

ISSN: 1300-915X

IOJPE

INTERNATIONAL ONLINE JOURNAL OF
PRIMARY EDUCATION



International Online Journal Of Primary Education

Volume 15 Issue 3

International Online Journal of Primary Education

ISSN: 1300-915X

SEPTEMBER 2026

Volume 15 – Issue 3

**Prof. Dr. Nergüz BULUT SERİN
Editor in Chief**

**Prof.Dr. Mehmet Engin DENİZ
Prof.Dr. Sinan OLKUN
Prof. Dr. Şule AYCAN
Assoc.Prof.Dr. Arzu GÜNGÖR LEUSHUIS
Editors**

Copyright © 2026 by author(s)

INTERNATIONAL ONLINE JOURNAL OF PRIMARY EDUCATION

All articles published in International Online Journal of Primary Education (IOJPE) are licensed under a [Creative Commons Attribution 4.0 International License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

IOJPE allows readers to read, download, copy, distribute, print, search, or link to the full texts of its articles and allow readers to use them for any other lawful purpose.

IOJPE does not charge authors an article processing fee (APF).

Published in TURKEY

Contact Address:

Prof. Dr. Nergüz BULUT SERİN

Editor in Chief

European University of Lefke

Lefke, Northern Cyprus TR-10 Mersin, Turkey, 99010

Message from the Editor,

I am very pleased to inform you that we have published the third issue in 2026. As an editor of International Online Journal of Primary Education (IOJPE), this issue is the success of our authors, very valuable reviewers who undertook the rigorous peer review of the manuscripts, and those of the editorial board who devoted their valuable time through the review process. In this respect, I would like to thank to all reviewers, researchers and the editorial board members. The articles should be original, unpublished, and not in consideration for publication elsewhere at the time of submission to International Online Journal of Primary Education (IOJPE). For any suggestions and comments on IOJPE, please do not hesitate to send me e-mail. The countries of the authors contributed to this issue (in alphabetical order): Tanzania and Türkiye.

Prof. Dr. Nergüz BULUT SERİN

Editor in Chief

Editor in Chief

PhD. Nergüz Bulut Serin, (European University of Lefke, North Cyprus)

 Orcid ID: [0000-0002-2074-3253](https://orcid.org/0000-0002-2074-3253) Scopus ID: [26656955100](https://scopus.org/26656955100) [Google Scholar](#)

Editors

PhD. Mehmet Engin Deniz, (Yıldız Teknik University, Turkey)

 Orcid Id: [0000-0002-7930-3121](https://orcid.org/0000-0002-7930-3121) Scopus ID: [660217798](https://scopus.org/660217798) [Google Scholar](#)

PhD. Sinan Olkun, (Ankara University, Turkey)

 Orcid Id: [0000-0003-3764-2528](https://orcid.org/0000-0003-3764-2528) Scopus ID: [8895882700](https://scopus.org/8895882700) [Google Scholar](#)

PhD. Arzu Güngör Leushuis, (Florida State University, United States)

 Orcid Id: [0000-0001-8197-121X](https://orcid.org/0000-0001-8197-121X) Scopus ID: [59673978900](https://scopus.org/59673978900) [Google Scholar](#)

PhD. Şule Aycan, (Muğla University, Turkey)

 Orcid Id: [0000-0001-8844-0438](https://orcid.org/0000-0001-8844-0438) Scopus ID: [6603100984](https://scopus.org/6603100984) [Google Scholar](#)

Linguistic Editors

PhD. Mehmet Ali Yavuz, (Cyprus International University, North Cyprus)

 Orcid Id: [0000-0002-7121-5194](https://orcid.org/0000-0002-7121-5194) Scopus ID: [57198107116](https://scopus.org/57198107116) [Google Scholar](#)

PhD. Nazife Aydınoglu, (Final International University, North Cyprus)

 Orcid Id: [0000-0002-0382-7092](https://orcid.org/0000-0002-0382-7092) [Scopus ID:](#) [Google Scholar](#)

Classroom Management in Primary Education

PhD. Chokri Kooli, (University of Ottawa, Canada)

 Orcid Id: [0000-0002-8211-8621](https://orcid.org/0000-0002-8211-8621) Scopus ID: [57210946891](https://scopus.org/57210946891) [Google Scholar](#)

PhD. Fahriye Altınay, (Near East University, North Cyprus)

 Orcid Id: [0000-0002-3861-6447](https://orcid.org/0000-0002-3861-6447) Scopus ID: [8350821700](https://scopus.org/8350821700) [Google Scholar](#)

PhD. Mehmet Durdu Karşlı, (Eastern Mediterranean University, North Cyprus)

 Orcid Id: [0000-0003-1239-4150](https://orcid.org/0000-0003-1239-4150) Scopus ID: [14063443600](https://scopus.org/14063443600) [Google Scholar](#)

Curriculum Development in Primary Education

PhD. Ali Ahmad Al-Barakat, (University of Sharjah, United Arab Emirates)

 Orcid Id: 0000-0002-2709-4962 Scopus ID: [35118761200](https://scopus.org/35118761200) [Google Scholar](#)

PhD. Arzu Güngör Leushuis, (Florida State University, United States)

 Orcid Id: [0000-0001-8197-121X](https://orcid.org/0000-0001-8197-121X) Scopus ID: [59673978900](https://scopus.org/59673978900) [Google Scholar](#)

PhD. Asuman Seda Saracaloğlu, (Adnan Menderes University, Turkey)

 Orcid Id: [0000-0001-7980-0892](https://orcid.org/0000-0001-7980-0892) Scopus ID: [26656925800](https://scopus.org/26656925800) [Google Scholar](#)

Computer Education and Instructional Technologies in Primary Education

PhD. Aytekin İşman, (Sakarya University, Turkey)

 Orcid Id: [0000-0003-0420-7976](https://orcid.org/0000-0003-0420-7976) Scopus ID: [10839357100](https://scopus.org/10839357100) [Google Scholar](https://scholar.google.com/citations?user=10839357100)

Ph.D. Ersun İşçiöğlu, (Eastern Mediterranean University, North Cyprus)

 Orcid Id: [0000-0002-0637-7912](https://orcid.org/0000-0002-0637-7912) Scopus ID: [36191794000](https://scopus.org/36191794000) [Google Scholar](https://scholar.google.com/citations?user=36191794000)

PhD. Halil İbrahim Yalın, (Cyprus International University, North Cyprus)

 Orcid Id: [0000-0002-6355-7661](https://orcid.org/0000-0002-6355-7661) Scopus ID: [23096832400](https://scopus.org/23096832400) [Google Scholar](https://scholar.google.com/citations?user=23096832400)

PhD. Sezer Kanbul, (Near East University, North Cyprus)

 Orcid Id: [0000-0002-4715-8089](https://orcid.org/0000-0002-4715-8089) Scopus ID: [26658005100](https://scopus.org/26658005100) [Google Scholar](https://scholar.google.com/citations?user=26658005100)

Ms Umut Tekgüç, (Bahçeşehir Cyprus University, North Cyprus)

 Orcid Id: [0000-0001-5974-5566](https://orcid.org/0000-0001-5974-5566) Scopus ID: [35300830300](https://scopus.org/35300830300) [Google Scholar](https://scholar.google.com/citations?user=35300830300)

PhD. Zehra Altınay, (Near East University, North Cyprus)

 Orcid Id: [0000-0002-6786-6787](https://orcid.org/0000-0002-6786-6787) Scopus ID: [8350821600](https://scopus.org/8350821600) [Google Scholar](https://scholar.google.com/citations?user=8350821600)

Educational Drama in Primary Education

PhD. Alev Önder, (Bahçeşehir University, Turkey)

 Orcid Id: [0000-0002-2736-4600](https://orcid.org/0000-0002-2736-4600) Scopus ID: [26656903100](https://scopus.org/26656903100) [Google Scholar](https://scholar.google.com/citations?user=26656903100)

Educational Psychology in Primary Education

PhD. Christina Athanasiades, (Psychology, Aristotle University of Thessaloniki, Greece)

 Orcid Id: [0000-0003-4916-9328](https://orcid.org/0000-0003-4916-9328) Scopus ID: [25652700900](https://scopus.org/25652700900) [Google Scholar](https://scholar.google.com/citations?user=25652700900)

PhD. Muhammad Sabil Farooq, (Nankai University Tianjin, P.R. China)

 Orcid Id: [0000-0001-7034-0172](https://orcid.org/0000-0001-7034-0172) Scopus ID: [57205442426](https://scopus.org/57205442426) [Google Scholar](https://scholar.google.com/citations?user=57205442426)

PhD. Rengin Karaca, (Dokuz Eylül University, Turkey)

 Orcid Id: [0000-0001-5955-0603](https://orcid.org/0000-0001-5955-0603) Scopus ID: [Google Scholar](https://scopus.org/)

PhD. Thanos Touloupis, (Aristotle University of Thessaloniki, Greece)

 Orcid Id: [0000-0003-2951-6919](https://orcid.org/0000-0003-2951-6919) Scopus ID: [56441555400](https://scopus.org/56441555400) [Google Scholar](https://scholar.google.com/citations?user=56441555400)

Fine Arts Education in Primary Education

PhD. Ayfer Kocabaş, (Dokuz Eylül University, Turkey)

 Orcid Id: [0000-0002-5566-212X](https://orcid.org/0000-0002-5566-212X) Scopus ID: [26656822800](https://scopus.org/26656822800) [Google Scholar](https://scholar.google.com/citations?user=26656822800)

Foreign Language Teaching in Primary Education

PhD. Nazife Aydınöğlu, (Final International University, North Cyprus)

 Orcid Id: [0000-0002-0382-7092](https://orcid.org/0000-0002-0382-7092) Scopus ID: [Google Scholar](https://scopus.org/)

PhD. İzzettin Kök, (Girne American University, North Cyprus)

 Orcid Id: [0000-0003-2229-8058](https://orcid.org/0000-0003-2229-8058) Scopus ID: [55127933400](https://scopus.org/55127933400) [Google Scholar](https://scholar.google.com/citations?user=55127933400)

PhD. Perihan Savaş, (Middle East Technical University Turkey)

 Orcid Id: [0000-0001-9839-3081](https://orcid.org/0000-0001-9839-3081) Scopus ID: [36337903400](https://scopus.org/36337903400) [Google Scholar](#)

PhD. Vahid Norouzi Lasari, (Charles University, Prague, Czech Republic)

 Orcid Id: [0000-0002-3359-2677](https://orcid.org/0000-0002-3359-2677) Scopus ID: [57700659200](https://scopus.org/57700659200) [Google Scholar](#)

Guidance and Counselling in Primary Education

PhD. Ferda Aysan, (Dokuz Eylül University, Turkey)

 Orcid Id: [0000-0003-1396-3183](https://orcid.org/0000-0003-1396-3183) Scopus ID: [6507300768](https://scopus.org/6507300768) [Google Scholar](#)

PhD. Mehmet Engin Deniz, (Yıldız Teknik University, Turkey)

 Orcid Id: [0000-0002-7930-3121](https://orcid.org/0000-0002-7930-3121) Scopus ID: [660217798](https://scopus.org/660217798) [Google Scholar](#)

PhD. Nergüz Bulut Serin, (European University of Lefke, North Cyprus)

 Orcid Id: [0000-0002-2074-3253](https://orcid.org/0000-0002-2074-3253) Scopus ID: [26656955100](https://scopus.org/26656955100) [Google Scholar](#)

Mathematics Education in Primary Education

PhD. Elizabeth Jakubowski, (Florida State University, United States)

 Orcid Id: [0000-0001-5348-2400](https://orcid.org/0000-0001-5348-2400) Scopus ID: [26428775500](https://scopus.org/26428775500) [Google Scholar](#)

PhD. Kakoma Luneta, (University of Johannesburg, South Africa)

 Orcid Id: [0000-0001-9061-0416](https://orcid.org/0000-0001-9061-0416) Scopus ID: [55893693500](https://scopus.org/55893693500) [Google Scholar](#)

PhD. Melih Turgut, (Norwegian University of Science and Technology (NTNU), Norway)

 Orcid Id: [0000-0003-3777-9882](https://orcid.org/0000-0003-3777-9882) Scopus ID: [26322098600](https://scopus.org/26322098600) [Google Scholar](#)

PhD. Moritz Herzog, (University of Wuppertal, Germany)

 Orcid Id: [0000-0002-6706-3351](https://orcid.org/0000-0002-6706-3351) Scopus ID: [57193900500](https://scopus.org/57193900500) [Google Scholar](#)

PhD. Nazan Sezen Yüksel, (Hacettepe University, Turkey)

 Orcid Id: [0000-0002-0539-3785](https://orcid.org/0000-0002-0539-3785) Scopus ID: [44661914200](https://scopus.org/44661914200) [Google Scholar](#)

PhD. Osman Cankoy, (Atatürk Teachers Academy, North Cyprus)

 Orcid Id: [0000-0002-4765-9297](https://orcid.org/0000-0002-4765-9297) Scopus ID: [8311142800](https://scopus.org/8311142800) [Google Scholar](#)

PhD. Sinan Olkun, (Ankara University, Turkey)

 Orcid Id: [0000-0003-3764-2528](https://orcid.org/0000-0003-3764-2528) Scopus ID: [8895882700](https://scopus.org/8895882700) [Google Scholar](#)

PhD. Stefan Haesen, (Thomas More University, Belgium)

 Orcid Id: [0000-0002-2769-9822](https://orcid.org/0000-0002-2769-9822) Scopus ID: [15519207100](https://scopus.org/15519207100) [Google Scholar](#)

Measurement and Evaluation in Primary Education

PhD. Bayram Bıçak, (Akdeniz University, Turkey)

 Orcid Id: [0000-0001-8006-4677](https://orcid.org/0000-0001-8006-4677) Scopus ID: [Google Scholar](#)

PhD. Emre Çetin, (Cyprus Social Sciences University, North Cyprus)

 Orcid Id: [0000-0001-5474-6164](https://orcid.org/0000-0001-5474-6164) Scopus ID: [55616695500](https://scopus.org/55616695500) [Google Scholar](#)

PhD. Gökhan İskifoğlu, (European University of Lefke, North Cyprus)

 Orcid Id: [0000-0001-8119-4254](https://orcid.org/0000-0001-8119-4254) Scopus ID: [55745026100](https://scopus.org/55745026100) [Google Scholar](#)

PhD. Selahattin Gelbal, (Hacettepe University, Turkey)

 Orcid Id: [0000-0001-5181-7262](https://orcid.org/0000-0001-5181-7262) Scopus ID: [15519291100](https://scopus.org/15519291100) [Google Scholar](#)

Music Education in Primary Education

PhD. Gulsen G. Erdal, (Kocaeli University, Turkey)

 Orcid Id: [0000-0002-5299-126X](https://orcid.org/0000-0002-5299-126X) Scopus ID: [Google Scholar](#)

PhD. Sezen Özeke, (Uludağ University, Turkey)

 Orcid Id: [0000-0001-9237-674X](https://orcid.org/0000-0001-9237-674X) Scopus ID: [26658548400](https://scopus.org/26658548400) [Google Scholar](#)

PhD. Şirin Akbulut Demirci, (Uludağ University, Turkey)

 Orcid Id: [0000-0001-8904-4920](https://orcid.org/0000-0001-8904-4920) Scopus ID: [Google Scholar](#)

Pre-School Education

PhD. Alev Önder, (Bahçeşehir University, Turkey)

 Orcid Id: [0000-0002-2736-4600](https://orcid.org/0000-0002-2736-4600) Scopus ID: [26656903100](https://scopus.org/26656903100) [Google Scholar](#)

PhD. Ilfa Zhulamanova, (University of Southern Indiana, United States)

 Orcid Id: [0000-0003-3333-4237](https://orcid.org/0000-0003-3333-4237) Scopus ID: [57191155432](https://scopus.org/57191155432) [Google Scholar](#)

PhD. Ithel Jones, (Florida State University, United States)

 Orcid Id: [0000-0002-0690-3070](https://orcid.org/0000-0002-0690-3070) Scopus ID: [8696931600](https://scopus.org/8696931600) [Google Scholar](#)

PhD. Rengin Zembat, (Maltepe University, Turkey)

 Orcid Id: [0000-0002-2377-8910](https://orcid.org/0000-0002-2377-8910) Scopus ID: [35955365300](https://scopus.org/35955365300) [Google Scholar](#)

Science Education in Primary Education

PhD. Oguz Serin, (European University of Lefke, North Cyprus)

 Orcid Id: [0000-0003-4739-605X](https://orcid.org/0000-0003-4739-605X) Scopus ID: [26656883200](https://scopus.org/26656883200) [Google Scholar](#)

PhD. Salih Çepni, (Uludağ University, Turkey)

 Orcid Id: [0000-0003-2343-8796](https://orcid.org/0000-0003-2343-8796) Scopus ID: [16642100700](https://scopus.org/16642100700) [Google Scholar](#)

PhD. Şule Aycan, (Muğla University, Turkey)

 Orcid Id: [0000-0001-8844-0438](https://orcid.org/0000-0001-8844-0438) Scopus ID: [6603100984](https://scopus.org/6603100984) [Google Scholar](#)

PhD. Uğur Serin, (Ministry of Education, Necip Fazıl Kısakürek Primary School, Buca, Turkey)

 Orcid ID: [0000-0002-4667-9472](https://orcid.org/0000-0002-4667-9472) Scopus ID: [36544438100](https://scopus.org/36544438100) [Google Scholar](#)

PhD. Woldie Belachew Balea, (Addis Ababa University, Ethiopia)

 Orcid Id: [0000-0002-7891-4385](https://orcid.org/0000-0002-7891-4385) Scopus ID: [57218449004](https://scopus.org/57218449004) [Google Scholar](#)

Social Sciences Education in Primary Education

PhD. Erdal Aslan, (Dokuz Eylül University, Turkey)

 Orcid Id: [0000-0002-9267-9852](https://orcid.org/0000-0002-9267-9852) Scopus ID: [56010448700](https://scopus.org/56010448700) [Google Scholar](#)

PhD. Z. Nurdan Baysal, (Marmara University, Turkey)

 Orcid Id: [0000-0002-3548-1217](https://orcid.org/0000-0002-3548-1217) Scopus ID: [36543669300](https://scopus.org/36543669300) [Google Scholar](#)

Sports Education in Primary Education

PhD. Erkut Konter, (Dokuz Eylül University, Turkey)

 Orcid Id: [0000-0003-1664-9077](https://orcid.org/0000-0003-1664-9077) Scopus ID: [28167487300](https://scopus.org/28167487300) [Google Scholar](#)

PhD. Metin Dalip, (State University of Tetova, Macedonia)

 Orcid Id: [0000-0002-7142-8931](https://orcid.org/0000-0002-7142-8931) Scopus ID: [Google Scholar](#)

PhD. Rana Varol, (Ege University, Turkey)

 Orcid Id: [0000-0002-9196-984X](https://orcid.org/0000-0002-9196-984X) Scopus ID: [57189325705](https://scopus.org/57189325705) [Google Scholar](#)

Special Education in Primary Education

PhD. Hakan Sari, (Necmettin Erbakan University, Turkey)

 Orcid Id: [0000-0003-4528-8936](https://orcid.org/0000-0003-4528-8936) Scopus ID: [8043728500](https://scopus.org/8043728500) [Google Scholar](#)

PhD. Hasan Avcıoğlu, (Cyprus International University, North Cyprus)

 Orcid Id: [0000-0002-3464-2285](https://orcid.org/0000-0002-3464-2285) Scopus ID: [54974732100](https://scopus.org/54974732100) [Google Scholar](#)

PhD. Muhammad Zaheer Asghar, (Universitat Oberta de Catalunya, Barcelona, Spain)

 Orcid Id: [0000-0003-2634-0583](https://orcid.org/0000-0003-2634-0583) Scopus ID: [57208667494](https://scopus.org/57208667494) [Google Scholar](#)

PhD. Tevhide Kargın, (Hasan Kalyoncu University, Turkey)

 Orcid Id: [0000-0002-1243-8486](https://orcid.org/0000-0002-1243-8486) Scopus ID: [7801652354](https://scopus.org/7801652354) [Google Scholar](#)

Turkish Language Teaching in Primary Education

PhD. Ahmet Pehlivan, (Eastern Mediterranean University, North Cyprus)

 Orcid Id: [0000-0002-5987-6475](https://orcid.org/0000-0002-5987-6475) Scopus ID: [36456968000](https://scopus.org/36456968000) [Google Scholar](#)

PhD. Yüksel Girgin, (Adnan Menderes University, Turkey)

 Orcid Id: [0000-0002-0515-6077](https://orcid.org/0000-0002-0515-6077) Scopus ID: [Google Scholar](#)

Journal Cover Designer

Eser Yıldızlar, (University of Sunderland, England)

ISSN: 1300-915X

Volume 15, Issue 3 (2025)

Table of Contents

Research Articles

LOCAL GAME-BASED ENGLISH LITERACY INSTRUCTION: EXPERIMENTAL EVIDENCE FROM TANZANIAN RURAL PRIMARY SCHOOLS

Wilfred MADATA, Jesse Julius NDABAKURANE, John Misana BISEKO

Page: 197-218

EFFECTS OF RECYCLING-BASED MATHEMATICAL TASKS ON PRESCHOOL CHILDREN'S PATTERN RECOGNITION

Feriha Hande İDİL

Page: 219-234

CLASSROOM-BASED WORKING MEMORY TRAINING IN CHILDREN: EVIDENCE FROM A QUASI-EXPERIMENTAL PRIMARY SCHOOL STUDY

Emire ULUĞ, Ayşe Dilek ÖĞRETİR ÖZÇELİK

Page: 235-252

UNIQUE CONTRIBUTIONS OF SUBITIZING AND SEQUENTIAL COUNTING TO ARITHMETIC COMPETENCE IN PRIMARY SCHOOL CHILDREN

İsa Birkan GÜLDENOĞLU, Yasemin ARSLAN, Ufuk ÖZKUBAT

Page: 253-265

ISSN: 1300-915X



LOCAL GAME-BASED ENGLISH LITERACY INSTRUCTION: EXPERIMENTAL EVIDENCE FROM TANZANIAN RURAL PRIMARY SCHOOLS

Wilfred MADATA

Department of Foreign Languages and Literature, University of Dodoma, Tanzania

ORCID: <https://orcid.org/0009-0008-3175-0548>

wilfdata2022@gmail.com

Jesse Julius NDABAKURANE

Department of Foreign Languages and Literature, University of Dodoma, Tanzania

ORCID: <https://orcid.org/0000-0001-9254-6825>

jessetuisabe@gmail.com

John Misana BISEKO

Department of Foreign Languages and Literature, University of Dodoma, Tanzania

ORCID: <https://orcid.org/0000-0001-8017-4489>

drbiseko.udom@yahoo.com

Received: March 23, 2026

Accepted: August 14, 2026

Published: September 30, 2026

Suggested Citation:

Madata, W., Ndabakurane, J. J., & Biseko, J. M. (2026). Local game-based English literacy instruction: Experimental evidence from Tanzanian rural primary schools. *International Online Journal of Primary Education (IOJPE)*, 15(3), 197-218. <https://doi.org/10.55020/iojpe.1914504>



Copyright © 2026 by author(s). This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Abstract

Literacy underachievement remains persistent in low- and middle-income countries, particularly in early-grade English as a Foreign Language instruction. While game-based learning has gained scholarly attention, existing research focuses heavily on digital games, with little is known about culturally grounded, non-digital alternatives. This study addressed this gap from a post-method perspective by advocating for context-sensitive pedagogies rooted in local practices. It examined the effectiveness of two Tanzanian local games: *rede* and *ruka kamba*, adapted for structured English literacy activities. Using a quasi-experimental pre-test–post-test design, pupils in the experimental group participated in game-based literacy instruction for five months, while the control group received traditional instruction. Implementation fidelity was assessed to evaluate the quality of the intervention and was found to be high (87.8%). Non-parametric employed in the analysis revealed significantly greater literacy improvements in the experimental groups. The study highlights the potential of culturally responsive pedagogies and offers a potentially replicable model for low-resource classrooms across sub-Saharan Africa.

Keywords: local games, game-based learning, English literacy, primary education, reading, writing

INTRODUCTION

Over the past two decades, global education systems have witnessed a substantial expansion in primary school enrolment (Sánchez et al., 2023). Despite this expansion, yet this progress masks a deeper crisis as millions of children, especially in low- and middle-income countries, complete schooling without acquiring basic literacy skills. This trend is evident in sub-Saharan Africa, where the proportion of primary school-age children enrolled rose from 61% in 2000 to 80% in 2020 (United Nations, 2023). Even with these notable gains in access to education, literacy outcomes remain alarmingly low worldwide. For instance, World Bank and UNESCO (2025) reports that hundreds of millions of children and adults still lack basic reading and writing skills. This indicates that access to schooling does not necessarily translate into literacy proficiency, particularly in English, which is often the language of instruction in many African countries.

In low-and middle-income countries, the literacy learning challenges has been remaining a defining feature of education systems. As reported by the World Bank (2025), approximately 70% of the 10-year-olds in these contexts cannot read and understand a simple age-appropriate text. The situation



is particularly critical in English, a condition commonly described as a *learning poverty*. In Africa, the rate reaches 89%, with country-level figures being even more alarming: 94% in Liberia, 87% in Malawi, and 86% in Zambia (World Bank, 2025). These figures underscore a systemic failure to ensure quality instruction, especially in Anglophone countries across the region where English serves as the medium of instruction. They suggest that the expansion of access to education has not been matched by corresponding improvements in instructional effectiveness.

This systemic failure is reflected across sub-Saharan Africa where the foundational literacy in English remains low in both reading and writing (Kanté, 2024). UNESCO, ADEA and African Union further reports that one in five children reaches the minimum proficiency level in reading by the end of primary school. This crisis is marked by a high prevalence of “non-readers.” For example, 51% of Grade 2 pupils in Ghana and up to 90% of Grade 3 pupils in northern Nigeria were reported to have received “zero scores” for English oral reading fluency (RTI International, 2015). Writing performance is another reported area with weaknesses in Sub-Saharan Africa. RTI International (2015) reports that one-third of adults in sub-Saharan Africa are unable to read or write in English. Furthermore, RTI reports that in Ghana, average writing composition scores were 37% for Grade 3 and 58% for Grade 5. These figures underscore a fundamental reality: improvements in English literacy depend not merely on school participation but critically on the quality and relevance of instruction.

In multilingual African education contexts, the dominance of English as a second or third language further complicates literacy acquisition. A clear example is Tanzania, where low English literacy presents a major challenge, especially among rural primary school pupils. This situation constrains progress and limits lifelong learning opportunities. Ngonyani et al. (2025) reveal that many learners face significant difficulties with reading comprehension, specifically in understanding main ideas, summarizing texts, and reading English passages fluently. National data show that only a small proportion of Standard 3 pupils in Tanzania can successfully read a Standard 2-level English story, with many failing to reach functional literacy even by the upper primary years (Uwezo, 2019). Contributing factors such as limited teacher proficiency in English, inadequate teaching resources, and persistent reliance on rote, teacher-centred methods (Taylor, 2024), are common across the country. Collectively, this evidence suggests that instructional delivery practices play a central role in shaping literacy outcomes rather than learner ability alone.

Post-Method and Foreign Language Teaching

The topic of methods and post-methods in English language teaching is widely debated globally, with researchers such as Paudel (2018) and Lusianov (2020) noting competing claims about the superiority of different methods. This contentious environment has prompted the emergence of post-method pedagogy, which emphasizes teacher autonomy in developing context-sensitive teaching approaches informed by their experienced and local context. Post-method pedagogy seeks to reconceptualize the theory-practice relationship in language education, highlighted by Lusianov's concept of “glocalization,” which reflects the blend of global English and localized practices. Kumaravadivelu (2003) outlines essential post-method principles: conducting needs analyses to assess students' needs, promoting cultural awareness to enhance sociocultural understanding, viewing learners as autonomous individuals, and encouraging learning through meaningful classroom activities rather than direct instruction. Implementing these principles is vital, especially in settings like rural primary schools in Tanzania, where attention to local contexts can strengthen language teaching and learning. Incorporating cultural values into language learning can foster motivation and a stronger sense of connection among students, making the learning experience more relevant and personalized.

Game-Based Learning and Local Games in English Language Instruction

Game-based learning (GBL) has gained recognition as a promising pedagogical approach for language instruction. A growing body of research has documented the effectiveness of digital tools such as Kahoot!, Duolingo, and Quizlet in improving vocabulary, motivation, and engagement in English as a Foreign Language (EFL) context (Alotaibi, 2024; Forsström et al., 2025). However, this



focus on digital technology may overlook the instructional realities of low-income settings, where digital resources are scarce and reliable internet access remains limited.

However, local games, with their deep historical and cultural roots, play important educational roles. They fulfil children's developmental needs, promoting mental, emotional, social, and moral growth by engaging them in expressive activities and encouraging direct social interaction (Palumbo et al., 2019). Educators incorporate these games into curricula to enhance skills such as language acquisition, motor coordination, and cooperation (Hussain & Cheong, 2022). In English language teaching, local games have emerged as an innovative alternative to traditional methods that often fail to engage learners (Kurnia, 2014). Grounded in cognitive development theory, they promote learning through participation and social interaction (Albertovna, 2025), while creating low-anxiety learning environments in which learners can practice language skills such as vocabulary and pronunciation. In resource-constrained contexts like Tanzania, local games are particularly valuable because they are accessible and inexpensive. They foster active participation, peer learning, and cultural relevance. Moreover, evidence from countries such as Indonesia and Turkey demonstrates that local games can improve vocabulary, speaking, and broader linguistic skills (Rusiana & Nuraeningsih, 2016; Kenanoğlu & Duran, 2021).

Theoretical Framework

This study was guided by Sociocultural Theory (SCT) developed by Vygotsky (1978), and the Game-Based Learning Instructional Model (GBLIM) developed by Vandercruysse and Elen (2017). The SCT views learning as a social and culturally mediated process. It emphasises interaction, language, collaboration, and scaffolding within the Zone of Proximal Development (ZPD) (Poehner & Lu, 2024). In this study, SCT explains how pupils developed their reading and writing skills through interaction with peers and teachers during local game-based activities. However, SCT offers limited guidance on how these processes can be systematically structured and implemented in an educational intervention.

GBLIM complements the SCT by providing a practical framework for organising game-based learning. This model comprises three elements namely the Game-Based Learning Environment (GBLE), the implementation context, and the learner. The GBLE provides space in which reading and writing activities take place. This element aligns with SCT, which explains that learning develops through social interaction and culturally meaningful activities. The implementation context refers to how games are used within the learning environment. Its contribution to this study lies in ensuring that games are not used randomly but are carefully planned and linked to specific learning goals. Through the implementation context, the researcher carefully planned how local games could be organised to support the development of English literacy skills. In addition to the GBLE and the implementation context, the learners is another important element of the model. This element considers the differences in pupils' abilities, confidence and prior knowledge that may influence learning. Each learner actively makes meaning through the game based on what they already know and how they engage with the activity. Prior knowledge also plays an important role in helping learners connect new information with what they already understand. Together the SCT and GBLIM are crucial to this study. SCT explains how learning occurs through social interaction, whereas GBLIM guides the systematic design and implementation of local games to enhance pupils' English reading and writing skills.

