Author(s): Christoforou, Christoforos; Ktisti, Christiana; Richardson, Ulla; Papadopoulos, Timothy C.

Title: Microgenetic Analysis of Reading Remediation : A Novel Computational Framework

Year: 2023

Version: Published version

Please cite the original version:

Christoforou, C., Ktisti, C., Richardson, U., & Papadopoulos, T. C. (2023). Microgenetic Analysis of Reading Remediation : A Novel Computational Framework. Advances in Cognitive Psychology, 19(3), 297-315. https://doi.org/10.5709/acp-0400-6

Microgenetic Analysis of Reading Remediation: A Novel Computational Framework

Christoforos Christoforou[1], *Christiana Ktisti*[2], *Ulla Richardson*[3], *and Timothy C. Papadopoulos*[2]

[1] Division of Computer Science, Mathematics and Science, St. John's University, New York

[2] Department of Psychology & Center for Applied Neuroscience, University of Cyprus, Cyprus

[3] Centre for Applied Language Studies, University of Jyväskylä, Finland

ABSTRACT

Reading intervention program efficacy is usually determined by comparing participants' performance to controls on dependent measures at pre-, mid-, and post-intervention assessments. However, little is known about how learning progresses during different stages of the intervention. This lack of knowledge can be attributed to the absence of appropriate computational frameworks to encode, analyze, and capture such dynamics. We propose a novel computational framework to capture learning process dynamics during the intervention by analyzing microgenetic data. The framework addresses the problem of encoding microgenetic data into a common data representation model, introduces four information-theoretic metrics to capture the instantaneous developmental learning stages of groups and individuals, and provides the mathematical model to analyze those metrics for the study of learning stages during the intervention. We used data from a longitudinal reading remediation study involving 56 Greek-speaking 6-year-old children to demonstrate the framework's utility. Results showed that the framework functions as a new tool to explore the modulation in learning stages during the intervention, better understand how reading occurs, and how reading disability may be adequately treated.

KEYWORDS

microgenetic analysis, computational models, reading remediation

INTRODUCTION

Remediation happens after instruction has failed. Reading intervention programs aim to promote reading and spelling in children with reading difficulties by tackling critical precursor linguistic or cognitive skills (Lyytinen et al., 2009). Factors such as the type and severity of reading difficulties, the learner's cognitive and linguistic characteristics, and the interaction between aptitude and remediation features may be necessary to predict remedial program effectiveness (Kearns & Fuchs, 2013; Papadopoulos & Kendeou, 2010). Remediation is also typically given to small groups (Carlson & Das, 1997) or on an intensive one-to-one basis (Elbaum et al., 2000) depending on either the student's ability level, intervention type, or grade (Suggate, 2010). Furthermore, intensive reading interventions typically comprise tasks of increasing difficulty (Papadopoulos et al., 2003).

Despite this knowledge, we know little about how children progress through a reading intervention and when they benefit the most from it. This lack of knowledge can be attributed to the absence of appropriate computational tools that encode, capture, and analyze the dynamics of learning progression during an intervention, facilitating the study of learning throughout the intervention. To bridge this gap, the current study presents a novel computational framework for promoting the study of learning in transition in the context of microgenetic analysis.

Microgenetic Analysis

Microgenetic analysis, initially introduced by Werner (1957) and later defined by Inhelder et al. (1976) as microgenèse, aims to capture the temporal dynamics of direct experience – that is to say, the fine-grained details of a sequence of events that occur in a time between the presentation of a stimulus and the formation of a response to that stimulus. More recently, Karmiloff-Smith (2013) referred to the term microdevelopment to shift the focus more onto the cognitive aspects of development rather than molecular genetics, as the original term might imply. Similarly, Siegler (2000) and Siegler and Svetina (2002) discussed development as a variable process. For example, children employ multiple strategies to solve problems at each point as learning occurs. Consequently, the relative frequency of any given choice changes based on learners' development and experience as they perform a task (Fazio & Siegler, 2013). Thus, microgenetic analysis offers the possibility of understanding behavior with an emphasis on the process rather than the product.

Corresponding author: Christoforos Christoforou, Division of Computer Science, Mathematics and Science, St. John's University, New York.

Email: christoc@stjohns.edu

DOI • 10.5709/acp-0400-6

Three essential attributes characterize the microgenetic method. First, observations cover the time of a quickly evolving ability. Second, the density of records during this period is high, focusing on the rate of progress. Finally, changing performance observations are examined thoroughly to denote the processes that lead to change (Miller & Coyle, 1999). Consequently, microgenetic analysis allows detailed data collection to focus on tractable changes at the smallest time scale. Data of this sort demonstrate that change is not unexpected. Instead, it includes multiple phases of high intraindividual performance variability and short-term regressions as individuals progress through tasks (Adolph et al., 2008).

The current study leveraged the microgenetic method to examine how efficiently a child executes a particular task and benefits from reading remediation, focusing on two concepts: performance and effort. This approach's primary advantage is the detailed study of a child's transition between stages or progression from struggling with reading to reaching average reading development levels. Furthermore, focusing on the process rather than the product of the intervention, which is the standard focus of research on the treatment of reading difficulties (e.g., Jamshidifarsani et al., 2019; Suggate, 2010), can better highlight (a) the effectiveness of an intervention program, (b) its compatibility with the participants' rate and level of development, and (c) the learning process on the participants' part. To validate the suggested approach, we provided results from a preliminary application of the method to an existing pilot data set based on the adaptation of two interactive reading remediation programs in Greek, GraphoGame and PREP[1].

Computer-Assisted Reading Intervention

As general education classrooms become more heterogeneous due partly to the integration of students with learning and developmental disabilities (Scammacca et al., 2016), teachers need to have instructional techniques designed to meet the individual needs of their students. Therefore, teachers usually ask for guidelines explicitly stating what they should do to maximize student learning outcomes (van Garderen et al., 2012). Individualized learning (technological) platforms that adapt to each learner's performance providing an appropriate level of learning challenges, can support and meet these requirements. Review studies, especially those focusing on reading instruction (e.g., Blok et al., 2002; National Institute of Child Health and Human Development, 2000) or intervention (e.g., Jamshidifarsani et al., 2019) conclude that computer-assisted reading interventions tend to be generally effective.

Indeed, advances in computer-assisted interventions, particularly game-based interventions (e.g., Lyytinen et al., 2021), offer innovative support in treating reading difficulties and provide significant knowledge about the potential to prevent and remediate reading difficulties. Their advantage over traditional reading remediation programs lies in the instantaneously flexible learning environment that promotes active and individual-oriented reading support (e.g., Savage et al., 2009; Wouters et al., 2013). Computer speech prompts, along with user-friendly graphics, animation, and direct visual and audio feedback to correct and incorrect responses, have led to the development of applications that are highly motivating to young readers (Mayer & Moreno, 2002; Saine et al., 2011).

Despite remarkable advances in game-based interventions, the mechanisms supporting skills enhancement during a reading intervention remain obscure, and the results of these studies are mixed. For instance, Blok et al. (2002) reported a raw overall effect size generated over 50 different experimental comparisons, equal to 0.25, a small effect on beginning readers' abilities. Similarly, Cheung and Slavin (2013) have found a small albeit significant effect (d = 0.14) in support of struggling readers. The low effect size was attributed to few studies that met the inclusion standards, many of which were small experiments, and the random assignment to conditions in more extensive studies. Finally, Wouters et al. (2013) reported a positive effect of a game-based intervention on learning over conventional intervention methods when multiple training sessions were offered to support children's learning. However, it was unclear what this number of sessions should be and what it depended on. In addition, although integrating new with prior knowledge when participants were asked to work out loud appeared to enhance the effectiveness of interventions, the findings did not support the assumption that game-based interventions are more effective than conventional intervention methods because they provide more motivation for participation. Thus, a more detailed investigation of the degree to which the participants can control the activities and the use of the methods they discover to maintain the intrinsic motivation to continue interacting with the game is deemed necessary. Therefore, the current evidence regarding how a child progresses through an effective and efficient remediation intervention is limited. For this reason, a microgenetic method can be a reliable means to provide a fine-grained analysis of children's learning progress.

Computational Frameworks and Reading Improvement

The development of computational simulations in reading research allows answering questions regarding intervention effectiveness for readers experiencing difficulties (Church et al., 2021). Such simulations can inform intervention studies' findings as they allow for systematically testing multiple hypotheses at an individual level and under a wide array of training conditions. For example, Harm et al. (2003) used a reading development model to simulate detailed aspects of the learning process. The authors attempted to explain empirical findings within an explanatory framework, allowing them to examine why specific interventions are more effective than others. Their simulations replicated the patterns of success and failure found in the literature, speaking for the additive effect of simultaneous training on phonology and orthography. Nonetheless, the analysis focused on the additional resources the reader would need to build high-quality representations from phonology to orthography and conversely.

