

SPECIAL ARTICLE OPEN



Development and pilot testing of the number farm: a numeracy screener for preschool-age children

John S. Hutton¹✉, Priya Singh², Sarah Reel³ and Richard F. Ittenbach⁴

© The Author(s) 2026

IMPACT:

- Numeracy and math skills in the US, key predictors of academic and vocational success, are low and decreasing.
- Numeracy is not captured by general developmental screening, and no efficient, direct measure feasible for pediatric settings currently exists.
- In this pediatric primary care pilot, The Number Farm performed well in terms of psychometric validity, family acceptance, and relationships with home reading and math routines.
- The book-based approach of The Number Farm is family-centered, inexpensive, and scalable within existing literacy and early childhood education programs.
- These findings align with AAP recommendations to conduct developmental screening during clinic visits and promote school readiness.

Pediatric Research; <https://doi.org/10.1038/s41390-026-05061-w>

INTRODUCTION

Numeracy is a developmental process beginning in infancy, fueled by genetic and environmental factors, which involves practical concepts such as number sense, quantity, and basic operations (e.g., addition).¹ These form a foundation for math, which comprises an array of more formal, abstract subjects (e.g., algebra, geometry).¹ Numeracy and math are both vital in the technology-based economy, yet difficulties are prevalent and increasing.² Many children arrive at kindergarten (typically age 5–6 in the US) with low numeracy skills and consequent math readiness, particularly those from socioeconomically marginalized backgrounds, which has worsened since the COVID-19 pandemic.²

The American Academy of Pediatrics (AAP) recommends school readiness promotion and developmental screening during well-visits.^{3,4} However, general developmental screening (e.g., Survey of Well-Being of Young Children, SWYC) does not assess numeracy⁵ and while several preschool numeracy screeners have been validated in early childhood education settings (e.g., K-SEALS, PENS), none have been studied or utilized in pediatric practice.^{6,7} Reasons include time constraints, cost, need for specialized kits, and non-synchrony with typical workflows. *The Number Farm* (TNF) is a children's book-based measure that screens numeracy skills that are most often learned in the preschool period. TNF was inspired by a book-based approach to emergent literacy screening validated during pediatric primary care.^{8,9} The objective of this study was to

pilot TNF during well-child visits for feasibility, family acceptance, and psychometric validity, particularly with respect to home math and reading environments.

METHODS

Numeracy measures

The Number Farm (TNF) is a 14-page illustrated, rhyming children's book (Fig. 1b). A Power Point-rendered version was pilot tested for clarity with families and colleagues prior to the book's publication (Fig. 1a). TNF screener includes 11 scripted items that refer to the book's content (e.g., covers, scenes) yet are distinct from its text (0–28 points). Numeracy domains and weightings were selected to mirror those in the validated Preschool Early Numeracy Screener (PENS) developed by Purpura et al., cited as vital for school readiness and success in math.⁶ TNF items were designed in consultation with numeracy experts and a review of relevant literature. Item distribution reflective of component skills and corresponding sections of the book is summarized in Table 1.

The PENS was used as the "gold standard" comparison measure, involves 25-items with an estimated administration time of 5 minutes, and has been validated for preschool age.^{6,10} The PENS assesses the same numeracy domains as TNF.

Participants/Setting

This study was based at an academic pediatric primary care clinic serving largely Medicaid-eligible families in Cincinnati, OH. Recruitment was by

¹Division of General and Community Pediatrics, The University of Texas Southwestern Medical Center, Dallas, TX, USA. ²Virginia Pediatric and Adolescent Center, Springfield, VA, USA. ³Division of Hematology/Oncology, University of Cincinnati College of Medicine Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ⁴Division of Biostatistics and Epidemiology, University of Cincinnati College of Medicine Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA. ✉email: John.Hutton@utsouthwestern.edu