Research Gap

Despite the recognised benefits of game-based learning in English learning, significant gaps remain in the literature. First, existing empirical research has focused largely on digital games, while culturally relevant non-digital games remain comparatively underexplored (Herrero-Martín et al., 2025; Sancar-Tokmak & Bostancı, 2025; Vita-Barrull et al., 2024). Second, most research on non-digital game-based learning has been conducted outside Africa, particularly in European and Asian contexts. In Tanzania, research on local games has mainly addressed other learning areas; for example, Mtebe and Isingo (2024) developed local games to enhance primary pupils' numeracy skills, rather than



English literacy. Tanzanian studies on English reading have instead examined conventional interactive strategies such as group reading, collaborative discussion, and questioning (Ngonyani et al., 2025). Thus, empirical evidence on integrating culturally familiar local games such as rede and ruka kamba into English literacy instruction remains limited. Third, few studies have used the structured intervention designs to evaluate the effectiveness of local games across multiple literacy skills. To address these gaps, this study investigated how local games can enhance English literacy among Standard Three pupils in rural primary schools in Tanzania. Guided by SCT and GBLIM, the study sought to answer two research questions:

1. To what extent are structured local games-based intervention effective in developing overall literacy skills among Standard Three pupils in Tanzania?
2. How effective are local games in enhancing basic, intermediate, and advanced reading and writing skills?

This study makes several contributions. First, it provides empirical evidence on the extent to which well-structured local games-based instructional approaches can improve literacy outcomes. Furthermore, it provides evidence on whether local games can offer a cost-effective alternative to technology-dependent and culturally mismatched interventions in resource-constrained educational settings such as those in sub-Saharan Africa. Second, it offers teachers practical strategies for engaging learners through culturally meaningful classroom activities, potentially reducing reliance on rote instructional practices. Third, by documenting a replicable intervention model, the study provides insights that policymakers and educators across sub-Saharan Africa can adapt to support ongoing efforts to address the region's learning crisis. Finally, it responds directly to calls for contextually responsive pedagogies by demonstrating how local games can be systematically integrated into English instruction in ways that align with both curricular goals and learners' cultural experiences.

METHOD

The Design and Approach

This study formed part of a broader project grounded in Design-Based Research (DBR). The broader project aimed to explore how local games could be identified, pedagogically adapted, and iteratively refined to support the development of English literacy skills among primary school pupils. DBR was particularly suitable because it allows researchers to design and test educational innovations in authentic learning environments while simultaneously generating practical and theoretical insights into teaching and learning processes.

However, the present paper reports specifically on the intervention phase of the broader DBR project. To evaluate the effectiveness of the intervention, a quasi-experimental pre-test–post-test design was employed. The design involved two experimental groups and one control group. A quasi-experimental design was employed in this study because of its ability to allow the testing of instructional innovations under real classroom conditions (Creswell & Creswell, 2017; Hyman, 1982). Furthermore, this design allowed the researchers to examine changes in learning outcomes following an intervention while maintaining the natural classroom structure typically found in school settings.

In Tanzania, many primary school pupils experience difficulties in reading and writing English, particularly in rural areas where educational resources are often limited. However, conducting an intervention study across the entire country would have been time-consuming and financially demanding. Therefore, the study was conducted in Chamwino District located in Dodoma Region, Tanzania. Conducting the study in this district enabled close monitoring of the intervention and regular engagement with participating schools. Chamwino District was selected because the previous studies have documented significant literacy challenges among pupils in rural schools in the district (Ndijuye & Beatus, 2022). In addition, a preliminary literacy assessment conducted by the researchers prior to the intervention confirmed substantial difficulties in English reading and writing among Standard Three pupils in the district. The study focused specifically on rural wards to capture the realities of education



in remote learning environments. Three public primary schools, each located in a different ward, were purposively selected based on their relatively low performance in the preliminary literacy assessment.

The initial sampling frame comprised 2,201 registered Standard Three pupils in the selected public primary schools. From this population, a preliminary literacy assessment was conducted to establish baseline literacy proficiency levels. The required sample size for this preliminary assessment was determined using Cochran's sample size formula:

$$n = \frac{Z_{1-\alpha/2}^2 P(1-P)}{d^2}$$

Where $Z=1.96$ for a 95% confidence level, $p=.05$ (estimated population proportion), and $d=.05$ (margin of error) (Cochran, 1977). The formula yielded an initial sample size of 385 pupils.

However, due to practical constraints related to school organisation and the need to work with intact classes, the final achievable sample was adjusted to 328 pupils. These pupils were drawn from six participating schools, with one intact Standard Three class selected from each school. All 328 pupils participated in the preliminary literacy assessment conducted as part of the broader study. The 328 pupils therefore provided the baseline evidence used to inform the selection of participants for the subsequent intervention phase.

Following the preliminary assessment, a sub-sample was selected for the intervention phase. The selection was conducted purposively at the class level to maintain natural classroom settings and minimise disruption to school routines. Specifically, three intact classes ($n = 176$ pupils) were selected from the initial pool of 328 pupils based on comparable literacy performance as established during the preliminary assessment. This ensured that the groups included in the intervention were broadly comparable in baseline literacy proficiency.

The smaller intervention sample was considered appropriate because the intervention required sustained researcher engagement with classroom activities and systematic observation of pupils' participation in local game-based English instruction. The sample size was also constrained by the time available for implementing the intervention and the logistical demands of conducting the study across the participating schools. Thus, the reduction from 328 pupils in the preliminary assessment to 176 pupils in the intervention phase reflected a deliberate methodological and logistical decision rather than participant attrition.

The selected classes were then assigned to study conditions as follows: School A ($n = 59$) formed Experimental Group 1, School B ($n = 43$) formed Experimental Group 2, and School C ($n = 74$) served as the Control Group. The variation in group size was resulted from the differences in the number of eligible Standard Three pupils enrolled in each school. It was not caused by selective allocation or participant exclusion. The participants were drawn from different schools to minimise risk of contamination which could affect the findings. Such contamination could have occurred if the experimental and control groups had been selected from the same school. Overall, all subsequent statistical analyses evaluating the effects of the intervention reported in this study were based on the 176 pupils in the intervention sample.

The pupils' ages ranged from 8 to 11 years. At the time of the study, the pupils had received approximately six months of formal English instruction. They had already been introduced to several foundational literacy skills, including letter recognition and writing, basic word reading and writing, sentence construction, text reading accuracy, and reading comprehension. In addition to learning English, pupils were fluent in their local language Chigogo and had already developed reading and writing skills in Kiswahili, the national language used as the medium of instruction in most primary school subjects.

Intervention Process

The intervention was conducted over a 20-week period (120 minutes or 2 hours per week), during which pupils in experimental groups participated in structured English literacy activities embedded



within local games. This duration was determined based on the need for sufficient exposure to achieve meaningful literacy gains, as meta-analyses have identified intervention length as a significant moderating factor in student outcomes (Dai et al., 2025). Research suggests that moderate, sustained engagement, typically exceeding brief exposures but not extending beyond optimal thresholds, is preferable for foundational skill development (Lei et al., 2022). The 20-week timeframe provided sufficient cumulative exposure to support literacy development across multiple skill levels without disrupting regular instructional time.

Two experimental schools (School A and School B) received English literacy instruction integrated with adapted local games, while the control school (School C) continued receiving traditional English instruction without the use of local games. The two selected games were *rede* and *ruka kamba*, which were chosen from a wider set of local games because they were consistently identified by pupils as highly enjoyable and frequently played.

As part of the intervention design, two culturally familiar local games, *rede* and *ruka kamba*, were pedagogically adapted to integrate structured English language learning activities. These adaptations were intended to embed literacy instruction within authentic, play-based contexts while maintaining the core mechanics of the original games.

Rede is traditionally played with two players positioned outside the playing area who throw a ball towards players positioned within a defined playing area. The players inside the area attempt to dodge or catch the ball. In the present study, English language tasks were embedded within the game structure. This adaptation replaced the traditional end activity, in which players fill a bottle with sand, and ensured that gameplay was systematically linked to specific literacy learning objectives.

Similarly, *ruka kamba* involves two players swinging a rope while others attempt to jump continuously over the rotating rope. Players who fail to maintain the rhythm and are struck by the rope are typically eliminated from the round. In this study, elimination was re-conceptualised as a learning opportunity: pupils who were struck by the rope were required to complete specific English language tasks before rejoining the game. This adaptation allowed pupils to remain engaged in the activity while reinforcing targeted literacy skills.

The integration of English activities into both games was carefully structured to align with the study's instructional objectives, including letter recognition, word reading, and basic writing skills. All activities were standardised and administered consistently across sessions to ensure treatment fidelity. Teachers were provided with detailed guidelines on how and when to introduce the tasks during gameplay, and regular monitoring was conducted to ensure that the intervention was implemented as intended.

In this intervention, players who successfully caught and redirected the ball in *rede*, as well as those struck by the rope during *ruka kamba*, were required to complete specific English language activities (see Table 1). This approach ensured that participation in the games was directly and consistently linked to opportunities for literacy practice. These embedded activities were designed to target seven core literacy skills: Letter reading, read basic words, text oral reading accuracy, reading comprehension, letter writing accuracy, word writing accuracy, and text writing accuracy. By integrating literacy tasks into culturally familiar games, the intervention aimed to enhance pupil engagement while reinforcing foundational and higher- order literacy skills.

Table 1. Literacy activities embedded in Rede and Ruka Kamba games

Literacy component	<i>Rede</i> game activities	<i>Ruka kamba</i> game activities
Letter recognition	Pronounce, arrange, match, and write letters using wooden planks.	Identify, pronounce, spell, match, and write letters while playing <i>ruka kamba</i> game.

**Table 1 (Continued).** Literacy activities embedded in Rede and Ruka Kamba games

Literacy component	Rede game activities	Ruka kamba game activities
Word/vocabulary development	Form, spell, pronounce, match, and write simple words using wooden planks.	Read, spell, match, and write simple words during rope-skipping activities.
Sentence construction	Arrange words into meaningful sentences and practise reading and writing simple sentences.	Arrange scrambled words into sentences and practise reading and writing simple sentences during game play.
Reading comprehension	Match sentences with pictures, answer comprehension questions, and sequence simple stories.	Follow oral instructions, identify key ideas, arrange picture sequences, and answer comprehension questions.

Source: Researcher's construct (2026).

Procedures

To measure pupils' literacy development, the researchers developed an English Language Literacy Assessment Test, which was administered before (pre-test) and after (post-test) intervention. The assessment consisted of two components: reading literacy and writing literacy. The reading assessment included four tasks, namely, letter reading, read basic words, text oral reading accuracy, and reading comprehension, whereas, the writing assessment included three tasks: Letter writing accuracy, word writing accuracy, and text writing accuracy. Each task was scored on a scale of 0 to 100, allowing for comparisons across literacy domains.

The construction of the two assessment sets involved three phases: item creation, content validation, and piloting. Previous studies and assessment of English language literacy, including (USAID, 2014; Uwezo Tanzania, 2019), as well as several school examinations, were reviewed to guide item development (Downing & Haladyna, 2006). Furthermore, the Standard Three pupils' English textbook (TIE, 2023) was reviewed to inform the content selection. Content validity which evaluates whether a test adequately measures the intended content construct (Obilor & Miwari, 2022) was determined by a panel of six seasoned English teachers and two experts in applied linguistics. After incorporating feedback from teachers and experts feedback, the reading and writing assessments were piloted with 102 pupils aged 8 to 9 years in a different area outside the main study sample. This pilot investigation sought to assess the clarity and understandability of the assessment instructions, the time required to complete the assessment, and the consistency of test administration. The findings indicated that the assessment sets were sufficiently robust and appropriate for assessing the targeted literacy skills. The assessment instructions were clear and easy for the pupils to understand. Moreover, the assessment duration was considered appropriate, suggesting that the test batteries were balanced in length and allowed for comprehension assessment without causing excessive participant fatigue.

Two types of reliability were assessed: test-retest and internal consistency reliability. Test-retest reliability was evaluated one week after the initial test, with around 45% of participants (n=46) randomly selected for reassessment. The reliability of scores between the initial and subsequent testing was assessed using Pearson correlation coefficients, which ranged from .72 to .98 across different tasks, indicating acceptable to excellent stability over time. The analysis of internal consistency reliability for the reading and writing assessment sets yielded a Cronbach's alpha values of .873. As stated by George and Mallery (2024), this value falls within the range indicative of strong internal consistency among the items.

The assessments were administered under conditions tailored to the nature of each task. Reading tasks were conducted individually in a calm environment to minimise distractions and anxiety, allowing pupils to focus entirely on the task. The duration of these sessions ranged from 10 to 38 minutes, depending on individual pupils' skill levels. In contrast, writing tasks were carried out within the classroom, where all pupils completed them simultaneously under standardised conditions to ensure consistency. Each writing session was completed within a fixed duration of 45 minutes.



Test administration was led by three qualified research assistants, the experienced English teachers familiar with the pupils' context. Two English teachers were responsible for ensuring that the procedures were followed consistently and efficiently across all sessions in the two intervention schools. To promote fairness and reduce individual bias, the two English teachers who participated in the intervention jointly assessed the pupils' scripts with the researcher, enabling a reliable and consistent evaluation of pupils' performance. Moreover, the third research assistant was responsible for observing the intervention and recording data to assess implementation fidelity.

Fidelity of implementation was monitored throughout the intervention period to assess the extent to which the literacy programme was delivered as intended. Implementation fidelity refers to the degree to which an intervention is delivered in accordance with its planned procedures (Mori et al., 2025). An observation checklist was developed, comprising 78 key indicators across three core dimensions: adherence to the implementation guide, duration of sessions, and quality of delivery. During implementation, one researcher was assigned the role of assessing how teachers implemented the intervention process and recorded the fidelity data.

Data Analysis

All data were analysed quantitatively using Stata. Descriptive statistics, including means, standard deviations, medians, interquartile ranges, and minimum and maximum values, were first computed to provide an overview of pupils' literacy performance across the three groups at both pre-test and post-test stages.

To ensure that any observed changes in literacy outcomes could be reasonably attributed to the intervention, implementation fidelity was assessed before conducting the inferential analyses. A structured observation checklist comprising 78 indicators across three dimensions—adherence to the implementation guide, session duration, and quality of delivery—was used. Each indicator was scored as fully implemented (1), partially implemented (.5), or not implemented (0), and an overall fidelity percentage was calculated. This step helped establish the consistency and quality of intervention delivery across the participating schools.

Subsequently, preliminary analyses were conducted to examine the assumptions underlying the statistical analyses. Normality was assessed using the Shapiro–Wilk test, and homogeneity of variances was evaluated using Levene's test with the Brown–Forsythe correction. As most literacy variables violated these assumptions, non-parametric methods were used for the subsequent inferential analyses.

Guided by the research questions, the analysis proceeded in two stages. First, to determine the extent to which the intervention improved overall literacy performance and address the first research question, the Wilcoxon Signed-Rank Test was employed to compare pre-test and post-test scores within each group. This test was suitable for detecting within-group changes in paired, non-normally distributed data. Second, to examine the differences in literacy outcomes across groups and skill levels in addressing the second research question, the Kruskal–Wallis test was applied. Where statistically significant differences were observed, Bonferroni-adjusted pairwise comparisons were conducted to identify the specific groups accounting for the observed differences, while controlling for Type I error. To complement these analyses, effect sizes (ϵ^2) were calculated to determine the magnitude and practical significance of the observed differences.

RESULTS

Preliminary Analyses: Computation of Implementation of Fidelity

Before conducting the main statistical analyses, the fidelity of implementation was examined to determine the extent to which the intervention was delivered as intended. The purpose of this analysis was to ensure that any observed improvements in pupils' literacy outcomes could reasonably be attributed to the intervention rather than variations in its implementation. The fidelity score was calculated by dividing the weighted total implementation score by the total number of expected behaviours and multiplying the result by 100.



A total of 62 behaviours were fully implemented ($62 \times 1 = 62$), while 13 behaviours were partially implemented ($13 \times 0.5 = 6.5$), resulting in a weighted implementation score of 68.5. Given that the total number of expected behaviours was 78, the overall fidelity score was calculated as:

$$\text{Fidelity Score} = \left(\frac{68.5}{78} \right) \times 100 = 87.8\%.$$

This result indicates a high level of adherence to the planned intervention procedures, suggesting that the local game-based activities were implemented consistently across the intervention period.

Assumption Testing

Prior to addressing the main research questions, preliminary analyses were conducted to examine the distributional properties of the literacy data and determine whether the assumptions required for parametric statistical tests were satisfied. Normality of the literacy scores was assessed using the Shapiro–Wilk test, while homogeneity of variances across groups was examined using Levene’s test with the Brown–Forsythe correction.

As shown in appendices A to C, the results indicated that most literacy measures at both pre-test and post-test stages deviated significantly from normality across the three groups. At the pre-test stage, nearly all literacy variables showed significant departures from normality, with the exception of letter writing accuracy in Experimental Group 1. In addition, text writing accuracy in Experimental Group 2 exhibited zero variance at the pre-test stage, which precluded the calculation of a normality statistic for this variable.

Furthermore, at the post-test stage, the distributional properties improved slightly for some variables. For example, letter reading met the normality assumption across all three groups, while reading comprehension and letter writing accuracy in Experimental Group 2 approximated normality. However, the majority of literacy measures continued to violate the normality assumption, particularly text writing accuracy across all groups. The homogeneity of variances tests also revealed significant differences in variance across groups for most literacy variables at both testing points (see appendices B and C). Given the widespread violations of both normality and homogeneity assumptions, non-parametric statistical procedures were therefore used in the subsequent analyses.

Impact of Game-Based Intervention on Overall Literacy Outcomes

To address the first research question aimed at evaluating the effectiveness of the structured game-based intervention, several statistical analyses were conducted to assess both the magnitude and practical significance of changes in literacy performance. As shown in Table 2, the analyses compared the pupils’ pre-test and post-test performance within each group, provided graphical representation of the results, and calculated the effect sizes to evaluate the practical significance of the intervention.

Table 2. Performance across various components of English reading and writing skills before and after intervention

Variable	Type of Test	Control Group (CG)	Experimental Group 1 (Exp1)	Experimental Group 2 (Exp2)
Letter sounds	Pre	M = 37.30 (SD = 35.97)	M = 20.34 (SD = 25.86)	M = 21.63 (SD = 22.88)
	Post	M = 46.08 (SD = 27.59)	M = 60.68 (SD = 29.47)	M = 71.40 (SD = 20.99)
Read basic words	Pre	M = 20.81 (SD = 25.74)	M = 6.44 (SD = 16.37)	M = 25.12 (SD = 25.30)
	Post	M = 27.03 (SD = 20.12)	M = 40.00 (SD = 27.85)	M = 57.44 (SD = 28.46)
Text oral reading accuracy	Pre	M = 7.03 (SD = 10.43)	M = 1.02 (SD = 4.02)	M = 4.88 (SD = 13.69)
	Post	M = 8.51 (SD = 11.90)	M = 18.64 (SD = 21.45)	M = 39.53 (SD = 27.43)
Reading comprehension	Pre	M = 5.27 (SD = 6.02)	M = 1.02 (SD = 4.43)	M = 2.09 (SD = 4.12)
	Post	M = 17.43 (SD = 17.70)	M = 11.19 (SD = 18.53)	M = 30.58 (SD = 19.52)
Letter writing accuracy	Pre	M = 15.54 (SD = 19.94)	M = 1.36 (SD = 5.07)	M = 15.81 (SD = 20.84)
	Post	M = 49.46 (SD = 30.92)	M = 61.02 (SD = 35.07)	M = 64.65 (SD = 25.67)



Table 2 (Continued). Performance across various components of English reading and writing skills before and after intervention

Variable	Type of Test	Control Group (CG)	Experimental Group 1 (Exp1)	Experimental Group 2 (Exp2)
Word writing accuracy	Pre	M = 2.16 (SD = 6.03)	M = 1.02 (SD = 4.43)	M = 1.87 (SD = 6.99)
	Post	M = 3.78 (SD = 5.90)	M = 35.25 (SD = 30.86)	M = 25.58 (SD = 26.57)
Text writing accuracy	Pre	M = 0.41 (SD = 2.59)	M = 0.51 (SD = 2.89)	M = 0.00 (SD = 0.00)
	Post	M = 0.14 (SD = 1.16)	M = 21.69 (SD = 23.42)	M = 19.53 (SD = 19.99)

Source: Field data (2026).

Note: Exp 1 = Experimental Group 1, Exp 2 = Experimental Group 2. Values outside the parenthesis = Mean (M), values inside the parenthesis = Standard Deviation (SD). In Exp 1 and Exp 2 pupils' English instruction was integrated with local games *rede* and *ruka kamba*.

The findings in Table 2 demonstrated that the local games were highly effective. Pupils in the experimental groups showed significant improvements across all literacy domains. The largest gains were observed in advanced reading and writing skills. In contrast, the control group exhibited only limited progress. Specifically, at pre-test stage (see Table 2), pupils across all three groups demonstrated low levels of English literacy performance. In Experimental Group 1, the mean scores were low across all reading components. Pupils achieved a mean score of 20.34 in letter reading and 6.44 in read basic words. Higher-level skills showed even lower performance, with mean scores of 1.02 for both text oral reading accuracy and reading comprehension. The median score for both components was 0. For writing, the performance was limited with letter writing accuracy (M = 1.36) and text writing accuracy (M = .51). Additionally, Experimental Group 2 demonstrated higher scores in letter reading (M = 21.63) and read basic words (M = 25.12). Performance on the higher-level reading and writing tasks remained minimal, with median scores of 0 for text oral reading accuracy, reading comprehension, and text writing accuracy.

In the Control Group, pre-test scores were somewhat higher for letter reading (M = 37.30) and read basic words (M = 20.81). However, higher-order literacy skills remained weak, with text oral reading accuracy (M = 7.03) and text writing accuracy (M = .41). Despite of some variation across groups, the overall pre-test results suggest that pupils entered the intervention phase with broadly comparable literacy abilities, particularly in the more advanced reading and writing tasks.

Following the intervention, improvements were evident across most literacy domains in the experimental groups. In Experimental Group 1, letter reading increased from 20.34 at pre-test to 60.68 at post-test, while read basic words improved from 6.44 to 40.00. Similarly, Text Oral Reading Accuracy increased from 1.02 to 18.64, and letter writing accuracy rose from 1.36 to 61.02, with text writing accuracy reaching 21.69 at post-test.

Experimental Group 2 demonstrated a similar pattern of improvement. letter reading increased from 21.63 to 71.40, while read basic words rose from 25.12 to 57.44. Text oral reading accuracy improved markedly from 4.88 to 39.53, and reading comprehension increased from 2.09 to 30.58. Writing skills also improved, with letter writing accuracy rising from 15.81 to 64.65 and text writing accuracy increasing from 0 to 19.53.

In contrast, the Control Group showed smaller improvements. letter reading increased from 37.30 to 46.08, while read basic words rose from 20.81 to 27.03. Text oral reading accuracy increased only slightly from 7.03 to 8.51, and word writing accuracy improved marginally from 2.16 to 3.78, while text writing accuracy remained very low (.41 to .14). Generally, the descriptive results indicate that pupils who participated in the local games-based instruction suggested a greater improvement in both reading and writing literacy compared with those who received traditional instruction.

Pre-Post Changes in Literacy Skills Across Groups

To examine changes in literacy skills within each group, Wilcoxon signed-rank tests were conducted to compare pre-test and post-test scores for the seven literacy components as presented in Table 3.



Table 3. Pre–post changes in reading and writing literacy scores across groups based on Wilcoxon signed-rank tests

Literacy aspect	Exp 1 (n=59) Wilcoxon z	p	Exp 2 (n=43) Wilcoxon z	p	Control G (n=74) Wilcoxon z	p
Letter reading	5.51	< .001	5.46	< .001	2.82	.005
Read basic words	6.02	< .001	5.26	< .001	3.53	< .001
Text oral reading accuracy	5.81	< .001	5.33	< .001	1.18	.24
Reading comprehension	4.51	< .001	5.67	< .001	5.25	< .001
Letter writing accuracy	6.59	< .001	5.55	< .001	6.87	< .001
Word writing accuracy	6.18	< .001	4.97	< .001	2.73	.006
Text writing accuracy	5.71	< .001	5.5	< .001	-.59	.558

Source: Field data (2026)

Note: n = number of participants, p = p- value, Exp 1 = Experimental Group 1, Exp 2 = Experimental Group 2

As presented in Table 3, both experimental groups suggested statistically significant improvements across all literacy components ($z = 4.51$ – 6.59 , all $p < .001$). In Experimental Group 1, significant gains were observed in letter reading, read basic words, text oral reading accuracy, reading comprehension and all writing-related skills. Similarly, Experimental Group 2 demonstrated significant improvements across the same seven literacy domains, indicating consistent learning gains following the intervention.

In contrast, the control group showed significant improvements in only some literacy components, including letter reading, basic word reading, reading comprehension, letter writing accuracy, and word writing accuracy. However, no statistically significant improvements were observed in text oral reading accuracy or text writing accuracy among pupils in the control group. These results suggest that while some literacy development occurred in the control group, the improvements were relatively limited and did not extend to more complex literacy tasks such as continuous text reading and writing.

Regarding the visual comparison of reading performance, Figure 1 below presents the mean z-scores of reading skills across experimental and control groups. The analysis of the mean z -scores for reading skills was included to standardise the scores across reading components. By analysing the mean z-score facilitated a direct comparison of pupils' relative performance across different reading literacy domains.

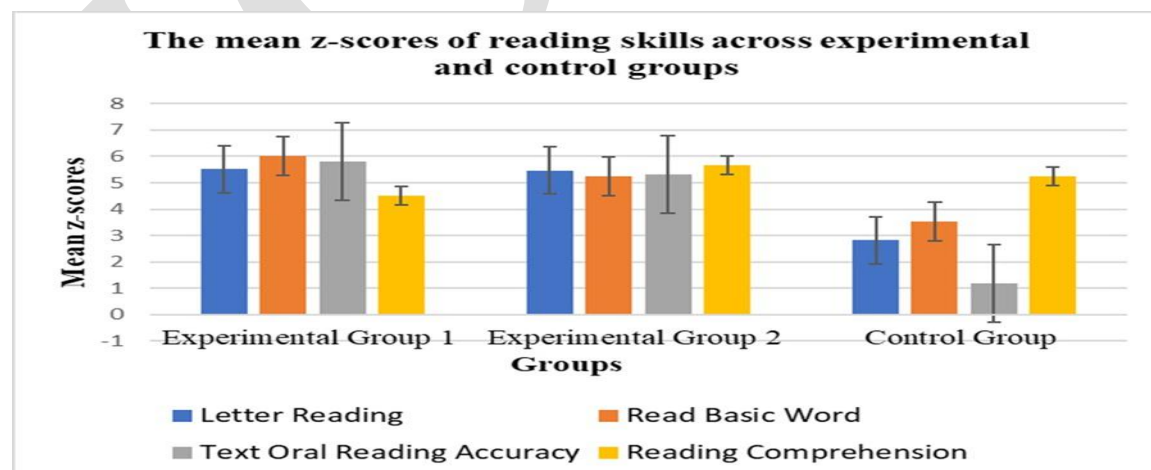


Figure 1. The mean z-scores of reading skills across experimental and control groups

Figure 1 illustrates the relative magnitude of improvements in reading-related literacy skills, including letter reading, read basic words, text oral reading accuracy, and reading comprehension. Both



experimental groups suggested substantially larger improvements compared with the control group, particularly in text oral reading accuracy and reading comprehension.

With regard to the visual comparison of the writing skills as shown in Figure 2, experimental groups shown larger improvements than the control group.

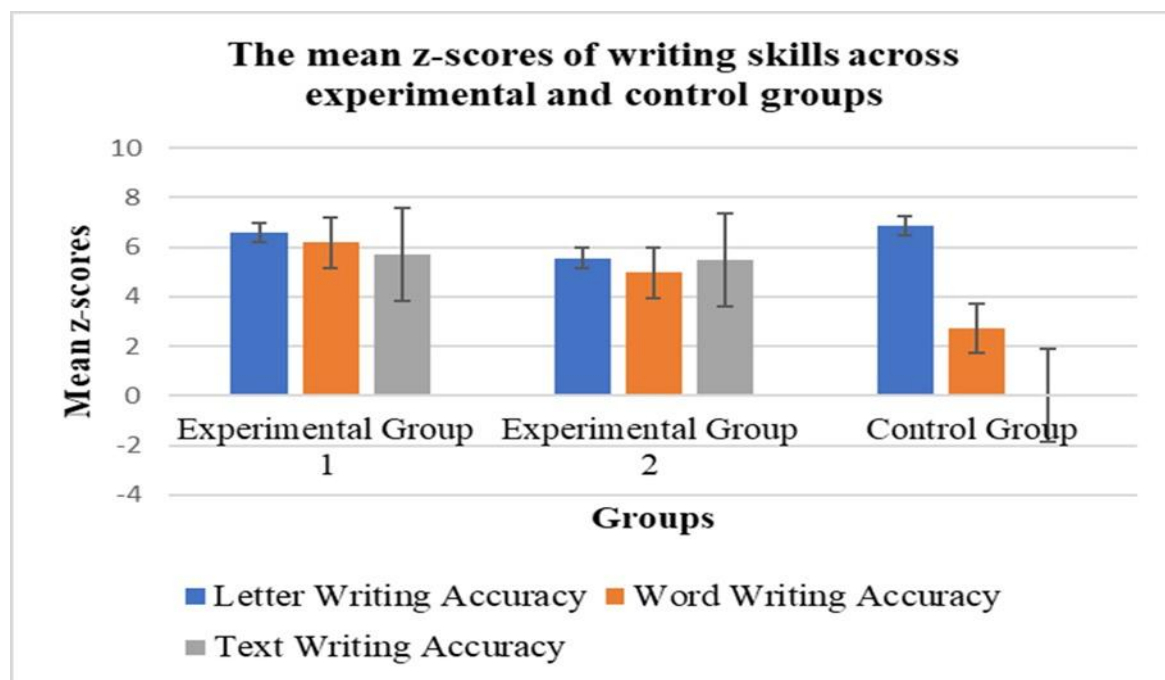


Figure 2. The mean z-scores of writing skills across experimental and control groups

Figure 2 presents the corresponding patterns for writing-related literacy skills, including letter writing accuracy, word writing accuracy, and text writing accuracy. Again, the experimental groups showed consistently larger improvements than the control group across all writing domains. These visual patterns are consistent with the statistical results, suggesting that the integration of local games into English instruction was associated with stronger literacy gains among pupils in the experimental groups.

Effect Sizes of Literacy Improvements

To assess the practical significance of the observed differences (see table 4 and Figure 3), effect sizes (ϵ^2) were calculated for the Kruskal–Wallis tests comparing the three groups. The results are presented in Table 3. At the pre-test stage, effect sizes were generally small or negligible, indicating that differences between the groups were relatively limited prior to the intervention.

However, at the post-test stage, larger effect sizes were observed across several literacy components, suggesting substantial improvements associated with the intervention. In particular the components that shown the large effect size include text oral reading accuracy ($\epsilon^2 = .25$), word writing accuracy ($\epsilon^2 = .34$), and text writing accuracy ($\epsilon^2 = .44$). These results indicate that the intervention had a strong impact on higher-order literacy skills involving connected-text reading and writing.

