $p = .049$).

Tabela 4

Modelos de regressão linear do desempenho em adição com tempo controlado e no TDE-II utilizando diferentes operacionalizações da comparação numérica não-simbólica

Variável	B	SE B	β	t	p
1) Variável de desfecho: Adição cronometrada					
Modelo 1.1: Escore total ($R^2 = .14$, R^2 ajust. = .09, AIC = 325.71, BIC = 336.26)					
(Intercepto)	7.14	7.85	—	0.91	0.367
NS não-simb. (total)	0.09	0.07	0.19	1.33	0.190
idade	-0.13	0.12	-0.14	-1.12	0.269
CPM	0.2	0.1	0.27	1.93	0.058
Modelo 1.2: Ponto de corte ROC ($R^2 = .11$, R^2 ajust. = .07, AIC = 327.27, BIC = 337.82)					
(Intercepto)	8.09	8.11	—	1	0.322
NS não-simb. (ROC)	-0.6	1.13	-0.07	-0.53	0.600
idade	-0.11	0.12	-0.12	-0.91	0.369
CPM	0.25	0.1	0.32	2.46	0.017*
Modelo 1.3: IES ($R^2 = .12$, R^2 ajust. = .07, AIC = 327.20, BIC = 337.75)					
(Intercepto)	4.83	8.95	—	0.54	0.591
NS não-simb. (IES)	3.26	5.56	0.07	0.59	0.560
idade	-0.11	0.12	-0.12	-0.91	0.365
CPM	0.26	0.1	0.34	2.64	0.011*
2) Variável de desfecho: TDE II					
Modelo 2.1: Escore total ($R^2 = .21$, R^2 ajust. = .17, AIC = 315.91, BIC = 326.47)					
(Intercepto)	2.89	7.24	—	0.4	0.691
NS não-simb. (total)	0.12	0.06	0.27	2.01	0.049*
idade	0.02	0.11	0.02	0.18	0.86
CPM	0.19	0.1	0.27	2.02	0.049*
Modelo 2.2: Ponto de corte ROC ($R^2 = .16$, R^2 ajust. = .12, AIC = 319.69, BIC = 330.24)					
(Intercepto)	3.99	7.62	—	0.52	0.603
NS não-simb. (ROC)	-0.66	1.06	-0.08	-0.62	0.538
idade	0.06	0.11	0.06	0.51	0.611
CPM	0.26	0.09	0.35	2.76	0.008**
Modelo 2.3: IES ($R^2 = .19$, R^2 ajust. = .15, AIC = 318.18, BIC = 328.73)					

(Intercepto)	-2.12	8.13	—	-0.25	0.800
NS não-simb. (IES)	6.98	5.17	0.16	1.35	0.182
idade	0.05	0.11	0.05	0.43	0.669
CPM	0.27	0.09	0.37	3	0.004**

Discussão

O presente estudo teve como objetivo estabelecer normas brasileiras para o *Numeracy Screener* (NS) e reunir evidências de validade para seu uso na avaliação de habilidades numéricas iniciais em crianças da educação infantil. De modo geral, os resultados indicam que o instrumento apresenta propriedades psicométricas adequadas e constitui uma medida sensível das diferenças individuais no processamento numérico precoce, especialmente no que se refere à comparação de magnitude simbólica.

Em consonância com estudos anteriores, o desempenho na comparação simbólica mostrou-se um preditor robusto do desempenho aritmético, tanto em análises transversais quanto longitudinais. Mesmo após o controle de variáveis relevantes, como idade e inteligência, a comparação simbólica manteve associação significativa com o desempenho em adição na educação infantil e com medidas subsequentes de desempenho matemático no ensino fundamental. Esse achado reforça a hipótese de que a compreensão do significado dos símbolos numéricos — isto é, a capacidade de mapear números para suas magnitudes correspondentes — constitui um componente central para o desenvolvimento da competência matemática (Rousselle & Noël, 2007; Lyons et al., 2014).

Por outro lado, a comparação não-simbólica apresentou associações mais fracas e inconsistentes com o desempenho aritmético. Embora essa habilidade esteja relacionada a sistemas cognitivos básicos e seja considerada um possível precursor do conhecimento simbólico (Feigenson et al., 2013), os resultados do presente estudo sugerem que sua contribuição para o desempenho matemático torna-se limitada quando consideradas

simultaneamente outras variáveis cognitivas e desenvolvimentais. Esse padrão é consistente com evidências que indicam que a relação entre habilidades não-simbólicas e matemática pode ser mediada por fatores como controle inibitório e acesso simbólico (Gilmore et al., 2013).