Likewise, computational simulations allow testing the different profiles of children with dyslexia, a factor that usually determines remedial effects. For example, Ziegler et al. (2008) used the computational dual-route model of reading (DRC; Coltheart et al., 2001) to investigate the dyslexia subtypes suggested in a model involving 9-year-old French-speaking readers. The authors simulated the reading performance of each participant with the DRC, using several tasks to estimate which of the DRC core processes were deficient for each person. This data was then used to add relative noise to individual deficiency and simulate the various impairments. The findings showed that children with dyslexia exhibited deficits in almost all DRC representational levels, from lexical to sub-lexical processes, speaking of the multidimensionality of the underlying deficits in dyslexia. Interestingly, the model produced incorrect simulations for a small group of three children identified as compensated dyslexics, indicating that not all behavioral data can be aligned with a computational model. Taken together, one would wonder how the detailed description of the skills required for reading (Harm et al., 2003) or the deficits involved in dyslexia (Ziegler et al., 2008) could map onto the components of a model that looks into how a child progresses through remediation and benefits from it.

Thus, what is still missing from the above models is a detailed account of what it takes (effort) to use additional resources to form phonological and orthographic representations and how the learner optimally uses them to overcome the possible multiple deficits associated with reading difficulties (performance). Therefore, the current study explored the readers' learning progress dynamics during the intervention to understand how the interaction between effort and performance can ensure that a child benefits from remediation.

GraphoGame Intervention as a Use Case

To test the utility of the proposed computational framework, we used available data from a remedial program focusing on phonological training (GraphoGame[2]). We chose this intervention because it can satisfy these learning conditions, particularly in languages with a transparent orthography (e.g., Lyytinen et al., 2009; Papadopoulos & Kendeou, 2010). Furthermore, such interventions aim to develop children's letter knowledge and reading ability regarding accuracy and fluency (Papadopoulos et al., 2015). In addition, we compared the efficacy of GraphoGame against PREP, a reading remediation program with a more cognitive focus (Papadopoulos et al., 2003), a short description of which is provided in Supplementary Material SM2.

Several training studies have demonstrated that specific phonological awareness training can positively impact early reading and beyond (e.g., Kjeldsen et al., 2014; Schneider et al., 1999). In addition, there is evidence that early reading programs that emphasize the relations between the phonological structure of spoken words and written language units can help close the gap between struggling readers and typically developing readers (Bus & Van Ijzendoorn, 1999; Lyytinen et al., 2021). It has also been demonstrated that promising, effective reading intervention programs combine direct training in phonological awareness with letter-sound training (Schneider et al., 2000).

GraphoGame is a child-friendly computer game that improves children's reading skills, emphasizing phonological skills and letter knowledge (Lyytinen et al., 2007). It is based on the assumption that the most predictive index of later reading difficulties that is most practical to implement is poor letter-sound knowledge. Therefore, in the GraphoGame application, acquiring alphabetic knowledge and facility with letter-sound relationships is essential to beginning reading (Richardson & Lyytinen, 2014; Lyytinen et al., 2007; Vanden Bempt et al., 2021). GraphoGame was initially developed within the Jyväskylä Longitudinal Dyslexia Study (e.g., Lyytinen et al., 2006) in Finnish, a language with a consistent orthography (Torppa et al., 2013), for children at risk for dyslexia[3].

The program provides practice in letter-sound relations, phonemic awareness, decoding skills, accuracy, and fluency and is delivered over the Internet (Richardson & Lyytinen, 2014; Ronimus et al., 2020). It focuses on the core issue of reading, that is, learning the connections between spoken and written language (Lyytinen et al., 2009), by providing an intensive adaptive learning environment with individualized repetition. It progresses from letter-sound relations to the phonological recoding and decoding stage, covering the core areas needed for fluent and accurate reading (Ahmed et al., 2020; Saine et al., 2011). The game incorporates a dynamic element in that it also adapts to the child's ability level and sets further levels according to this ability. Intervention data are recorded on a server, and online recordings enable researchers to monitor every individual's responses. Usually, GraphoGame is delivered over four weeks or longer in daily 20-minute sessions on a one-to-one basis. Recent research reviewing the effectiveness of GraphoGame speaks for its suitability mainly as a learning platform in a classroom, where high adult interaction produces an average positive effect size (g = 0.48, McTigue et al., 2019).

The Finnish orthography in which the program was initially devised is similar to Greek regarding syllabic and orthographic consistency. Specifically, both orthographies have a simple syllabic structure characterized by a predominance of open CV syllables with few initial or final consonant clusters. They are also based on consistent one-to-one mappings between graphemes and phonemes (Seymour et al., 2003). Children in transparent orthographies can read accurately relatively early with adequate teaching (Aro & Wimmer, 2003; Papadopoulos et al., 2021). However, intensive and individual training is necessary for children at risk for or exhibit reading difficulties to become adequate readers in the Finnish or Greek language context.

The Current Study

Reading intervention program efficacy is traditionally determined by comparing participants' performance to controls (i.e., untreated children with reading difficulties or typically developing counterparts) on linguistic and cognitive measures before and after remediation. Yet, a computerized implementation of remedial programs enables recording microgenetic data during the intervention, such as logging details about individuals' specific actions on each intervention task. Such information can, in turn, allow researchers to gain valuable insights into understanding the learning progress dynamics of the readers during the intervention, as well as individual (or group) gain variation on different elements of the intervention. These insights can further inform the design of reading intervention programs tailored to the individual's progress dynamics.

The current study proposes a methodological framework for encoding and modelling microgenetic data. In particular, the framework addresses the problem of encoding heterogeneous microgenetic data into a common informative data representation model. In doing so, we used data obtained during GraphoGame–part of a broader study investigating the effectiveness of computer-based interventions (*ReaDI-STANCE*, Papadopoulos, 2019) – to demonstrate how the framework could be applied. Therefore, the primary aim was to describe a mathematical model to visualize and analyze learning developmental stages during the intervention. We formalized a unified encoding model of the microgenetic data. Finally, we proposed four metrics to characterize the readers' developmental stages and learning progress dynamics during the intervention.

To the best of our knowledge, no other study has examined learning progress in reading remediation using microgenetic methods and computer applications focusing on parameters of effort and performance, as the current study does. Consequently, the cognitive effort through which children acquire the skills and knowledge while working on specific reading tasks has received little attention. For this reason, the current study explored how learning is discovered on the child's part during remediation by observing learning occurring within a subject over time.

METHODS

Participants

Participants were Grade 1 students recruited from public primary schools in Cyprus at the end of January (5th month of the school year). Because a formal diagnosis of dyslexia is rare in Cyprus, to identify the children with dyslexia, we first asked teachers to nominate children experiencing reading difficulties with no sensory, intellectual, or attentional problems. Research has shown that teachers' judgements about their students' reading levels are generally confirmed by the children's subsequent reading scores (e.g., Virinkoski et al., 2018). After obtaining parental consent, nominated children were tested on reading fluency and general cognitive ability measures to ensure they met the inclusionary criteria for reading difficulties, as described in the DSM-V (American Psychiatric Association, 2013).

Fifty-six children (30 males, 26 females; $M_{age} = 6.91$, $SD = 0.48$) who scored at least one *SD* below their respective age group mean on two reading fluency tasks (word reading fluency and phonemic decoding fluency; ERS-AB; Papadopoulos et al., 2009[4]), and within the average range on verbal (Vocabulary Wechsler Intelligence Scale for Children-Third Edition, Wechsler, 1992; Greek standardization: Georgas et al., 1997) and nonverbal ability (Nonverbal Matrices from the DN-CAS, Naglieri & Das, 1997; Greek standardization: Papadopoulos, Georgiou, Kendeou, & Spanoudis, 2009) met the inclusionary criteria and were included in the intervention groups. Following a randomized controlled trial (RCT) design[5], the 56 children with reading difficulties were assigned to a 5-week intervention focusing on phonological (GraphoGame; GG, $n = 14$) training, cognitive (PREP, $n = 14$) training, or the two combined (PREP-to-GG or GG-to-PREP; $n = 14$ in each group).

A chronological age control (CA-C) group ($n = 17$) of typically developing readers receiving no remediation also participated in the study. All participants were native-Greek speakers Caucasian from middle to upper-middle socioeconomic backgrounds (based on the schools' location and reports from the teachers). Groups were matched for age, gender, parental education levels, and nonverbal and verbal ability.