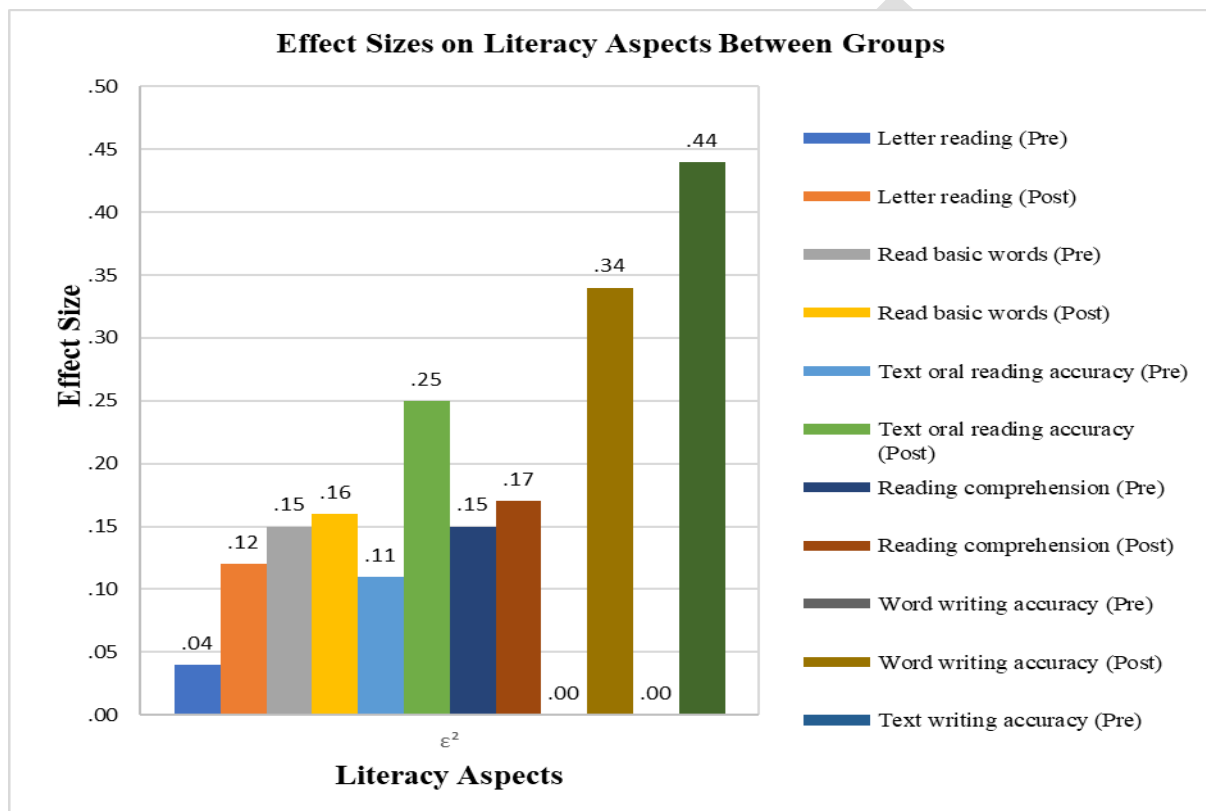
Table 4. Kruskal–Wallis test effect size results for literacy skills at pre-test and post-test

Literacy Aspect	Test	χ^2 (with ties)	ϵ^2	Magnitude
Letter reading	Pre	8.94	.04	Small
Letter reading	Post	22.77	.12	Medium
Read basic words	Pre	28.55	.15	Medium
Read basic words	Post	29.16	.16	Medium
Text oral reading accuracy	Pre	21.67	.11	Medium


Table 4 (Continued). Kruskal–Wallis test effect size results for literacy skills at pre-test and post-test

Literacy Aspect	Test	χ^2 (with ties)	ϵ^2	Magnitude
Text oral reading accuracy	Post	44.98	.25	Large
Reading comprehension	Pre	27.85	.15	Medium
Reading comprehension	Post	31.77	.17	Medium–Large
Word writing	Pre	1.92	.00	Negligible
Word writing	Post	60.60	.34	Large
Text writing	Pre	1.39	.00	Negligible
Text writing	Post	77.39	.44	Large

Source: Field data (2026).


Figure 3. Effect sizes of literacy aspects

Differential Effects of Local Games Across Literacy Skill Levels

This section addresses the second research question, which examined the effectiveness of local games in enhancing basic, intermediate, and advanced reading and writing skill levels. The analysis compared the outcomes of the two experimental groups, each of which integrated a different local game, with those of the control group at the pre-test and post-test stages. Kruskal–Wallis tests with Bonferroni-adjusted pairwise comparisons were conducted to compare Experimental Group 1 (Exp2), Experimental Group 2 (Exp2), and the control groups (see Table 5) at both pre-test and post-test across the seven literacy components.

At pre-test, the letter reading scores did not differ significantly across the three groups. Regarding post-test, both experimental Group 1 and Experimental Group 2 showed higher scores than the control group, indicating improvements on basic reading performance after intervention. Similar patterns were observed in basic word reading, where both experimental groups showed marked gains and clearly outperformed the control group following the intervention. At the text level, substantial improvements were also evident. In text oral reading accuracy, both experimental groups demonstrated notable enhancements in reading performance, further confirming the effectiveness of



local games in supporting more complex reading processes. Reading comprehension, a key indicator of higher-order literacy skills, showed significant gains as well, with Experimental Group 2 in particular displaying noticeable improvement in post-test results.

Table 5. Bonferroni post-hoc comparisons for literacy outcomes (n=176)

Variable	Type of Test	F (2, 173)	p-value	Significant Bonferroni Comparisons (Mean Rank Difference, 95 % CI)
Letter reading	Pre	4.66	< .001	CG vs Exp1: 25.03 [4.63, 45.42]
	Post	12.93	< .001	CG vs Exp1: -26.17[-46.21, -6.13]; CG vs Exp2: -44.85 [-66.87, -22.83]
Read basic words	Pre	16.86	< .001	Exp2 vs Exp1: 46.67 [25.60, 67.74]; Exp3 vs Exp1: 34.60 [16.26, 52.94]
	Post	17.3	< .001	Exp2 vs Exp1: 30.49 [7.99,52.99]; CG vs Exp 2: -52.27 [-73.79, -30.75]
Text oral reading accuracy	Pre	12.22	< .001	CG vs Exp 1:30.61 [15.38, 45.83]; CG vs Exp 2: -63.47 [-83.30, -43.63]
	Post	29.92	< .001	Exp 2 vs Exp 1: 39.96 [19.22, 60.70]; CG vs Exp 2: -63.47 [-83.30, -43.63]
Reading comprehension	Pre	16.37	< .001	CG vs Exp 1: 35.47 [20.10, 50.84]; CG vs Exp 2: 23.81 [6.93, 40.70]
	Post	19.18	< .001	Exp 2 vs Exp 1: 56.46 [34.41, 78.50]; CG vs Exp 2: -34.00 [-55.08, -12.92]
Letter writing accuracy	Pre	21.59	< .001	Exp 2 vs Exp 1: 41.15 [21.98, 60.31]; CG vs Exp1: 41.68 [25.00, 58.36]
	Post	4.56	.012	CG vs Exp 1: -22.48 [-43.42, -1.54]
Word writing accuracy	Pre	0.96	.386	None
	Post	45.83	< .001	CG vs Exp 1: -62.62 [-79.41, -45.84]; CG vs Exp 2: -50.17 [-68.61, -31.73]
Text writing accuracy	Pre	0.69	.503	None
	Post	68.57	< .001	CG vs Exp 1: -58.24 [-72.59, -43.88]; CG vs Exp 2: -64.04 [-79.81, -48.27]

Note: Exp1 = Experimental Group 1, Exp 2= Experimental Group 2, CG = Control group.

Moreover, the positive effects of local games extended beyond the reading to writing skills. That is to say, letter writing accuracy also differed significantly across the groups at post-test, with both experimental groups outperforming the control group. This suggests that the intervention supported the development of foundational writing abilities. Although word writing accuracy showed no significant differences among groups at the pre-test stage, post-test results revealed higher scores among the experimental groups, reflecting gains observed following the intervention period. Similarly, text writing accuracy demonstrated impressive improvements, with both experimental groups significantly outperforming the control group in the post-test. Overall, the findings highlight the potential of local games to strengthen basic, intermediate, and advanced reading and writing skills. The results suggest that integrating local games into literacy instruction can create an engaging learning environment that supports broader literacy development among Standard Three pupils.

DISCUSSION, CONCLUSION, and RECOMMENDATIONS

This study investigated the English literacy performance of primary school pupils following a five-month intervention that integrated the Tanzanian local games into English instruction. Specifically, the games *rede* and *ruka kamba* were incorporated into structured English language activities aimed at supporting the development of pupils' reading and writing skills through play-based learning. The study addressed two research questions: (1) What is the effect of a structured intervention integrating *rede* and *ruka kamba* on the development of English literacy skills? and (2)



Does the effectiveness of local game-based instruction vary across different levels of literacy skill complexity?

In relation to the first research question, the findings indicated that pupils who participated in the local games-based intervention achieved better literacy outcomes than those who received traditional instruction. At pre-test, pupils in both experimental and control groups demonstrated low levels of literacy performance, particularly on more complex reading and writing tasks. At post-test, however, pupils in both experimental groups performed better than those in the control group across the assessed literacy components. These findings suggest that the structured integration of *rede* and *ruka kamba* into English instruction contributed to improved overall literacy performance.

The differences between pre-test and post-test performance may be related to the opportunities for active participation, repeated practice, and social interaction provided by the intervention. The literacy tasks embedded within the familiar local games, allowed pupils to practice English through structured and interactive activities rather than relying on conventional classroom exercises. These findings are consistent with previous studies showing that play-based and game-supported approaches can enhance learner engagement, motivation, and cognitive involvement in literacy learning, particularly among young learners and second-language learners (Kangas et al., 2023; Nicolopoulou et al., 2015; Plass et al., 2020). The use of familiar activities may also have supported pupils' participation by providing a familiar social and learning environment in which they could engage with English.

The fidelity of implementation further strengthens the interpretation of the intervention findings. The intervention achieved an implementation fidelity level of 87.8%, while classroom observations indicated that teachers implemented most of the prescribed instructional behaviours either fully or partially. This consistency suggests that the intervention was delivered largely as intended reducing the likelihood that substantial differences in instructional delivery accounted for the observed group differences (Kingsbery et al., 2025; Mori et al., 2025). Nevertheless, fidelity data alone cannot establish which individual features of the intervention produced the observed improvements. Possible mechanisms include repeated practice, peer interaction, learner engagement, movement, and collaborative language use. Further is needed to determine the relative contribution of these mechanisms to pupils' literacy development (Gee, 2003; Mayer, 2021).

The magnitude of the observed differences also provides important evidence in relation to the first objective. Pre-test effect sizes were negligible to small, whereas post-test analyses showed medium to large effects across the assessed literacy domains. Particularly strong effects were observed in text oral reading accuracy and reading comprehension. This pattern suggests that the intervention was associated with meaningful differences in literacy performance that extended beyond basic recognition skills. These findings are consistent with research suggesting that game-based learning can produce stronger learning outcomes when game activities are deliberately aligned with the specific instructional objectives (Mayer, 2021; Plass et al., 2020).

With regard to the second specific objective, the study examined the effectiveness of local games in enhancing basic, intermediate, and advanced reading and writing skills. The findings showed that the effects of the intervention differed across various literacy components. Improvements were observed in foundational skills, connected-text reading and writing, and higher-order skills such as reading comprehension and text writing. This variation suggests that the effectiveness of local games should not be assessed solely in terms of overall literacy performance but also in relation to the complexity of the specific literacy skills being targeted.

At the basic level, the intervention targeted skills such as letter-reading knowledge, read basic words, and letter writing accuracy. Pupils initially demonstrated weaknesses in these areas. The repeated practice provided through the games may have helped pupils consolidate their foundational literacy knowledge. Because the activities gave pupils repeated opportunities to respond and practice, they could strengthen basic literacy skills while participating in meaningful and interactive classroom activities. This finding is consistent with previous research showing that game-based learning can



provide opportunities for repeated practice while sustaining pupils' engagement in learning (Kangas et al., 2023; Plass et al., 2020).

At the intermediate level, the intervention involved skills requiring pupils to move beyond isolated letters and words towards connected language use. Text oral reading accuracy, for example, requires pupils to coordinate word recognition, decoding, fluency, and oral language processing. The relatively strong effects observed for this component may reflect the opportunities for repeated oral language practice created by the interactive nature of the games. Peer participation may have provided additional opportunities for pupils to hear and produce English during activities. Such opportunities are important in second-language contexts where exposure to English may be limited outside the classroom.

The findings for advanced literacy skills were particularly notable. The reading comprehension and text writing accuracy require pupils to integrate several linguistic and cognitive processes. Reading comprehension involves constructing meaning from connected text, whereas text writing requires pupils to generate, organize, and express ideas in written language. The experimental groups showed substantial improvement in these areas, with strong effects observed for reading comprehension and text writing accuracy. This pattern suggests that the intervention provided opportunities for pupils to engage with English beyond basic decoding and word-level recognition and to use language in more complex ways.

The stronger effects observed for some advanced skills contrast with meta-analytic findings reported by Kim et al. (2019), which showed larger effects for foundational literacy skills and smaller effects for reading comprehension and oral language outcomes. The difference may be partly explained by contextual and pedagogical factors. Many interventions examined by Kim et al. involved digital or skills-focused games, whereas the present study employed culturally familiar, non-digital games. *Rede* and *ruka kamba* incorporated social interaction, movement, repetition, and collaborative participation. These features may have created opportunities for extended language use, interaction, and meaning-making, which are relevant to the development of more complex literacy skills.

Differences between the two experimental groups across several literacy components provide further insight into the second objective. Although both games produced better outcomes than the control condition, their effects were not identical across all skills. This pattern suggests that game structure and the demands of the embedded literacy tasks may influence the type of learning it supports. The differences in movement, turn-taking, repetition, sequencing, and peer interaction may create different opportunities for practicing reading and writing processes. This interpretation is consistent with research emphasizing the importance of aligning game characteristics and instructional activities with specific learning objectives (Mayer, 2021; Plass et al., 2020).

The findings can also be interpreted through a sociocultural perspective. The cultural familiarity of *rede* and *ruka kamba* may have facilitated pupils' participation because pupils understood the basic activities and social associated with the games. This familiarity could have allowed pupils to devote greater attention to the English literacy tasks embedded within the activities. The games may also have functioned as culturally meaningful mediational tools that enabled pupils to interact with peers, receive assistance, and practice language collaboratively. Such processes are consistent with sociocultural perspective on learning, which emphasize participation, interaction, scaffolding, and culturally mediated activity (Rogoff, 2003).

The multidimensional assessment used in the study further clarifies the pattern of effects across literacy levels. The assessment distinguished between letter sounds, basic word reading, text oral reading accuracy, reading comprehension, letter writing accuracy, word writing accuracy, and text writing accuracy. Examining these components separately revealed that the magnitude of the intervention effects differed according to the specific literacy process being assessed. The stronger effects for some higher-order skills may reflect the greater opportunities the games provided for extended language use, social interaction, and meaning-making.



Taken together, the findings related to the second objective indicate that local games were associated with improvements across different levels of literacy complexity, although the magnitude of the effects varied across individual skills. The variation appears to be related not only to the level and nature of the literacy skills involved but also to the structure of the game and the linguistic and cognitive demands of the associated tasks. Thus, the findings support the need to examine local games in relation to the specific literacy skills they are intended to develop, rather than treating game-based instruction as a uniform intervention.

Contribution of the Research

The findings of this study provide empirical evidence that the structured integration of Tanzanian local games can improve English literacy outcomes among Standard Three pupils in rural public primary schools. In particular, the use of *rede* and *ruka kamba* was associated with improved literacy performance. The intervention was associated with statistically significant and educationally meaningful gains across the assessed literacy domains. Significantly, the strongest effects were not limited to foundational skills but also extended to more complex areas such as oral reading accuracy, reading comprehension, and text writing accuracy. This pattern suggests that culturally familiar games can provide opportunities for pupils to engage with English through repeated practice, social interaction, and meaningful language use in interactive learning contexts.

The findings also offer a broader perspective on educational innovation in resource-constrained educational settings. They show that local cultural practices need not be viewed as barriers to formal learning; instead, they can be drawn upon as meaningful resources for classroom instruction. Familiar play activities can be transformed into structured learning experiences that build on pupils' existing cultural knowledge while introducing new academic content. This asset-based perspective is consistent with sociocultural approaches to learning, which emphasize the role of social interaction, cultural tools, and learners' existing experiences in the development of knowledge and skills (Rogoff, 2003). The study therefore provides a contextually grounded contribution to the growing literature on game-based learning and early literacy development.

Limitations of the Research

This study has five limitations that should be considered when interpreting the findings. First, the study was conducted in a single district, Chamwino District, in Dodoma Region, Tanzania. The study population was drawn primarily from the Gogo community. Therefore, the findings may not be generalisable to other learning contexts with different sociocultural, linguistic, and educational characteristics. Any transfer of the findings to such contexts should therefore be made with caution. Second, the study focused on rural public primary schools and a specific sample of participants, which may limit the representativeness of the findings within Tanzania more broadly. Caution is therefore needed when extending the findings to urban schools or other educational settings. Moreover, the study involved only Standard Three pupils and English language teachers, limiting the applicability of the findings to other educational levels and groups of educational stakeholders.

The study was also limited by the duration and scope of the intervention. Because the study did not include a sufficiently long follow-up period, it was not possible to determine whether the observed gains were sustained over time. Furthermore, the study focused on reading and writing skills and did not assess listening or speaking. This limitation restricts the extent to which the findings can be interpreted as evidence of overall English language proficiency. Future research could address these limitations by using larger and more diverse samples, including participants from different educational levels and sociocultural contexts, extending the intervention period and follow-up, and examining all four English language skills.

Conclusions

In conclusion, this study demonstrates that integrating culturally familiar local games, particularly *rede* and *ruka kamba*, into English literacy instruction can support improved reading and writing outcomes among rural Tanzanian primary school learners. The results revealed marked and statistically significant gains in pupils' literacy performance, particularly in oral reading, reading



comprehension, and text writing, while the high implementation fidelity (87.8%) provides further support for the consistency of the intervention delivery. These gains indicate that local games can support the development of literacy skills beyond basic letter and word recognition. Generally, the findings suggest that non-digital game-based instruction offers a practical, low-cost, and contextually relevant approach to strengthening early English literacy in resource-constrained settings. However, further research involving more diverse sociocultural contexts and longer follow-up periods is needed to determine the broader applicability and long-term effects of this approach.

Recommendations

Based on the findings, teachers, education policymakers and curriculum developers should consider the integration of local games useful for developing English literacy skills, particularly in resource-constrained rural schools. Teachers should align the game activities with specific reading and writing skills. Moreover, the professional development and locally relevant instructional materials should also be strengthened to support effective implementation and adaptation of game-based activities. In addition, further research should be conducted to examine the local game-based English approach across diverse sociocultural contexts and over longer follow-up periods to determine the sustainability and broader applicability of the observed literacy gains.

Funding

No funding was received from any individual or institution for this research.

Ethics and Conflict of Interest

As the authors of this article, we have adhered to the ethical guidelines throughout the research process, and there are no conflicts of interest among us. The authors declare that this manuscript is original, has not been published or submitted elsewhere, and that all authors approved its final version. The research complied with ethical principles and there are no conflicts of interest among the authors.

Author Contribution

All authors contributed equally to the research. All authors read and approved of the final manuscript.

Data Availability

The data that support the findings of this study are available on request from the corresponding author.

Corresponding Author

Correspondence to Wilfred MADATA, wilfdata2022@gmail.com

REFERENCES

- Alotaibi, M. S. (2024). Game-based learning in early childhood education: a systematic review and meta-analysis. *Frontiers in Psychology*, 15(April). <https://doi.org/10.3389/fpsyg.2024.1307881>
- Cochran, W. G. (1977). *Sampling Techniques*. 3rd Edition. John Wiley & Sons, New York.
- Creswell, J., & Creswell, J., (2017). *Research design: Qualitative, quantitative, and mixed methods approach*. Sage publications.
- Dai, W. A., Xu, W., & Xing, Q. W. (2025). Gamified learning impact: A meta-analysis of game element combinations on students' learning outcomes. *Educational technology research and development*, 73(4), 2617-2643. <https://doi.org/10.1007/s11423-025-10493-y>
- Downing, S. M., & Haladyna, T. M. (2006). *Handbook of test development*. Lawrence Erlbaum Associates.
- Forsström, S., Njå, M., Munthe, E., Álvarez-Galván, J.-L., & Houldsworth, L. (2025). The impact of digital technologies on students' learning. In *OECD Education Working Papers* (Issue 335). https://www.oecd.org/en/publications/the-impact-of-digital-technologies-on-students-learning_9997e7b3-en.html
- Gee, J. P. (2003). What video games have to teach us about learning and literacy. *Computers in entertainment (CIE)*, 1(1), 20-20. <https://doi.org/10.1145/950566.950595>
- George, D., & Mallery, P. (2024). *IBM SPSS statistics 29 Step by Step: A simple guide and reference*. Routledge.



- Herrero-Martín, M., Aldavero, C., Lahuerta, A., Carro, L., & Herrero, A. J. (2025). Developing key competencies in primary education using modern board games. *International Journal of Game-Based Learning (IJGBL)*, 15(1), 1-17. <https://doi.org/10.4018/IJGBL.393673>
- Hyman, R. (1982). Quasi-experimentation: Design and analysis issues for field settings. *Journal of Personality Assessment*, 46(1), 96–97. https://doi.org/10.1207/s15327752jpa4601_16
- Hussain, B., & Cheong, J. P. G. (2022). Improving gross motor skills of children through traditional games skills practiced along the contextual interference continuum. *Frontiers in Psychology*, 13, 986403. <https://doi.org/10.3389/fpsyg.2022.986403>
- Lei, H., Wang, C., Chiu, M. M., & Chen, S. (2022). Do educational games affect students' achievement emotions? Evidence from a meta-analysis. *Journal of Computer Assisted Learning*, 38(4), 946-959. <https://doi.org/10.1111/jcal.12664>
- Kanté, A. (2024). Senegal's chairmanship of the African union. *Yearbook on the African Union*, 56. Leiden Boston: Brill
- Kangas, J., Lastikka, A., & Arvola, O. (2023). Inclusive play: Defining elements of playful teaching and learning in culturally and linguistically diverse ECEC. *Education Sciences*, 13(9), 956. <https://doi.org/10.3390/educsci13090956>
- Kenanoğlu, D., & Duran, M. (2021). The effect of traditional games on the language development of pre-school. *Asian Journal of Education and Training*, 7(1), 74–81. <https://doi.org/10.20448/journal.522.2021.71.74.81>
- Kim, Y. G., Lee, H., & Zuilkowski, S. S. (2019). Impact of literacy interventions on reading skills in low- and middle-income countries: A Meta-Analysis. *Child Development*, 00(0), 1–23. <https://doi.org/10.1111/cdev.13204>
- Kingsbery, C. R., Beach, K. D., Gesel, S. A., Wood, C. L., Washburn, E. K., & Connors, T. (2025). The impact of multilevel coaching on pre-service teachers' fidelity of implementation of an evidence-based reading intervention. *Education Sciences*, 15(2), 244. <https://doi.org/10.3390/educsci15020244>
- Kumaravadivelu, B. (2003). *Beyond methods: Macrostrategies for language teaching*. Recuperado de: http://www.ugr.es/~isanz/archivos_m3thodology/kumaravidelulibrocap1.pdf.
- Kurnia, D. M. (2014). Integrating traditional games into the EYL classroom. *TEFLIN (Teaching English as a Foreign Language in Indonesia) International Proceeding*, 516-518.
- Lusianov, J. D. (2020). Post-method era and glocalization in language teaching and learning. *4th International Conference on Language, Literature, Culture, and Education (ICOLLITE 2020)*, 509(Icollite), 360–366.
- Mayer, R. E. (2021). *Multimedia learning (3rd ed)*. Cambridge University Press.
- Mori, Y., Ståhlberg, T., Zhang, X., Mishina, K., Herkama, S., Korpilahti-Leino, T., ... & Sourander, A. (2025). Let's learn about emotions program: Acceptability, fidelity, and students' mental well-being outcomes for Finnish primary school children. *Children*, 12(9), 1251. <https://doi.org/10.3390/children12091251>
- Mtebe, J. S., & Iningo, C. R. (2024). Developing local games for enhancing numeracy skills in primary schools in Tanzania: A participatory approach. *East African Journal of Science, Technology and Innovation*, 5(2), 1-16. <https://doi.org/10.37425/6mc17170>
- Ndijuye, L. G., & Beatus, J. (2024). The rural and urban divide in early literacy acquisition in Tanzania: The mediating roles of home and school contexts. *Education 3-13*, 52(8), 1267-1282. <https://doi.org/10.1080/03004279.2022.2145851>
- Ngonyani, E., Matete, R. E., & Kashaigili, J. J. (2026). Reading comprehension difficulties in English language for primary school pupils in Tanzania. *Educational Dimension*, 14, 72-84. <https://doi.org/10.55056/ed.873>
- Nicolopoulou, A., Cortina, S. K., Ilgaz, H., Cates, B. C., & Sá, A. B. De. (2015). Using a narrative- and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence. *Early Childhood Research Quarterly*, 31, 147–162. <https://doi.org/10.1016/j.ecresq.2015.01.006>
- Obilor, E. I., & Miwari, G. U. (2022). Content validity in educational assessment. *International Journal of Innovative Education Research*, 10(2), 57-69.
- Palumbo, C., Ambretti, A., & Kourkoutas, E. (2019). Traditional games body and movement. *Journal of Sports Science*, 7, 33–37. <https://doi.org/10.17265/2332-7839/2019.01.005>
- Paudel, P. (2018). Teachers' perception on postmethod pedagogy in EFL classes of Nepal. *Prithvi Academic Journal*, 1(1), 46-57. <https://doi.org/10.3126/paj.v1i1.25899>
- Plass, J. L., Mayer, R. E., & Homer, B. D. (2020). *Handbook of game-based learning*. Mit Press.
- Poehner, M. E., & Lu, X. (2024). Sociocultural theory and corpus-based English language teaching. *TESOL Quarterly*, 58(3), 1256–1263. <https://doi.org/10.1002/tesq.3282>
- Rogoff, B. (2003). *The cultural nature of human development*. Oxford University Press.



- RTI International. (2015). *Status of early grade reading in Sub-Saharan*. Washington, DC: U.S. Agency for International Development.
- Rusiana, R., & Nuraeningsih, N. (2016). Teaching English to young learners through traditional games. *Language Circle: Journal of Language and Literature*, 10(2), 193-200.
- Sancar-Tokmak, H., & Bostancı, S. (2025). Improving elementary students' English-speaking skills through non-digital gamification: An action research study. *British Educational Research Journal*, 51(6), 2655–2674. <https://doi.org/10.1002/berj.4197>
- Sánchez, J. B., Lai, M. S., Sawyer, C., & Gerland, P. (2023). *Population, education and sustainable development: interlinkages and select policy implications*. <https://doi.org/10.18356/27081990-152>
- Tanzania Institute of Education. [TIE] (2023). *English for primary schools: Pupil's book, standard three* (3rd ed.). Tanzania Institute of Education.
- Taylor, B. (2024). Tanzanian affairs issue 137. <https://www.tzaffairs.org/category/issue-number/issue-137/>
- United Nations, Department of Economic and Social Affairs. (2023). *Population, education and sustainable development: Interlinkages and select policy implications* (UN DESA Policy Brief No. 152). New York, NY: United Nations.
- USAID. (2014). *PhilEd Data: Analytical support for early grade reading in the Philippines*. <https://shared.rti.org/content/philed-data-analytical-support-early-grade-reading-philippines-2014-summary-results>
- Uwezo (2019). *Are our children learning?. Uwezo Uganda eighth learning assessment report*. Kampala: Twaweza East Africa.
- Vita-Barrull, N., Estrada-Plana, V., March-Llanes, J., Sotoca-Orgaz, P., Guzmán, N., Ayesa, R., & Moya-Higueras, J. (2024). Do you play in class? Board games to promote cognitive and educational development in primary school: A cluster randomized controlled trial. *Learning and Instruction*, 93, 101946. <https://doi.org/10.1016/j.learninstruc.2024.101946>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- World Bank. (2025). *New report offers evidence-based solutions to address global literacy crisis among children*. <https://www.worldbank.org/en/news/press-release/2025/10/30/new-report-offers-evidence-based-solutions-to-address-global-literacy-crisis-among-children>.
- World Bank & UNESCO (2025). *Learning poverty: Share of children at the end-of-primary age below minimum reading proficiency adjusted by out-of-school children (%)*. World Bank Data Catalog. <https://data.worldbank.org/indicator/SE.LPV.PRIM>

About the Authors:

Wilfred MADATA

Wilfred MADATA is a Ph.D candidate at the Department of Foreign Languages and Literature at the University of Dodoma, Tanzania, specialising in Applied Linguistics. His research interests include English language teaching, literacy development, game-based learning, and culturally responsive pedagogy, with particular emphasis on improving reading and writing outcomes among English learners.

Jesse Julius NDABAKURANE

Jesse Julius NDABAKURANE is an academic staff in the Department of Foreign Languages and Literature at the University of Dodoma, Tanzania. He holds a PhD in Linguistics and conducts research on language and education, multilingual pedagogy, English language learning, curriculum development, and language-supportive approaches to teaching and learning.

John Misana BISEKO

John Misana BISEKO is an academic staff in the Department of Foreign Languages and Literature at the University of Dodoma, Tanzania. His research interests include vocabulary learning, English language education, reading comprehension, language pedagogy, curriculum materials, and language-in-education policy in multilingual contexts.