As análises de acurácia diagnóstica reforçam a superioridade da comparação simbólica como ferramenta de triagem. A parte simbólica do NS apresentou boa capacidade discriminativa para identificar crianças com baixo desempenho em adição, com valores elevados de sensibilidade, especificidade e valor preditivo negativo. Esse perfil indica que o instrumento é particularmente eficaz para descartar casos sem risco, característica desejável em contextos de triagem educacional. Em contraste, a parte não-simbólica apresentou desempenho mais modesto e instável, com menor capacidade de discriminação, o que limita sua utilidade como ferramenta isolada para identificação de risco.

Outro achado relevante refere-se à forma de operacionalização do desempenho no NS. Medidas contínuas, como o escore total e o escore de eficiência inversa (IES), mostraram-se preditores mais consistentes do desempenho matemático subsequente do que classificações dicotômicas baseadas em pontos de corte. Esse resultado sugere que a variabilidade individual no nível e na eficiência do processamento numérico fornece uma caracterização mais informativa das habilidades iniciais do que abordagens baseadas em limiares arbitrários.

Adicionalmente, a análise de desempenho em relação ao nível de acaso revelou que uma proporção significativa de crianças apresentou desempenho abaixo do esperado na tarefa simbólica, o que pode refletir dificuldades iniciais na compreensão de números indo-arábicos. No entanto, essa classificação não se mostrou um preditor significativo do

desempenho matemático posterior, indicando que medidas mais graduais de desempenho são mais sensíveis para capturar diferenças individuais relevantes.

No que se refere às análises longitudinais, os resultados evidenciam continuidade entre habilidades numéricas iniciais e desempenho matemático posterior, corroborando a literatura que destaca o papel da numeracia precoce como base para o aprendizado formal da matemática (Duncan et al., 2007). A associação consistente entre comparação simbólica e desempenho no TDE II e na tarefa de adição cronometrada sugere que o NS pode ser utilizado como um indicador precoce de risco para dificuldades matemáticas.

A análise de atrito longitudinal indicou um padrão levemente seletivo, com maior retenção de crianças com melhor desempenho inicial. No entanto, análises de sensibilidade demonstraram que esse viés não comprometeu as associações observadas, reforçando a robustez dos resultados.

Do ponto de vista aplicado, os achados deste estudo têm implicações importantes para o contexto educacional brasileiro. A escassez de instrumentos breves, de baixo custo e com normas locais para avaliação de habilidades numéricas iniciais representa um desafio para a identificação precoce de dificuldades em matemática. Nesse cenário, o NS destaca-se como uma ferramenta viável, de fácil aplicação e potencialmente útil para uso em larga escala por professores e profissionais da educação.

Apesar de suas contribuições, o estudo apresenta algumas limitações. A perda amostral ao longo do seguimento longitudinal reduz o tamanho da amostra na onda 2, o que pode impactar o poder estatístico das análises. Além disso, a ausência de normas específicas para algumas das medidas utilizadas, como a tarefa de adição do PRONUMERO, limita a interpretação dos resultados em termos normativos.

Futuras pesquisas devem buscar replicar esses achados em amostras maiores e mais diversificadas, bem como investigar a sensibilidade do NS a intervenções

educacionais. Adicionalmente, seria relevante examinar de forma mais detalhada os mecanismos cognitivos que mediam a relação entre comparação simbólica e desempenho matemático, incluindo o papel de funções executivas e habilidades linguísticas.

Em síntese, os resultados do presente estudo reforçam o papel central do processamento numérico simbólico no desenvolvimento matemático e indicam que o *Numeracy Screener* constitui uma ferramenta promissora para a avaliação e triagem precoce de habilidades numéricas no contexto brasileiro.

Referências

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing* (7th ed.). AERA.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Prentice Hall.
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., & Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven: Escala Especial*. Centro Editor de Testes e Pesquisas em Psicologia.
- Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D. (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology*, 117, 12–28.
<https://doi.org/10.1016/j.jecp.2013.08.010>
- Cantlon, J. F. (2012). Math, monkeys, and the developing brain. *Proceedings of the National Academy of Sciences*, 109(Supplement 1), 10725–10732.
<https://doi.org/10.1073/pnas.1201893109>

Caron, T. A. (2007). *Math is ruthlessly cumulative: The truth about fluency and the promise of differentiated practice*. Renaissance Learning.