Procedure

The remediation phase commenced immediately after the screening that led to the group assignment. The training period consisted of 25 (30 min) sessions in 5 weeks, with five sessions per week. Trained graduate psychology students or special education teachers delivered remediation during school hours in a quiet room. All participating schools were equipped with Windows 10 Desktop PCs with high-speed internet access (nearly 100 Mbps). Treatment fidelity was ensured in two ways. First, all trainers received a detailed manual explaining the intervention procedure and were asked to follow it as instructed. Second, the research group was responsible for the training fidelity through daily communication and weekly debriefings with the trainers. Exposure times in minutes were logged on a university server and sent daily to the research group. Trainers whose children did not complete the daily routine according to the advised session durations were contacted and encouraged to increase participation times. Outcomes were assessed in several cognitive, linguistic, reading, and orthographic processing measures before (T1), during (T2), and after treatment (T3) in Grade 1. In the context of this study, we report the findings on the reading fluency measures in Supplementary Material SM1. The study was carried out per the Cyprus National Bioethics Committee recommendations (ΕΕΒΚ/ΕΠ/2011/10). It also received approval from the Ministry of Education 7.15.01.23/21). Parental consent and school consent were obtained before the initial assessment.

ADAPTATION OF THE GRAPHOGAME INTERVENTION IN GREEK

The design of the training content of the GraphoGame intervention adapted into Greek was based on research findings relevant to the acquisition of letter knowledge and phonological awareness in Greek (e.g., Manolitsis & Tafa, 2011; Papadopoulos et al., 2012). The Greek GraphoGame comprised 240 levels across 12 tasks, including multiple-choice trials. Because each letter in Greek represents a distinct phoneme[6], the game started by introducing these correspondences. Using a synthetic phonics approach, the game began by presenting phonetically and visually distinct grapheme-phoneme correspondences as vowels (e.g., /α/, /ε/, /ο/, /a/, /e/, /o/) after which it moved to give correspondences that were phonetically less distinguishable, as consonants (e.g., /μ/, /ν/, /λ/, /m/, /n/, /l/). Next, it introduced larger sublexical units, such as syllables or rimes, before introducing words. Training material included syllables consisting of two- (e.g., /ta/, /τα/) to four-letter (e.g., /stra/, /στρα/) syllables, and one- (e.g., /to/, /το/) to five-syllable (e.g., /sokolataki/, /σοκολατάκι/) words. The expectation was that word decoding would be achieved by knowing the individual letters' sounds and arriving at the written words by combining the letters correctly. Participants were asked to pair an audio segment (phoneme, syllable, word) with the corresponding visual representation as quickly as possible from 2 to 9 written options presented on the computer screen depending on the particular task (see also Richardson & Lyytinen, 2014, for a detailed description of the program). The difficulty level increased within each and across tasks based on letter complexity.

The game continually logged the participants' performance on accuracy and time measures and progressed according to the participants' level of attainment. A new trial immediately followed each answer. Learning material was provided in subsequent trials to help participants achieve 80% correct responses on each level before moving to the next. This approach offered sufficient challenge and ample opportunity for success, facilitating game engagement. Participants received immediate auditory and visual feedback on their responses. The turnaround in each game was short, providing rewards after approximately one minute of training time. Finally, we note that all participants completed the tasks of their assigned program in the same order until they concluded the intervention, albeit each participant advanced to a different level of the intervention.

Computational Framework for Microgenetic Analysis of Reading Remediation

The proposed computational framework comprises a microgenetic-data encoding model. The model focuses on encoding heterogeneous microgenetic measures into a common informative data representation and a developmental learning stage model that uses this common encoding to generate mathematical metrics to analyze and characterize the readers' developmental stages during an intervention.

THE MICROGENETIC DATA ENCODING MODEL

The level of detail and the format of microgenetic data vary considerably between tasks and participants during an intervention. For example, microgenetic data could be as detailed as logging particular key presses and mouse movements or logging time and accuracy on a second-by-second basis during the task. Moreover, the type of measurements recorded during different intervention tasks may differ. Furthermore, the variation in each task's difficulty level and the program's structure, such as GraphoGame, generate many missing values. This heterogeneity in recorded data constitutes a challenge in developing a unified encoding scheme of microgenetic data for understanding learning progress dynamics.

Motivated by the *rate-level* proposal first advocated by Zigler (1969) as a cognitive-developmental approach for studying individual differences in intelligence, we propose an encoding model that introduces performance and effort concepts. Zigler and Balla (1982) and Zigler and Hodapp (1986) have argued that the development of children with intellectual deficits does not differ from that of typically developing counterparts, except that it progresses at a slower rate and attains a lower level (or asymptote). Based on this assertion, the model we propose assumes that, whatever the underlying format of the raw microgenetic data may be, there is a mapping from the raw data to a performance score and an effort score for each participant and each task (level) of GraphoGame. Intuitively, performance corresponds to a rating of how well a participant executes a particular task, and effort corresponds to the energy (or resources) a participant allocates to the specific task. An example of a performance score could be the number of correct (individual) answers during a task's execution. Similarly, an example of an effort score could be the total time taken and the number of attempts on the participant's part to complete the task. It is up to the researcher to define this mapping depending on the microgenetic data available in each task. However, there is considerable flexibility in defining each mapping, which can vary from task to task or level to level. The resulting performance-effort space captures information about the learning progress dynamics of every individual within a developmental framework.

In the current study, we defined each intervention program's performance and effort scores separately to handle the diversity in microgenetic data recorded in each paradigm. In the case of the GraphoGame intervention, the program allowed each participant to retake a task multiple times until the accuracy threshold for the task was met before the participant could advance to the next task. The software recorded data about each attempt's accuracy score and corresponding response time. With these microgenetic data available, we define performance as the accuracy score achieved by the participant at its best attempt on a given task and effort as the total response time across all attempts. The PREP intervention followed a different delivery protocol for each task. In brief, each participant performed a series of tasks, each corresponding to a game-like activity. A participant repeated the same task (i.e., type of game-like activity) multiple times but at increasing difficulty. The task terminated when a participant failed to reach the accuracy threshold two times in a row on the same difficulty level of the task (we refer the interested reader to the PREP manual, Das, 1999, for the details on the delivery protocol for PREP). The software recorded data about each attempt's accuracy score and corresponding response time. Based on the recorded micro-genetic data of the PREP intervention, we defined the performance score as the accuracy at each level divided by the participant's response time to that level and the effort score as the sum of response times across all attempts at that level. We hypothesized that the proposed mappings of performance and effort scores encode information about the learning progress dynamics as measured by the available micro-genetic measurements.

However, the scores of total performance and effort measures are unsuitable for comparisons across groups, tasks, or intervention levels. First, the overall scores vary widely between the tasks due to differences in the task nature or difficulty level. Second, there are missing values on noncompleted tasks. Our method employs a rank score transformation to achieve score comparability, compensate for the missing values in microgenetic data, and accommodate flexibility in the performance and effort metrics specification. Specifically, for each task (or task/level pair) of the intervention, each participant is assigned a performance-rank and an effort-rank corresponding to that participant's relative ranking compared to the performance and (and respectively effort) scores of all other participants under the same task. The performance-rank and effort-rank of each participant on each task are obtained by simply ordering the participants based on their performance-scores (and, respectively, effort-scores) in descending order. A participant's rank on a specific task is the participant's position in the resulting ordering. If there is a tie in matching scores among two or more participants, those ties are resolved based on the relative ranking of those participants in the immediately preceding task. When a participant does not complete a particular task (i.e., fails to reach that level in the intervention; thus, there are missing values in the performance and effort scores), the corresponding score of the participant is set to a negative constant value (i.e., negative one). Due to this attrition handling procedure, participants who concluded the intervention at an earlier task are assigned to lower ranks (and remain at that rank for all subsequent tasks), and participants who advanced further in the intervention are assigned higher ranks. Another side-effect of this attrition and the tie-breaking procedure is that participants have rank values for all tasks (even if that participant does not complete a given task), thus handling the problem of missing values in the raw scores. For example, the first participant who fails to reach the *i*th task in the intervention is ranked last and will remain last in all subsequent tasks.

More formally, for a given individual n and a particular intervention task t, the resulting microgenetic data encoding model is defined through the variables $P_{t,n}$ which corresponds to the relative ranking of the individual n during the execution of a task t based on the participant's performance score, and $E_{t,n}$ that corresponds to the relative ranking of the same individual based on the participant's effort score and considering the tie-breaking and attrition handling producers. The overall process of defining the microgeneric data encoding model is illustrated in Figure 1, Panel A.