APPENDICES

Appendix A: Normality testing using Shapiro-Wilk test

Group	n	W	p-value	Interpretation
Exp1	59	.895	< .001	Violated
Exp2	43	.856	< .001	Violated
Control G	74	.944	.003	Violated
Exp1	59	.685	< .001	Violated
Exp2	43	.887	< .001	Violated
Control G	74	.864	< .001	Violated
Exp1	59	.639	< .001	Violated
Exp2	43	.628	< .001	Violated
Control G	74	.848	< .001	Violated
Exp1	59	.578	< .001	Violated
Exp2	43	.895	< .001	Violated
Control G	74	.928	< .001	Violated
Exp1	59	.999	1	Met
Exp2	43	.830	< .001	Violated
Control G	74	.863	< .001	Violated
Exp1	59	.510	< .001	Violated
Exp2	43	—	—	Zero variance
Control G	74	.501	< .001	Violated
Exp1	59	.652	< .001	Violated
Exp2	43	.569	< .001	Violated
Control G	74	.822	< .001	Violated
Exp1	59	.977	.316	Met
Exp2	43	.970	.322	Met
Control G	74	.984	.488	Met
Exp1	59	.923	.001	Violated
Exp2	43	.963	.178	Met
Control G	74	.960	.019	Violated
Exp1	59	.867	< .001	Violated
Exp2	43	.971	.345	Met
Control G	74	.873	< .001	Violated
Exp1	59	.753	< .001	Violated
Exp2	43	.952	.069	Met
Control G	74	.913	< .001	Violated
Exp1	59	.910	< .001	Violated
Exp2	43	.962	.167	Met
Control G	74	.938	< .001	Violated
Exp1	59	.928	< .001	Violated
Exp2	43	.866	< .001	Violated
Control G	74	.84	< .001	Violated
Exp1	59	.906	< .001	Violated
Exp2	43	.811	< .001	Violated
Control G	74	.514	< .001	Violated



Appendix B: Homogeneity testing for pre-test using levene's test with Brown Forsythe modification

Variable	Test Type	df1	df2	p-value	Homogeneity
	Levene's (Brown–Forsythe)	2	173	.001	No
LR Pre	Levene's (Brown–Forsythe)	2	173	< .001	No
RBW Pre	Levene's (Brown–Forsythe)	2	173	.003	No
TORA Pre	Levene's (Brown–Forsythe)	2	173	< .001	No
RC Pre	Levene's (Brown–Forsythe)	2	173	< .001	No
LWA Pre	Levene's (Brown–Forsythe)	2	173	.519	Yes
WWA Pre	Levene's (Brown–Forsythe)	2	173	.540	Yes
TWA Pre					

Note: LR=Letter reading, RBW=Read basic words, TORA = Text oral reading accuracy, RC = reading comprehension, LWA = Letter writing accuracy, WWA = Word writing accuracy, TWA = Text writing accuracy

Appendix C: Homogeneity testing for post-test using Levene's test with Brown Forsythe modification

Variable	Test Type	df1	df2	p-value	Homogeneity
	Levene's (Brown–Forsythe)	2	173	.013	No
LR_Post	Levene's (Brown–Forsythe)	2	173	.010	No
RBW_Post	Levene's (Brown–Forsythe)	2	173	< .001	No
TORA_Post	Levene's (Brown–Forsythe)	2	173	.346	Yes
RC_Post	Levene's (Brown–Forsythe)	2	173	.059	Yes
LWA_Post	Levene's (Brown–Forsythe)	2	173	< .001	No
WWA_Post	Levene's (Brown–Forsythe)	2	173	< .001	No
TWA_Post					



EFFECTS OF RECYCLING-BASED MATHEMATICAL TASKS ON PRESCHOOL CHILDREN'S PATTERN RECOGNITION

Feriha Hande İDİL

İzmir Provincial Directorate of National Education, İzmir, Türkiye

ORCID: <https://orcid.org/0000-0002-6205-7278>

ferihahande.idil@meb.gov.tr

Received: June 6, 2026

Accepted: September 7, 2026

Published: September 30, 2026

Suggested Citation:

İdil, F. H. (2026). Effects of recycling-based mathematical tasks on preschool children's pattern recognition. *International Online Journal of Primary Education (IOJPE)*, 15(3), 219-234. <https://doi.org/10.55020/iojpe.1965127>



Copyright © 2026 by author(s). This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Abstract

Pattern recognition is an important component of early mathematical development and provides a foundation for later algebraic and relational thinking. This study examined whether recycling-based mathematical tasks improve preschool children's pattern recognition performance more effectively than traditional manipulative-based instruction. A randomized pre-test/post-test experimental design was used with 60 preschool children assigned to experimental and control groups. The experimental group participated in an eight-week intervention involving structured mathematical tasks with recycled materials, organized according to the Eco-Mathematical Embodied Learning framework, whereas the control group received traditional patterning instruction using standard classroom manipulatives. Children's pattern recognition performance was assessed using a researcher-developed assessment, and teacher-rated engagement scores were used to document participation. Results showed that the experimental group demonstrated greater improvement than the control group. Repeated measures ANOVA indicated a significant group by time interaction, and teacher-rated engagement scores were higher in the experimental group. The findings suggest that recycled materials, when used within a structured instructional sequence, may serve as accessible, sensory-rich resources for supporting early mathematical learning in sustainability-related classroom practices.

Keywords: early childhood mathematics, pattern recognition, recycled materials, sustainability education, preschool children

INTRODUCTION

Early mathematical development provides an important foundation for children's later learning, problem solving, and participation in formal mathematics education. Among early mathematical competencies, pattern recognition has a central role because it supports children's ability to identify regularities, compare relationships, extend sequences, and develop early forms of algebraic and relational thinking (Bakker et al., 2024; Rizos & Gkrekas, 2024). Research has shown that preschool children's patterning abilities are associated with later mathematical achievement and broader cognitive development, making pattern recognition a meaningful focus for early childhood mathematics instruction (Coolen et al., 2023; Ling & Loh, 2023). Evidence from primary school samples reinforces this continuity. A recent meta-analysis of 81 effect sizes reported a moderate positive association between patterning skills and mathematical achievement across early childhood and the primary grades (Zhang et al., 2026), and a large-scale study of first- to fourth-grade students showed that pattern recognition continues to develop, and to pose difficulty, well beyond the preschool years (Reuter & Reinhold, 2025). Strengthening pattern recognition before school entry is therefore relevant not only for preschool mathematics but also for the mathematical demands children meet in primary education.

In preschool classrooms, patterning activities are commonly supported through traditional classroom manipulatives such as blocks, beads, counters, and colored shapes, and prior research has emphasized that the transfer of learning depends partly on how knowledge is connected to contexts, materials, and later applications (Barnett & Ceci, 2002). These materials can provide useful opportunities for hands-on learning; however, they may also present mathematical ideas in relatively uniform and decontextualized ways. In contrast, everyday materials may offer children more familiar, authentic,



and sensory-rich opportunities to explore mathematical relationships. Recycled materials such as bottle caps, cardboard tubes, egg cartons, fabric scraps, and packaging materials vary naturally in color, shape, size, texture, and sound. These properties can support classification, comparison, sequencing, and pattern construction while also connecting mathematical activity to children's everyday experiences.

The use of recycled materials may also help situate early mathematics instruction within sustainability-related classroom practices, which are increasingly emphasized in preschool education as part of meaningful, developmentally appropriate learning experiences (Ohlsson et al., 2024a; Watt & Frydenberg, 2025). Early childhood education increasingly emphasizes the importance of helping young children engage in environmental issues through concrete, age-appropriate experiences. When children use discarded or reusable materials in mathematical tasks, they encounter mathematics not only as an abstract skill but also as part of meaningful activity connected to reuse, waste reduction, and care for the environment. Such integration may be especially valuable in early childhood settings, where learning is often grounded in play, sensory exploration, social interaction, and familiar objects.

Although research has separately examined embodied mathematics learning, material-based instruction, and sustainability education in early childhood, fewer studies have investigated how recycled materials can be systematically used to support specific mathematical competencies such as pattern recognition. In particular, limited experimental evidence is available on whether recycling-based mathematical tasks can produce stronger gains in preschool children's pattern recognition than traditional manipulative-based instruction. Addressing this gap is important because recycled materials may offer a low-cost and accessible way to enrich early mathematics learning while connecting classroom activities to sustainability-related practices.

This study draws on the Eco-Mathematical Embodied Learning (EMEL) framework as an instructional design structure for organizing recycling-based mathematical tasks. In this study, EMEL is used to sequence children's experiences across four phases: environmental engagement, mathematical mapping, cognitive construction, and environmental integration. Rather than treating recycled materials only as substitutes for traditional classroom manipulatives, the framework positions them as authentic learning resources that can support embodied engagement, pattern recognition, and sustainability-related meaning making.

The purpose of this study was to examine whether recycling-based mathematical tasks designed within the EMEL framework improve preschool children's pattern recognition performance more effectively than conventional instruction using standard classroom manipulatives. The study also examined children's engagement during recycling-based mathematical activities through structured teacher observations. By focusing on pattern recognition in preschool classrooms, this study aimed to contribute to early childhood mathematics education by exploring how low-cost, familiar, and sustainability-related materials can be used to support children's mathematical learning.

Literature Review

This section reviews the literature informing the design and focus of the present study. It first discusses pattern recognition as a foundational component of early mathematical development. It then considers material-based and embodied approaches to early mathematics learning, with particular attention to the role of authentic and sensory-rich materials. The next section examines sustainability-related learning in early childhood education and the potential of recycled materials to connect mathematical activity with everyday environmental practices. Finally, it presents the Eco-Mathematical Embodied Learning framework as the instructional design structure used to organize the recycling-based mathematical tasks in this study.

Pattern Recognition in Early Childhood Mathematics

Pattern recognition is a foundational component of early mathematical thinking. It involves children's ability to notice regularities, identify repeating units, compare relationships, complete sequences, and extend patterns into new cases. These abilities are closely related to classification, seriation,



correspondence, relational reasoning, and early algebraic thinking (Papic et al., 2011; Björklund & Pramling, 2014). For preschool children, patterning activities provide opportunities to move beyond simple matching toward more flexible forms of mathematical reasoning, such as predicting what comes next, explaining a pattern rule, and recognizing the same structure across different materials or representations.

Pattern recognition is also developmentally appropriate for preschool classrooms because it can be explored through play, movement, talk, visual arrangements, and hands-on manipulation. Children can construct repeating patterns, growing patterns, and spatial patterns using objects that differ in color, shape, size, or texture. These activities support children's attention to structure and regularity, which are important for later mathematical learning.

Recent work has also highlighted the value of sequenced, multimodal, play-based experiences for supporting young children's attribute-based reasoning and classification. For example, Hainaut (2026) found that coordinated multimodal guided play supported pre-kindergarten children's movement from global recognition toward more analytic reasoning focused on defining attributes. This is relevant to pattern recognition because children's ability to attend to attributes is central to identifying, constructing, and extending patterns. For this reason, pattern recognition provides a useful context for examining how instructional materials shape early mathematical thinking. When children use objects to build or extend patterns, they must attend to both visible attributes and underlying structures. This makes patterning an appropriate focus for investigating whether recycled materials, with their natural variation and everyday familiarity, can support early mathematical learning differently from uniform traditional classroom manipulatives.

Material-Based and Embodied Mathematics Learning

Embodied perspectives on mathematics education emphasize that mathematical understanding develops through action, perception, gesture, spatial experience, and interaction with physical tools and materials (Abrahamson et al., 2022; Alberto et al., 2022; Shvarts & Abrahamson, 2023; Way & Ginns, 2024). From this perspective, young children learn mathematical ideas not only through verbal explanation or symbolic representation, but also by building understanding through touching, moving, comparing, arranging, and transforming objects. This view is especially relevant in early childhood education, where learning is often grounded in sensory experience, play, and social interaction.

Manipulatives are commonly used in early mathematics instruction because they allow children to represent mathematical relationships concretely. Traditional classroom manipulatives can be helpful because they are durable, visually clear, and designed for repeated classroom use. However, their uniformity may sometimes limit opportunities for children to work with naturally varied materials. Recycled materials offer a different type of material experience, extending embodied and action-based views of mathematics learning beyond purpose-designed manipulatives toward familiar classroom materials (Alberto et al., 2022; Abrahamson et al., 2022). Variations in texture, size, color, shape, sound, and prior use can invite children to compare attributes, make decisions about relevant and irrelevant features, and construct patterns across less standardized objects. In this sense, recycled materials may provide a rich context for embodied mathematical activity.

The authenticity of materials may also influence children's engagement with mathematical tasks. Situated learning perspectives suggest that learning becomes more meaningful when it is connected to familiar objects, everyday practices, and socially relevant purposes (Lave & Wenger, 1991). Recycled materials are recognizable objects from children's homes, classrooms, and communities. When these materials are used in mathematical activities, children may connect patterning tasks to familiar experiences with consumption, reuse, sorting, and caring for shared spaces. This connection can make mathematical exploration more meaningful than activities that rely solely on purpose-designed materials. Recent early childhood research has also emphasized the importance of structured, hands-on experiences that connect cognition, action, and meaning. In a study on toy repair as engineering learning, Peleg and Levy (2026) showed that structured repair activities supported young children's functional understanding, design-oriented problem solving, and emotionally engaged



learning while also introducing values of sustainability and care. Although their study focused on engineering rather than mathematics, it supports the broader argument that authentic materials and meaningful object-based activity can bridge cognitive, affective, and practical dimensions of early learning. This perspective aligns with the present study's use of recycled materials as meaningful resources for mathematical exploration.

Sustainability-Related Learning in Early Childhood

Sustainability education in early childhood has moved beyond simple nature appreciation toward broader attention to children's participation in environmentally meaningful practices. Research in early childhood education suggests that young children can engage with sustainability through concrete, developmentally appropriate experiences such as sorting materials, reusing objects, caring for classroom environments, and discussing everyday choices related to waste and resources (Hedefalk et al., 2015; Ernst & Burcak, 2019). Experimental and practice-oriented studies have also shown that sustainability-related early childhood experiences can support children's environmental understanding, social-emotional learning, and teacher-guided classroom participation (Kornilaki et al., 2022; Ohlsson et al., 2024b; Watt & Frydenberg, 2025). Such activities can help children begin to understand environmental responsibility in ways that are accessible and connected to daily life.

However, sustainability-related activities are often implemented separately from core academic learning, although recent preschool research emphasizes the need for intentional pedagogical planning and integrated learning environments in education for sustainability (Ohlsson et al., 2024b). Mathematics activities may be treated as one part of the curriculum, while environmental education is addressed through separate projects or thematic activities. This separation can limit opportunities for integrated learning. Recycling-based mathematical tasks offer one way to connect these domains by using sustainability-related materials and practices to support mathematical thinking. For example, children can sort recycled materials by attributes, create repeating or growing patterns with reused objects, and apply patterning skills to organize classroom recycling systems.

Recent studies suggest that sustainability in early childhood education is increasingly approached through pedagogical design, teacher practice, and meaningful classroom participation. Koşan and Güler-Yıldız (2026), for example, examined professional development designed to support preschool teachers' sustainability-related views and practices, showing that sustainability can be strengthened through intentional pedagogical planning. In addition, research on early childhood repair activities suggests that reuse and care for objects can provide concrete ways for children to participate in sustainability-related practices (Peleg & Levy, 2026). These studies support the relevance of positioning recycled materials not only as inexpensive resources, but also as objects that can connect classroom learning with sustainability-related meanings.

In this study, sustainability is not treated as a separate outcome measured independently from mathematics performance. Rather, it provides a meaningful classroom context for organizing mathematical activity. This distinction is important because the primary outcome of the study is children's pattern recognition performance. Sustainability-related practices are used to enrich the learning context, support engagement, and position recycled materials as authentic resources for early mathematical exploration.

The EMEL Framework as an Instructional Design Structure

The Eco-Mathematical Embodied Learning framework was used in this study as an instructional design structure for sequencing recycling-based mathematical tasks. The framework brings together three complementary ideas: children learn mathematical concepts through embodied interaction with materials (Abrahamson et al., 2022; Alberto et al., 2022); learning becomes more meaningful when it is connected to familiar and socially relevant contexts (Barnett & Ceci, 2002; Lave & Wenger, 1991); and sustainability-related practices can provide concrete experiences for young children (Kornilaki et al., 2022; Ohlsson et al., 2024b). In this study, the framework guided how recycled materials were introduced, explored, mathematically structured, and connected to classroom practices.



The framework includes four phases. In the Environmental Engagement Phase, children explore recycled materials through sensory engagement, discussion of material origins, and classification activities. This phase is intended to activate children's familiarity with materials and support attention to attributes such as color, shape, size, texture, and function. In the Mathematical Mapping Phase, children begin to connect these material attributes to mathematical structures by creating simple patterns such as AB and AAB sequences or arranging objects by size, color, or quantity. In the Cognitive Construction Phase, children work with more complex patterning tasks, including completing, extending, and constructing patterns using varied recycled materials. In the Environmental Integration Phase, children apply patterning ideas to sustainability-related classroom practices, such as organizing recycling systems or explaining patterns to peers. Using EMEL as an instructional structure helped ensure that the intervention moved gradually from informal exploration toward more explicit mathematical reasoning and application. The framework also helped distinguish the experimental intervention from traditional manipulative-based instruction. In the experimental group, recycled materials were not used only as replacement objects; they were embedded within a sequence of tasks that connected sensory exploration, pattern construction, mathematical explanation, and sustainability-related classroom activity. This study therefore examines whether such a structured use of recycled materials can support preschool children's pattern recognition more effectively than conventional instruction using standard classroom manipulatives.

Research Questions

This study used a randomized pre-test/post-test experimental design supplemented by structured teacher observations. Guided by the Eco-Mathematical Embodied Learning framework as an instructional design structure, the study addressed the following research questions:

1. To what extent do recycling-based mathematical tasks improve preschool children's pattern recognition performance compared with traditional manipulative-based instruction?
2. How do children's engagement levels during recycling-based mathematical tasks compare with those observed during traditional manipulative-based mathematics instruction?

The first research question examined the effect of the recycling-based intervention on children's pattern recognition performance through pre-test and post-test comparisons between the experimental and control groups. The second research question examined children's engagement during mathematics activities using structured teacher observation scores. Together, these questions focused on both learning outcomes and observed engagement while maintaining pattern recognition as the primary outcome of the study.

METHOD

Participants

The participants were 60 preschool children aged 4 to 6 years who were enrolled in three early childhood education centers. The centers had not previously implemented a systematic recycling-based mathematics curriculum. Children were assigned to either the experimental group ($n = 30$) or the control group ($n = 30$). Stratified random assignment was used to support balance across age and gender.

The experimental group included 16 girls and 14 boys, with a mean age of 5.2 years ($SD = .7$). The control group included 15 girls and 15 boys, with a mean age of 5.1 years ($SD = .8$). The age range was 4.1 to 6.0 years in the experimental group and 4.0 to 5.9 years in the control group. Table 1 presents the demographic characteristics of the participants.

**Table 1.** Demographic characteristics of participants

Characteristic	Experimental Group (n=30)	Control Group (n=30)
Gender (Girls/Boys)	16/14	15/15
Mean Age (years)	5.2 (SD = .7)	5.1 (SD = .8)
Age Range (years)	4.1 - 6.0	4.0 - 5.9

Inclusion criteria required that children were enrolled in preschool, were between 4 and 6 years of age, had no prior systematic exposure to recycling-based mathematics activities, and followed typical developmental trajectories according to school records and teacher reports. Children with diagnosed developmental or learning disabilities that required individualized instructional adaptations were not included in the final sample. Children who missed more than 20% of the intervention sessions were also excluded from the final analysis to support consistency in intervention exposure.

Research Design

The study used a randomized pre-test/post-test experimental design to examine the effects of recycling-based mathematical tasks on preschool children's pattern recognition performance. The independent variable was the type of mathematics instruction. Children in the experimental group participated in recycling-based mathematical tasks structured through the Eco-Mathematical Embodied Learning framework, whereas children in the control group received conventional patterning instruction using standard classroom manipulatives. The primary dependent variable was children's pattern recognition performance, measured through a researcher-developed pattern recognition assessment administered before and after the intervention. Children's engagement during mathematics activities was also documented through structured teacher observations.

The intervention lasted eight weeks and was organized according to the four phases of the EMEL framework. The Environmental Engagement Phase was implemented during weeks 1–2, the Mathematical Mapping Phase during weeks 3–4, the Cognitive Construction Phase during weeks 5–6, and the Environmental Integration Phase during weeks 7–8. This sequence was designed to move children gradually from informal exploration of recycled materials toward more explicit pattern construction, pattern extension, and application of patterning ideas within sustainability-related classroom practices.

In the experimental condition, recycling-based mathematical tasks were designed around four instructional principles. First, environmental embeddedness was supported by introducing tasks through brief discussions about where materials came from, how they are commonly used, and how they can be reused. Second, multi-sensory engagement was encouraged through children's tactile, visual, and auditory exploration of recycled materials. Third, material familiarity was emphasized by using objects drawn from children's everyday environments, such as bottle caps, cardboard tubes, egg cartons, packaging materials, and fabric scraps. Fourth, mathematical progression was built into the intervention by increasing the complexity of patterning tasks across the eight-week sequence.

During the Environmental Engagement Phase, children participated in activities such as "Treasure Sort," in which they sorted recycled materials by origin or observable attributes, such as household use, color, size, texture, and shape. The mathematical focus of this phase was classification and attribute recognition. In another activity, "Eco-Detective," children identified regularities, and repeated structures already present in recycled materials, such as ridges on bottle caps, compartments in egg cartons, or repeated shapes in packaging. During the Mathematical Mapping Phase, children began to use recycled materials to construct more explicit mathematical patterns. In "Green Sequences," children created AB and AAB patterns using sorted recycled materials. In "Recycling Line-Up," children arranged materials by size, color, or quantity to create simple growing patterns. The mathematical focus of this phase was connecting material attributes to pattern structures, including repeating units, order, and gradual change. During the Cognitive Construction Phase, the tasks involved more complex patterning demands. In "Pattern Puzzles," children completed and



extended patterns such as ABC, ABBC, and growing patterns using varied recycled materials. In “Eco-Engineering,” children used pattern principles to design simple structures or arrangements with recycled materials. The mathematical focus of this phase was pattern completion, pattern extension, spatial organization, and relational reasoning. During the Environmental Integration Phase, children applied patterning ideas to classroom-based sustainability practices. In “Community Helpers,” children created patterned labels, signs, or arrangements for classroom recycling systems. In “Pattern Ambassadors,” children explained simple patterning ideas using recycled materials in peer-sharing activities. The mathematical focus of this phase was transfer, mathematical communication, and the application of pattern recognition in meaningful classroom contexts.

The control condition addressed similar patterning objectives but used traditional classroom manipulatives rather than recycled materials. Children in the control group completed patterning activities involving blocks, counters, beads, or colored shapes. These activities focused on identifying, completing, and extending patterns, but they did not include the structured sustainability-related context or the four-phase recycled-material sequence used in the experimental condition.

Materials and Instruments

Children’s pattern recognition performance was assessed using a 30-item pattern recognition assessment developed for the present study based on existing research on early childhood patterning (Papic et al., 2011; Wijns et al., 2019). The assessment included three subscales: pattern identification, pattern completion, and pattern extension, with 10 items in each subscale. Pattern identification items required children to recognize repeated units in visual or object-based sequences. Pattern completion items required children to identify the missing element in a partially completed pattern. Pattern extension items required children to continue a repeating or growing pattern according to its underlying structure.

For example, a pattern identification item asked children to identify what repeated in a sequence such as red circle–blue circle–red circle–blue circle. A pattern completion item asked children to select the missing element in a sequence such as triangle–square–triangle–. A pattern extension item asked children to continue a growing pattern such as one block–two blocks–three blocks–. Each item was scored on a 0–3 scale. A score of 0 indicated an incorrect or no response; 1 indicated a partially correct response showing limited recognition of the pattern; 2 indicated a correct response with minimal support or limited explanation; and 3 indicated a correct and independent response demonstrating recognition of the underlying pattern structure. Each subscale therefore ranged from 0 to 30. Total pattern recognition scores ranged from 0 to 90, with higher scores indicating stronger pattern recognition performance. To reduce overlap between the assessment and the intervention, the assessment items used neutral visual and object-based prompts rather than the recycled materials used in the experimental activities. This distinction was intended to ensure that the post-test measured children’s pattern recognition performance rather than familiarity with the specific materials used during the intervention. Evidence for content validity was established through expert review. Two experts in early childhood mathematics education evaluated the pattern recognition assessment in terms of developmental appropriateness, clarity of instructions, alignment with the three intended subscales, and suitability for children aged 4 to 6 years. Revisions were made based on their feedback to improve item clarity, visual presentation, and balance across difficulty levels.

Children’s engagement during mathematics activities was documented using a structured teacher observation checklist developed for the present study. The checklist was informed by observational approaches to involvement and engagement in early childhood education, particularly the Leuven Involvement Scale for Young Children (Laevers, 1994). The Leuven approach conceptualizes involvement as the degree to which children are focused, interested, mentally active, and engaged in an activity. In the present study, the checklist included 10 behavioral indicators rated on a 10-point scale, focusing on sustained attention, active participation, peer interaction, persistence, responsiveness to teacher prompts, and involvement with task materials.



The teacher observation checklist was used as a supplementary process measure rather than as a primary outcome measure. Teachers recorded children's engagement during designated mathematics activity sessions in both experimental and control conditions. The engagement scores were used to compare children's observed participation in recycling-based and traditional manipulative-based mathematics activities and were interpreted in relation to the pattern recognition findings. Before the intervention, teachers were oriented to the indicators and scoring procedures to support consistent use of the checklist across activity sessions. Internal consistency was examined for the total pattern recognition assessment, its three subscales, and the engagement checklist in the present sample before conducting the main analyses. Cronbach's alpha coefficients indicated acceptable to good reliability for all measures. For the total pattern recognition assessment (30 items), alpha was .87 at pre-test and .89 at post-test. At the subscale level (10 items each), alpha values were .78 (pre-test) and .81 (post-test) for pattern identification, .80 and .84 for pattern completion, and .82 and .85 for pattern extension. The 10-item teacher observation checklist yielded an alpha of .86 across the eight-week implementation period. All coefficients exceeded the conventional threshold of .70, indicating that both instruments produced internally consistent scores in the present sample.

For the experimental intervention, a diverse set of recycled materials was used to support patterning activities. These materials included plastic containers of different sizes and colors, bottle caps varying in size, color, and texture, cardboard packaging pieces, fabric scraps, paper tubes, egg cartons, and selected natural materials such as pinecones, shells, and stones. Materials were selected based on their suitability for classification, comparison, sequencing, and pattern construction. Before use, all materials were cleaned, inspected, and organized. Items with sharp edges, unsafe surfaces, or small detachable parts that could pose choking hazards were removed. The remaining materials were grouped according to relevant mathematical attributes, including color, shape, size, texture, quantity, and potential use in repeating or growing patterns.

The control group used traditional classroom manipulatives commonly available in preschool classrooms, including blocks, counters, beads, and colored shapes. These materials were used to address similar patterning objectives, including pattern identification, completion, and extension, but without the recycled-material context or sustainability-related task sequence used in the experimental condition.

Data Collection and Analysis

Data were collected at three levels: children's pattern recognition performance, teacher-recorded engagement during mathematics activities, and demographic information. Pattern recognition data were collected at two time points. The pre-test was administered during the week before the intervention began, and the post-test was administered during the week following the completion of the eight-week intervention. The pattern recognition assessment was administered individually to each child in a quiet space adjacent to the classroom. Scripted instructions and standardized administration procedures were used at both time points to ensure consistency across children and groups.

The researchers who administered and scored the pattern recognition assessment were blind to the children's group assignments. This procedure was used to reduce potential assessment bias. Children's responses were scored according to the predetermined 0–3 scoring rubric. A score of 0 indicated an incorrect or no response; 1 indicated a partially correct response showing limited recognition of the pattern; 2 indicated a correct response with minimal support or limited explanation; and 3 indicated a correct and independent response demonstrating recognition of the underlying pattern structure. Subscale scores were calculated separately for pattern identification, pattern completion, and pattern extension. A total pattern recognition score was calculated by summing the scores across all 30 items.

Teacher-recorded engagement observations were conducted weekly throughout the intervention period during designated mathematics activity sessions. Before the intervention began, teachers were oriented to the observation checklist and scoring procedures. They were instructed to rate children's engagement based on observable behaviors, including sustained attention, active participation,



persistence, peer interaction, responsiveness to prompts, and involvement with task materials. Engagement scores were recorded for children in both experimental and control groups during comparable mathematics activity periods.

All quantitative data were analyzed using SPSS version 27.0. Descriptive statistics were calculated to summarize demographic characteristics, pre-test and post-test pattern recognition scores, subscale scores, and teacher-recorded engagement scores. Internal consistency coefficients were examined for the pattern recognition assessment and the engagement checklist in the present sample before the main analysis. Cronbach's alpha was computed for the total assessment, for each of its three subscales, and for the observation checklist; these coefficients are reported in the Materials and Instruments section. Independent samples t-tests were used to examine baseline equivalence between the experimental and control groups. Paired-samples t-tests were conducted to examine within-group changes from pre-test to post-test.

Repeated measures ANOVA was used to examine the group $\times$ time interaction and determine whether changes in pattern recognition performance differed between the experimental and control groups. Multiple regression analysis was conducted to examine whether group assignment predicted post-test pattern recognition scores after controlling for pre-test scores, age, and gender. Effect sizes were reported using Cohen's d for t-tests and partial eta squared (η^2) for ANOVA. Statistical significance was set at $p < .05$.

RESULTS

This section presents the quantitative findings of the study in relation to children's pattern recognition performance and teacher-rated engagement during mathematics activities. First, baseline equivalence between experimental and control groups is reported. Next, pre-test and post-test changes in overall pattern recognition performance are presented, followed by the results of the repeated measures ANOVA examining the group $\times$ time interaction. Subscale-level findings for pattern identification, pattern completion, and pattern extension are then reported. Finally, regression results and teacher-rated engagement scores are presented as supplementary evidence regarding children's participation during mathematics activities.

Preliminary Analyses and Baseline Equivalence

The demographic characteristics of the participants are reported in Table 1 in the Participants section. The experimental and control groups were comparable in terms of age and gender distribution. Before the intervention, an independent-samples t-test was conducted to examine whether the groups differed in pattern recognition performance. The results showed no statistically significant difference between the experimental group ($M = 45.3$, $SD = 7.2$) and the control group ($M = 44.8$, $SD = 6.9$), $t_{(58)} = .27$, $p = .785$. This finding indicated baseline equivalence in pattern recognition performance before the intervention.

Effects on Total Pattern Recognition Performance

Descriptive statistics for pre-test and post-test pattern recognition scores are presented in Table 2. The experimental group increased from a pre-test mean of 45.3 ($SD = 7.2$) to a post-test mean of 60.5 ($SD = 6.8$). The control group increased from a pre-test mean of 44.8 ($SD = 6.9$) to a post-test mean of 47.2 ($SD = 7.1$).

Paired-samples t tests were conducted to examine within-group changes from pre-test to post-test. The experimental group showed a statistically significant increase in pattern recognition scores, $t_{(29)} = 9.87$, $p < .001$, Cohen's $d = 1.80$. The increase observed in the control group was not statistically significant, $t_{(29)} = 1.65$, $p = .110$, Cohen's $d = .30$. These results indicate that children in the experimental group showed greater improvement in pattern recognition performance than children in the control group.


Table 2. Pre-test and post-test pattern recognition scores by group

Group	Pre-test M	Pre-test SD	Post-test M	Post-test SD
Experimental (n = 30)	45.3	7.2	60.5	6.8
Control (n = 30)	44.8	6.9	47.2	7.1

Note. Total pattern recognition scores ranged from 0 to 90. Higher scores indicate stronger pattern recognition performance.