Clements, D. H., & Sarama, J. (2011). *Learning and teaching early math: The road to fluency*. Routledge.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, 103(4), 469–479.
<https://doi.org/10.1016/j.jecp.2009.01.010>

Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Rev. ed.). Oxford University Press.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>

Elliott, L., Feigenson, L., Halberda, J., & Libertus, M. E. (2018). Early symbolic number knowledge and the approximate number system predict children's math abilities. *Developmental Psychology*, 54(7), 1207–1221.
<https://doi.org/10.1037/dev0000516>

Feigenson, L. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives*, 7(2), 74–79.
<https://doi.org/10.1111/cdep.12019>

- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8(7), 307–314. <https://doi.org/10.1016/j.tics.2004.05.002>
- Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 10-year longitudinal study. *Psychological Science*, 22(3), 304–308. <https://doi.org/10.1177/0956797610397236>
- Geary, D. C., van Marle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science*, 29(2), 191–205. <https://doi.org/10.1177/0956797617729817>
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., Simms, V., & Inglis, M. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS ONE*, 8(6), e67374. <https://doi.org/10.1371/journal.pone.0067374>
- Gomides, M., Lopes-Silva, J., Moura, R., Fumagalli de Salles, J., & Haase, V. (2022). *PRONUMERO: Bateria de avaliação do processamento numérico e do cálculo* [Assessment battery for numerical processing and calculation] (Manual). Vetor Editora.
- Haase, V. G., Júlio-Costa, A., Wood, G., & Fischer, F. (2014). Mathematical learning disabilities and difficulties in Brazil: An overview. *Research in Developmental Disabilities*, 35(10), 2531–2542. <https://doi.org/10.1016/j.ridd.2014.06.015>
- Halberda, J., Mazocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, 455, 665–668. <https://doi.org/10.1038/nature07246>

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, 103(1), 17–29.

<https://doi.org/10.1016/j.jecp.2008.04.001>

LeFevre, J. A., Skwarchuk, S. L., Smith-Chant, B. L., Fast, L., Kamawar, D., & Bisanz, J. (2010). Home numeracy experiences and children's math performance in the early school years. *Canadian Journal of Behavioural Science*, 42(2), 57–66.

<https://doi.org/10.1037/a0014532>

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, 17(5), 714–726. <https://doi.org/10.1111/desc.12152>

Marcelino, L., de Sousa, Ó., & Lopes, A. (2017). Predictive relation between early numerical competencies and mathematics achievement in first grade Portuguese children. *Frontiers in Psychology*, 8, 1103. <https://doi.org/10.3389/fpsyg.2017.01103>

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS ONE*, 6(9), e23749. <https://doi.org/10.1371/journal.pone.0023749>

Nosworthy, N., Bugden, S., Archibald, L., Evans, B., & Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS ONE*, 8(7), e67918. <https://doi.org/10.1371/journal.pone.0067918>

OECD. (2013). *OECD Skills Outlook 2013: First results from the survey of adult skills*. OECD Publishing.

Ouyang, X., Zou, X., & Zhang, X. (2024). General cognitive and numerical precursors of mathematical difficulty in kindergarten children at risk for dyscalculia. *Learning and Individual Differences, 114*, 102514.

<https://doi.org/10.1016/j.lindif.2024.102514>

Paula, J. J., Schlottfeldt, C. G. M. F., Malloy-Diniz, L. F., & Mizuta, G. A. A. (2018). Normas brasileiras do Teste das Matrizes Progressivas Coloridas de Raven para crianças de 5 a 11 anos. *Psicologia: Teoria e Pesquisa, 34*, e34410.

Parsons, S., & Bynner, J. (2005). *Does numeracy matter more?* National Research and Development Centre for Adult Literacy and Numeracy.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*(10), 542–551.

<https://doi.org/10.1016/j.tics.2010.07.001>

Purpura, D. J., & Lonigan, C. J. (2013). Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal, 50*(1), 178–209.

<https://doi.org/10.3102/0002831212465332>

Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The role of the cardinality principle in the development of symbolic numerical competencies. *Mathematical Thinking and Learning, 15*(2), 93–107.