DEVELOPMENTAL LEARNING STAGE MODEL ON MICROGENETIC DATA

In this section, we propose a model to characterize readers' learning dynamics during the intervention program based on the microgenetic data encoding model defined by the variables $P_{t,n}$ and $E_{t,n}$ (i.e., performance-rank and effort-rank) above.

Histogram Profiles of Performance and Effort Ranks. First, we introduce the concept of the histogram profile (HP). For a sub-group of participants (G) and a subset of tasks (T), we consider the histogram H(G,T) of all rank values (either performance-rank or effort-rank) attained by all participants in the subgroup G during the subset of tasks T. The histogram is an approximate representation of the frequency of different rankings of all participants and tasks in the group/task sub-group of interests. Next, we apply a series of mathematical operations on histogram H to transform it into a probability distribution over rankings. We treat the histogram H as an aggregation of Dirac delta functions[7] and apply a convolution operator on H using a Gaussian kernel[8]. This operation results in a smooth but un-normalized continuous function over rankings. To convert the resulting function into a probability distribution, we divide it with a normalizing constant z (i.e., z is the area under the unnormalized function over its domain). We term this probability distribution as the histogram profile HP(G,T) of the group.

Next, we apply a series of mathematical operations on histogram H (a convolution with a Gaussian kernel, followed by normalization, dividing the convolution result with an appropriate constant) to obtain a smoothed estimate of the probability distribution over the rankings. We term this probability distribution as the histogram profile HP(G,T) of the group.

We note that the histogram profile HP carries all available information about the group's overall achievement during the tasks' execution. For example, had the participants in group G achieved the highest possible rankings during the task, their HP would have been skewed toward the left of the distribution's domain (i.e., high rankings). On the other hand, had they achieved the lowest possible rankings, their HP would have been skewed toward the right of the distribution's domain (lower ranks). Similarly, if the group had no particular achievement trend in the intervention, their HP would follow a uniform distribution. We denote the HP for these three specific cases as $HP^{(opt)}$ and $HP^{(worst)}$, respectively, and we note that all three can be expressed as analytic formulas. We will refer to these three HPs as benchmarked HPs because they constitute measurable milestones reflecting the achievement stages of a group. Figure 1, Panel B illustrates the calculation of HP, $HP^{(opt)}$, and $HP^{(worst)}$ histogram profiles.

Intuitively, a distance or similarity measure between the measured HP and the benchmarked HPs would reflect the degree to which the group's achievement deviates from or approaches a favorable or unfavorable benchmark. For example, the closer the HP of a group during a subset of the task is to $HP^{(opt)}$, and the further away it is from $HP^{(worst)}$, the better the group's achievement level. Moreover, such similarity measures across subgroups and task levels provide information about the achievement level dynamics. We propose using Jensen-Shannon's divergence as a similarity measure in our computational framework.

Developmental Learning Stage Metric for a Group. Typically, learning intervention programs define the order in an increasing difficulty level in which intervention tasks and levels are administered. For example, the GraphoGame intervention defines an ordering of the tasks and levels in increasing difficulty by allowing the participant to construct letters into syllables gradually, then small words, and then larger words. Moreover, we are typically interested in the progress of a small subgroup of participants with common characteristics; for example, participants with similar cognitive, linguistic, reading, or orthographic processing skills scores were obtained on pre-, mid-, or post-intervention assessments. Given such a subgroup of interest, Gi, and a set of consecutive task/level pairs, Ti, we define the metric of the developmental learning stage (DLS) in terms of either performance-ranking or effort-ranking as follows:

$$DLS(G_i, T_i) = \log\left(\frac{\sqrt{J\left(HP(G_i,T_i), HP^{\{worst\}}(G_i,T_i)\right)}}{\sqrt{J\left(HP(G_i,T_i), HP^{\{opt\}}(G_i,T_i)\right)}}\right)$$

where the HP(*Gi*, *Ti*) corresponds to the Histogram Profile of the group estimated based on participants' performance on tasks in T_i. The function *J* is Jensen-Shannon's divergence[9] between the two probability distributions. Intuitively, DLS describes how the group performance is more similar to the best possible performance or worst. If the distances of HP to $HP^{(worst)}$ and HP to $HP^{(opt)}$ are the same, the DLS equals 0. If the distance of the observed HP to $HP^{(opt)}$ is greater than that of HP to $HP^{(worst)}$, the DLS would have a negative value. If the opposite holds, the DLS will have a positive value, reflecting the group's proximity to the

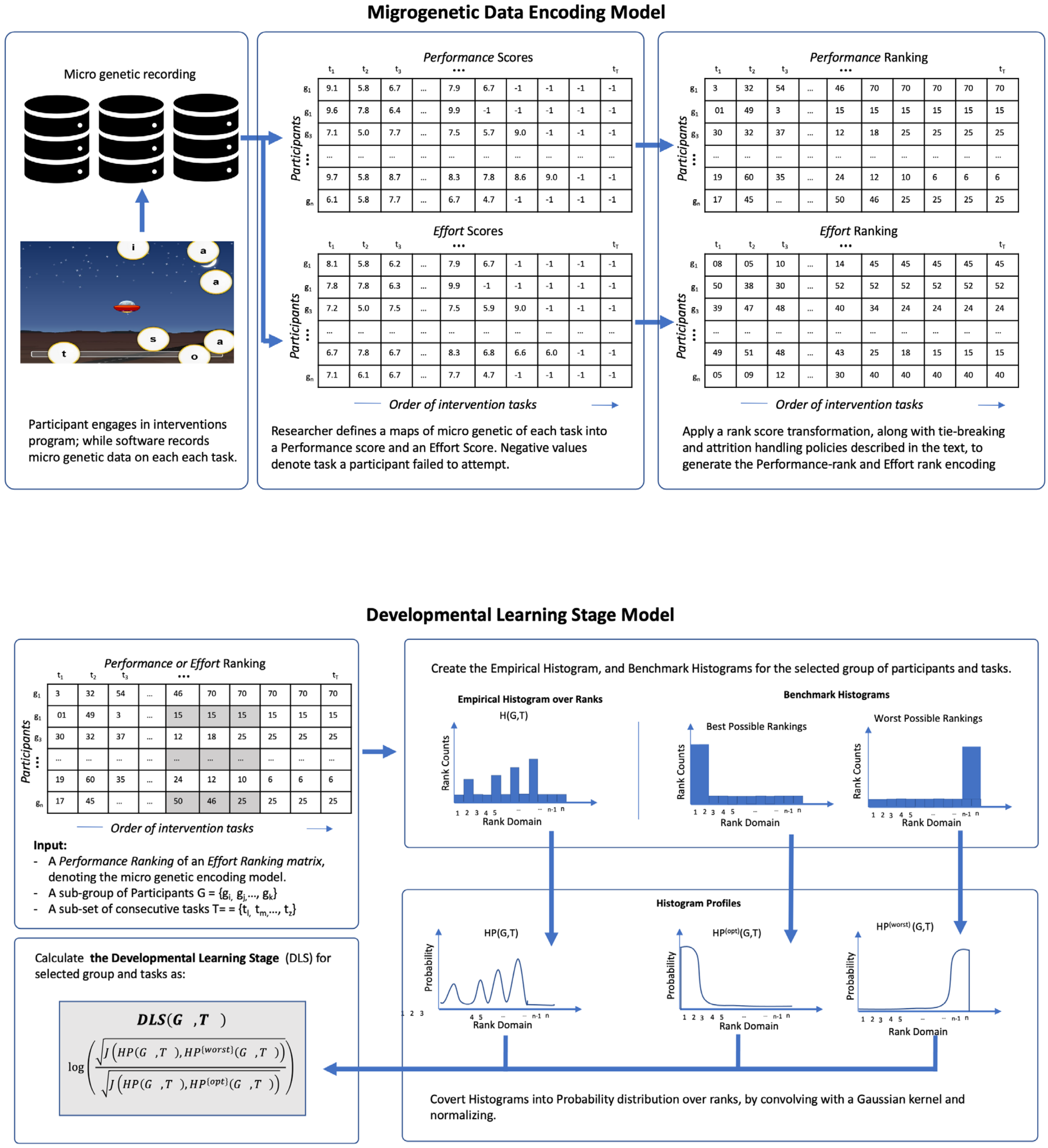

FIGURE 1.