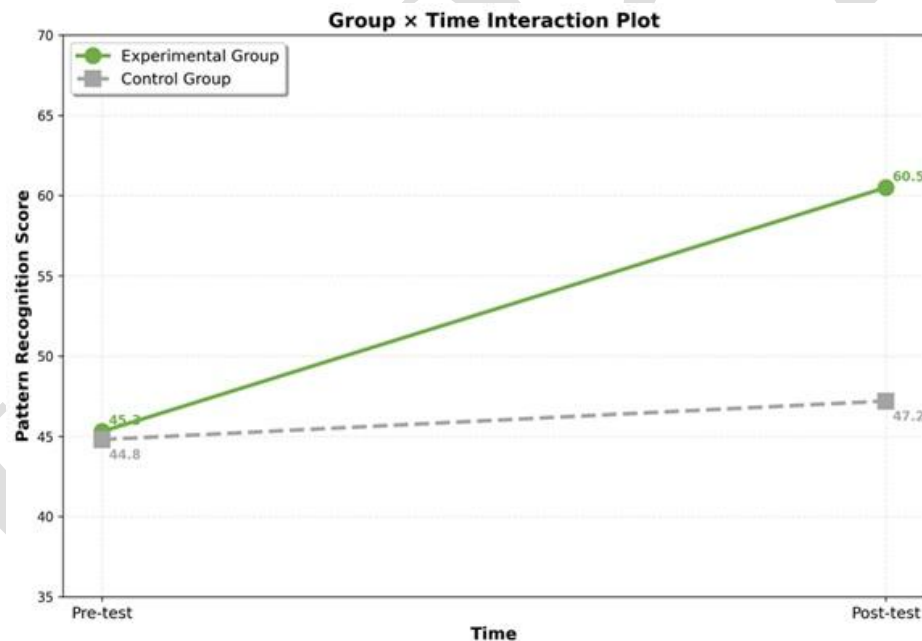
Group × Time Interaction

A repeated measures ANOVA was conducted to determine whether changes in pattern recognition performance differed between the experimental and control groups over time. The analysis revealed a significant main effect of group, $F_{(1, 58)} = 25.43$, $p < .001$, $\eta^2 = .305$, and a significant main effect of time, $F_{(1, 58)} = 37.87$, $p < .001$, $\eta^2 = .395$. More importantly, the group × time interaction was statistically significant, $F_{(1, 58)} = 32.15$, $p < .001$, $\eta^2 = .356$. This interaction indicated that the magnitude of change from pre-test to post-test differed significantly between the experimental and control groups.

Table 3. Repeated measures ANOVA summary for pattern recognition scores

Source	df	F	p	η^2
Group	1, 58	25.43	< .001	.305
Time	1, 58	37.87	< .001	.395
Group × Time	1, 58	32.15	< .001	.356

Note. η^2 = partial eta squared. The group × time interaction represents the differential change in pattern recognition scores from pre-test to post-test across the experimental and control groups.


Figure 1. Group × time interaction for pattern recognition scores

Mean pattern recognition scores are shown for the experimental and control groups at pre-test and post-test. The steeper increase in the experimental group indicates a larger gain in pattern recognition performance following the recycling-based mathematical task intervention.

Subscale-Level Pattern Recognition Outcomes

Subscale-level descriptive statistics are presented in Table 4. The experimental group showed increases across all three pattern recognition subscales. Pattern identification scores increased from 15.2 (SD = 2.4) to 19.3 (SD = 2.1), pattern completion scores increased from 14.7 (SD = 2.6) to 20.2 (SD = 2.4), and pattern extension scores increased from 15.4 (SD = 2.8) to 21.0 (SD = 2.5). In the control group, changes were smaller across the three subscales. Pattern identification scores increased from 15.0 (SD = 2.3) to 16.1 (SD = 2.5), pattern completion scores increased from 14.9 (SD = 2.5) to



16.1 (SD = 2.6), and pattern extension scores changed only slightly from 14.9 (SD = 2.7) to 15.0 (SD = 2.8).

Table 4. Pattern recognition subscale scores by group and time

Subscale	Experimental Pre-test M (SD)	Experimental Post-test M (SD)	Control Pre-test M (SD)	Control Post-test M (SD)
Pattern identification	15.2 (2.4)	19.3 (2.1)	15.0 (2.3)	16.1 (2.5)
Pattern completion	14.7 (2.6)	20.2 (2.4)	14.9 (2.5)	16.1 (2.6)
Pattern extension	15.4 (2.8)	21.0 (2.5)	14.9 (2.7)	15.0 (2.8)

Note. Each subscale score ranged from 0 to 30. Higher scores indicate stronger performance in the corresponding pattern recognition dimension. Values are presented as means with standard deviations in parentheses.

Predictors of Post-Test Pattern Recognition Performance

A multiple regression analysis was conducted to examine whether group assignment predicted post-test pattern recognition scores after controlling for pre-test score, age, and gender. The regression model was statistically significant, $F_{(4, 55)} = 28.15$, $p < .001$, and explained 67% of the variance in post-test scores, $R^2 = .67$, adjusted $R^2 = .65$. Group assignment was a significant predictor of post-test pattern recognition performance, $B = 9.80$, $SE = 1.16$, $t = 8.42$, $p < .001$. Pre-test score was also a significant predictor, $B = .45$, $SE = .09$, $t = 4.87$, $p < .001$. Neither age, $B = .12$, $SE = .11$, $t = 1.05$, $p = .299$, nor gender, $B = .08$, $SE = .11$, $t = .70$, $p = .487$, was a significant predictor. Unstandardized coefficients and their standard errors are reported in Table 5.

Table 5. Regression analysis predicting post-test pattern recognition scores

Predictor	B	SE	t	p
Group assignment	9.80	1.16	8.42	< .001
Pre-test score	.45	.09	4.87	< .001
Age	.12	.11	1.05	.299
Gender	.08	.11	.70	.487

Note. $R^2 = .67$, adjusted $R^2 = .65$, $F(4, 55) = 28.15$, $p < .001$. Group assignment was coded to compare the experimental and control conditions. Gender was entered as a control variable. B = unstandardized regression coefficient; SE = standard error of B.

Teacher-Rated Engagement During Mathematics Activities

Teacher-rated engagement scores were compared between the experimental and control groups. As shown in Table 6, the experimental group had a higher mean engagement score ($M = 8.2$, $SD = 1.1$) than the control group ($M = 6.3$, $SD = 1.3$). This difference was statistically significant, $t_{(58)} = 6.93$, $p < .001$, Cohen's $d = 1.67$. These results suggest that children in the experimental group were rated as more engaged during mathematics activities than children in the control group.

Table 6. Teacher-rated engagement scores by group

Group	Mean	SD
Experimental (n = 30)	8.2	1.1
Control (n = 30)	6.3	1.3

Note. Engagement was rated by teachers using a structured observation checklist. Scores were based on a 10-point scale, with higher scores indicating greater observed engagement during mathematics activities.

DISCUSSION, CONCLUSION, and RECOMMENDATIONS

The purpose of this study was to examine whether recycling-based mathematical tasks support preschool children's pattern recognition performance more effectively than traditional manipulative-based instruction. The findings indicated that children in the experimental group showed greater improvement from pre-test to post-test than children in the control group. The significant group $\times$ time interaction suggests that the recycling-based intervention was associated with stronger gains in pattern recognition performance. These findings provide initial evidence that recycled materials, when used within a structured instructional sequence, may serve as meaningful and sensory-rich resources for early mathematics learning. The improvement observed in the experimental group is consistent with embodied perspectives on early mathematics learning, which emphasize the role of action,



perception, gesture, and material interaction in children's mathematical understanding (Shvarts & Abrahamson, 2023; Way & Ginns, 2024). In the present study, children worked with recycled materials that varied naturally in color, shape, size, texture, and form. These variations may have encouraged children to compare attributes, attend to regularities, and construct patterns across less uniform materials. Unlike traditional classroom manipulatives, which are often designed to be visually consistent and predictable, recycled materials provided opportunities for children to identify relevant pattern features while managing natural variation. This may have supported more flexible forms of pattern recognition.

The subscale findings provide further insight into how children's pattern recognition developed during the intervention. The experimental group showed gains across pattern identification, pattern completion, and pattern extension. The increases in pattern completion and pattern extension are particularly important because these tasks require children to move beyond recognizing surface-level repetition and to reason about the underlying structure of a pattern. Previous research suggests that early patterning is closely related to relational reasoning, structural awareness, and later mathematical development (Mulligan & Mitchelmore, 2009; Papic et al., 2011; Wijns et al., 2019). The stronger gains observed in the experimental group suggest that recycling-based tasks may be especially useful for supporting children's ability to complete and extend patterns. The findings also align with research emphasizing the value of play-based, multimodal, and material-rich learning experiences in early childhood education. Young children often develop mathematical ideas through active exploration, visual comparison, movement, and interaction with concrete materials. Recent work has shown that structured multimodal experiences can support children's attention to attributes and their movement toward more analytic forms of reasoning (Hainaut, 2026). In this study, the recycling-based tasks required children to sort, compare, arrange, complete, and extend patterns using familiar objects. Such activities may have helped children connect mathematical thinking with everyday material experiences.

The sustainability-related context of the intervention may also have contributed to children's engagement with mathematical tasks, as recent early childhood sustainability research indicates that meaningful environmental practices can support children's participation, attitudes, and social-emotional engagement (Kornilaki et al., 2022; Ohlsson et al., 2024b; Watt & Frydenberg, 2025). Recycled materials were not used simply as low-cost substitutes for traditional classroom manipulatives; rather, they were embedded in activities involving reuse, sorting, and classroom-based sustainability practices. This connection may have made the tasks more meaningful for children by linking mathematical activity to familiar objects and everyday environmental routines. Research in early childhood education increasingly suggests that young children can participate in sustainability-related practices through concrete and developmentally appropriate classroom experiences (Hedefalk et al., 2015; Ernst & Burcak, 2019). However, in the present study, sustainability was used as a meaningful instructional context rather than as a separately measured outcome. Therefore, the findings should be interpreted primarily in relation to children's pattern recognition performance and engagement.

Teacher-rated engagement scores were higher in the experimental group than in the control group. This finding suggests that children were perceived by their teachers as more engaged during recycling-based mathematical activities. The tactile variety, novelty, and everyday familiarity of recycled materials may have contributed to sustained attention and participation. This interpretation is consistent with early childhood research showing that meaningful object-based activities can support children's active involvement, problem solving, and affective engagement (Peleg & Levy, 2026). At the same time, engagement was measured through teacher ratings rather than independent observational coding. Therefore, this result should be interpreted as supplementary evidence about children's participation during the intervention.

The regression findings further support the association between group assignment and post-test pattern recognition performance after controlling for pre-test scores, age, and gender. Group



assignment remained a significant predictor, suggesting that the observed differences in post-test scores were not explained only by initial performance or demographic characteristics. The non-significant effects of age and gender indicate that the intervention may have provided accessible entry points for children across the 4- to 6-year age range represented in the sample. This is important for early childhood classrooms, where children often differ in developmental level, language use, attention span, and prior mathematical experiences. The study contributes to early childhood mathematics education by showing how recycled materials can be used within a structured sequence of mathematical tasks. The EMEL framework was used in this study as an instructional design structure rather than as a fully validated theoretical model. Its four phases helped organize the intervention from initial material exploration to explicit pattern construction and application in classroom practices. This structure may be useful for teachers who wish to integrate mathematical learning with familiar, low-cost, and sustainability-related materials. The findings suggest that carefully sequenced recycling-based tasks can support pattern recognition while also providing a meaningful context for classroom participation. Taken together, the findings indicate that recycled materials can enrich early mathematics learning when they are embedded within a structured and developmentally appropriate instructional sequence.

This study suggests that recycling-based mathematical tasks may provide a low-cost, accessible, and developmentally appropriate approach for enriching early childhood mathematics instruction. When carefully planned, recycled materials can support children's pattern recognition while also situating mathematical learning within meaningful classroom practices related to reuse, care for materials, and sustainability. The study also offers a practical example of how early mathematics instruction can be connected to sustainability-related classroom practices without treating sustainability as a separate measured outcome. Because patterning knowledge remains linked to mathematical outcomes well into the primary years (Borriello et al., 2023), preschool gains of this kind are best understood as part of a longer developmental trajectory rather than as isolated outcomes. At the same time, recent reviews caution that gains from training general skills do not transfer automatically to broader mathematical achievement (Fyfe & Borriello, 2025), which supports the view that recycling-based tasks should be embedded in mathematics instruction rather than treated as a substitute for it.

Several suggestions can be offered based on these findings. First, early childhood teachers may use familiar materials such as bottle caps, cardboard tubes, egg cartons, packaging pieces, and fabric scraps to design patterning tasks involving classification, comparison, sequencing, completion, and extension. Such materials may enrich early mathematics activities because they are concrete, sensory-rich, accessible, and connected to children's everyday experiences. Second, sustainability-related classroom practices may be used as a meaningful context for early mathematical learning. In this study, sustainability served as a context for organizing mathematical tasks around reuse, sorting, and care for materials. Therefore, the findings do not show that children's environmental awareness increased. Instead, they suggest that pattern recognition activities can be situated within sustainability-related practices in developmentally appropriate ways. For early childhood teachers, this means that mathematical learning and sustainability-related routines do not need to be treated as separate classroom domains. Third, the EMEL framework may be used as an instructional planning structure. Its four-phase sequence—environmental engagement, mathematical mapping, cognitive construction, and environmental integration—offers a way to move from informal exploration of recycled materials toward explicit mathematical reasoning and application. For example, children may first explore and classify materials, then use their attributes to create simple patterns, later complete and extend more complex patterns, and finally apply patterning ideas in classroom routines. This progression may help teachers design tasks that are not only hands-on but also mathematically focused and developmentally progressive, consistent with research emphasizing intentional pedagogical design in preschool sustainability education (Ohlsson et al., 2024b).

Finally, careful material selection and classroom organization are important for implementing recycling-based mathematical tasks. Materials should be clean, safe, age-appropriate, and sufficiently varied to support mathematical comparison and pattern construction. Teachers need to consider which



material attributes are mathematically relevant, such as color, shape, size, texture, quantity, and order. Family participation in collecting reusable materials may further support the approach, but such participation should be organized carefully to ensure safety, hygiene, and consistency across activities.

Limitations of the Research

Several limitations should be considered when interpreting the findings. First, the sample size was relatively small and included children from three preschool centers. Although the experimental and control groups were comparable at pre-test, larger studies are needed to examine whether similar results can be obtained across different early childhood settings. Second, the intervention lasted eight weeks and did not include a delayed post-test. Therefore, the study does not provide evidence about the long-term maintenance of children's pattern recognition gains. Future research should examine whether the gains observed after recycling-based mathematical tasks are sustained over longer periods. Third, the study focused only on pattern recognition. Although patterning is an important component of early mathematical development, the findings cannot be generalized directly to other mathematical domains. Future research should examine whether recycling-based mathematical tasks can support areas such as number sense, measurement, geometry, spatial reasoning, and data representation. Different mathematical domains may require different types of recycled materials, task structures, and teacher guidance.

A further limitation concerns the possible influence of a novelty effect. Children in the experimental group worked with recycled materials that were new to them and that differed from their usual classroom manipulatives in texture, sound, shape, and origin, whereas children in the control group used standard and highly familiar materials. It is therefore possible that part of the higher engagement and the larger performance gains observed in the experimental group reflected the unfamiliarity and sensory richness of the materials rather than the eco-mathematical pedagogy itself. This is consistent with what is commonly described as a novelty or Hawthorne effect, and the present design cannot separate the two influences. Future research could address this issue by including a second comparison condition in which children use novel but non-recycled materials, by extending the intervention so that initial novelty declines, or by measuring engagement repeatedly across the implementation period to examine whether group differences persist once the materials are no longer new.

A further limitation concerns measurement. The pattern recognition assessment and teacher-rated engagement checklist were developed for the present study and informed by existing literature. Although expert review was used to support content validity, further validation studies are needed with larger samples. Future research should examine the psychometric properties of these instruments in greater depth, including internal consistency, inter-rater reliability where applicable, and evidence of construct validity. Finally, engagement was assessed through teacher ratings. These scores provided useful supplementary information about children's observed participation, but they may also be influenced by teacher perceptions. Future studies should include independent observers, video-based coding, or systematic field notes to examine children's engagement processes more deeply. Research should also examine implementation conditions, including teacher preparation, implementation fidelity, classroom discourse, and children's mathematical explanations. The implications of this study should therefore be interpreted as promising but preliminary.

Funding

No funding was received from any individual or institution for this research.

Ethics and Conflict of Interest

Ethical approval for the study was obtained from the Human Research Ethics Committee of Dokuz Eylül University with decision number 36 dated December 3, 2025. The official approval document was issued by Dokuz Eylül University Buca Faculty of Education on December 16, 2025, with document number E-10042736-204.01.07-1644970. The research was conducted in accordance with ethical principles, and the author declares that there is no conflict of interest. Written informed



consent was obtained from the children's parents or legal guardians before participation. Children's participation was voluntary, and all data were anonymized to protect confidentiality. All recycled materials used in the intervention were cleaned, inspected, and selected according to age-appropriate safety criteria. The author declares that there is no conflict of interest regarding the publication of this study.

Data Availability

The data that support the findings of this study are available on request from the author.

Corresponding Author

Correspondence to Feriha Hande İDİL, ferihahande.idil@meb.gov.tr

REFERENCES

- Abrahamson, D., Dutton, E., & Bakker, A. (2022). Towards an enactivist mathematics pedagogy. In S. A. Stolz (Ed.), *The body, embodiment, and education: An interdisciplinary approach* (pp. 156–182). Routledge. <https://doi.org/10.4324/9781003142010-9>
- Alberto, R., Shvarts, A., Drijvers, P., & Bakker, A. (2022). Action-based embodied design for mathematics learning: A decade of variations on a theme. *International Journal of Child-Computer Interaction*, 32, Article 100419. <https://doi.org/10.1016/j.ijcci.2021.100419>
- Bakker, M., Torbeyns, J., Verschaffel, L., & De Smedt, B. (2024). Cognitive characteristics of children with high mathematics achievement before they start formal schooling. *Child Development*, 95(6), 2062–2081. <https://doi.org/10.1111/cdev.14140>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Björklund, C., & Pramling, N. (2014). Pattern discernment and pseudo-conceptual development in early childhood mathematics education. *International Journal of Early Years Education*, 22(1), 89–104. <https://doi.org/10.1080/09669760.2013.809657>
- Borriello, G. A., Grenell, A., Vest, N. A., Moore, K., & Fyfe, E. R. (2023). Links between repeating and growing pattern knowledge and math outcomes in children and adults. *Child Development*, 94(2), e103–e118. <https://doi.org/10.1111/cdev.13882>
- Coolen, I. E. J. I., Omont-Lescieux, S., & Knops, A. (2023). Now you see it, now you don't: Cognitive skills and their contributions to mathematics across early development. *Journal of Cognition*, 6(1), Article 43. <https://doi.org/10.5334/joc.309>
- Ernst, J., & Burcak, F. (2019). Young children's contributions to sustainability: The influence of nature play on curiosity, executive function skills, creative thinking, and resilience. *Sustainability*, 11(15), Article 4212. <https://doi.org/10.3390/su11154212>
- Fyfe, E. R., & Borriello, G. A. (2025). The (in)effectiveness of training domain-general skills to support early math knowledge. *Child Development Perspectives*, 19(1), 54–60. <https://doi.org/10.1111/cdep.12526>
- Hainaut, B. (2026). Multimodal play-based learning and geometric reasoning in pre-K: A naturalistic mixed-methods study. *Early Childhood Education Journal*, 54, 4303–4311. <https://doi.org/10.1007/s10643-026-02184-x>
- Hedefalk, M., Almqvist, J., & Östman, L. (2015). Education for sustainable development in early childhood education: A review of the research literature. *Environmental Education Research*, 21(7), 975–990. <https://doi.org/10.1080/13504622.2014.971716>
- Kornilaki, E. N., Skouteris, H., & Morris, H. (2022). Developing connections between healthy living and environmental sustainability concepts in Cretan preschool children: A randomized trial. *Early Child Development and Care*, 192(11), 1685–1698. <https://doi.org/10.1080/03004430.2021.1925664>
- Koşan, Y., & Güler-Yıldız, T. (2026). From theory to practice: Empowering preschool teachers for sustainability via pedagogical systems theory based professional development. *Early Childhood Education Journal*, 54, 3161–3175. <https://doi.org/10.1007/s10643-025-02086-4>
- Laevers, F. (1994). *The Leuven Involvement Scale for Young Children: LIS-YC manual*. Centre for Experiential Education.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.



- Ling, M. K. D., & Loh, S. C. (2023). Relationships between cognitive pattern recognition and specific mathematical domains in mathematics education. *International Journal of Mathematical Education in Science and Technology*, 54(2), 159-179. <https://doi.org/10.1080/0020739X.2021.1949059>
- Mulligan, J., & Mitchelmore, M. (2009). Awareness of pattern and structure in early mathematical development. *Mathematics Education Research Journal*, 21(2), 33–49. <https://doi.org/10.1007/BF03217544>
- Ohlsson, A., Borg, F., & Gericke, N. (2024). Education for sustainability in preschool: Swedish preschool teachers' perspectives. *Cogent Education*, 11(1), Article 2353477. <https://doi.org/10.1080/2331186X.2024.2353477>
- Ohlsson, A., Gericke, N., & Borg, F. (2024). Implementing education for sustainability in preschool: Teaching strategies and learning environments. *Journal of Outdoor and Environmental Education*. Advance online publication. <https://doi.org/10.1007/s42322-024-00181-5>
- Papic, M. M., Mulligan, J. T., & Mitchelmore, M. C. (2011). Assessing the development of preschoolers' mathematical patterning. *Journal for Research in Mathematics Education*, 42(3), 237–268. <https://doi.org/10.5951/jresmetheduc.42.3.0237>
- Peleg, A., & Levy, S. T. (2026). Repairing toys as engineering learning: Tool use, problem-solving, and emotions in early childhood. *Early Childhood Education Journal*, 54, 4015–4027. <https://doi.org/10.1007/s10643-026-02130-x>
- Reuter, D., & Reinhold, F. (2025). Investigating the complexity of pattern recognition in primary students. *Educational Studies in Mathematics*, 121(2), 229–260. <https://doi.org/10.1007/s10649-025-10464-3>
- Rizos, I., & Gkrekas, N. (2024). Pattern recognition among primary school students: The relationship with mathematical problem-solving. *Contemporary Mathematics and Science Education*, 5(2), Article ep24010. <https://doi.org/10.30935/conmaths/14689>
- Shvarts, A., & Abrahamson, D. (2023). Coordination dynamics of semiotic mediation: A functional dynamic systems perspective on mathematics teaching/learning. *Constructivist Foundations*, 18(2), 220-234.
- Watt, B. A., & Frydenberg, E. (2025). Early childhood education for sustainability: Outcomes for social and emotional learning. *Australasian Journal of Early Childhood*, 50(2), 131–145. <https://doi.org/10.1177/18369391241287939>
- Way, J., & Ginns, P. (2024). Embodied learning in early mathematics education: Translating research into principles to inform teaching. *Education Sciences*, 14(7), 696. <https://doi.org/10.3390/educsci14070696>
- Wijns, N., Torbeyns, J., Bakker, M., De Smedt, B., & Verschaffel, L. (2019). Four-year olds' understanding of repeating and growing patterns and its association with early numerical ability. *Early Childhood Research Quarterly*, 49, 152–163. <https://doi.org/10.1016/j.ecresq.2019.06.004>
- Zhang, Y., Ching, B. H.-H., Qiu, T. H.-Q., Ma, T. C.-I., & Huang, M. (2026). The relation between patterning skills and children's mathematical achievement: A meta-analysis. *Educational Psychology Review*, 38, Article 25. <https://doi.org/10.1007/s10648-026-10126-7>

About the Author:

Feriha Hande İDİL

Feriha Hande İdil, Ph.D., works as a researcher and project manager at the İzmir Provincial Directorate of National Education, İzmir, Türkiye. Her research interests centre on mathematics education, and in particular on the development of mathematical thinking in the preschool years. She studies pattern recognition, concrete and material-based learning, sustainability-oriented learning environments, and out-of-school learning. She also conducts research at the intersection of science education and educational technology, and takes part in education projects at the national level.



CLASSROOM-BASED WORKING MEMORY TRAINING IN CHILDREN: EVIDENCE FROM A QUASI-EXPERIMENTAL PRIMARY SCHOOL STUDY

Emire ULUĞ

Assistant Professor, Child Development Program, Eldivan Vocational School of Health Services,
Çankırı Karatekin University, Çankırı, Türkiye
ORCID: <https://orcid.org/0000-0002-4051-5119>
emireulug@karatekin.edu.tr

Ayşe Dilek ÖĞRETİR ÖZÇELİK

Professor, Department of Early Childhood Education, Gazi Faculty of Education,
Gazi University, Ankara, Türkiye
ORCID: <https://orcid.org/0000-0002-6380-4757>
ogretir@gazi.edu.tr

Received: May 12, 2026

Accepted: September 11, 2026

Published: September 30, 2026

Suggested Citation:

Uluğ, E., & Öğretir-Özçelik, A. D. (2026). Classroom-based working memory training in children: Evidence from a quasi-experimental primary school study. *International Online Journal of Primary Education (IOJPE)*, 15(3), 235-252. <https://doi.org/10.55020/iojpe.1949761>



Copyright © 2026 by author(s). This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Abstract

This study investigated the effects of a classroom-based Working Memory Training Program on working memory and emotion regulation in second-grade students. A quasi-experimental non-equivalent pretest–posttest control-group design was used. The sample consisted of 49 students from two intact second-grade classrooms in the same public school, with 22 students in the experimental group and 27 in the control group. The experimental group received the Working Memory Training Program for eight weeks, with one session per week, while the control group continued regular classroom instruction. Data were collected using the Working Memory Scale and the Children's Emotion Regulation Scale–Adult Form. Both groups completed pretest and posttest assessments, and the experimental group completed a three-month follow-up assessment to examine the maintenance of gains. Data were analyzed using descriptive statistics, ANCOVA models controlling for pretest scores, bootstrap confidence intervals, and correlation analyses. Results showed that, after controlling for pretest scores, the experimental group had significantly higher posttest scores in verbal working memory, visual working memory, and total working memory. The experimental group also showed significantly higher adjusted posttest scores in anger, excitement, and total emotion regulation, whereas group differences in fear and sadness were not significant. Follow-up findings indicated that gains were generally maintained. Positive correlations were found between working memory and emotion regulation.

Keywords: working memory, emotion regulation, primary education, intervention program, quasi-experimental

INTRODUCTION

Working memory (WM) is a limited-capacity cognitive system that temporarily stores and processes information, enabling individuals to perform complex cognitive tasks such as reasoning, understanding, and learning (Baddeley, 2000; Bruning et al., 2014). Several models have advanced our understanding of WM, including Baddeley and Hitch's (1974) multicomponent model, Cowan's (2005) embedded processes model, and Engle's (2002) attention control framework. These models emphasize WM as both a storage and executive control system critical for adaptive cognition.

Theoretical Framework

Beyond its cognitive functions, WM plays a significant role in emotion regulation (ER), which refers to the processes by which individuals monitor, evaluate, and modify their emotional responses (Gross, 1998). According to Gross's process model, cognitive strategies such as reappraisal require substantial WM resources to maintain emotionally relevant information, reinterpret it, and inhibit dominant responses. Similarly, attentional control theories (Miyake et al., 2000; Kane & Engle, 2002)



suggest that WM supports ER by allocating attention and suppressing irrelevant emotional distractors. Therefore, WM can be conceptualized as a cognitive pathway to ER.

Consistent with this view, meta-analytic evidence shows modest associations between executive functions (including updating, blocking, and replacing) and cognitive reappraisal, suggesting that cognitive control processes are significant components of at least some ER strategies (Toh et al., 2024). More specifically, the ability to update has been proposed as a building block for adaptive ER because it allows individuals to replace dominant negative interpretations with alternative meanings that support reappraisal (Pe et al., 2013).

Empirical research supports these theoretical connections. Individuals with higher working memory capacity are more likely to use adaptive emotion regulation strategies, exhibit lower levels of negative mood, and display greater emotional resilience (Barkus, 2020; Hendricks & Buchanan, 2016; Pe et al., 2013; Xiu et al., 2016). Neuroimaging studies also reveal overlapping neural networks in the prefrontal cortex underlying both working memory and emotion regulation (Schmeichel & Tang, 2015; Schweizer et al., 2019). Conversely, mood states affect working memory performance: negative mood impairs tasks requiring selective attention and updating (Osaka et al., 2013; Hou & Cai, 2022), while positive mood increases planning and flexibility (Phillips et al., 2002).

Developmentally, WM and ER are closely intertwined in childhood; rapid advances in cognitive control coincide with the emergence of self-regulation skills (Malagoli & Usai, 2018; Zaehring et al., 2018). Children with stronger WM exhibit better impulse control, frustration tolerance, and adaptive emotional responses (Kahl et al., 2021). Given that maladaptive ER contributes to psychopathology (Urry & Gross, 2010), clarifying the cognitive foundations of ER in childhood is of both theoretical and practical importance. The relationship between WM and self-regulation is theoretically and empirically supported; however, the main question in intervention studies is whether WM training improves not only performance on target tasks but also broader functioning.

The current literature suggests that WM training yields reliable, direct gains on working memory outcomes (e.g., performance on trained or closely related WM tasks). However, evidence for transfer to broader, nontrained outcomes, such as emotion regulation, is generally small, inconsistent, and context-dependent (Melby-Lervåg & Hulme, 2016; Kassai et al., 2019). Recent child-focused studies have likewise reported improvements in working memory following training (Berger et al., 2025; Hong & Zhou, 2025).

However, there are also arguments suggesting that ER may be a more plausible target for transfer from WM training than many other distal outcomes because ER is a cognitively demanding process requiring the maintenance and updating of mental representations and the suppression of dominant responses (Gross, 1998, 2015; Miyake et al., 2000). Narrative and systematic syntheses focusing on WM training and ER note that gains in ER-related outcomes have been reported in different populations; however, there is considerable heterogeneity in training protocols, outcome measures, and methodological rigor (Barkus, 2020; Papé & Maree, 2023). For example, Barkus's transdiagnostic review summarizes studies reporting improvements in mood and ER-related indicators after standardized WM training and proposes two possible explanations: (a) overlap between cognitive WM, affective WM, and ER systems, and (b) increased cognitive efficiency that may increase available resources under emotional load (Barkus, 2020). Complementing this narrative, a meta-analytic review of emotional working memory training (eWMT) reports improvements in reappraisal and reductions in depressive and anxiety symptoms; however, it highlights the heterogeneity and limited number of available studies (Ou et al., 2026). Taken together, this evidence supports a cautious conclusion: WM-based training may affect ER under some conditions; however, the effects likely depend on the training content (e.g., emotional vs. neutral WM demands), population characteristics, outcome measurement, and implementation context.

The early school years are a period of rapid development in cognitive control and self-regulation, making them theoretically favorable for interventions; however, evidence regarding the transfer of



WM training to ER in children, particularly in school-based settings, is limited and heterogeneous (Papé & Maree, 2023). Theoretically, WM and ER share common cognitive control processes such as executive attention, updating, and inhibition (Engle, 2002; Gross, 1998, 2015; Miyake et al., 2000; Toh et al., 2024); however, meta-analyses show that direct effects of WM training on WM performance are more consistent than transfer effects on broader outcomes (Kassai et al., 2019; Melby-Lervåg & Hulme, 2016). However, executive function training may produce more pronounced far-transfer effects in neurodevelopmental or clinical samples, and WM/inhibition-targeted protocols may reduce mood dysregulation (Bombonato et al., 2024; Groves et al., 2022). These findings underscore the need to test whether WMTP produces transfer to ER in typically developing children in a real school context by evaluating both WM and ER outcomes within the same design (Barkus, 2020; Papé & Maree, 2023).

In summary, three conclusions emerge from the literature most closely related to this study. First, WM and ER are theoretically linked through their shared dependence on cognitive control processes (Engle, 2002; Gross, 2015; Miyake et al., 2000; Toh et al., 2024). Second, WM training can produce reliable, direct improvements in WM outcomes; however, transfer to broader outcomes is not guaranteed and should be addressed as an empirical question (Kassai et al., 2019; Melby-Lervåg & Hulme, 2016). Third, evidence regarding the effects of WM training on ER in children appears limited and mixed, particularly in school-based and classroom-focused contexts; this increases the need for studies that evaluate both WM and ER outcomes together within the same design (Barkus, 2020; Papé & Maree, 2023).