Received: 21 November 2025 Revised: 15 April 2026 Accepted: 23 April 2026

Published online: 14 May 2026

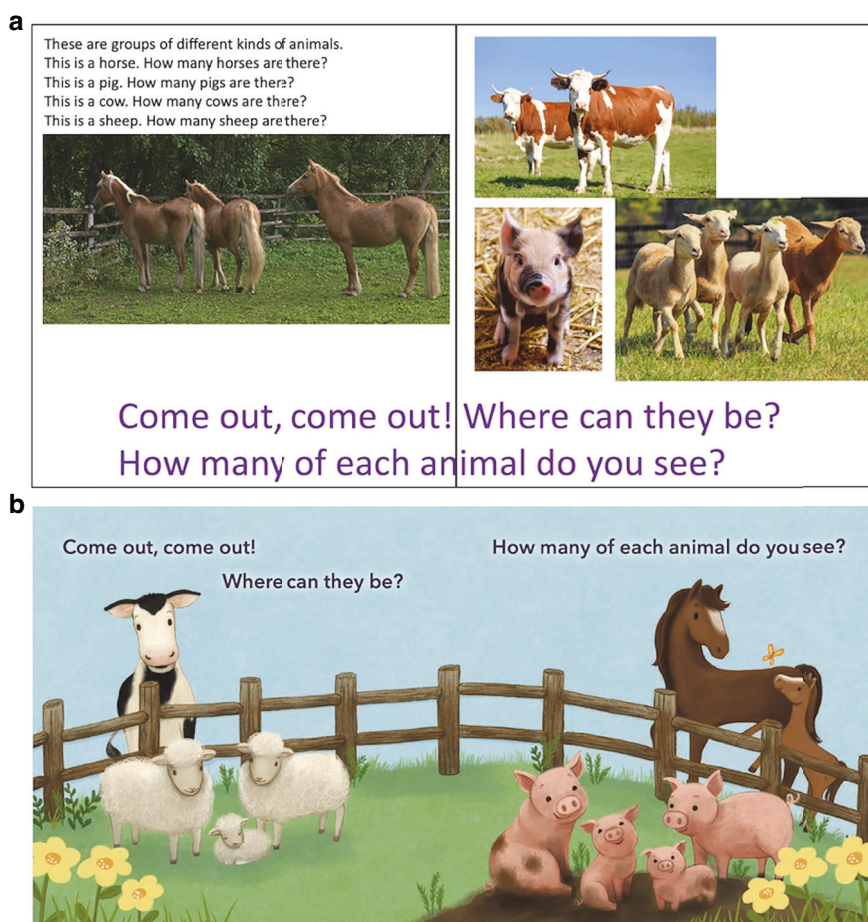


Fig. 1 Sample content from *The Number Farm*. Top (a) is from the mockup version and bottom (b) is from the published book, for Item 5 (1:1 counting, sets). The 11-item scripted assessment is based on the book's covers and interior illustrations but is distinct from its text.

convenience sampling of families presenting for well-visits meeting these criteria: child age 36 to 59 months, no developmental disorder likely to confer cognitive delay (e.g., Trisomy 21, autism), parent/child fluent in English, no acute infectious illness. The study was approved by the sponsoring institution's Institutional Review Board. Prior to screening, written informed consent was obtained, and families were provided with a gift card after data collection.

Screening and assessments

Pediatric residents collected all data in the exam room before or after the physician conducted the visit, mindful of clinic flow. Following informed consent, TNF and PENS were administered to the child in predetermined, random order, noting screening time and barriers (e.g., interruption by staff). Three-item, Likert-based TNF impression surveys developed by the study team based on prior work were then administered to the child and parent.^{8,9} Next, surveys of home math activities and attitudes (Likert-based, developed by the study team), home reading environment (StimQ READ¹¹), screen time (ScreenQ¹²) and demographics (e.g., sex, preschool enrollment) were administered to the parent.

Statistical analyses

Descriptive statistics were computed for all variables. TNF items were analyzed using modern theory Rasch analyses,¹³ providing estimates of smoothness, modality, difficulty, polarity, response sufficiency, and item-fit. A preliminary estimate of TNF reliability was calculated using Cronbach's alpha (α_c), and concurrent validity was estimated using Pearson correlation coefficients between TNF and PENS scores. Pearson coefficients were also computed between TNF total score and demographic, home reading and math environment, and digital media use measures. Analyses were conducted using SAS v9.4 and Winsteps v5.6.4 software.