Panel A: The process of calculating the microgenetic data encoding model: Researchers specify customized mapping between the diverging formats of microgenetic data into performance and effort scores. A rank score transformation, along with tie-breaking and attrition handling policies described in the text, generate the performance-rank and effort rank space, encoding the learning process dynamics of microgenetic data. Panel B: The process of calculating the developmental learning stage model: Performance ranking (or effort ranking matrix), along with a selection of a group of participants and tasks are used to calculation the empirical and benchmark histograms. The histograms are converted into histogram profiles functions by convolution with a Gaussian kernel and appropriate normalization. The histogram profiles capture the probability density over ranks for the empirical and benchmark histograms. The development learning stage for the selected group is then calculated as a function of the histogram profiles HP, $HP^{(opt)}$, and $HP^{(worst)}$. The function J (defined as the square root of the Jensen-Shannon divergence) is a measure of the distance between the different histogram profiles.

best possible performance it could have achieved. The calculation of the DLS is illustrated in Figure 1, Panel B.

Developmental Stage Metric for an Individual. It is of interest to know the degree to which each participant contributes to the group's DLS during a subset of tasks. To this end, we propose the individual's developmental learning stage (iDLS) metric. In particular, for a specific participant g in the group G_i and a set of consecutive tasks T_i, we define the contribution of participant g to the instantaneous developmental stage metric as

$$iDLS(g|\,G_i,T_i) = DLS(G_i,T_i) - DLS(G_i - g,T_i)$$

where $G_i - g$ corresponds to the set of all participants in group G_i after information from participant *g* has been removed and replaced with a uniform distribution. Consequently, the iDLS can be considered a measure of how a participant's absence from the group would affect the group's DLS. Finally, we note that both DLS and iDLS metrics can be evaluated regarding performance-rank or effort-rank measures.

Developmental Learning Stage Dynamics for Group and Individuals. Both metrics (DLS and iDLS) proposed in the previous sections are static. They capture group or individual performance and effort information for a fixed instance during the intervention. Often, one is interested in modelling how the performance and effort of either a group of participants or a participant change during the intervention. We note that the time instance modelled by DLS and iDLS during the intervention is specified by selecting the tasks in set *T*. Given a sequence $S=\{T_1, T_{-2},\ldots, T_M\}$ of the task's set, the sequence of DLS (and equivalently iDLS) evaluated on *S* captures the variation of instantaneous learning dynamics during the intervention and, thus, it can be interpreted as a model of the developmental learning stage dynamics.

FOR ANALYZING THE DEVELOPMENTAL LEARNING STAGE METRICS

The previous section introduced the model that captures the developmental learning stages during intervention for specific individuals or groups and a subset of tasks based on microgenetic data. These metrics can be defined as individual performance or effort during the intervention. This section describes several methods for establishing the relation of the proposed metrics to reading fluency measures and how to analyze these data to gain insights into the reading performance of individuals or groups during the interventions. For clarity, we introduce these methods in the GraphoGame intervention use-case context, but we note that the methods can be applied to any generic reading remediation program.

Dynamic Correlation Trace. Our proposed model's first step of analysis involved examining the relationship between the proposed DLS metrics at each intervention level and reading assessment performance scores at the post-intervention (Time 3) reading assessment measures. In particular, for a group of participants G and a sequence of consecutive tasks $T=[T_1,\ldots T_n]$ that portrays the intervention tasks' progression, we calculate the iDLS(G,Tt) correlation to post-intervention reading measures for each Tt. We termed the resulting sequence of the correlation coefficient a *dynamic correlation trace*. This captures the dynamics (i.e., the change) throughout the intervention in the association strength between the DLS metrics and reading assessment measures obtained post-intervention. In the context of the GraphoGame use-case, we consider phonemic decoding and word reading fluency as post-intervention reading assessment measures. However, we note that the same model applies to any assessment measure obtained by the experimenter. We performed a permutation test that modelled the null hypothesis of no correlation to establish significant correlation levels for the dynamic correlation trace values.

The dynamic correlation trace offers a way to visualize the changes in correlation between the proposed DLS metric and the intervention's efficacy on different reading measures. Moreover, it can help determine when the intervention group's DLS metric starts to better reflect the intervention program's potential effects, allowing an experimenter to pinpoint when the intervention becomes effective. Besides, by calculating the dynamic correlation trace for different reading assessment scores, an experimenter can study how various tasks impact the underlying reading skills. Notably, the statistically significant correlation values of the dynamic correlation trace validate the connection between the proposed DLS metric and actual reading performance measures.

Relation to Reading Remediation Effects. Reading performance scores obtained pre- (T1), mid- (T2), and post- (T3) intervention capture information about reading measures at those specific moments in time. Therefore, we consider those moments as anchor time points during the intervention that, taken together, capture the reading remediation effects of an individual in terms of the corresponding performance measure. The proposed iDLS metric aims to capture the learning dynamics between those anchor points, allowing for a more granular assessment of the remediation effects at every task. To establish the relationship between the proposed metric and the reading remediation effects, we explore how iDLS scores can be inferred from the reading performance at the three anchor time points at specific moments during the intervention. We argue that the degree to which iDLS can be interpolated by the reading performance measurement at the anchor time points denotes the amount of information that remedial effects on that measurement are reflected in the iDLS. Moreover, the coefficients associated with the three anchor points give insights into which anchor points of reading performance contribute the most to this association.

Towards that, we examined the degree to which the reading performance scores obtained pre- (T1), mid- (T2), and post- (T3) intervention predict the proposed iDLS metric at different tasks/times during the interventions. Thus, we model the reading remediation program effects, reflected in the variance of the instantaneous DLS metric and how those change throughout the intervention. Specifically, we model the iDLS metric for a given participant g during a task T, as follows:

$$i\widehat{DLS(g,T)} = \beta_1 R^{(pre)} + \beta_2 R^{(mid)} + \beta_3 R^{(post)} + \epsilon$$

where the R^{pre}, R^{mid} and R^{post} correspond to the absolute reading assessment scores obtained at pre-, mid-, and post-intervention. For example, as reading assessment scores in the GraphoGame use-case, we considered word reading fluency and phonemic decoding fluency at pre-, mid-, and post-intervention. In this model, we considered the instantaneous DLS of the participants *g* at any given task during the intervention as the dependent variable. The independent variables are the reading performance scores measured pre-, mid-, and post-intervention. We note that the model can be estimated for iDLS scores derived in terms of Performance-ranking or Effort-ranking scores of the individuals during the intervention. We estimated the model using a 10-fold cross-validation procedure, where nine blocks were used to estimate $i\widehat{DLS}(g,t)$ for every participant n and every task t. The model fit is then calculated using R^2 scores between the actual performance iDLS and the estimated performance $i\widehat{DLS}$ of the group. The modulation of R2 describes the percentage of variance in iDLS that the pre-, mid-, and post-intervention reading performance scores can explain. Moreover, the β values represent the impact of these scores on reading performance.

Group Differences in Developmental Learning Stages. Furthermore, we employ the proposed iDLS metric to identify group differences based on the overall performance and effort at specific times during the intervention. In particular, we examined whether grouping participants based on their cognitive, linguistic, reading, and orthographic processing skills differed in their instantaneous DLS at intervention intervals of interest. In this analysis, we divided the participants into three groups (high, medium, and low) based on their actual reading fluency gains from the intervention at T3. Then, we compared the iDLS scores of these new groups, using the aggregated score of every participant along with Tasks 30-40 (which correspond to the window of tasks around the third quartile of all tasks administered). We calculated and compared the iDLS scores for performance-ranking and effort-ranking data. Moreover, we tested for differences regarding performance and effort across intervention groups. In particular, we compared the aggregated iDLS scores between the GG and PREP-to-GG groups. The aggregated score for this comparison was calculated with tasks 1-10 (i.e., the beginning of the GraphoGame intervention). This selection's motivation was to test if exposing participants to the PREP intervention before the GraphoGame intervention would affect their performance during the GraphoGame intervention.