This study evaluated the effects of a classroom-based Working Memory Training Program (WMTP) on both WM and ER in second-grade students. Three research questions guided the study: (1) Do the WMTP and control groups differ in baseline-adjusted posttest verbal, visual, and total WM scores? (2) Do the groups differ in baseline-adjusted posttest ER scores for anger, fear, excitement, sadness, and total ER? (3) Are WM and ER outcomes positively associated? WM outcomes were the direct targets of the program, whereas ER outcomes were considered theoretically plausible secondary transfer outcomes because transfer from WM training is complex and context-sensitive (Barkus, 2020; Melby-Lervåg & Hulme, 2016; Papé & Maree, 2023).

METHOD

Study Design

This study employed a classroom-based quasi-experimental design with a nonequivalent pretest-posttest control-group structure. Two intact second-grade classrooms from the same public primary school participated: one classroom received the WMTP (intervention group) and the other continued with regular classroom instruction as a business-as-usual control group. Because allocation occurred at the classroom level, the design represents a clustered (classroom-based) implementation rather than individual randomization. Such nonequivalent pretest-posttest classroom-based quasi-experimental designs are commonly used in educational field settings when random assignment is not feasible and are particularly relevant for evaluating interventions under real-world constraints (Buluş, 2021; Gopalan et al., 2020; Handley et al., 2018).

The study included three measurement points. At pretest and posttest, both groups completed (a) the Working Memory Scale, which assesses working memory across multiple subdomains in children aged 5–10 years (Ergül et al., 2018), and (b) the Emotion Regulation Scale–Adult Form (Rydell et al., 2003; Turkish adaptation reported by Tatlı-Harmanacı & Güngör-Aytar, 2023). A follow-up (retention/maintenance) assessment was administered three months after posttest and included the same two measures, but it was conducted only with the intervention group to evaluate the maintenance of gains; therefore, follow-up results are interpreted as within-group maintenance evidence rather than as a between-group causal comparison. Detailed information on the measures and scoring procedures is provided below.



Given the nonequivalent-group structure, baseline equivalence could not be assumed. Accordingly, the primary between-group posttest comparisons were planned using baseline-adjusted models, estimating posttest group differences while controlling for pretest scores. Baseline adjustment (e.g., ANCOVA-style adjustment) is widely recommended in pretest–posttest designs because it improves precision and helps account for initial differences between groups (Vickers & Altman, 2001; Van Breukelen, 2006). At the same time, statistical adjustment cannot fully eliminate selection bias if intact groups differ systematically on unmeasured factors; therefore, causal interpretations should remain cautious in quasi-experimental contexts (Gopalan et al., 2020; Shadish et al., 2002).

Participants

Participants were second-grade students from two intact classrooms in a public primary school in Central Anatolia, Türkiye (2022–2023 academic year). The study used a feasibility-based sample determined by the availability of these classrooms. One classroom was designated as the experimental group (Working Memory Training Program; WMTP) and the other as the business-as-usual control group, reflecting implementation under natural school conditions.

Initially, 52 children were enrolled (experimental $n = 23$, control $n = 29$). Three students discontinued participation (experimental: 1; control: 2) and were excluded from all analyses, including baseline, yielding a final analytic sample of 49 children (experimental $n = 22$, control $n = 27$). No a priori power analysis was conducted because the sample size was constrained by classroom availability; results are therefore interpreted using effect sizes and confidence intervals in addition to p -values.

Eligibility required classroom enrollment, parental consent, and child assent. Both groups completed pretest and posttest assessments; a follow-up assessment was administered only to the experimental group and treated as within-group maintenance evidence. Make-up sessions were offered to support intervention attendance.

Intervention: Working Memory Training Program (WMTP)

WMTP is a classroom-based, game-mediated intervention developed by the authors (the research team) for second-grade students. The program was designed to promote active participation, accommodate individual differences, and support collaborative learning through child-centered instructional practices aligned with progressive education principles. The WMTP is grounded in Information Processing Theory and the multicomponent working memory model (Baddeley & Hitch, 1974; Baddeley, 2000). Accordingly, the program targets working memory components through four instructional modules: visuospatial sketchpad, phonological loop, episodic buffer, and central executive.

The program was developed using the Taba–Tyler curriculum development model, beginning with a needs analysis and proceeding through specification of goals and objectives, selection and organization of content, design of learning experiences, and iterative refinement. The needs analysis combined an analytical review of relevant programs and research with a democratic component that included interviews with a child development expert, a program development expert, and three classroom teachers. Stakeholders consistently emphasized that, for second-grade students, working-memory activities should be short, concrete, highly engaging, game-based, and supported by visual and multisensory elements; experts also underscored the importance of structured repetition and flexibility to accommodate individual differences. Accordingly, WMTP was structured using a modular programming approach, and within key modules (particularly the visuospatial and phonological components) learning experiences were arranged in a spiral progression, revisiting core skills while gradually increasing complexity.

Importantly, the WMTP did not include explicit emotion regulation instruction. Instead, it was designed to strengthen working-memory processes with the hypothesis that improved cognitive control would indirectly support children's emotion regulation capacities.

WMTP consisted of 21 developmentally appropriate activities delivered across eight weekly sessions. Each session typically included two to three consecutive activities and lasted approximately one to



two class periods, depending on classroom pacing. The session plan (see Table 1) mapped activities and strategies to the targeted working memory components. Across sessions, children practiced strategies embedded in classroom tasks, including rehearsal/inner repetition, grouping/chunking, segmentation, storytelling/narrative elaboration, and mnemonic supports. Activities were implemented through educational games, question–answer interactions, and, when appropriate, computer-assisted tasks, with an emphasis on active, guided practice rather than passive instruction. Table 1 summarizes the eight-session structure of the WMTP, including the targeted working memory components, strategies, and activities.

Table 1. Structure of the WMTP

Session	Working Memory Component	Strategy	Activities
1	Working Memory (Introduction)	—	Treasure Hunt – Getting to Know Game; Touch the Number I Say
2	Visuospatial Domain	Rehearsal (inner repetition); Chunking (grouping)	Remembering What I Said; Shopping Basket; Find What’s Missing
3	Phonological Loop	Chunking (grouping)	I Went to the Market Game; Pumpkin Game; Listen, Grasp, Draw
4	Episodic Buffer	Storytelling; Chunking (grouping)	Story Cubes; Waiter Game
5	Visuospatial Domain	Elaboration / Meaning-making	Cup Game; Fast Seer Wins
6	Phonological Loop	Rehearsal (inner repetition)	Catch the Rhythm; My Teacher Says That; Camel–Gücüye
7	Visuospatial Domain	Mnemonics	Touch the Color I Say; Fast Colors Game; Encoding
8	Central Executive	Elaboration / Linking (integration)	Reverse Cup Game; LEGO Construction; Divided Attention

To strengthen the validity and applicability of the program's content, it was reviewed by five experts in child development and education, and a pilot implementation was conducted with five second-grade students. The feedback was used to improve the activity instructions, pacing, and materials.

The program was delivered by the researcher, who had completed formal training and certification for administering the Working Memory Scale. Fidelity was supported through standardized session plans and instructions, monitoring of attendance and dose, and make-up opportunities for missed sessions. The control group did not receive the WMTP during the study; after completion of data collection, the training materials were provided to the control-group teacher, but subsequent implementation was not monitored as part of this study.

Measures

Working memory assessment

Working memory was assessed with the Working Memory Scale (WMS) developed for children aged 5–10 years (Ergül et al., 2018). The WMS comprises four domains and nine subtests: verbal short-term memory (Digit Recall, Word Recall, Nonword Recall), visual short-term memory (Dot/Pattern Matrix, Block Recall), verbal working memory (Backward Digit Recall, First Word Recall), and visual working memory (Odd-One-Out, Spatial Discrimination/Span). Subtests are administered in sequences that increase in length, typically with two trials per level; administration proceeds as long as the child meets the continuation rule and is discontinued when the stopping criterion is reached (i.e., failure on both trials at a given level). Each correctly recalled sequence is scored 1, whereas an incorrect sequence is scored 0. For the present study, subtests were aggregated into modality-based composites: Verbal Memory (sum of all verbal short-term and verbal working memory subtests) and Visual Memory (sum of all visual short-term and visual working memory subtests). A total score was computed as the sum of the Verbal and Visual composites and was used as the primary overall index in analyses. In the present study, the total WMS score demonstrated high internal consistency (Cronbach’s $\alpha = .91$).

**Emotion regulation measure (adult form)**

Emotion regulation was measured using the adult-report version of the Emotion Regulation Scale for Children (Rydell et al., 2003; Turkish adaptation by Tatlı-Harmancı & Güngör-Aytar, 2023). The ERS-A includes 40 items covering four emotion domains (anger, fear, excitement/enthusiasm, and sadness) and yields scores reflecting emotion regulation (and, separately, emotion intensity). Parents rated each item on a 5-point Likert scale (1–5); item responses were summed according to the scoring guidelines (including reverse-coded items, where applicable) to compute a total emotion regulation score, with higher scores indicating better overall regulation as defined by the scale. In line with the study aims, analyses used only the total emotion regulation score; the emotion intensity dimension was not analyzed because it was not targeted by the intervention, thereby maintaining a parsimonious set of outcomes. Internal consistency for the ERS-A total score in the present sample was acceptable (Cronbach's $\alpha = .82$).

Demographics

Demographic and background information was obtained via a Personal Information Form. Variables included child age and gender, maternal and paternal education levels, number of siblings, birth order, and preschool education attendance. These data were used to describe the sample and, where relevant, in adjusted analyses to account for potential confounding influences in group comparisons and association tests.

Procedure

Following ethical approval from the Gazi University Scientific Publication and Ethics Committee (No. E.585142; 09/02/2023) and official implementation permission from the Ministry of National Education (No. E.74258403; 12/04/2023), the study was conducted in a public primary school located in a central district in Central Anatolia, Türkiye. Two intact second-grade classrooms from the same school participated. Because individual randomization was not feasible in the natural school context, group allocation occurred at the classroom level: One classroom was designated as the WMTP intervention group and the other served as a business-as-usual control group continuing routine instruction without any additional cognitive training activities or WMTP materials.

After obtaining written parental informed consent and child assent, baseline assessments were administered. Children in both groups completed the WMS individually in a quiet room at school (Support Education Room) using standardized instructions, with each assessment lasting approximately 20–30 minutes. In the same week as the child pretest assessments, parents completed the Emotion Regulation Scale–Adult Form face-to-face with the researcher at school.

The intervention began in the first week following the pretest. The WMTP was delivered over eight consecutive weeks, with one session per week during the school day (typically after lunch) in the intervention classroom. Sessions followed a standardized structure (brief warm-up/attention-focusing activity, core working memory tasks, brief reflection and strategy emphasis, and closure) and were delivered by the first author. Attendance was monitored throughout the program, and make-up sessions were offered when needed to minimize missed exposure. The control group continued with regular classroom activities during this period.

Immediately after the final WMTP session (i.e., in the last week of the intervention period), posttest assessments were administered. Posttest procedures mirrored baseline procedures: children were assessed individually in the same setting and under the same standardized conditions. During the same posttest week, parents again completed the Emotion Regulation Scale–Adult Form face-to-face with the researcher.

To examine maintenance of gains, a follow-up assessment was conducted approximately three months after the end of the intervention for the intervention group only. The follow-up was framed as within-group maintenance/retention evidence rather than a between-group causal comparison because the control group did not complete follow-up assessments. Follow-up procedures replicated the earlier assessment conditions: children were assessed individually in the same quiet room using the same



measures and standardized instructions, and parents completed the Emotion Regulation Scale–Adult Form face-to-face with the researcher during the same follow-up week.

Initially, 52 children were enrolled (intervention $n=23$, control $n=29$). Three children discontinued participation (intervention: 1; control: 2) and were excluded from analyses because they did not complete the measurement sequence, resulting in a final analytic sample of 49 children (intervention $n=22$, control $n=27$).

Data Analysis

All analyses were conducted using IBM SPSS Statistics (Version 26). Prior to hypothesis testing, data were screened for accuracy, missingness, outliers, and assumption violations. There were no missing data for the primary study variables; therefore, analyses were conducted using complete cases. Descriptive statistics (means, standard deviations, and ranges) were computed for each outcome by group and measurement occasion (pretest, posttest, and, when available, follow-up). Normality was evaluated using Shapiro–Wilk tests and inspection of skewness and kurtosis values and Q–Q plots. Homogeneity of variances was assessed with Levene’s test. The significance level was set at $\alpha = .05$ (two-tailed). Along with p-values, effect sizes and confidence intervals were reported to support interpretation.

Because participants were assigned at the classroom level (two intact classes) and baseline equivalence could not be assumed, posttest program effects were estimated using analysis of covariance (ANCOVA). Separate ANCOVA models were fitted for each working memory outcome (verbal, visual, and total) and each emotion regulation outcome (anger, fear, excitement, sadness, and total). In each model, the corresponding posttest score served as the dependent variable, group (WMTP vs. business-as-usual control) was the fixed factor, and the corresponding pretest score was entered as a covariate to adjust for baseline differences. Baseline-adjusted posttest group means (estimated marginal means) and adjusted mean differences (WMTP – control) were reported. Because group assignment occurred at the classroom level and only two intact classrooms were included, classroom-level influences could not be modeled separately from intervention effects. Therefore, all student-level analyses should be interpreted with caution, particularly with regard to causal inference.

Given the quasi-experimental design and the possibility of heterogeneity of variances in some outcomes, bootstrap bias-corrected and accelerated (BCa) 95% confidence intervals for adjusted mean differences were computed as a robustness check using 2,000 resamples. Partial eta squared (η^2) was reported for the omnibus group effect in each ANCOVA.

Follow-up assessments were administered only to the WMTP group; therefore, follow-up results were not used for between-group causal inference. Maintenance of gains within the WMTP group was examined descriptively (pretest–posttest–follow-up means and standard deviations) and, where appropriate, with repeated-measures analyses across the three time points (pretest, posttest, follow-up). These follow-up findings were interpreted as evidence of maintenance/retention within the intervention group and were reported in the Supplementary Material to keep the main Results focused on primary between-group estimates.

Bivariate associations between working memory indices (verbal, visual, total) and emotion regulation outcomes (subscales and total) were examined using Pearson product–moment correlations (two-tailed). To evaluate whether the association between total working memory and total emotion regulation remained after accounting for relevant child and family characteristics, partial correlations were computed while controlling for child age, number of siblings, birth order, child gender, mother’s education, father’s education, and preschool education attendance. Correlation coefficients (r), corresponding p-values, and degrees of freedom were reported.

RESULTS

This section presents the quantitative findings regarding the effectiveness of WMTP. The results are organized to examine (i) the program’s effects on working memory performance (verbal, visual, and



total); (ii) emotion regulation indicators; and (iii) the associations between working memory and emotion regulation. Descriptive statistics are reported first, followed by inferential analyses aligned with the study research questions. When appropriate, effect sizes and confidence intervals are provided to strengthen the reporting.

Effect of WMTP on Working Memory

This subsection examines the effect of WMTP on working memory performance. First, pretest and posttest means for the experimental and control groups, along with the follow-up assessment administered to the experimental group, are summarized. Next, between-group differences in posttest working memory outcomes are tested using ANCOVA while controlling for pretest scores. Because the follow-up assessment was collected only in the experimental group, follow-up findings are interpreted descriptively as an indicator of maintenance rather than as a between-group program effect.

Table 2. Pretest, posttest, and follow-up means (SD) for working memory in the experimental and control groups

Outcome	Group	Pre-test Mean (SD)	Post-test Mean (SD)	Follow-up Mean (SD)*
Verbal Memory	Experimental	12.32 (3.97)	16.50 (4.73)	17.27 (5.61)
	Control	16.04 (5.03)	17.59 (5.81)	—
Visual Memory	Experimental	5.68 (4.13)	8.59 (4.55)	9.23 (4.14)
	Control	7.30 (3.58)	8.30 (3.52)	—
Total WM	Experimental	18.00 (7.30)	25.09 (8.48)	26.50 (9.00)
	Control	23.33 (7.20)	25.74 (8.12)	—

*Follow-up was administered only to the experimental group.

Descriptive results indicate that the experimental group showed increases from pretest to posttest in verbal memory, visual memory, and total working memory, and that these gains were generally maintained at follow-up. However, because the control group had higher pretest means on some sub-dimensions (particularly verbal memory and total working memory), the groups were not fully equivalent at baseline. Therefore, evaluating program effects based on between-group posttest differences while controlling for pretest scores was considered more appropriate.

At follow-up, the experimental group's scores generally showed a stable or increasing trend; detailed follow-up results are provided in the Supplementary Material (see Supplementary Table S1).

Following the descriptive findings, the effect of WMTP on posttest working memory outcomes was examined using ANCOVA with pretest scores as covariates. For each sub-dimension, we report the group effect (F , p , and η^2), baseline-adjusted (estimated marginal) posttest means, and bootstrap BCa 95% confidence intervals for the adjusted mean differences (Table 3).

Table 3. ANCOVA results for posttest working memory outcomes (adjusted for pretest scores) and Bootstrap BCa CI

Outcome	Levene p	Group effect (df)	F	p	η^2	Adjusted Posttest Mean (SE) Exp	Adjusted Posttest Mean (SE) Ctrl	Adj. Mean Diff (Exp–Ctrl) BCa 95% CI
Verbal Memory	.327	Group (1,46)	14.793	<.001	.243	18.653 (.525)	15.838 (.470)	2.815 ** [1.410, 4.133]
Visual Memory	.039	Group (1,46)	6.311	.016	.121	9.359 (.494)	7.671 (.445)	1.688 * [.314, 3.136]
Total	.381	Group (1,46)	21.676	<.001	.320	28.154 (.761)	23.245 (.682)	4.909 ** [2.806, 7.058]

Note. η^2 = partial eta squared. Adjusted means are estimated marginal means from ANCOVA evaluated at the sample mean of the covariate. The 95% confidence intervals for adjusted mean differences are bootstrap BCa intervals based on 2,000 resamples. * $p < .05$. ** $p < .01$.



After controlling for pretest scores, the experimental group scored significantly higher than the control group on posttest verbal working memory, visual working memory, and total working memory. Effect sizes were $\eta^2 = .243$ for verbal working memory, $\eta^2 = .121$ for visual working memory, and $\eta^2 = .320$ for total working memory, with the largest effect observed for total working memory. In addition, the bootstrap BCa 95% confidence intervals for the adjusted mean differences did not include zero, supporting the direction and consistency of the between-group effects.

Effect of WMTP on Emotion Regulation

In this section, the effects of WMTP on children's emotion regulation are reported. First, descriptive statistics (pretest, posttest, and follow-up means and standard deviations) are presented for the experimental and control groups. Next, to evaluate the program effect on posttest outcomes while accounting for baseline differences, posttest emotion regulation scores were compared between groups using ANCOVA with pretest scores entered as a covariate. Along with the omnibus group effects (F, p, and partial eta squared), baseline-adjusted (estimated marginal) posttest means and bootstrap BCa 95% confidence intervals for the adjusted mean differences are reported. Follow-up (maintenance) data were collected only in the experimental group and are therefore reported in the Supplementary Material as descriptive information rather than as a between-group test. Table 4 presents these descriptive statistics.

Table 4. Pretest, posttest, and follow-up means (SD) for emotion regulation in the experimental and control groups

Subscale	Group	Pretest Mean (SD)	Posttest Mean (SD)	Follow-up Mean (SD)*
Anger	Experimental	16.73 (3.99)	19.77 (4.18)	19.41 (3.82)
	Control	23.30 (4.79)	23.81 (1.57)	—
Fear	Experimental	17.82 (4.50)	19.77 (3.90)	19.68 (3.80)
	Control	24.19 (3.55)	24.19 (2.04)	—
Excitement	Experimental	17.86 (4.38)	20.55 (4.21)	20.55 (4.21)
	Control	23.59 (3.57)	22.48 (2.64)	—
Sadness	Experimental	18.45 (3.16)	19.45 (4.21)	19.36 (3.42)
	Control	23.22 (4.60)	23.04 (2.82)	—
Total	Experimental	70.86 (13.58)	79.55 (14.40)	79.00 (13.29)
	Control	94.30 (4.78)	93.52 (5.27)	—

*The follow-up assessment was administered only to the experimental group.

As shown in Table 4, the experimental group increased from pretest to posttest across all subscales and the total score. In the control group, the pattern was less consistent: anger scores increased slightly, fear scores remained unchanged, whereas excitement, sadness, and total emotion regulation scores decreased from pretest to posttest. At follow-up, scores in the experimental group remained close to their posttest levels across all outcomes, suggesting maintenance of the observed posttest gains.

Because the groups differed at baseline on emotion regulation scores (Table 4), posttest outcomes were compared using ANCOVA with corresponding pretest scores entered as covariates. For each outcome, we report the group effect (F, p, η^2), baseline-adjusted (estimated marginal) posttest means, and bootstrap BCa 95% confidence intervals for the adjusted mean differences (Exp – Ctrl) (Table 5).

Table 5. Baseline-adjusted posttest comparisons for Emotion Regulation: ANCOVA and Bootstrap BCa CIs to CIs

Subscale	Levene p	Group effect F(1,46)	p	η^2	Adjusted Posttest Mean (SE) Exp	Adjusted Posttest Mean (SE) Ctrl	Adj. Mean Diff (Exp–Ctrl) BCa 95% CI
Anger	.046	6.12	.017*	.117	22.31 (.69)	20.92 (.62)	1.39 [.25, 2.74]
Fear	.388	3.04	.088	.062	22.21 (.68)	21.18 (.60)	1.03 [–.18, 2.31]



Table 5 (Continued). Baseline-adjusted posttest comparisons for Emotion Regulation: ANCOVA and Bootstrap BCa CIs to CIs

Subscale	Levene p	Group effect F _(1,46)	p	η^2	Adjusted Posttest Mean (SE) Exp	Adjusted Posttest Mean (SE) Ctrl	Adj. Mean Diff (Exp–Ctrl) BCa 95% CI
Excitement	.382	5.27	.026*	.103	22.97 (.77)	21.39 (.69)	1.58 [.19, 3.06]
Sadness	.229	1.41	.241	.030	21.23 (.75)	20.47 (.68)	.76 [–.59, 2.20]
Total	.110	5.84	.020*	.113	88.12 (2.21)	83.94 (2.00)	4.18 [–.76, 7.96]

Note. η^2 = partial eta squared. Adjusted means are estimated marginal means from ANCOVA evaluated at the sample mean of the covariate. Confidence intervals for adjusted mean differences (Exp–Ctrl) are bootstrap BCa 95% CIs based on 2,000 resamples.

After controlling for pretest scores, significant group differences favoring the experimental group were found for anger, $F_{(1, 46)} = 6.12$, $p = .017$, $\eta^2 = .117$; excitement, $F_{(1, 46)} = 5.27$, $p = .026$, $\eta^2 = .103$; and total emotion regulation, $F_{(1, 46)} = 5.84$, $p = .020$, $\eta^2 = .113$. The baseline-adjusted posttest mean was higher in the experimental group than in the control group for anger (22.31 vs. 20.92; adjusted mean difference = 1.39, BCa 95% CI [0.25, 2.74]), excitement (22.97 vs. 21.39; adjusted mean difference = 1.58, BCa 95% CI [0.19, 3.06]), and total emotion regulation (88.12 vs. 83.94; adjusted mean difference = 4.18, BCa 95% CI [–.76, 7.96]). In each case, the bootstrap confidence interval excluded zero, consistent with the ANCOVA results.

In contrast, the adjusted between-group differences were not statistically significant for fear, $F_{(1, 46)} = 3.04$, $p = .088$, $\eta^2 = .062$, or sadness, $F_{(1, 46)} = 1.41$, $p = .241$, $\eta^2 = .030$. Although the adjusted posttest means were numerically higher in the experimental group for both fear (22.21 vs. 21.18) and sadness (21.23 vs. 20.47), the corresponding bootstrap BCa 95% confidence intervals included zero (fear: adjusted mean difference = 1.03, 95% CI [–.18, 2.31]; sadness: adjusted mean difference = .76, 95% CI [–.59, 2.20]), providing no statistically significant evidence of between-group differences for these outcomes.

Levene's test suggested potential heterogeneity of variances for the anger subscale ($p = .046$), whereas it was not significant for the remaining outcomes (Table 5). Accordingly, the bootstrap BCa confidence intervals are emphasized as a robustness check for the adjusted mean differences.

Relationship Between Working Memory and Emotion Regulation

This section examines the associations between working memory (verbal, visual, and total) and emotion regulation (anger, fear, excitement, sadness, and total). Pearson product–moment correlation coefficients (two-tailed) were computed to evaluate bivariate relationships. In addition, partial correlations were conducted for the association between total working memory and total emotion regulation, controlling for child age, number of siblings, birth order, child gender, mother's education, father's education, and preschool education attendance. Table 6 summarizes these bivariate and partial correlations.

Table 6. Pearson correlations between working memory and emotion regulation (total) and partial correlation for total scores

Working memory	Anger	Fear	Excitement	Sadness	Emotion Regulation (Total)	Partial r (Total ER)*
Verbal Memory	.27	.37**	.35*	.39**	.37**	.26
Visual Memory	.24	.32*	.34*	.36*	.34*	.25
Total WM	.29*	.40**	.39**	.43**	.41**	.31*

Note. N = 49. Values are Pearson correlation coefficients (two-tailed). * $p < .05$. ** $p < .01$.

*Partial correlations are reported for the association between each working memory index and total emotion regulation, controlling for age, number of siblings, birth order, gender, mother's education, father's education, and preschool education attendance (df = 40).

As shown in Table 6, verbal working memory was positively correlated with fear ($r = .37$, $p < .01$), excitement ($r = .35$, $p < .05$), sadness ($r = .39$, $p < .01$), and total emotion regulation ($r = .37$, $p < .01$), whereas its correlation with anger was not significant ($r = .27$, $p = .058$). Visual working memory



showed significant positive correlations with fear ($r = .32, p < .05$), excitement ($r = .34, p < .05$), sadness ($r = .36, p < .05$), and total emotion regulation ($r = .34, p < .05$), but was not significantly related to anger ($r = .24, p = .104$). Total working memory was positively correlated with anger ($r = .29, p < .05$), fear ($r = .40, p < .01$), excitement ($r = .39, p < .01$), sadness ($r = .43, p < .01$), and total emotion regulation ($r = .41, p < .01$). Finally, the association between each working memory index and total emotion regulation remained positive after controlling for covariates (verbal WM: partial $r = .26, p = .099$; visual WM: partial $r = .25, p = .107$; total WM: partial $r = .31, p < .05$; $df = 40$), indicating that only the total working memory–total emotion regulation association remained statistically significant after covariate adjustment.

DISCUSSION, CONCLUSION, and RECOMMENDATIONS

This study investigated whether a WMTTP implemented in a classroom setting was associated with gains in WM and ER in second-grade children and whether these areas were interrelated. Overall, the findings pointed to a clearer and more pronounced pattern for the primary cognitive goals of the intervention than for the broader socio-emotional outcomes. After accounting for baseline differences, the intervention group demonstrated higher posttest scores than the control group in verbal, visual, and total working memory, and follow-up scores for the experimental group showed that these gains were maintained over time. Emotion regulation outcomes were also generally positive, but these should be interpreted more cautiously as potential transfer effects rather than direct effects of a program that explicitly targets emotional processes.

The clearest effects in the present study were observed for working memory outcomes, which is consistent with the literature indicating that working memory training tends to produce its most reliable benefits on targeted or near-transfer measures. Bharadwaj et al. (2022) demonstrated in their meta-analysis that working memory training yielded significant post-intervention improvements in verbal and visuospatial working memory in school-age children, with only limited evidence for more distal transfer. Likewise, Vernucci et al. (2023) reported that process-based training improved working memory performance in typically developing school-age children, while also noting that broader transfer effects were less uniform. Extending this pattern to ecologically valid school settings, Berger et al. (2025) found substantial immediate and lasting gains in children's working memory capacity following a randomized educational intervention. Hong and Zhou (2025) further showed that computerized working memory training can enhance children's working memory capacity, although these gains are shaped by the characteristics of the training task itself. Against this background, the present findings, showing clearer improvements in verbal, visual, and total working memory than in more distal outcomes, fit well with the broader conclusion that working memory training most robustly affects the cognitive functions it directly engages.

The effects of the intervention were not equally strong across all working memory outcomes, and this pattern is consistent with previous research. In the present study, gains were more pronounced for total and verbal working memory than for visual working memory. Similar variation has been reported in the literature. For example, Bharadwaj et al. (2022) found post-intervention improvements in both verbal and visuospatial working memory, but only verbal working memory showed evidence of maintenance at follow-up, suggesting that some working memory domains may be more responsive or more stable over time than others. Likewise, Vernucci et al. (2023) concluded that process-based training can improve untrained working memory performance in typically developing school-age children, although the size of these gains may differ across outcomes. Dunning et al. (2013) and Henry et al. (2014) reported a similar pattern, with clearer gains on working memory measures than on broader cognitive or academic outcomes. Taken together, these studies suggest that the stronger gains observed here in verbal and total working memory, relative to visual working memory, are not unusual. Rather, they likely reflect differences in how closely the training activities matched the outcome measures, as well as variation in the cognitive demands and developmental sensitivity of different working memory components.



The emotion regulation findings in the present study are best interpreted as promising but secondary transfer effects rather than direct outcomes of the intervention. This interpretation is consistent with Barkus's (2020) review, which showed that working memory training can support several emotion-related outcomes, although these effects are typically less consistent than gains on targeted cognitive measures. Similarly, Tzuriel et al. (2024) found that non-computerized working memory training in typically developing preschool children improved not only some working memory outcomes but also some self-regulation measures, whereas transfer did not extend uniformly across all domains. Likewise, Zhao and Zhang (2024) reported that working memory training in children with ADHD improved adaptive cognitive emotion-regulation strategies and reduced maladaptive strategies, with effects maintained at follow-up. In the same vein, Bombonato et al. (2024) concluded that executive-function training can yield far-transfer effects in children but also emphasized substantial heterogeneity across studies. Taken together, this literature suggests that improvements in emotion regulation following working memory training are theoretically plausible and empirically supported under some conditions, but they should still be interpreted cautiously because they are more variable and context-sensitive than gains on targeted working memory outcomes.

A plausible explanation for these transfer effects is that working memory and emotion regulation rely on partially overlapping self-regulatory processes, especially controlled attention, goal maintenance, and the inhibition of less adaptive responses. Wolfe and Bell (2007) provided early developmental evidence that working memory is closely linked to attentional and regulatory processes, while Raver and Blair (2016) argued that children's adaptation depends on an integrated system of cognitive, emotional, and behavioral regulation. Similarly, Zelazo and Carlson (2012) emphasized that both "cool" and "hot" executive functions depend on top-down control, even when expressed in affectively different contexts. More recent studies support the functional significance of this overlap. Groves et al. (2020, 2022) showed that poorer working memory is associated with greater emotion dysregulation, whereas better-developed working memory predicts better emotion regulation. Experimental findings also point in the same direction: Xie et al. (2021) and Li et al. (2020) reported that executive-function training improved both cognitive control and emotion-related functioning in young children. Within this framework, the present pattern, showing clearer gains in anger, excitement, and total emotion regulation, but not in fear or sadness, may suggest that the WMTP primarily supported specific aspects of emotion regulation rather than producing uniform improvements across emotional domains. Accordingly, the nonsignificant findings for fear and sadness should not be taken as contradicting the overall findings, but rather as indicating that transfer to emotion regulation may be selective, with some emotional domains appearing more sensitive than others to improvements in children's underlying cognitive-control systems.