RESULTS

Sample characteristics

Thirty-one children participated, mean age = 50 months (± 4 ; 36–71), 17 boys and 14 girls. Barriers to screening were rare (under 10%) and most often included interruption by clinic staff to initiate the visit. In each of these instances, screening was successfully completed after the physician had finished. TNF administration time was under 10 minutes.

Screeener total scores and TNF item-level analysis

Mean TNF score was 11 (± 8 ; 0–26) and PENS score was 10 (± 8 ; 0–25). TNF and PENS scores were not normally distributed (Shapiro-Wilk $p = 0.03$ and 0.02 , respectively), just below the $p = 0.05$ threshold.

TNF item-level statistics are provided in Table 2. Rasch estimates of item difficulty ranged from 2.72 (less difficult; item 5, numeral names) to 1.26 (more difficult; item 10, addition) and were evenly distributed between harder and easier items, with a slight concentration of moderately difficult items (1, 9, 12). Item fit statistics were within generally accepted ± 2 standard deviations, suggesting no outliers likely to influence the distribution of scores. Point-measure correlations ranged from 0.35 (item 1, concept of number) to 0.88 (item 5, 1-1 counting of sets) and were generally between 0.52 and 0.70 across all items, suggesting moderate relationships between each item and the entire TNF scale. Items 4 (set comparison) and 5 (ordinality) represent exceptions, with relatively high point-measure correlations of $r = 0.82$ and $r = 0.88$, respectively.

Table 1. The number farm skill and scoring distribution

Section	Skill Domain	Item #	Points	Percent
Front cover	Concept of number	1	1	4%
Front cover	Numeral names	2, 3	5	18%
Spread 1	Verbal counting	4	3	11%
Spread 2	1-1 counting (sets)	5	4	14%
Spread 3	Set-numeral mapping (quantities)	6	4	14%
Spread 4	Exact set comparison	7	3	11%
Spread 5	Number order (ordinality)	8	2	7%
Spread 6	Numeral comparison (cardinality)	9	2	7%
Spread 7	Addition (story problems)	10	2	7%
Back cover	Number writing (child's age)	11	2	7%
Total			28	100%

Front and back covers and each of 7 interior page spreads assess an early numeracy component skill expected during the preschool age range. A scripted administration form mirrors yet is distinct from the book's narrative. Points are awarded for correct responses, and the total possible score is 28 points.

Table 2. Item analysis and summary statistics for the number farm (Rasch Analysis)

Item	M(SD)	Difficulty	Standard error	Infit z	Outfit z	Point-measure correlation
10 Addition	0.5 (0.6)	1.26	0.35	-0.43	-0.21	0.56
9 Cardinality	0.5 (0.8)	0.93	0.32	-0.31	0.38	0.57
1 Concept of numbers	0.5 (0.5)	0.83	0.31	-0.36	1.60	0.35
11 Age number writing	0.6 (0.9)	0.73	0.31	-1.23	-1.02	0.73
8 Ordinality	0.6 (0.7)	0.64	0.30	-0.58	-0.26	0.61
2 Number naming open	0.6 (0.8)	0.55	0.29	-1.82	-1.07	0.74
3 Number naming specific	1.1 (0.8)	-0.51	0.25	-1.08	-0.31	0.72
7 Exact set comparison	1.1 (1.0)	-0.57	0.24	1.23	1.05	0.53
6 Set-number mapping	1.2 (1.5)	-0.74	0.24	1.36	1.16	0.74
4 Verbal counting	1.5 (1.3)	-1.23	0.23	-0.83	-0.59	0.82
5 1-1 counting of sets	2.5 (1.6)	-2.72	0.23	-0.27	-0.47	0.88

Terms used to describe each numbered skill domain are for respective items and organized by item difficulty level, where negative is easier and positive is harder.