RESULTS

Dynamic Correlation Trace Results

As part of the dynamic correlation trace analysis, we calculated the correlation values between the iDLS scores on performance-rank measurements during different stages of the GraphoGame intervention and post-intervention reading scores for the three intervention groups (GG, PREP-to-GG, and GG-to-PREP). The results are displayed in Figure 2. The top row shows the correlation trace to phonemic decod-

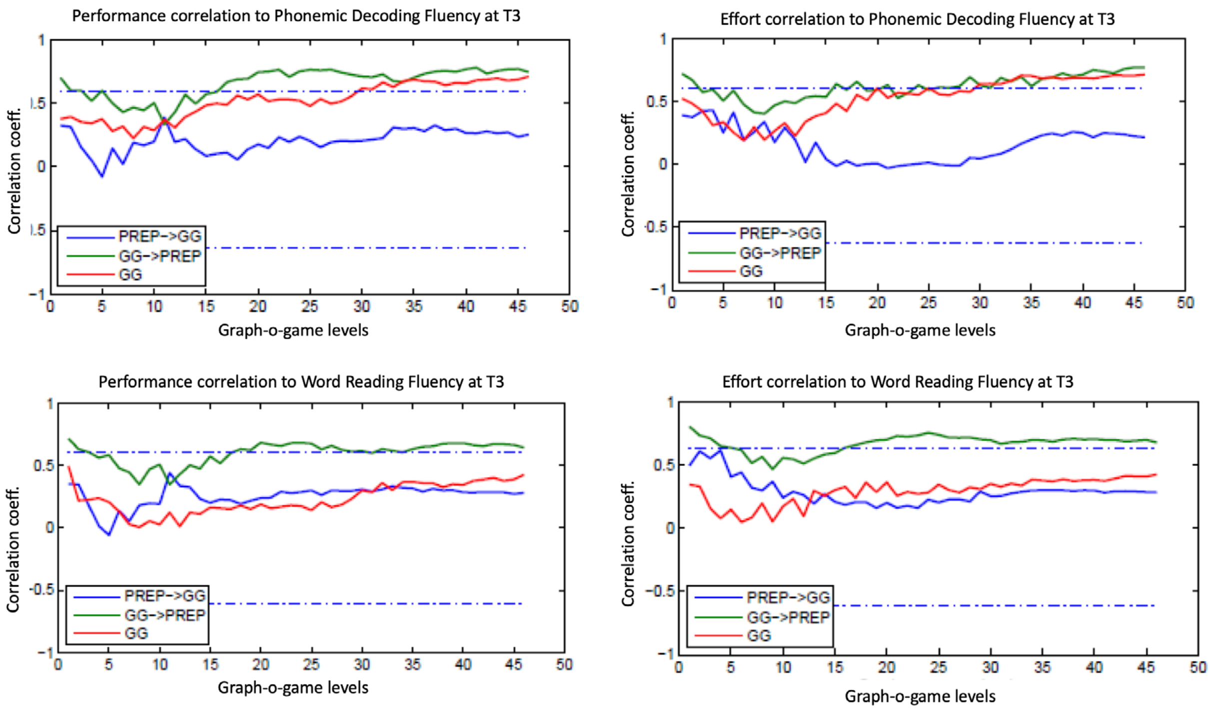

FIGURE 2.

Correlation values between the instantaneous performance dynamics and effort scores during different stages of the Graphogame intervention and post intervention reading scores. The top row shows the correlation to phonemic decoding fluency scores. The bottom row shows the correlation to word reading fluency scores, both obtained by the participants at T3 (post-intervention). The red line shows the correlation of participants in the Graphogame group, the green line shows the correlation of participants in the GG+PREP group, and blue line shows the correlation of participants in the PREP+GG group. The dotted blue line shows the threshold for .05 significance level.

ing fluency scores, and the bottom row shows the correlation trace to word reading fluency scores participants obtained post-intervention. Red lines show the correlation in participants in the GraphoGame group, green lines show the correlation in participants in the GG-to-PREP group, and blue lines show the correlation in participants in the PREP-to-GG group. The dotted blue line indicates the threshold for .05 significant levels. The correlation is defined based on the score values obtained within each intervention group throughout the program and shows how the final group performance correlates with these values. Moreover, to establish significant correlation levels, we performed a permutation test and modelled the null hypothesis of no correlation (i.e., by randomizing the labels assignment of the group and the performance level of each participant)

At the earliest stage of the intervention (< 16 tasks), none of the groups showed a significant correlation between their phonemic decoding fluency and their corresponding instantaneous performance metric (top left Figure 2). However, after the 16th task, the GG-to-PREP group crossed the significance line ($p < .05$), with the correlation reaching a peak after the 24th task of the intervention ($r = 0.8$). The GraphoGame group's performance followed a similar pattern, with its correlation increasing after the 15th task but reaching significance ($p < .05$) after the 30th task.

The correlation patterns between the instantaneous performances and word reading fluency showed a somewhat different picture (bottom left Figure 2). At the early stages of the intervention (< 16 tasks), none of the groups showed a significant correlation between their word reading fluency and their instantaneous performance metric. However, the GG-to-PREP group reached a significant correlation after the 16th task and retained significance throughout the intervention. None of the other groups showed significant correlations between their instantaneous performance dynamic and word reading or phonemic decoding fluency at any intervention stage.

The correlation trace to post-intervention reading scores provides unique insights into understanding the intervention's learning dynamics and overall effect. Such insights are missing in traditional pre/post-reading measures. At the very least, the results suggest that, for certain groups, the proposed instantaneous performance metrics can carry information that predicts (to some degree) the intervention impact. Moreover, the predictive power of these metrics is modulated by the different stages of the intervention.

Relation to Reading Remediation Effects

Next, we examined how the proposed iDLS metrics predict the overall performance of an individual during the intervention based on the reading scores obtained pre- (T1), mid- (T2), and post- (T3) intervention. This analysis quantified which factors (i.e., reading fluency or phonemic decoding) impact the proposed iDLS metric and how this relation is modulated over time. The results of this analysis are shown in Figure 3. The predictive model based on the phonemic decoding fluency and the instantaneous performance score (top Figure 3) revealed that the explained variance (measured in r^2) followed a cumulative pattern after the 15th task ($r^2 = 0.2$). Also, by the 45th task, the phonemic decoding fluency predicted 45% of the variance ($r^2 = 0.45$) in instantaneous performance scores. Moreover, the model's β coefficients showed that the factors T2 (mid-assessment) and T3 (post-assessment) carried almost all of the model's predictive power when the original phonemic decoding fluency score (T1) had little to no predictive power. These results suggest that participants' instantaneous performance was better described by their underlying reading performance on phonemic decoding fluency measured post-intervention (i.e., T2 and T3). The

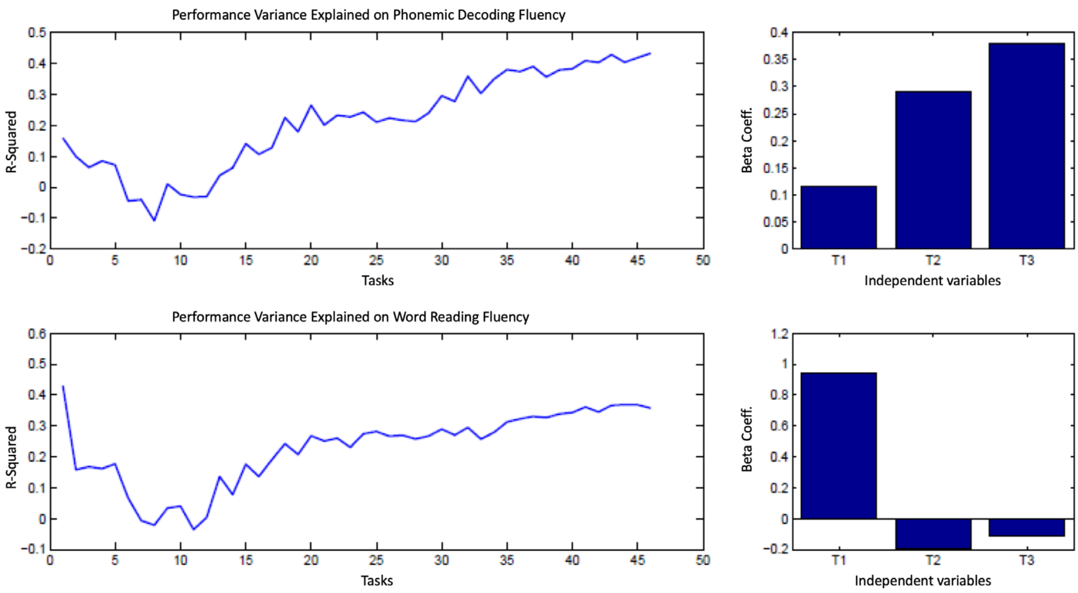

FIGURE 3.
General linear model prediction analysis based on iDLS.

initial learning state carried little information regarding participants' performance during the intervention. Thus, the proposed iDLS metric is informative of the progression of the phonemic decoding fluency performance during interventions and can be considered an index of the phonemic decoding fluency at different intervention times.