The correlational findings further support the view that working memory and emotion regulation are developmentally linked through shared self-regulatory processes. In the present study, verbal WM was positively associated with fear, excitement, sadness, and total ER, whereas its association with anger was not statistically significant. Visual WM was positively associated with fear, excitement, sadness, and total ER, but not with anger. Total WM, by contrast, was positively associated with all ER outcomes, and its relation with total ER remained significant even after controlling for age, number of siblings, birth order, gender, parental education, and preschool education attendance. Wolfe and Bell (2007) showed that early attentional and regulatory processes are closely tied to later working memory performance, while Raver and Blair (2016) conceptualized self-regulation as an integrated system in which cognitive, emotional, and behavioral control processes jointly support children's adaptation. Similarly, Zelazo and Carlson (2012) argued that both cool and hot executive functions rely on top-down control, making it plausible that stronger working memory would be associated with better emotion regulation. More direct support comes from Groves et al. (2020, 2022), who reported that poorer working memory was linked to greater emotion dysregulation, whereas stronger working memory predicted better emotional functioning. Taken together, the present findings suggest that the WM-ER association is most meaningfully understood at the level of general



regulatory capacity: children with stronger working memory may be better able to maintain goals, direct attention, and inhibit less adaptive responses when managing emotional demands.

A key strength of this study is that it contributes at both theoretical and contextual levels. First, WMTP demonstrated its most pronounced effects on its targeted outcomes: verbal, visual, and total working memory, reinforcing confidence in the specificity and structural fit of the intervention. Second, the study contributes not only by evaluating intervention-related outcomes but also by examining the association between working memory and emotion regulation, thus offering theoretically relevant insight into how cognitive and emotional self-regulation may be connected. Third, the emotion regulation findings should be interpreted cautiously, but the presence of positive patterns across various indicators of emotion regulation provides preliminary evidence that the effects of school-based working memory training may extend beyond narrowly defined cognitive performance. Another strength is the inclusion of follow-up data for the intervention group, allowing for at least an initial examination of sustainability rather than limiting interpretation to immediate posttest effects. Finally, the implementation of the program in a real school setting enhances the practical significance of the findings by demonstrating that a structured, developmentally appropriate working memory intervention can be delivered under ordinary educational conditions. Taken together, these features strengthen the study's contribution by positioning it not merely as an educational trial but as a contextually grounded and theoretically informed investigation into how a school-based cognitive intervention can support children's broader self-regulatory functioning.

Limitations

These findings should be evaluated in light of several methodological limitations, while also considering the strengths of the study. First, the study used a quasi-experimental design with existing classrooms in a natural school setting. ANCOVA adjusted for observed pretest differences, but it cannot eliminate selection bias or unmeasured confounding. The study also included only two classrooms from a single public school and a relatively small sample, which limits generalizability. Accordingly, the adjusted between-group differences should be interpreted as preliminary quasi-experimental evidence rather than definitive causal estimates. The fact that follow-up measurement was available only for the experimental group also means that the follow-up findings should be interpreted as within-group evidence of the maintenance of gains. Emotion regulation in this article was assessed using adult-reported measures. This choice was intended to provide an alternative assessment source by capturing adults' observations of children's emotion regulation patterns in everyday contexts; however, the future use of child self-report together with other measurement sources may contribute to a more comprehensive examination of these findings. Finally, because emotion regulation was not the primary goal of the intervention, it is more appropriate to consider the results in this area as potential secondary/transfer effects rather than direct effects.

Conclusion

In conclusion, this study suggests that a classroom-based WMTP can support children's working memory development and may also have positive implications for emotion regulation, although these findings warrant greater caution. The strongest effects were observed for the cognitive outcomes directly targeted by the intervention, while gains related to emotion regulation should be considered possible secondary or transfer effects. However, the correlational findings between working memory and emotion regulation suggest that these two areas may be linked through common self-regulation processes. In this respect, the study demonstrates that an intervention conducted in a real school context can have both theoretical and practical value and underscores the importance of considering cognitive support programs in the early school years in relation to broader developmental outcomes.

Recommendations

The findings of this study suggest that structured working memory activities in early elementary school can be considered a viable school-based approach that not only supports cognitive skills but also contributes to children's broader foundation for self-regulation. However, such programs are better viewed not as direct emotion regulation interventions but primarily as supportive practices that



target cognitive control processes and may provide secondary gains in emotion regulation. In terms of implementation, age-appropriate program structure, sufficient implementation time, careful program preparation, and close monitoring of the implementation process are important. Future research could contribute to the field by re-examining these findings in different school contexts and with more diverse samples, addressing follow-up measures more comprehensively, and supporting emotion regulation findings with multiple sources of information. Furthermore, studies examining the conditions under which the relationship between working memory and emotion regulation becomes more pronounced could help clarify these findings.

Funding

No specific funding was received for this study.

Ethics and Conflict of Interest

The study received ethical approval from the Scientific Publication and Ethics Committee of Gazi University (No. E.585142, dated 09/02/2023) and official implementation permission from the Ministry of National Education (No. E.74258403, dated 12/04/2023). The research was conducted in accordance with relevant ethical principles. Written informed consent was obtained from parents, and child assent was obtained from participating children. The authors declare that there are no conflicts of interest regarding the research, authorship, and/or publication of this article.

Author Contribution

This article is derived from the first author's doctoral dissertation titled "Examining the Effects of a Working Memory Training Program on Emotion Regulation and Working Memory Skills of Second-Grade Students" completed at Gazi University in December 2025 under the supervision of the second author. Emire Uluğ: Conceptualization, methodology, investigation, data curation, formal analysis, development and implementation of the intervention, and writing—original draft. Ayşe Dilek Öğretir Özçelik: Supervision, methodology, validation, and writing—review and editing. Both authors read and approved the final manuscript.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Corresponding Author

Correspondence to Emire ULUĞ, emireulug@karatekin.edu.tr

REFERENCES

- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417–423. [https://doi.org/10.1016/S1364-6613\(00\)01538-2](https://doi.org/10.1016/S1364-6613(00)01538-2)
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–90). [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Barkus, E. (2020). Effects of working memory training on emotion regulation: Transdiagnostic review. *PsyCh Journal*, 9(2), 258–279. <https://doi.org/10.1002/pchj.353>
- Berger, E. M., Fehr, E., Hermes, H., Schunk, D., & Winkel, K. (2025). The impact of working-memory training on children's cognitive and noncognitive skills. *Journal of Political Economy*, 133(2), 492–521. <https://doi.org/10.1086/732884>
- Bharadwaj, S. V., Yeatts, P., & Headley, J. (2022). Efficacy of cogmed working memory training program in improving working memory in school-age children with and without neurological insults or disorders: A meta-analysis. *Applied Neuropsychology: Child*, 11(4), 891–903. <https://doi.org/10.1080/21622965.2021.1920943>
- Bombonato, C., Del Lucchese, B., Ruffini, C., Di Lieto, M. C., Brovedani, P., Sgandurra, G., Cioni, G., & Pecini, C. (2024). Far transfer effects of trainings on executive functions in neurodevelopmental disorders: A systematic review and metanalysis. *Neuropsychology Review*, 34, 98–133. <https://doi.org/10.1007/s11065-022-09574-z>
- Bruning, R., Schraw, G., & Norby, M. (2014). *Bilişsel psikoloji ve öğretim* [in Turkish]. (ZN Ersözlü and R. Ülker, Trans.). Ankara: Nobel Academic Publishing.



- Buluş, M. (2021). Sample size determination and optimal design of randomized/non-equivalent pretest-posttest control-group designs. *Adiyaman University Journal of Educational Sciences*, 11(1), 48-69. <https://doi.org/10.17984/adyuebd.941434>
- Cowan, N. (2005). Working-memory capacity limits in a theoretical context. In C. Izawa & N. Ohta (Eds.), *Human learning and memory: Advances in theory and application: The 4th Tsukuba International Conference on Memory* (pp. 155–175). Lawrence Erlbaum Associates Publishers.
- Dunning, D. L., Holmes, J., & Gathercole, S. E. (2013). Does working memory training lead to generalized improvements in children with low working memory? A randomized controlled trial. *Developmental science*, 16(6), 915-925. <https://doi.org/10.1111/desc.12068>
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, 11(1), 19–23. <https://doi.org/10.1111/1467-8721.001>
- Ergül, C., Yılmaz, Ç. Ö., & Demir, E. (2018). 5-10 yaş grubu çocuklara yönelik geliştirilmiş çalışma belleği ölçeğinin geçerlik ve güvenirliği [Validity and reliability of the working memory scale for children aged 5-10 years]. *Journal of Theory and Practice in Education*, 14(2), 187-214. <https://doi.org/10.17244/eku.427280>
- Gopalan, M., Rosinger, K., & Ahn, J. B. (2020). Use of quasi-experimental research designs in education research: Growth, promise, and challenges. *Review of Research in Education*, 44(1), 218–243. <https://doi.org/10.3102/0091732X20903302>
- Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, 2(3), 271–299. <https://doi.org/10.1037/1089-2680.2.3.27>
- Gross, J. J. (2015). Emotion regulation: Current status and future prospects. *Psychological Inquiry*, 26(1), 1–26. <https://doi.org/10.1080/1047840X.2014.940781>
- Groves, N. B., Kofler, M. J., Wells, E. L., Day, T. N., & Chan, E. S. (2020). An examination of relations among working memory, ADHD symptoms, and emotion regulation. *Journal of Abnormal Child Psychology*, 48(4), 525-537. <https://doi.org/10.1007/s10802-019-00612-8>
- Groves, N. B., Wells, E. L., Soto, E. F., Marsh, C. L., Jaisle, E. M., Harvey, T. K., & Kofler, M. J. (2022). Executive functioning and emotion regulation in children with and without ADHD. Research on child and adolescent psychopathology, 50(6), 721-735. <https://doi.org/10.1007/s10802-021-00883-0>
- Handley, M. A., Lyles, C. R., McCulloch, C., & Cattamanchi, A. (2018). Selecting and improving quasi-experimental designs in effectiveness and implementation research. *Annual Review of Public Health*, 39, 5–25. <https://doi.org/10.1146/annurev-publhealth-040617-014128>
- Hendricks, M. A., & Buchanan, T. W. (2016). Individual differences in cognitive control processes and their relationship to emotion regulation. *Cognition and Emotion*, 30(5), 912–924. <https://doi.org/10.1080/02699931.2015.1032893>
- Henry, L., Messer, D., & Nash, G. (2014). Testing for near and far transfer effects with a short, face-to-face adaptive working memory training intervention in typical children. *Infant and Child Development*, 23(1), 84–103. <https://doi.org/10.1002/icd.1816>
- Hong, Y., & Zhou, R. (2025). Can working memory training improve working memory capacity in typically developing children?: Training methods reveal specific effects. *Biological Psychology*. 202, 109140. <https://doi.org/10.1016/j.biopsycho.2025.109140>
- Hou, T. Y., & Cai, W. P. (2022). What emotion dimensions can affect working memory performance in healthy adults? A review. *World Journal of Clinical Cases*, 10(2), 401–417. <https://doi.org/10.12998/wjcc.v10.i2.401>
- Kahl, T., Grob, A., & Möhring, W. (2021). Does emotion regulation compensate deficits in various executive functions in children's and adolescents' mathematical achievement?. *Learning and Individual Differences*, 89, 102034. <https://doi.org/10.1016/j.lindif.2021.102034>
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review*, 9(4), 637–671. <https://doi.org/10.3758/BF03196323>
- Kassai, R., Futo, J., Demetrovics, Z., & Takacs, Z. K. (2019). A meta-analysis of the experimental evidence on the near- and far-transfer effects among children's executive function skills. *Psychological Bulletin*, 145(2), 165–188. <https://doi.org/10.1037/bul0000180>
- Li, Q., Liu, P., Yan, N., & Feng, T. (2020). Executive function training improves emotional competence for preschool children: The roles of inhibition control and working memory. *Frontiers in psychology*, 11, 347. <https://doi.org/10.3389/fpsyg.2020.00347>



- Malagoli, C., & Usai, M. C. (2018). The effects of gender and age on inhibition and working memory organization in 14-to 19-year-old adolescents and young adults. *Cognitive Development*, 45, 10-23. <https://doi.org/10.1016/j.cogdev.2017.10.005>
- Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Au et al.(2014) and Karbach and Verhaeghen (2014). *Psychonomic Bulletin & Review*, 23(1), 324-330. <https://doi.org/10.3758/s13423-015-0862-z>
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*, 41(1), 49–100. <https://doi.org/10.1006/cogp.1999.0734>
- Osaka, M., Yaoi, K., Minamoto, T., & Osaka, N. (2013). When do negative and positive emotions modulate working memory performance? *Scientific Reports*, 3, Article 1375. <https://doi.org/10.1038/srep01375>
- Ou, J., Mou, H., Zhou, T., & Wang, Y. (2026). Does emotional working memory training ameliorate anxiety and depression? A meta-analytic review. *Brain Sciences*, 16(1), 30. <https://doi.org/10.3390/brainsci16010030>
- Papé, T. L., & Maree, D. J. (2023). A systematic exploratory review investigating the relationship between working memory and emotion regulation: implications for working memory training. *South African Journal of Psychology*, 53(4), 548-564. <https://doi.org/10.1177/00812463231186347>
- Pe, M. L., Raes, F., & Kuppens, P. (2013). The cognitive building blocks of emotion regulation: Ability to update working memory moderates the efficacy of rumination and reappraisal on emotion. *PLOS ONE*, 8(7), e69071. <https://doi.org/10.1371/journal.pone.0069071>
- Phillips, L. H., Bull, R., Adams, E., & Fraser, L. (2002). Positive mood and executive function: Evidence from Stroop and fluency tasks. *Emotion*, 2(1), 12–22. <https://doi.org/10.1037/1528-3542.2.1.12>
- Raver, C. C., & Blair, C. (2016). Neuroscientific insights: Attention, working memory, and inhibitory control. *The Future of Children*, 95-118. <http://www.jstor.org/stable/43940583>
- Rydell, A. M., Berlin, L., & Bohlin, G. (2003). Emotionality, emotion regulation, and adaptation among 5- to 8-year-old children. *Emotion*, 3(1), 30–47. <https://doi.org/10.1037/1528-3542.3.1.30>
- Schmeichel, B. J., & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science*, 24(2), 93–98. <https://doi.org/10.1177/0963721414555178>
- Schweizer, S., Satpute, A. B., Atzil, S., Field, A. P., Hitchcock, C., Black, M., Barrett, L. F., & Dalgleish, T. (2019). The impact of affective information on working memory: A pair of meta-analytic reviews of behavioral and neuroimaging evidence. *Psychological Bulletin*, 145(6), 566–609. <https://doi.org/10.1037/bul0000193>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin. Boston, New York.
- Tatlı-Harmancı, S., & Güngör-Aytar, A. (2023). Çocuklar için duygu düzenleme ölçeği çocuk formu (çddö) ve yetişkin formunun (ÇDDÖ-YF) Türkçeye uyarlanması [Study of Turkish adaption of the emotion regulation scale for children child form (ÇDDÖ) and adult Form (ÇDDÖ-YF)]. *Milli Eğitim Journal*, 52(237), 71–106. <https://doi.org/10.37669/milliegitim.1072215>
- Toh, W. X., Keh, J. S., Gross, J. J., & Carstensen, L. L. (2024). The role of executive function in cognitive reappraisal: A meta-analytic review. *Emotion*, 24(7), 1563. <https://doi.org/10.1037/emo0001373>
- Tzuriel, D., Weiss, T., & Kashy-Rosenbaum, G. (2024). The effects of working memory training on working memory, self-regulation, and analogical reasoning of preschool children. *British Journal of Educational Psychology*, 94(4), 1132-1160. <https://doi.org/10.1111/bjep.12709>
- Urry, H. L., & Gross, J. J. (2010). Emotion regulation in older age. *Current Directions in Psychological Science*, 19(6), 352-357. <https://doi.org/10.1177/0963721410388395>
- Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline had more power in randomized studies and more bias in nonrandomized studies. *Journal of Clinical Epidemiology*, 59(9), 920–925. <https://doi.org/10.1016/j.jclinepi.2006.02.007>
- Vernucci, S., Canet-Juric, L., & Richard's, M. M. (2023). Effects of working memory training on cognitive and academic abilities in typically developing school-age children. *Psychological Research*, 87(1), 308–326. <https://doi.org/10.1007/s00426-022-01647-1>
- Vickers, A. J., & Altman, D. G. (2001). Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123–1124. <https://doi.org/10.1136/bmj.323.7321.1123>



- Wolfe, C. D., & Bell, M. A. (2007). The integration of cognition and emotion during infancy and early childhood: Regulatory processes associated with the development of working memory. *Brain and Cognition*, 65(1), 3–13. <https://doi.org/10.1016/j.bandc.2006.01.009>
- Xie, J., Liu, S., & Fang, P. (2021). Cognitive training improves emotion regulation in Chinese preschool children. *Pediatrics International*, 63(11), 1303–1310. <https://doi.org/10.1111/ped.14661>
- Xiu, L., Zhou, R., & Jiang, Y. (2016). Working memory training improves emotion regulation ability: Evidence from HRV. *Physiology & Behavior*, 155, 25–29. <https://doi.org/10.1016/j.physbeh.2015.12.004>
- Zaehring, J., Falquez, R., Schubert, A. L., Nees, F., & Barnow, S. (2018). Neural correlates of reappraisal considering working memory capacity and cognitive flexibility. *Brain Imaging and Behavior*, 12, 1529–1543. <https://doi.org/10.1007/s11682-017-9788-6>
- Zelazo, P. D., & Carlson, S. M. (2012). Hot and cool executive function in childhood and adolescence: Development and plasticity. *Child Development Perspectives*, 6(4), 354–360. <https://doi.org/10.1111/j.1750-8606.2012.00246.x>
- Zhao, D., & Zhang, J. (2024). The effects of working memory training on attention deficit, adaptive and non-adaptive cognitive emotion regulation of Chinese children with Attention Deficit/Hyperactivity Disorder (ADHD). *BMC Psychology*, 12(1), 59. <https://doi.org/10.1186/s40359-024-01539-6>

About the Authors:

Emire ULUĞ

Emire Uluğ is an Assistant Professor in the Child Development Program at Eldivan Vocational School of Health Services, Çankırı Karatekin University, Türkiye. She holds a PhD in Child Development and Education from Gazi University. Her research interests include child development, working memory and executive functions, emotion regulation, parenting, and the use of artificial intelligence in education and child development. Her research particularly focuses on cognitive and socio-emotional development in children and the development and evaluation of educational interventions designed to support these skills.

Ayşe Dilek ÖĞRETİR ÖZÇELİK

Ayşe Dilek Öğretir Özçelik is a Professor in the Department of Early Childhood Education, Gazi Faculty of Education, Gazi University, Türkiye. She holds a PhD in Educational Sciences. She has extensive experience in teaching, research, graduate supervision, and academic projects in the field of child development and education. Her research interests include early childhood education, child development, parenting attitudes and practices, child psychology, and children's social and emotional development.

APPENDICES

Supplementary Table S1.

Follow-up outcomes within the intervention group (pretest–posttest–follow-up)

Outcome	Pretest M (SE) [95% CI]	Posttest M (SE) [95% CI]	Follow-up M (SE) [95% CI]	F(df1, df2)	p	η^2	Pre- Post p	Post- FU p	Pre- FU p
Total working memory	18.00 (1.56) [14.76, 21.24]	25.09 (1.81) [21.33, 28.85]	26.50 (1.92) [22.51, 30.49]	64.59 (2, 42)	< .001	.755	< .001	.103	< .001
Verbal working memory*	2.73 (.34) [2.02, 3.44]	4.00 (.40) [3.17, 4.83]	4.36 (0.55) [3.22, 5.50]	11.46 (1.504, 31.577)	< .001	.353	< .001	.913	.004
Visual working memory*	3.09 (.57) [1.90, 4.28]	5.00 (.51) [3.94, 6.06]	4.77 (.46) [3.80, 5.74]	13.80 (1.268, 26.629)	< .001	.397	.002	.782	.004
Emotion regulation –total (adult form)	70.86 (2.90) [64.84, 76.88]	79.55 (3.07) [73.17, 85.93]	79.00 (2.83) [73.11, 84.89]	30.03 (1.083, 22.734)	< .001	.588	< .001	.921	< .001

Note. FU = follow-up. Pairwise comparisons are Bonferroni-adjusted. Follow-up data were available only for the intervention group; therefore, maintenance was evaluated within the intervention group and could not be tested



comparatively against the control group.

*Greenhouse–Geisser corrected degrees of freedom were reported due to sphericity violations.

Follow-up (retention) analyses conducted within the intervention group ($n = 22$) indicated significant time effects for all outcomes, suggesting meaningful change from pretest to posttest and the follow-up period. Total working memory showed a significant effect of Time, $F_{(2, 42)} = 64.59$, $p < .001$, $\eta^2 = .755$; Bonferroni-adjusted comparisons indicated a significant increase from pretest to posttest ($p < .001$) and from pretest to follow-up ($p < .001$), while the posttest–follow-up difference was not significant ($p = .103$), supporting maintenance of gains. Verbal working memory also demonstrated a significant time effect with Greenhouse–Geisser correction, $F_{(1.504, 31.577)} = 11.46$, $p < .001$, $\eta^2 = .353$; scores improved from pretest to posttest ($p < .001$) and remained higher at follow-up relative to pretest ($p = .004$), with no significant change from posttest to follow-up ($p = .913$), indicating maintenance. Similarly, visual working memory showed a significant corrected time effect, $F_{(1.268, 26.629)} = 13.80$, $p < .001$, $\eta^2 = .397$; pretest-to-posttest ($p = .002$) and pretest-to-follow-up ($p = .004$) improvements were significant, whereas posttest and follow-up did not differ ($p = .782$), again suggesting maintained gains. Finally, adult emotion regulation (total) displayed a significant corrected time effect, $F_{(1.083, 22.734)} = 30.03$, $p < .001$, $\eta^2 = .588$; scores increased from pretest to posttest ($p < .001$) and from pretest to follow-up ($p < .001$), with no significant posttest–follow-up difference ($p = .921$), consistent with maintenance. Because follow-up data were available only for the intervention group, these retention findings reflect within-group maintenance and cannot be interpreted as comparative long-term effects relative to the control group.



UNIQUE CONTRIBUTIONS OF SUBITIZING AND SEQUENTIAL COUNTING TO ARITHMETIC COMPETENCE IN PRIMARY SCHOOL CHILDREN

Birkan GÜLDENOĞLU

Prof. Dr., Ankara University,

Faculty of Educational Sciences, Department of Special Education, Ankara, Türkiye

ORCID: <https://orcid.org/0000-0002-9629-1505>

guldenoglu@ankara.edu.tr

Yasemin ARSLAN

Dr., İstanbul Medeniyet University, Faculty of Education, Department of Special Education, İstanbul, Türkiye

ORCID: <https://orcid.org/0000-0002-1021-9028>

yasemin.arslan@medeniyet.edu.tr

Ufuk ÖZKUBAT

Assoc. Prof., Gazi University, Faculty of Education, Department of Special Education, Ankara, Türkiye

ORCID: <https://orcid.org/0000-0002-9626-5112>

ufukozkubat@gazi.edu.tr

Received: June 9, 2026

Accepted: September 16, 2026

Published: September 30, 2026

Suggested Citation:

Güldenoğlu, B., Arslan, Y., & Özkubat, U. (2026). Unique contributions of subitizing and sequential counting to arithmetic competence in primary school children. *International Online Journal of Primary Education (IOJPE)*, 15(3), 253-265. <https://doi.org/10.55020/iojpe.1967208>



Copyright © 2026 by author(s). This is an open access article under the [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).

Abstract

Early numerical skills are considered critical foundations for the development of later mathematical competence. However, relatively few studies have examined the unique contributions of different foundational numerical processes within the same analytical framework. The present study investigated the relative roles of subitizing and sequential counting in explaining variability in early arithmetic competence among primary school children. A sample of 133 first-grade students completed computerized tasks assessing subitizing accuracy, forward counting, backward counting, and arithmetic competence. Correlation and hierarchical regression analyses were conducted to examine the relationships among these variables. Results indicated that both subitizing and sequential counting were significantly associated with arithmetic competence. However, when the shared variance among numerical skills was considered, backward counting emerged as the strongest independent predictor of arithmetic competence, accounting for the largest proportion of unique variance. Additional analyses examining reaction time and inverse efficiency scores further revealed that greater efficiency in numerical processing, particularly in backward counting, was associated with higher arithmetic performance. These findings suggest that early arithmetic competence may depend less on the recognition of small quantities and more on the ability to flexibly manipulate symbolic number sequences. From both developmental and educational perspectives, the results highlight the importance of symbolic sequence manipulation as a key mechanism linking early numerical representations to emerging arithmetic skills.

Keywords: subitizing, sequential counting, backward counting, arithmetic competence, numerical cognition

INTRODUCTION

Mathematical competence acquired in early childhood is a strong predictor of later academic and life outcomes, including employability, financial literacy, and health management (Geary, 2011; Dowker, 2023). Despite its critical importance, a substantial proportion of children experience persistent difficulties in acquiring foundational numerical skills. These challenges highlight the importance of understanding the cognitive mechanisms underlying early mathematical development and the factors that contribute to variability in children's arithmetic performance (Bruine de Bruin & Slovic, 2021; Kucian & von Aster, 2015).

Two foundational cognitive skills, subitizing and sequential counting, are widely recognized as key components of early numerical development and have been consistently associated with children's



emerging arithmetic competence. Subitizing refers to the rapid and accurate identification of small quantities, typically up to four items, without deliberate counting (Butterworth, 2010; Trick & Pylyshyn, 1994). This ability is supported by the Object Tracking System (OTS), a cognitive mechanism that enables the parallel individuation of a limited number of visual objects (Feigenson et al., 2004; vanMarle et al., 2016). Early numerical cognition has also been linked to approximate quantity processing mechanisms such as the Approximate Number System (ANS), which supports the estimation of larger quantities (Dolfi et al., 2024; Wilkey et al., 2020).

Subitizing can be broadly distinguished into perceptual and conceptual forms. Perceptual subitizing involves the rapid recognition of small quantities without sequential counting, whereas conceptual subitizing relies on grouping strategies that allow individuals to identify somewhat larger quantities through structured patterns (Clements & Sarama, 2014). Performance in subitizing tasks may vary depending on stimulus arrangement and visual complexity (Ashkenazi et al., 2013; Schindler et al., 2022). By enabling rapid identification of exact quantities, subitizing supports the mapping between perceptual quantities and verbal number labels.

Complementing subitizing, sequential counting represents a symbolic process involving the ordered recitation and application of number words according to principles such as stable order and one-to-one correspondence (Dowker, 2023). Through repeated engagement with number sequences, children learn to coordinate symbolic number words with quantities, enabling counting to function as an important developmental bridge toward arithmetic skills (Clements & Sarama, 2014; Estevez-Perez et al., 2019). While forward counting primarily involves retrieving a memorized sequence, backward counting requires the flexible manipulation of that sequence in reverse order. This process places greater demands on working memory and cognitive control and may therefore reflect a more advanced level of symbolic number processing that is closely related to arithmetic problem solving.

Arithmetic competence is typically conceptualized as a multidimensional construct involving numerical magnitude understanding, symbolic number knowledge, procedural skills, and arithmetic fact retrieval. Early numerical skills, such as subitizing and sequential counting, play a central role in this developmental progression by supporting children's ability to represent quantities, access number sequences, and coordinate numerical relationships during problem solving. Disruptions or delays in these foundational processes may therefore contribute to persistent variability in arithmetic performance. Empirical evidence supports this framework. Longitudinal research has shown that early subitizing and counting abilities are associated with later mathematics achievement (Gray & Reeve, 2014; Hannula-Sormunen et al., 2015; Lemaire & Callies, 2009). Furthermore, symbolic number processing, particularly symbolic number comparison and enumeration, has been identified as an important predictor of arithmetic development (Aragon et al., 2023). These findings highlight the role of early numerical representations and symbolic access processes in supporting children's emerging arithmetic competence (De Smedt et al., 2013; Mussolin et al., 2010; Rousselle & Noël, 2007). From this perspective, arithmetic competence can be understood as the outcome of an integrative process in which early quantitative representations become coordinated with symbolic numerical knowledge. Within this framework, subitizing supports the rapid identification of exact quantities, whereas sequential counting provides access to ordered symbolic number representations that enable children to represent numerical relationships and develop procedural strategies for solving arithmetic problems.

The relationship between foundational numerical skills and arithmetic performance may also be influenced by domain-general cognitive mechanisms such as executive functioning, working memory, and processing speed (Ashkenazi et al., 2009; Geary et al., 2007; Mazzocco & Grimm, 2013). For example, Cui et al. (2024) demonstrated that visual form perception fully mediated the relationship between subitizing and arithmetic ability, highlighting the role of perceptual and attentional processes in numerical development. However, relatively few studies have examined the unique and combined contributions of these foundational numerical skills within the same analytical framework.



Despite the substantial body of research on early numerical cognition, important gaps remain regarding how foundational numerical skills jointly contribute to children's developing arithmetic competence. Much of the existing evidence has been generated in Western European and North American contexts, leaving other educational settings comparatively underrepresented. Investigating these relationships in different linguistic and educational contexts provides an opportunity to examine the generalizability of existing theoretical models. In Türkiye, research examining the combined roles of subitizing and sequential counting in early arithmetic development remains limited. Examining these foundational processes within the same analytical framework may therefore contribute to a clearer understanding of how early numerical representations support the emergence of arithmetic competence. Previous research with Turkish children has highlighted the developmental importance of number concepts, counting principles, and the use of counting strategies in mathematical problem solving (Olkun et al., 2013; Olkun et al., 2014). Building on this evidence, examining subitizing and sequential counting together may provide a more differentiated understanding of the numerical processes associated with arithmetic competence among Turkish-speaking children. Such evidence may also contribute to the international literature by extending research on foundational numerical development to a comparatively underrepresented educational and linguistic context.

The present study examines the unique contributions of two foundational numerical skills (subitizing and sequential counting) to arithmetic competence in primary school children. Drawing on developmental models of numerical cognition, the study conceptualizes early arithmetic competence as emerging from the integration of non-symbolic quantity processing and symbolic number knowledge. Within this framework, subitizing reflects children's ability to rapidly recognize exact quantities, whereas sequential counting reflects their ability to access and flexibly manipulate ordered number sequences. Although previous research has established that both skills are associated with later mathematical achievement, relatively few studies have examined their relative contributions within the same analytical framework. It is therefore expected that both subitizing and sequential counting will be positively associated with children's arithmetic competence; however, because manipulating number sequences places greater demands on working memory and cognitive control, sequential counting, particularly backward counting, may account for additional variance in arithmetic performance beyond that explained by subitizing alone.

Based on these theoretical considerations, the present study addresses the following research questions:

RQ1. To what extent are subitizing and sequential counting associated with children's arithmetic competence?