Reliability and validity

For reliability, internal consistency was $r_{C\alpha} = 0.94$, considered "excellent."¹⁴ For concurrent validity, the Pearson correlation between TNF and PENS total scores was 0.94 ($p < 0.001$). Figure 2 is a scatter plot of TNF and PENS scores.

TNF scores were strongly correlated with child age ($r_p = 0.72$, $p < 0.001$) and preschool enrollment ($r_p = 0.38$, $p < 0.05$) but not with child sex. TNF scores were positively correlated with reported number of children's books at home ($r_p = 0.48$), days/week of shared reading ($r_p = 0.64$) and frequency of playing math-related games ($r_p = 0.39$; all p 's < 0.05). Reported parental enjoyment of math was negatively correlated with parental difficulties in math ($r_p = -0.41$, $p < 0.05$). There were trends towards negative correlation between TNF scores and reported parental difficulties in math ($r_p = -0.28$, $p = 0.12$) and positive correlation with reported parental enjoyment of math ($r_p = 0.24$, $p = 0.20$). TNF scores were not significantly correlated with digital media use (ScreenQ score).

Survey responses/satisfaction

Ninety-seven percent of children reported their experience with TNF as "Fun." Parent impression of TNF was also favorable, with 100% rating screening as "helpful," 75% rating time required as "just right" (10% too much), 94% stating that they learned new information about the child's skills, and 97% that they were

inspired to work more with their child in math at home. Eighteen parents (58%) were at least somewhat surprised by their child's performance. Of these, 15 (83%) were surprised at how well their child performed, and 3 (17%) were surprised at how poorly (17%).

A majority of parents reported that they were somewhat familiar with the math skills needed for kindergarten (55%, 10% Not at All), and 52% were somewhat worried about their child's math readiness (35% Not at All).

DISCUSSION

With trusted, frequent access to families during early childhood, pediatricians are poised to conduct developmental screening and provide guidance promoting school readiness, as recommended by the AAP.⁴ However, while early literacy screening has shown promise,⁸ there is currently no established approach to numeracy screening in pediatric practice. Further, commonly used screening measures (e.g., Ages & Stages Questionnaire, SWYC) do not specifically assess numeracy skills. This reflects an important gap and opportunity, as a large and increasing number of children arrive at kindergarten (typically age 5-6 in the US) with low math readiness, particularly those from socioeconomically disadvantaged households.^{4,15,16} The goal of this study was to test a children's book-based approach to numeracy screening during pediatric primary care.