The predictive model based on the word reading fluency and the instantaneous performance score (bottom Figure 3) showed a similar pattern but a slower development rate. By the 45th task, the word reading fluency score explained 36% of the variance ($r^2 = 0.36$) in instantaneous performance scores during the task. However, the model's β coefficients for the three factors showed that the T1 score carried all the predictive power ($r^2 = 0.92$), while the scores at T2 and T3 had negative predictive power. These results suggest that the participants' word reading skills before remediation better explained the performance at the end of the intervention. Thus, the GraphoGame intervention did not modulate participants' underlying reading stage regarding word reading fluency to the same degree as in phonemic decoding fluency. This finding is sensible considering GraphoGame's training objectives, as the intervention emphasizes phonological skills and letter knowledge enhancement (Lyytinen et al., 2007). Therefore, more opportunities for learning to read naturally are expected to facilitate the transfer of solid phonemic decoding skills to other related (word) reading material. Interestingly, phonemic decoding and word reading models failed to predict the instantaneous performance between the first and the tenth tasks.

Group Performance Differences Results

Finally, we examined how the groups differed regarding their iDLS performance scores. In doing so, we divided the participants into three groups (high, medium, and low) based on their actual reading fluency gains from the intervention at T3. Then, we compared the iDLS scores of these three groups, using the aggregated score of every participant across Tasks 30-40 (which correspond to the window of tasks around the third quartile of all tasks administered).

Group comparison based on iDLS measures showed a decreasing pattern (Figure 4, left) in the aggregated iDLS scores among the three groups where the high gainers scored comparatively higher in iDLS, low gainers scored lower in iDLS scores, and mid-gainers achieved a value in-between. The analysis showed that these group differences were statistically significant ($F = 5.64$, $p < .01$). Moreover, we tested for differences regarding performance across intervention groups. In particular, we compared the aggregated iDLS scores between the GG and PREP-to-GG groups. The aggregated score for this comparison was calculated with Tasks 1-10 (i.e., the beginning of the GraphoGame intervention). This selection's motivation was to test if exposing participants to the PREP intervention before the GraphoGame intervention would affect their performance during the GraphoGame intervention. Indeed, results showed that the PREP-to-GG group's aggregated iDLS score was higher than that of the GG group (Figure 4, right). Again, these group differences were statistically significant ($F = 6.27$, $p < .05$).

DISCUSSION

In the current study, we proposed a novel generic framework for analyzing microgenetic data to explore the learning progress dynamics and readers' developmental stages during an intervention. Our model was theoretically motivated by Zigler's cognitive-developmental approach for studying individual differences and examined the contribution of microgenetic analysis to establishing a link between reading intervention and the changes in performance and effort during training. Our microgenetic results showed differences between the experimental groups regarding how their performance is modulated at different stages of the intervention and their predictive power regarding their final scores. Also, the findings showed that designing and implementing intervention schemes in reading research has never been easy, and it will not become so unless we understand how children progress through an intervention. We discuss these findings below.

Microgenetic Analysis

The current study's foremost challenge was developing and validating a microgenetic method to collect data on how the anticipated improvement (i.e., reading fluency performance) is produced in the participant-treatment interaction. Microgenetic analysis of the learning situation and a participant's responses during an intervention is necessary if we wish to establish a link between the theory of the linguistic or cognitive functions underpinning a remediation program and the changes in performance and effort that occur during training. Our results showed

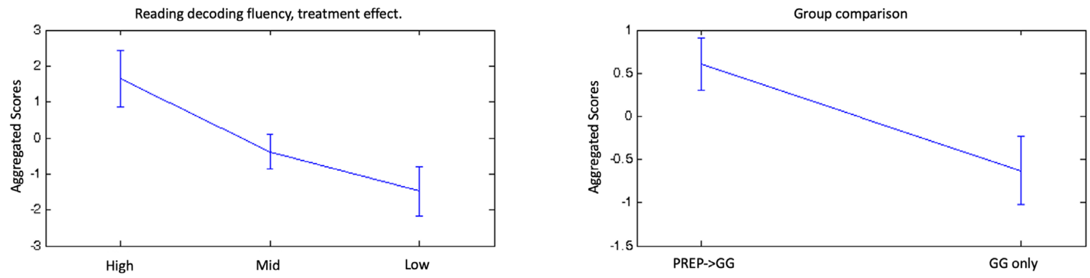

FIGURE 4.

Left: Group comparisons between, high- mid- and low-performing groups and terms of reading decoding fluency. Right: Group differences between the PRES->GG group and the GG only group, in terms of the aggregated iDLS metric.

that the proposed microgenetic analysis framework could help explore such interactions. We elaborate on these points next.

First, our results confirmed that the new iDLS metrics proposed under the microgenetic data analysis framework, calculated mid-way into the intervention, captured information relevant to the actual gains (regarding reading fluency) post-intervention. This finding suggests that the metrics can monitor learning progress dynamics at different stages during the intervention. The microgenetic data analysis framework introduced a correlation trace analysis to achieve that. The correlation trace between the instantaneous performance and effort scores at each level of the intervention and reading scores at post-intervention provided insights into the effects of the GraphoGame remedial program, allowing to generate hypotheses on the optimal duration. The results showed that the combined GG-to-PREP group exhibited a significant correlation between instantaneous performance and fluency scores (for phonemic decoding and word reading fluency). The GraphoGame intervention should be administered for at least 16 tasks for a significant effect. Moreover, the GraphoGame intervention could be terminated by the 30th task (1/8 of the entire program duration) without weakening the outcome in reading decoding fluency. Legitimately, GraphoGame treatment could have the same impact as the complete treatment if administered for a shorter period (see Lyytinen & Richardson, 2013, for a similar argument). Also, shortening the GraphoGame intervention duration freed the cognitive intervention from any fatigue effect, which positively impacted the overall treatment. Therefore, the results also suggest that the observed differences in correlation trace patterns between the combined GG-to-PREP and the GraphoGame groups are modulated by the impact of PREP intervention on word reading fluency. Thus, metrics generated by the proposed framework can be considered effective indicators of reading gains (i.e., phonemic decoding fluency) due to intervention, monitor such gains throughout the interventions, and generate insights towards determining the most effective duration of intervention.

We have also implemented a general linear model and reported findings on the predictive power of each independent variable. Results of the GraphoGame intervention revealed that participants' performance during the intervention is better described by their underlying performance in phonemic decoding fluency measured at the post-intervention assessment. Results also showed that the performance in word reading fluency before the remediation affected the performance during the GraphoGame intervention, especially in more challenging tasks. Regarding the phonemic decoding fluency skills, these can be attributed partly to the theoretical underpinnings of GraphoGame and partly to the transparency of the Greek language or a combination of the two. GraphoGame aims to enhance accuracy (including phonological awareness and orthographic knowledge) and automaticity, represented by decoding and fluency (Richardson & Lyytinen, 2014). Also, given that in a transparent orthographic system, the letter-sound connections can be drilled efficiently and without complications (Lyytinen et al., 2009) and that phonemic decoding accuracy and fluency are strongly predicted by phonological skills in Greek (Papadopoulos et al., 2020), it does not come as a surprise that GraphoGame alone, or in combination with cognitive training, may also lead to efficient pseudoword decoding. The transparency of the Greek language allows young readers to use the phonological representations of any grain-size units (rhyme, syllable, or phoneme) that are available to them (Papadopoulos et al., 2012), enabling even children who show insufficient phonological processing at school entry to gradually tackle their difficulties with phonological processing and find means to compensate for poor reading performance (Papadopoulos et al., 2009). Thus, the finding that participants showed more notable attainment on phonemic decoding than real word reading can be reliably attributed to the remediation rather than classroom instruction. For the real word reading to continue to be enhanced, additional learning opportunities and continuous interaction with reading material would be necessary (Vaessen & Blomert, 2010).

Children must also have some necessary word reading skills before receiving the program to gain from remediation on word reading fluency. Thus, by providing remediation at the grapho-phonemic level, GraphoGame may have created a basis for future independent learning and a cognitive and linguistic foundation on which phonological processes and reading can be further built (Lyytinen et al., 2009). Therefore, the proposed framework, combined with an appropriate experiment design, can help identify the factors that modulate participants' performance during the intervention and generate insights into the prerequisite skills expected by participants to be most effective.