RQ2. Do subitizing and sequential counting make unique contributions to explaining variability in arithmetic performance when examined within the same analytical model?

METHOD

Participants

The study sample consisted of 133 first-grade students enrolled in public primary schools in Ankara, Türkiye. Participants were selected according to several inclusion criteria. First, all children were native Turkish speakers to ensure linguistic consistency in the administration of numerical tasks. Second, students with any documented sensory, intellectual, or developmental disabilities (e.g., hearing, visual, or intellectual impairments) were excluded from the study based on school records and consultations with school counselors.

Teachers provided information about students' general mathematics performance and categorized them as average or low performers based on classroom assessments. This classification was used only to ensure variability in arithmetic performance and did not represent a formal diagnostic categorization.



To document the demographic characteristics of the sample, a background information form was completed for each participant. This form included information on students' gender as well as family socioeconomic indicators such as parental education levels and household income. The final sample was balanced in terms of gender and represented a range of socioeconomic backgrounds.

Stimuli and Tasks

The study employed three computerized paradigms designed to assess subitizing, sequential counting, and arithmetic competence. The computerized paradigms were developed as part of a larger research project supported by TÜBİTAK. All tasks were administered using specialized experimental software that presented stimuli on a computer screen and automatically recorded response accuracy and reaction times (in milliseconds). Instructions were pre-recorded and delivered audibly to ensure consistency in task administration.

Subitizing Task

Subitizing ability was assessed using a computerized adaptation of the task developed by Ashkenazi et al. (2013). In this task, participants were required to rapidly and accurately report the number of dots presented on the screen. Dot arrays ranged from 1 to 5 items and were displayed in both regular (structured) and scattered (random) spatial arrangements to capture variability in enumeration strategies.

Each trial began with a blank screen presented for 250 ms, followed by a fixation cross (+) displayed for 300 ms to orient participants' attention. After an additional 500 ms blank interval, the dot array appeared on the screen for 200 ms. Participants were instructed to report the number of dots as quickly and accurately as possible. Two practice trials were administered prior to the main task to ensure that participants understood the instructions.

The task consisted of 14 experimental trials presented in randomized order. Stimuli were generated using the NASCO program developed by Guillaume et al. (2020). All dots were displayed in black against a white background. Reaction time was calculated as the interval between stimulus onset and the participant's response. Accuracy scores were computed by assigning one point for each correct response and zero points for incorrect responses.

Sequential Counting Task

Sequential counting skills were assessed using computerized forward and backward counting tasks adapted from a commonly used counting sequence paradigm available through the Mathsframe educational resource. Participants were required to identify missing numbers within numerical sequences ranging from 1 to 99. The sequences progressed in rhythmic intervals of 1, 2, 3, 4, 5, or 10, thereby assessing children's ability to process and manipulate ordered number sequences.

Sequences were presented randomly and included items from both single-digit (1–9) and two-digit (10–99) number ranges. Each trial followed a standardized presentation sequence consisting of a 250 ms blank screen, a 300 ms fixation cross, and a 500 ms blank interval prior to stimulus presentation. The sequence remained visible on the screen until the participant responded.

Stimuli consisted of black numerical symbols displayed in Arial Black, 32-point font, against a white background within a black circular frame. Performance was measured using response accuracy, with one point assigned for each correct response and zero points for incorrect responses.

Arithmetic Competence Task

Arithmetic competence was assessed using two tasks: Operation Accuracy and Missing Operation. The Operation Accuracy task, adapted from Bellon et al. (2016), required participants to judge the correctness of basic arithmetic operations involving addition and subtraction. Each operation type included both correct and incorrect equations, with incorrect items generated by modifying the correct result by ± 1 . Participants indicated whether the presented operation was correct or incorrect.



The Missing Operation task was developed for the present study to assess children's ability to identify missing components within simple arithmetic expressions. Participants were required to determine the appropriate operation or value needed to complete the equation correctly.

Each trial followed a standardized presentation procedure consisting of a 250 ms blank screen, followed by a 300 ms fixation cross (+). The arithmetic problem was then displayed and remained visible until a response was made or a maximum response window of 20 seconds elapsed. If no response was provided within this time frame, the item was automatically coded as incorrect. Performance on the arithmetic task was evaluated using accuracy scores.

Validity and Reliability

Content validity of the paradigms was established through expert evaluation. A structured item specification table describing the objectives, instructions, and item construction procedures for each paradigm was submitted to three academics specializing in dyscalculia and learning disabilities. Experts evaluated each item with respect to theoretical relevance, clarity of instructions, alignment with the intended construct, and appropriateness of task difficulty on a five-point Likert scale (1 = not suitable to 5 = highly suitable).

Items were retained if they met the predefined criteria of a mean expert rating above 4.25, a standard deviation below 1.00, and a coefficient of variation under 25%. Items requiring revision were modified according to expert feedback prior to final task implementation.

Internal consistency reliability for each paradigm was calculated using the Kuder-Richardson Formula 20 (KR-20). The reliability coefficients indicated acceptable to high internal consistency across tasks: subitizing (.80), forward counting (.79), backward counting (.82), and arithmetic competence (.84). These values indicate that the paradigms provided reliable measurements of the targeted numerical constructs.

Procedure

Pilot Study

A pilot study was conducted with 15 first-grade students from two public primary schools in Ankara, Türkiye. Participants were identified by their classroom teachers as either average-performing ($n = 7$) or low-performing ($n = 8$) in mathematics. The sample included eight girls and seven boys, and none of the participants had any documented developmental or sensory disabilities.

Prior to participation, written informed consent was obtained from the students' parents. All pilot assessments were conducted individually using the computerized testing interface. Sessions took place in quiet rooms provided by the schools to ensure a distraction-free environment. During the administration, the participant and the researcher sat side by side, with students responding to the on-screen tasks while the researcher monitored the procedure.

The pilot study focused on evaluating the clarity of task instructions, the comprehensibility of the stimuli, and the overall administration time. Observations made during this phase informed minor adjustments to the paradigms and administration procedures prior to the main data collection.

Main Data Collection

The main data collection was conducted following ethical approval and official authorization from the Turkish Ministry of National Education. Individual assessment sessions were organized in designated quiet areas within the participating schools, such as empty classrooms, counseling rooms, or libraries.

Participation was voluntary, and all students were informed about the purpose and procedures of the study before the assessment began. Each session lasted approximately 20–25 minutes. A standardized administration protocol was followed for all participants. At the beginning of each session, the experimenter provided brief explanations and demonstrations of the tasks to ensure that participants understood the instructions. Participants were then asked to complete the tasks independently using the computerized interface.



The paradigms were administered in a fixed order: (1) subitizing, (2) sequential counting, and (3) arithmetic competence tasks. Upon completion of the assessment session, participants were thanked for their participation.

RESULTS

Descriptive statistics, correlation analyses, and hierarchical regression analyses were conducted to examine the relationships between subitizing, sequential counting, and arithmetic competence. Descriptive statistics for the primary study variables are presented in Table 1.

Table 1. Descriptive statistics for subitizing, sequential counting, and arithmetic measures

Variable	n	Mean	SD	Min	Max
Subitizing accuracy	133	10.13	2.95	0	14
Forward counting accuracy	133	5.35	2.52	0	10
Backward counting accuracy	133	4.95	2.80	0	10
Operation accuracy	133	6.83	1.27	3	8
Missing operation accuracy	132	8.05	3.11	0	14
Arithmetic competence	132	14.87	3.91	4	22
Subitizing RT (ms)	133	66,236.09	22,176.86	17,700	179,400
Forward counting RT (ms)	133	150,118.80	82,430.21	44,500	481,500
Backward counting RT (ms)	133	170,858.65	108,649.46	55,500	750,900
Operation RT (ms)	133	90,209.02	41,924.49	28,400	230,500
Missing operation RT (ms)	132	225,137.12	110,492.78	73,100	653,800

Note. RT = reaction time (milliseconds). RT values represent the total response time across trials. Arithmetic competence scores were calculated as the sum of operation accuracy and missing operation accuracy.

To address the first research question, Pearson correlation analyses were conducted to examine the associations between subitizing proficiency, sequential counting performance, and children's arithmetic competence. The correlation matrix for the main study variables is presented in Table 2.

Table 2. Pearson correlations among subitizing, sequential counting, and arithmetic competence

Variable	1	2	3	4
1. Subitizing accuracy	—			
2. Forward counting accuracy	.349***	—		
3. Backward counting accuracy	.431***	.787***	—	
4. Arithmetic competence	.412***	.629***	.708***	—

As shown in Table 2, arithmetic competence was positively associated with subitizing ($r = .412$), forward counting ($r = .629$), and backward counting ($r = .708$), with the strongest association observed for backward counting. In addition to these associations, significant relationships were observed among the predictor variables themselves. Subitizing accuracy was moderately correlated with both forward ($r = .349$) and backward counting ($r = .431$), while forward and backward counting showed a strong correlation ($r = .787$), suggesting that these skills represent closely related components of early numerical competence. Prior to conducting the regression analyses, potential multicollinearity among predictor variables was examined. Although forward and backward counting were strongly correlated, variance inflation factor (VIF) values ranged between 1.23 and 2.84, which are well below the commonly accepted threshold of 5. These values indicate that multicollinearity did not pose a threat to the stability of the regression estimates.

To examine the unique contributions of subitizing and sequential counting to arithmetic competence, a hierarchical multiple regression analysis was conducted. In the first model, subitizing accuracy was entered as the sole predictor to evaluate its independent contribution to arithmetic competence. In the second model, forward counting accuracy was added to examine whether sequential counting performance explained additional variance beyond subitizing proficiency. In the final model, backward counting accuracy was included to determine whether this more cognitively demanding



counting skill provided further explanatory power. The results of the hierarchical regression analysis are presented in Table 3.

Table 3. Hierarchical multiple regression predicting arithmetic competence

Model	Predictor	B	SE B	β	t	p	R ²	ΔR^2	VIF
1	Subitizing accuracy	.548	.106	.412	5.16	< .001	.170	—	1.00
2	Subitizing accuracy	.285	.094	.215	3.03	.003	.436	.266	1.15
	Forward counting accuracy	.856	.110	.553	7.81	< .001			1.15
3	Subitizing accuracy	.167	.090	.126	1.87	.064	.528	.091	1.23
	Forward counting accuracy	.284	.153	.183	1.85	.066			2.64
	Backward counting accuracy	.710	.143	.509	4.98	< .001			2.84

Note. Dependent variable = arithmetic competence. B = unstandardized regression coefficient; β = standardized regression coefficient; ΔR^2 = change in explained variance across hierarchical steps.

In the first model, subitizing accuracy was entered as the sole predictor. This model was statistically significant, $F_{(1,130)} = 26.62$, $p < .001$, explaining 17.0% of the variance in arithmetic competence ($R^2 = .170$). Subitizing accuracy was a significant predictor of arithmetic competence ($\beta = .412$, $p < .001$).

In the second model, forward counting accuracy was added to the regression equation. The inclusion of forward counting significantly improved the model, $\Delta R^2 = .266$, F change $_{(1,129)} = 60.99$, $p < .001$, increasing the total explained variance to 43.6% ($R^2 = .436$). Both subitizing accuracy ($\beta = .215$, $p = .003$) and forward counting accuracy ($\beta = .553$, $p < .001$) were significant predictors at this stage.

In the final model, backward counting accuracy was included. This addition produced a further significant increase in explained variance, $\Delta R^2 = .091$, F change $_{(1,128)} = 24.75$, $p < .001$, with the full model explaining 52.8% of the variance in arithmetic competence ($R^2 = .528$). In the final model, backward counting accuracy emerged as the strongest predictor ($\beta = .509$, $p < .001$), whereas the effects of subitizing ($\beta = .126$, $p = .064$) and forward counting ($\beta = .183$, $p = .066$) were no longer statistically significant. Semi-partial correlations were examined to estimate the unique contribution of each predictor to arithmetic competence. Results indicated that backward counting accounted for the largest proportion of unique variance in arithmetic competence ($sr^2 = .091$). In contrast, the unique contributions of subitizing ($sr^2 = .013$) and forward counting ($sr^2 = .013$) were comparatively small. These findings suggest that although subitizing and forward counting are associated with arithmetic competence, backward counting provides the strongest independent contribution when all predictors are considered simultaneously.

Additional analyses examined RT measures to explore processing efficiency in early numerical tasks. Reaction time values represented the cumulative response time across trials. Pearson correlation analyses were conducted to examine whether processing speed in early numerical tasks was associated with arithmetic competence. The results are presented in Table 4.

Table 4. Pearson correlations between reaction time measures and arithmetic competence

Variable	1	2	3	4
1. Subitizing RT	—			
2. Forward counting RT	.318***	—		
3. Backward counting RT	.259**	.593***	—	
4. Arithmetic competence	-.362***	-.062	.039	—

Note. RT = reaction time (milliseconds). ** $p < .01$, *** $p < .001$.

As shown in Table 4, subitizing reaction time demonstrated a significant negative correlation with arithmetic competence ($r = -.362$, $p < .001$), indicating that slower subitizing responses were associated with lower arithmetic performance. In contrast, reaction times for forward counting



($r = .062$, $p = .482$) and backward counting ($r = .039$, $p = .657$) were not significantly associated with arithmetic competence.

Additionally, significant positive correlations were observed among the RT measures themselves. Subitizing RT was moderately correlated with forward counting RT ($r = .318$, $p < .001$) and backward counting RT ($r = .259$, $p = .003$), while forward and backward counting RT were strongly correlated ($r = .593$, $p < .001$), suggesting that processing speed across numerical tasks may reflect a partially shared cognitive component.

To provide a more integrated assessment of performance that accounts for both response accuracy and processing speed, inverse efficiency scores (IES) were calculated for the subitizing and sequential counting tasks (Bruyer & Brysbaert, 2011). IES values were computed by dividing reaction time by the proportion of correct responses. Pearson correlation analyses were then conducted to examine the relationships between efficiency scores and children's arithmetic competence. The results are presented in Table 5.

Table 5. Pearson correlations between inverse efficiency scores and arithmetic competence

Variable	1	2	3	4
1. Subitizing efficiency (IES)	—			
2. Forward counting efficiency (IES)	.264**	—		
3. Backward counting efficiency (IES)	.277**	.631***	—	
4. Arithmetic competence	-.395***	-.440***	-.505***	—

Note. IES = Inverse Efficiency Score (reaction time divided by accuracy). Higher IES values indicate lower processing efficiency. ** $p < .01$, *** $p < .001$.

DISCUSSION, CONCLUSION, and RECOMMENDATIONS

The present study examined how two foundational numerical skills (subitizing and sequential counting) contribute to variability in early arithmetic competence. Overall, the findings suggest that although both perceptual quantity recognition and symbolic number-sequence knowledge are associated with children's arithmetic performance, their contributions are not equivalent. Subitizing was significantly related to arithmetic competence, indicating that early quantity representations remain an important component of numerical development (De Smedt et al., 2013; Feigenson et al., 2004; Schneider et al., 2017). However, when the shared variance among numerical skills was considered, backward counting emerged as the strongest independent predictor of arithmetic competence. This pattern suggests that the ability to flexibly manipulate symbolic number sequences may play a particularly central role in the transition from early quantity representations to formal arithmetic processing (Dowker, 2023). In this sense, variability in early arithmetic competence appears to depend less on the mere recognition of quantities and more on children's capacity to operate on symbolic number sequences in a flexible and coordinated manner.

The observed association between subitizing and arithmetic competence is consistent with developmental models suggesting that early numerical understanding begins with the rapid recognition of small quantities (Butterworth, 2010; Trick & Pylyshyn, 1994). Children who more efficiently recognized small numerosities demonstrated stronger arithmetic performance, supporting the view that early quantity representations provide a cognitive foundation for later numerical learning (Gray & Reeve, 2014; Hannula-Sormunen et al., 2015). However, the relatively small unique contribution of subitizing when other numerical skills were considered suggests that perceptual quantity recognition alone may not be sufficient to explain variability in arithmetic competence.

Recent international research further supports the importance of foundational numerical processes in the development of arithmetic competence. For example, Leuenberger et al. (2024) demonstrated that instructional support targeting conceptual subitizing and non-counting strategies can enhance first-grade students' numerical and computation competencies, while recent research has continued to identify counting and subitizing as important components of early mathematical development (Cruz et al., 2024). At the same time, evidence suggests that arithmetic performance depends on a broader



combination of numerical abilities, including counting knowledge and number-line representations (Chan & Mazzocco, 2024). Viewed alongside these findings, the present results extend the international literature by showing that, although subitizing is associated with arithmetic competence, the ability to manipulate symbolic number sequences, particularly through backward counting, may provide additional explanatory value for individual differences in first-grade arithmetic performance.

Sequential counting represents a critical step in the transition from perceptual quantity recognition to symbolic numerical processing (vanMarle et al., 2016). Forward counting requires children to access and reproduce the ordered sequence of number words, thereby supporting the coordination of symbolic representations with quantity information. In the present study, forward counting was strongly correlated with arithmetic competence, indicating that children who demonstrated greater fluency in accessing number sequences also tended to perform better in arithmetic tasks. However, when considered alongside other numerical predictors, forward counting did not uniquely explain additional variance in arithmetic performance. This pattern suggests that while forward counting reflects an important stage in the acquisition of symbolic numerical knowledge, its role may primarily involve the retrieval of memorized number sequences rather than the flexible manipulation of those sequences during problem solving (Geary, 2011). In this sense, forward counting may reflect children's access to symbolic numerical knowledge, whereas successful arithmetic performance appears to require the ability to actively transform and reorganize those symbolic representations.

In contrast to forward counting, backward counting emerged as the strongest independent predictor of arithmetic competence. Backward counting requires children to actively manipulate the number sequence in reverse order, a process that places greater demands on working memory, cognitive control, and the internal representation of the mental number line (Lemaire & Callies, 2009). Backward counting may reflect children's ability to mentally traverse the mental number line, a process closely aligned with the cognitive mechanisms underlying subtraction and arithmetic transformation. The substantial unique contribution of backward counting observed in this study suggests that arithmetic competence may depend less on the ability to retrieve number sequences and more on the ability to flexibly manipulate them, a pattern consistent with evidence showing that numerical ordering abilities predict arithmetic competence (Dowker, 2023; Lyons & Beilock, 2011). From a developmental perspective, this finding supports the idea that early arithmetic skills are closely linked to children's capacity to operate on symbolic number representations rather than merely accessing them. The results therefore highlight backward counting as a particularly informative indicator of children's readiness to engage in more complex arithmetic reasoning. Taken together, these findings suggest a developmental progression in early numerical competence, in which perceptual quantity recognition provides the initial foundation for numerical understanding, forward counting supports the acquisition of symbolic number sequences, and backward counting reflects a more advanced stage involving the flexible manipulation of symbolic numerical representations that directly supports arithmetic reasoning. From this perspective, backward counting may function as a behavioral indicator of children's emerging capacity to manipulate symbolic number structures, a cognitive process that appears to be closely aligned with the demands of early arithmetic reasoning.

In addition to accuracy-based analyses, the present study also examined the role of processing efficiency in early numerical performance. Reaction time analyses indicated that slower responses in the subitizing task were associated with lower arithmetic competence, suggesting that efficiency in rapid quantity recognition may reflect broader cognitive processing efficiency in early numerical tasks. Furthermore, analyses based on inverse efficiency scores revealed consistent relationships between efficiency in numerical processing and arithmetic competence, particularly for backward counting. Children who processed backward counting sequences more efficiently demonstrated substantially stronger arithmetic performance. These findings support the view that early mathematical competence is influenced not only by the accuracy of numerical representations but also by the efficiency with which these representations can be accessed and manipulated during cognitive processing (Wilkey et al., 2020). From a cognitive perspective, this pattern aligns with process-oriented accounts of numerical development suggesting that arithmetic performance depends



not only on the availability of numerical representations but also on the efficiency with which these representations can be coordinated during real-time problem solving (De Smedt et al., 2013). These results suggest that the development of arithmetic competence may depend not only on what numerical representations children possess, but also on how efficiently those representations can be mobilized during cognitive processing.

The linguistic and educational context should also be considered when interpreting these findings. Previous research with Turkish children has highlighted the importance of counting processes in the development of number concepts and mathematical problem solving (Olkun et al., 2013). Research with Turkish first-grade students has also demonstrated the developmental importance of counting principles during the early stages of formal mathematics education (Olkun et al., 2014). Although the present study was not designed to directly examine linguistic or cross-cultural effects, the strong association between sequential counting, particularly backward counting, and arithmetic competence provides further evidence from Turkish-speaking first-grade children. From an international perspective, these findings contribute evidence from a comparatively underrepresented educational and linguistic context and suggest that the role of sequential counting in arithmetic development warrants further examination across languages and educational systems.

From an educational perspective, the findings suggest that early mathematics instruction should emphasize not only counting proficiency but also the flexible manipulation of number sequences (Clements & Sarama, 2014). Although early mathematics activities frequently focus on forward counting and number identification, instructional tasks that require children to manipulate numerical sequences, such as backward counting, may more effectively support the development of arithmetic reasoning. Engaging students in such sequence-manipulation activities may strengthen number-line representations and provide an important bridge between early counting skills and the emergence of flexible arithmetic strategies (Kucian & von Aster, 2015).

Despite its contributions, several limitations should be considered when interpreting the findings. First, the cross-sectional design limits the ability to draw conclusions about developmental trajectories or causal relationships among numerical skills and arithmetic competence. Longitudinal research would provide stronger evidence regarding how these relationships evolve over time. Second, arithmetic competence was assessed using experimental paradigms rather than standardized mathematics assessments, which may limit the direct comparability of the findings with classroom-based achievement measures. Future studies could integrate standardized mathematics assessments to further examine the generalizability of the results. Finally, the present study focused on a sample of first-grade students in a single educational context. Replication across different age groups and educational systems would help clarify the extent to which the observed relationships reflect broader developmental patterns in early numerical cognition. Future longitudinal studies across diverse educational contexts are needed to clarify how early numerical skills develop into formal arithmetic competence.

Funding

This study was part of a national project supported by the Scientific and Technological Research Institution of Turkey (Project number: 123K646).

Ethics and Conflict of Interest

This study was conducted in accordance with the principles of ethical research involving human participants. Ethical approval was obtained from the Ankara University Ethics Committee (Approval Date: July 17, 2023; Meeting No. 16; Decision No. 159). In addition, the necessary official permissions to conduct the study in schools were obtained from the relevant authorities affiliated with the Turkish Ministry of National Education. Written informed consent was obtained from the parents or legal guardians of all participating children, and participation was voluntary. Participants were informed about the study procedures and their right to withdraw from the study at any time. All data were treated confidentially and used solely for scientific research purposes. We have no conflict of



interest to disclose. The authors declare that there are no conflicts of interest regarding the research, authorship, and/or publication of this article.

Author Contribution

Prof. Guldenoglu conceptualized the study and wrote the manuscript. Dr. Arslan collected the data. Assoc. Prof. Ozkubat contributed to the study design and provided support for the literature review. All authors reviewed and approved the final version of the manuscript.

Data Availability

The data supporting the findings of this study were collected as part of a national project supported by the Scientific and Technological Research Institution of Turkey (Project number: 123K646). Due to the terms of the project agreement, the data are not publicly available but may be made available by the authors upon reasonable request.

Corresponding Author

Correspondence to Birkan GÜLDENOĞLU, guldenoglu@ankara.edu.tr

REFERENCES

- Alrefaei, M. M. (2024). Number sense deficits in children with developmental dyscalculia, dyslexia, co-occurring disorder and their typically developing peers. *Applied Neuropsychology: Child*, 1–8. <https://doi.org/10.1080/21622965.2024.2364729>
- Aragon, E., Canto-López, M. C., Aguilar, M., Menacho, I., & Navarro, J. (2023). Longitudinal study of symbolic and non-symbolic magnitude processing and its relationship with mathematical achievement, *Revista de Psicodidáctica*, 28(1), 44–50. <https://doi.org/10.1016/j.psicoe.2022.08.001>
- Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2009). Numerical distance effect in developmental dyscalculia. *Cognitive Development*, 24(4), 387–400. <https://doi.org/10.1016/j.cogdev.2009.09.006>
- Ashkenazi, S., Mark-Zigdon, N., & Henik, A. (2013). Do subitizing deficits in developmental dyscalculia involve pattern recognition weakness? *Developmental Science*, 16(1), 35–46. <https://doi.org/10.1111/j.1467-7687.2012.01190.x>
- Bruine de Bruin, W., & Slovic, P. (2021). Low numeracy is associated with poor financial well-being around the world. *Plos one*, 16(11), e0260378. <https://doi.org/10.1371/journal.pone.0260378>
- Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, 51(1), 5–13. <https://doi.org/10.5334/pb-51-1-5>
- Butterworth B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in cognitive sciences*, 14(12), 534–541. <https://doi.org/10.1016/j.tics.2010.09.007>
- Chan, J. Y.-C., & Mazzocco, M. M. M. (2024). New measures of number line estimation performance reveal children's ordinal understanding of numbers. *Journal of Experimental Child Psychology*, 245, Article 105965. <https://doi.org/10.1016/j.jecp.2024.105965>
- Clements, D. H., & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach* (2nd ed.). Routledge.
- Cui, J., Lu, J., Feng, J., Zuo, Y., Cui, Z., & Zhou, X. (2024). The unique role of visual form perception in children's language ability. *Psychological Development and Education*, 40(6), 782–791. <https://doi.org/10.16187/j.cnki.issn1001-4918.2024.06.03>
- Cruz, J., Alves, D., Carvalho, M., Abreu Mendes, S., Rodrigues, B., & Cadime, I. (2024). Assessment of math abilities before school entry: A tool development. *Frontiers in Education*, 8, Article 1347143. <https://doi.org/10.3389/feduc.2023.1347143>
- De Smedt, B., Noël, M. P., Gilmore, C., & Ansari D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48–55. <http://dx.doi.org/10.1016/j.tine.2013.06.001>
- Dolfi, S., Decarli, G., Lunardon, M., De Filippo De Grazia, M., Gerola, S., Lanfranchi, S., Cossu, G., Sella, F., Testolin, A., & Zorzi, M. (2024). Weaker number sense accounts for impaired numerosity perception in dyscalculia: Behavioral and computational evidence. *Developmental Science*, 27(6), e13538. <https://doi.org/10.1111/desc.13538>
- Dowker, A. (2023). The componential nature of arithmetical cognition: Some important questions. *Frontiers in Psychology*, 14, 1188271. <https://doi.org/10.3389/fpsyg.2023.1188271>



- Estevez-Perez, N., Castro-Canizares, D., Martinez-Montes, E., & Reigosa-Crespo, V. (2019). Numerical processing profiles in children with varying degrees of arithmetical achievement. *Acta Psychologica*, 198, 102849. <https://doi.org/10.1016/j.actpsy.2019.05.001>
- 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 five-year longitudinal study. *Developmental Psychology*, 47(6), 1539–1552. <https://doi.org/10.1037/a0025510>
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, 78 (4), 1343-1359. <https://doi.org/10.1111/j.1467-8624.2007.01069.x>
- Gray, S. A., & Reeve, R. A. (2014). Preschoolers' dot enumeration abilities are markers of their arithmetic competence. *Plos One*, 9(4), e94428. <https://doi.org/10.1371/journal.pone.0094428>
- Hannula-Sormunen, M. M., Lehtinen, E., & Rasanen, P. (2015). Preschool children's spontaneous focusing on numerosity, subitizing, and counting skills as predictors of their mathematical performance seven years later at school. *Mathematical Thinking and Learning*, 17(2), 155-177. <https://doi.org/10.1080/10986065.2015.1016814>
- Kucian, K., & von Aster, M. (2015). Developmental dyscalculia. *European Journal of Pediatrics*, 174(1), 1–13. <https://doi.org/10.1007/s00431-014-2455-7>
- Lemaire, P., & Callies, S. (2009). Children's strategies in complex arithmetic. *Journal of experimental child psychology*, 103(1), 49–65. <https://doi.org/10.1016/j.jecp.2008.09.007>
- Leuenberger, D., Diener, M., Wehren-Müller, M., Hofmann-Villiger, A., Vogt, F., & Moser Opitz, E. (2024). Fostering computation competence with non-counting strategies and conceptual subitizing in Grade 1: An intervention study in inclusive classrooms. *Journal für Mathematik-Didaktik*, 45, Article 12. <https://doi.org/10.1007/s13138-024-00236-6>
- Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, 121(2), 256–261. <https://doi.org/10.1016/j.cognition.2011.07.009>
- Mazzocco, M. M. M., & Grimm, K. J. (2013). Growth in rapid automatized naming from grades K to 8 in children with math or reading disabilities. *Journal of Learning Disabilities*, 46(6), 517-533. <https://doi.org/10.1177/0022219413477475>
- Mussolin, C., Mejias, S. & Noel, M.P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition*, 115(1), 10-25. <https://doi.org/10.1016/j.cognition.2009.10.006>
- Olkun, S., Çelik, E., Sönmez, M. T., & Can, D. (2014). İlköğretim birinci sınıf Türk öğrencilerinde sayma ilkelerinin gelişimi [Development of counting principles in Turkish first graders]. *Başkent University Journal of Education*, 1(2), 115–125.
- Olkun, S., Fidan, E., & Babacan Özer, A. (2013). The development of number concept and the use of counting in problem solving of 5–7-year-olds. *Education and Science*, 38(169), 236–248. <https://doi.org/10.15390/ES.2013.1169>
- Rousselle, L., & Noel, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102(3), 361–395. <https://doi.org/10.1016/j.cognition.2006.01.005>
- Schindler, M., Doderer, J. H., Simon, A. L., Schaffernicht, E., Lilienthal, A. J., & Schäfer, K. (2022). Small number enumeration processes of deaf or hard-of-hearing students: A study using eye tracking and artificial intelligence. *Frontiers in Psychology*, 13, 909775. <https://doi.org/10.3389/fpsyg.2022.909775>
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental science*, 20(3), 10.1111/desc.12372. <https://doi.org/10.1111/desc.12372>
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision. *Psychological Review*, 101(1), 80–102. <https://doi.org/10.1037/0033-295X.101.1.80>
- vanMarle, K., Chu, F., Mou, Y. et. al. (2016). Attaching meaning to the number words: contributions of the object tracking and approximate number systems. *Developmental Science*, 21(1), e12495. <https://doi.org/10.1111/desc.12495>
- Wilkey, E. D., Pollack, C., & Price, G. R. (2020). Dyscalculia and typical math achievement are associated with individual differences in number-specific executive function. *Child Development*, 91(2), 596–619. <https://doi.org/10.1111/cdev.13194>



About the Authors:

Birkan GÜLDENOĞLU

Prof. Dr. İsa Birkan Güldenoğlu is a faculty member in the Department of Special Education at Ankara University, Türkiye. His research interests include special education, early numeracy skills, mathematical cognition, and children's mathematical development.

Yasemin ARSLAN

Dr. Yasemin Arslan is a researcher in the Department of Special Education at İstanbul Medeniyet University, Türkiye. Her research interests include special education, early numeracy, mathematical learning, and children's cognitive and academic development.

Ufuk ÖZKUBAT

Assoc. Prof. Dr. Ufuk Özkubat is a faculty member in the Department of Special Education at Gazi University, Türkiye. His research interests include special education, early numeracy skills, mathematical cognition, learning difficulties, and academic development in children.