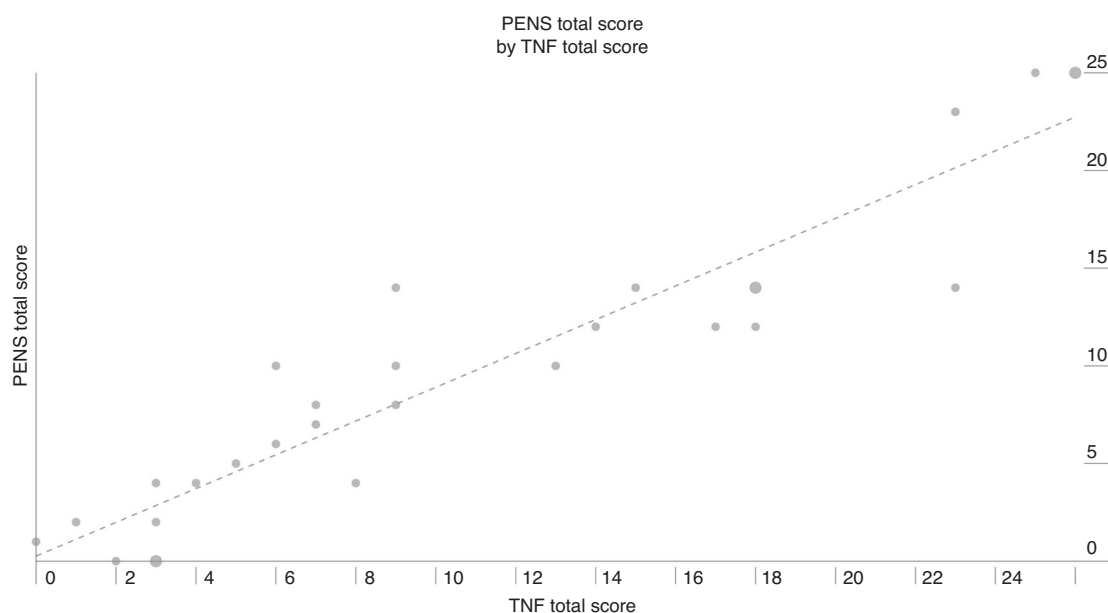


Fig. 2 Scatter plot of *The Number Farm* (TNF) vs. PENS total scores ($n = 31$).

Psychometric properties of TNF were strong, particularly for a new measure capturing a complex cognitive domain. We attribute this to an evidence-based conceptual model involving well-defined component numeracy skills,¹⁷ and a book-based approach that has been successful in screening emergent literacy skills at this age.⁸ TNF and PENS total scores were highly correlated ($r_p = 0.94$), suggesting strong concurrent validity, and internal consistency was excellent ($r_{C\text{-}a} = 0.94$), suggesting high coherence among screener items.

In terms of internal psychometrics, Rasch analyses of TNF items were also promising. TNF item and total scores were all positively correlated, with all response options selected by no less than 6% of the sample (an accepted standard), and monotonic ordering for each response option (i.e., full credit, partial credit, no credit, in a uniform direction). Item difficulty was balanced between easier and harder items, with addition (item 10) and numeral comparison (item 9) scoring most difficult, and 1-1 counting (item 5) and verbal counting (item 4) scoring easiest (Table 2). This is consistent with expected skill trajectories, where more abstract concepts and mental manipulation thereof (i.e., concept of number size, addition) were harder, while rote concepts (e.g., counting aloud, counting small groups) were easier.^{1,17} Item 5 (1-1 counting of sets) was by far the easiest item, potentially attributable to the small number of animals in each set tested (1 to 4 per set). Item revision to optimize these metrics is planned for a subsequent version of TNF.

Item-measure correlations were largely moderately positive, below 0.80, suggesting measure coherence and that each item contributed uniquely to the TNF total score. Exceptions were items 4 (verbal counting) and 5 (counting sets), which were also the easiest. This finding may reflect the influence of overall cognitive development on total score, rather than mastery of specific skills, evidenced by the strong correlation between TNF score and child age. Revision of these items, as mentioned above, may help minimize this effect.

TNF scores were strongly correlated with child age and preschool enrollment but not significantly with sex, consistent with developmental maturation, established benefits of preschool,^{4,18} and lagging sex disparities in math abilities.¹⁹ They were marginally positively correlated with reported parental enjoyment of math and marginally negatively correlated with parental reported difficulties in math, likely reflecting varying

parental abilities and/or educational experiences. They were significantly, positively correlated with the reported number of children's books at home, days/week of shared reading and playing math-related games, each reflecting empowering ways for parents to help the child at home.¹⁸ They were not correlated with digital media use, surprising considering the potential displacement of parent-child interaction and recent evidence of negative effects.²⁰ However, such evidence is scant in terms of risks or benefits of digital media, warranting further study.

Responses to parent and child surveys were strongly favorable, suggesting that TNF screening may be feasible, useful, and enjoyable during well-visits. Administration time was under 10 minutes, which is excessive for most clinics, yet attributable to the research stage of this work. Published work involving literacy screening has found that screening time improves substantially with practice, which seems a reasonable investment, especially if reimbursable as developmental surveillance,³ when utilized by embedded school readiness staff,^{5,9} in specialty clinics where visits are often longer, and/or as a pediatric resident teaching tool.²¹ Further refinement of TNF to improve efficiency is planned.