<https://doi.org/10.1080/10986065.2013.768384>

Rodrigues, S. D., & Ciasca, S. M. (2020). Tradução e adaptação para o português (brasileiro) da bateria de aferição de competências matemáticas (BAC-MAT). *Revista Psicopedagogia, 37*(113), 168-182. <https://doi.org/10.5935/0103-8486.20200012>

8486.20200012

Rodrigues, L. R., Santos, A. A. A., & Haase, V. G. (2023). Early numeracy and school performance: Evidence from a Brazilian sample. *Psicologia: Reflexão e Crítica*, 36(1), 9. <https://doi.org/10.1186/s41155-023-00254-9>

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology*, 46(5), 995–1007.
<https://doi.org/10.1037/a0018880>

Rousselle, L., & Noël, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs nonsymbolic number magnitude processing. *Cognition*, 102(3), 361–395.
<https://doi.org/10.1016/j.cognition.2006.01.005>

Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, 108(3), 662–674.
<https://doi.org/10.1016/j.cognition.2008.05.007>

Silva, C. B., & Ciasca, S. M. (2022). Mathematics learning difficulties: Identification and intervention in Brazilian public schools. *Revista Brasileira de Educação Especial*, 28, 137–152. <https://doi.org/10.1590/S1413-653828222001>

Spaepen, E., Gunderson, E. A., Gibson, D. J., & Levine, S. C. (2018). Meaning before order: Cardinal principle knowledge predicts improvement in understanding the successor principle and exact ordering. *Cognition*, 180, 265–274.
<https://doi.org/10.1016/j.cognition.2018.07.008>

Stein, L. M., Giacomoni, C. H., & Fonseca, R. P. (2019). *Teste de Desempenho Escolar – 2ª Edição (TDE II): Manual para aplicação e interpretação*. Hogrefe.

Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., & van Lieshout, E. C. D. M. (2017). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learning and Instruction*, 50, 1–13. <https://doi.org/10.1016/j.learninstruc.2016.11.001>

Apêndice 1 - Score Frequencies and Percentiles for the Symbolic and nonsymbolic Comparison Parts of the Task

Score	Symbolic part			nonsymbolic part		
	Simple frequency	Cumulative frequency	Percentile	Simple frequency	Cumulative frequency	Percentile
0	0	0	0	0	0	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	0	0
5	0	0	0	0	0	0
6	0	0	0	0	0	0
7	0	0	0	0	0	0

8	1	1	1	0	0	0
9	1	2	2	0	0	0
10	1	3	3	0	0	0
11	4	7	8	0	0	0
12	1	8	9	0	0	0
13	1	9	10	0	0	0
14	1	10	11	0	0	0
15	1	11	12	0	0	0
16	3	14	16	0	0	0
17	0	14	16	0	0	0
18	1	15	17	0	0	0
19	1	16	18	0	0	0
20	4	20	22	0	0	0
21	2	22	24	0	0	0
22	4	26	29	1	1	1
23	0	26	29	2	3	2

24	0	26	29	1	4	2
25	0	26	29	1	5	3
26	1	27	30	2	7	4
27	0	27	30	3	10	6
28	0	27	30	5	15	9
29	2	29	32	3	18	11
30	0	29	32	4	22	13
31	0	29	32	5	27	16
32	3	32	36	3	30	18
33	2	34	38	7	37	23
34	2	36	40	8	45	27
35	3	39	43	9	54	33
36	5	44	49	8	62	38
37	3	47	52	10	72	44
38	4	51	57	6	78	48
39	5	56	62	5	83	51

40	1	57	63	14	97	59
41	1	58	64	5	102	62
42	6	64	71	6	108	66
43	2	66	73	11	119	73
44	3	69	77	7	126	77
45	4	73	81	5	131	80
46	1	74	82	3	134	82
47	3	77	86	7	141	86
48	2	79	88	6	147	90
49	2	81	90	3	150	91
50	0	81	90	5	155	95
51	1	82	91	6	161	98
52	0	82	91	1	162	99
53	1	83	92	0	162	99
54	1	84	93	2	164	100
55	1	85	94	0	164	100

56	5	90	100	0	164	100
----	---	----	-----	---	-----	-----

Material Suplementar

O Link a seguir direciona para o site oficial para download e instruções de impressão dos livretos do Numeracy Screener: <https://www.numeracyscreener.org/grade-k-diy-instructions.html>