This study has limitations. The sample size was small and collected at an academic primary care clinic serving largely socioeconomically disadvantaged, primarily Black families, limiting generalizability, though this is a vital demographic in terms of school readiness.⁴ TNF was administered by pediatric residents functioning as clinical research coordinators. Future studies are needed to explore TNF screening during usual clinic workflows (e.g., via medical assistants, clinicians).²¹ This study also has strengths, including randomization of TNF/PENS screening to minimize fatigue effects, robust psychometric analyses, and leveraging of a successful book-based approach to literacy screening.^{8,9} TNF is family-centered, inexpensive (\$3.00/book, given to the child to take home), requires minimal training, and is scalable within existing programs (e.g., Reach Out and Read²²), with substantial advantages relative to kit-based measures. Altogether, this pilot provides compelling data that TNF may be an appealing option for early numeracy screening in pediatric settings in the preschool age range (3–5 years old), complementing existing school-based tools. More expansive studies are warranted, with a goal to help align efforts by pediatricians and early childhood educators in this vital school readiness domain,

including timely screening with familiar, reliable tools and tailored guidance for families to help their child succeed.

For more information about *The Number Farm* or to order for research or clinical use, visit <http://www.blumanateeepress.com/nonprofit/> or email press@blumanateebooks.com or John.Hutton@utsouthwestern.edu.

DATA AVAILABILITY

Data and other materials utilized in this study will be freely available to any researcher wishing to use them for non-commercial purposes, without breaching participant confidentiality. Written queries, including intended use of these materials, should be sent to the corresponding author for consideration.

REFERENCES

- Clements, D.H., Sarama, J. *Early Childhood Mathematics Research: Learning Trajectories for Young Children*. Routledge, (2009).
- National Assessment of Educational Progress at Grades 4 and 8: Mathematics (National Center for Education Statistics, 2024).
- American Academy of Pediatrics Bright Futures National Center. *Bright Futures: Guidelines for Health Supervision of Infants, Children, and Adolescents (4th edition)*. 3rd ed. American Academy of Pediatrics (2017).
- Williams, P.G., Lerner, M.A. School Readiness. *Pediatrics*. **144**, <https://doi.org/10.1542/peds.2019-1766>
- Bode, S.M., Denny, S., Masonheimer, A., Irwin, M.K., Hutton, J. Comparison of literacy and developmental screening of preschool-aged children during primary care. *Pediatrics*. **156** <https://doi.org/10.1542/peds.2024-069941>
- Purpura, D. J. & Lonigan, C. J. Early Numeracy Assessment: The Development of the Preschool Numeracy Scales. *Early Educ Dev* **26**, 286–313, <https://doi.org/10.1080/10409289.2015.991084> (2015).
- Kass, C. E. Kaufman Survey of Early Academic and Language Skills (K-SEALS). *Diagnostique* **24**, 135–144 (1999).
- Hutton, J.S. et al. The Reading House: a children's book for emergent literacy screening during well-child visits. *Pediatrics*. 2019;143
- Bode, S.M., Denny, S., Irwin, M.K., Hutton, J. A primary care-based approach to improving kindergarten readiness. *Pediatrics*. **155**. <https://doi.org/10.1542/peds.2024-069159C>
- Purpura, D. J., Reid, E. E., Eiland, M. D. & Baroody, A. J. Using a Brief Preschool Early Numeracy Skills Screener to Identify Young Children With Mathematics Difficulties. *School Psychol. Rev.* **44**, 41–59 (2015).
- Bellevue Project for Early Language Literacy and Education Success. STIMQ Home Cognitive Environment. Bellevue Project for Early Language, Literacy, and Education Success (BELLE). 2022. <http://pediatrics.med.nyu.edu/developmental/research/the-belle-project/stimq-cognitive-home-environment>
- Hutton, J. S., Huang, G., Sahay, R. D., DeWitt, T. & Ittenbach, R. F. A novel, composite measure of screen-based media use in young children (ScreenQ) and associations with parenting practices and cognitive abilities. *Pediatr. Res.* **87**, 1211–1218, <https://doi.org/10.1038/s41390-020-0765-1> (2020).
- Smith E.V., Smith R.M.. *Rasch measurement: advanced and specialized applications*. JAM Press (2007).
- Peterson R.A. A Meta-analysis of Cronbach's Coefficient Alpha *J. Consum. Res.* 21 (1994).
- Solano, I.S., Weyer, M. *Closing the Opportunity Gap in Early Childhood Education*. 25. (2017).
- Duncan, G. J. et al. School readiness and later achievement. *Dev. Psychol.* **43**, 1428–1446, <https://doi.org/10.1037/0012-1649.43.6.1428> (2007). Nov.
- Braeuning, D., Ribner, A., Moeller, K. & Blair, C. The Multifactorial Nature of Early Numeracy and Its Stability. *Front. Psychol.* **11**, 518981. <https://doi.org/10.3389/fpsyg.2020.518981> (2020).
- Young, J. M., Reed, K. E., Rosenberg, H. & Kook, J. F. Adding family math to the equation: Promoting Head Start preschoolers' mathematics learning at home and school. *Early Childhood Res. Q.* **63**, 43–58 (2023).
- Kersey, A. J., Braham, E. J., Csumitta, K. D., Libertus, M. E. & Cantlon, J. F. No intrinsic gender differences in children's earliest numerical abilities. *NPJ Sci. Learn* **3**, 12, <https://doi.org/10.1038/s41539-018-0028-7> (2018).
- Li, X. et al. Screen Time and Standardized Academic Achievement Tests in Elementary School. *JAMA Netw. Open* **8**, e2537092, <https://doi.org/10.1001/jamanetworkopen.2025.37092> (2025). Oct 1.
- Bruton, L., Nunes, D. & Quadri, M. Scoping review of pediatric primary care educational literature. *Clin. Pediatr.* **64**, 538–545, <https://doi.org/10.1177/00099228241274905> (2025).
- Klass, P., Dreyer, B. P. & Mendelsohn, A. L. Reach out and read: literacy promotion in pediatric primary care. *Adv. Pediatr.* **56**, 11–27, <https://doi.org/10.1016/j.yapd.2009.08.009> (2009). S0065-3101(09)00010-3].

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Laura Justice, PhD, from the Crane Center for Early Childhood Research and Policy at The Ohio State University for consultation in the development of *The Number Farm* screening measure.

AUTHOR CONTRIBUTIONS

John S. Hutton conceived, wrote, and coordinated the publication of *The Number Farm*, developed and collaborated in the refinement of the screening measure, designed all aspects of the study, mentored the pediatric residents administering the study, collaborated in analyses, drafted the initial manuscript and subsequent revisions, and approved the final manuscript as submitted. Priya Singh collaborated in study design and data collection during pediatric well-visits, assisted with drafting the manuscript, and approved the final manuscript as submitted. Sarah Reel collaborated in study design and data collection during pediatric well-visits, assisted with drafting the manuscript, and approved the final manuscript as submitted. Richard F. Ittenbach conducted psychometric analyses for this study, consulted in the development of *The Number Farm* screening instrument, reviewed and revised the manuscript, and approved the final manuscript as submitted.

FUNDING

This study was funded by the Tecklenburg Fund for Literacy at Cincinnati Children's Hospital and internal funds from UT Southwestern Medical Center (Hutton).

COMPETING INTERESTS

John Hutton conceived, wrote, and edited *The Number Farm*, the children's book and screening tool featured in this study. Dr. Hutton is the founder of Blue Manatee Press, the publisher of *The Number Farm*, though he receives no salary for this role. *The Number Farm* is intended to be distributed at low-cost as a screening tool to clinics, preschools, and non-profit organizations. Royalties, if any, are reinvested into the development of other books. The other authors have no financial relationships to disclose.

INFORMED CONSENT

This study was approved by the sponsoring institution's Institutional Review Board. Prior to screening, written informed consent was obtained from the child's parent.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to John S. Hutton.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026