The proposed microgenetic data analysis framework can provide additional insights into understanding differences among intervention groups at different learning stages. For example, we demonstrated that using the proposed framework allowed us to examine whether administering the cognitive intervention earlier could impact a group's performance at the early stages of the phonological intervention. Our results show a significant difference between the PREP-to-GG and GG groups, with the former having much higher iDLS scores than the latter. These findings provide additional evidence that the distal (cognitive) processes (i.e., successive and simultaneous processing) support the development of proximal (linguistic) processes (e.g., phonological) and, thus, reading (for more information on the relationship between distal and proximal processes, see Supplementary Material SM1). Such findings align with traditional reading development models, showing that distal cognitive processes, such as information processing abilities, predict word reading through proximal cognitive skills, such as phonological awareness (e.g., Das et al., 2000; Papadopoulos et al., 2020).

Consequently, the present findings provide additional evidence for two significant aspects of the contemporary literature. First, they demonstrate that game-based interventions directing children's attention to goal-oriented behaviors, from distal to proximal or top-down processes, can systematically enhance reading development (see Verwimp et al., 2023, for a similar argument). Second, they showcase how computational accounts of various aspects of task performance at a microgenetic level can better inform learning-curve analysis models (e.g., Stafford & Vaci, 2022), thus further allowing research on game-based interventions to contribute to the broader literature on skill development.

The proposed framework can facilitate the design of adaptive, personalized, computer-assisted reading intervention programs that account for each child's heterogeneity, individual differences, and learning progression (Verwimp et al., 2021). Such individualized interventions could help establish more effective intervention strategies optimized to each individual's cognitive deficit profile and learning trajectory. Current trends in reading research underscore that heterogeneity and individual differences in dyslexia profiles can be explained systematically only with a personalized computational model of learning trajectories (Perry et al., 2019). Accordingly, personalized and targeted intervention strategies can only achieve optimal remediation. Thus, insights from the proposed framework can inform the implementation of customized remedial applications and maximize the efficacy of remediation to benefit children with dyslexia.

In conclusion, we proposed a novel methodological framework for examining learning progress dynamics in reading remediation using microgenetic data. The framework addressed the problem of encoding microgenetic data into a common data representation model, introduced four information-theoretic metrics to capture the instantaneous developmental learning stages of groups and individuals, and provided the mathematical model to analyze those metrics to study learning stages during the intervention. Although the GraphoGame intervention was used in the current study as a case for validating the model, the current approach is not tied to a particular intervention. Instead, this approach could be used with other intensive and focused remedial programs in reading or psychological research involving progressively difficult requirements. Our findings demonstrated the proposed framework's ability to capture the learning stage dynamics during the intervention. The suggested model can indeed provide unique insights into exploring learning progress dynamics. Thus, the proposed framework offers a starting point for further research to study the modulation in learning stages during an intervention and better understand how reading occurs and how reading disability may be adequately treated.

FOOTNOTES

[1] With the article's main objective being the development of the computational model, the presentation is focused on the first program for reasons of economy.

[2] Data was derived from the ReaDI-STANCE project focusing on treating reading difficulties in a group of 6-year-old Greek learners. Information regarding group comparisons on reading achievement is reported in Supplementary Material SM1. For further details, the interested reader may contact the last author.

[3] GraphoGame was developed through the GraphoLearn initiative, a global academic effort dedicated to creating evidence-based tools for literacy acquisition. To this day, GraphoLearn continues to research and develop language-specific versions of GraphoGame.

[4] Reading fluency was assessed with two tasks, Word Reading and Phonemic Decoding, taken from the standardized Early Reading Skills Assessment Battery (ERS-AB; Papadopoulos et al., 2009). In both tasks, the instruction to the participants was to read as fast as possible a list of given words as follows. Before each task, a short practice list of 5 words/non-words was presented. Participants' score was the number of correct words or non-words read within a 60-s time limit. The reported Cronbach's α for the Word Reading task in the standardization sample was .88, and for the non-word reading task, .92 in Grade 1. The Word Reading test consisted of 80 words forming a 2 × 2 × 2 factorial design of frequency (high/low), orthographic regularity (regular/exception), and length (bisyllable/trisyllable). The words included mainly nouns with a few adjectives and verbs. The Phonemic Decoding test consisted of 45 pronounceable non-words derived from real words after changing two or three letters. The task started with one-syllable words and ended with five-syllabic words, with the majority being two- and three-syllabic words (25 and 12 words, respectively).

[5] In the current study, we used only the data from GraphoGame and those that combined GraphoGame with PREP. Our purpose was to demonstrate the application of the microgenetic method, not to speak about the group intervention effects. Information about the PREP cognitive remediation is provided in Supplementary Material SM2.

[6] Protopapas and Vlahou (2009) have reported a consistency of 98% in the feedforward condition, from orthography to phonology.

[7] Histograms can be represented as the sum of Dirac delta functions http://science-memo.blogspot.com/2013/11/demystify-dirac-delta-function-for-data.html

[8] A Gaussian kernel is a function derived from the normal probability distribution, and is centered at zero. It is often used as a kernel in the convolution operation to smooth an input function (Shapiro & Stockman, 2001).

[9] Jensen-Shannon divergence (JSD) is a popular method of measuring the similarity between two probability distributions. The square root of the JSD is a distance function (Endres & Schindelin, 2003).

ACKNOWLEDGEMENTS

The authors declare that the research was conducted without commercial or financial relationships construed as a potential conflict of interest.

This research was supported by a Cyprus Research Promotion Foundation Grant and the European Regional Development Fund (ERDF): EXCELLENCE HUBS/1216/0508 granted to Timothy C. Papadopoulos.

DATA AVAILABILITY

The datasets presented in this article are not readily available because further data analysis is currently in progress.

REFERENCES

Adolph, K. E., Robinson, S. R., Young, J. W., & Gill-Alvarez, F. (2008). What is the shape of developmental change? *Psychological Review, 115*(3), 527–543. https://doi.org/10.1037/0033-295X.115.3.527

Ahmed, H., Wilson, A., Mead, N., Noble, H., Richardson, U., Wolpert, M. A., & Goswami, U. (2020). An evaluation of the efficacy of Graphogame rime for promoting English phonics knowledge in poor readers. *Frontiers in Education, 5*, 132. https://doi.org/10.3389/

feduc.2020.00132

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*(4), 621–635. https://doi.org/10.1017/S0142716403000316

Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research, 72*(1), 101–130. https://doi.org/10.3102/00346543072001101

Bus, A. G., & Van Ijzendoorn, M. H. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*(3), 403–414. https://doi.org/10.1037/0022-0663.91.3.403

Carlson, J., & Das, J. P. (1997). A process approach to remediating word-decoding deficiencies in Chapter 1 children. *Learning Disability Quarterly, 20*(2), 93–102. https://doi.org/10.2307/1511217

Cheung, A. C., & Slavin, R. E. (2013). Effects of educational technology applications on reading outcomes for struggling readers: A best-evidence synthesis. *Reading Research Quarterly, 48*(3), 277–299. https://doi.org/10.1002/rrq.50

Church, J. A., Grigorenko, E. L., & Fletcher, J. M. (2021). The role of neural and genetic processes in learning to read and specific reading disabilities: Implications for instruction. *Reading Research Quarterly, 58(2)*, 203–219. https://doi.org/10.1002/rrq.439

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Das, J. P. (1999). *Professor PREP: PASS Reading Enhancement Program-Administration Manual.* Sarka Educaitonal Resources Inc.

Das, J. P., Parrila, R. K., & Papadopoulos, T. C. (2000). Cognitive education and reading disability. In A. Kozulin & Y. Rand (Eds.). *Experience of mediated learning: An impact of Feuerstein's theory in education and psychology* (pp. 274–291). Pergamon Press.

Elbaum, B., Vaughn, S., Hughes, M. T., & Moody, S. W. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of Educational Psychology, 92(4)*, 605–619. https://doi.org/10.1037/0022-0663.92.4.605

Endres, D. M., & Schindelin, J. E. (2003). A new metric for probability distributions. *IEEE Transactions on Information Theory, 49*(7), 1858-1860. https://doi.org/10.1109/TIT.2003.813506

Fazio, L. K., & Siegler, R. S. (2013). Microgenetic learning analysis: A distinction without a difference. *Human Development, 56*(1), 52–58. https://doi.org/10.1159/000345542

Georgas, J., Paraskevopoulos, I. N., Besevegis, E., & Giannitsas, N. D. (1997). *The Hellenic WISC-III.* Psychometric Laboratory, University of Athens.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading, 7*(2), 155–182. https://doi.org/10.1207/S1532799XSSR0702_3

Inhelder, B., Ackermann, E., Blanchet, A., Karmiloff-Smith, A., Kilcher, H., & Robert, M. (1976). Des structures cognitives aux rocedures de découverte. *Archives de Psychologie, 44*,(171) 52–72.

Jamshidifarsani, H., Garbaya, S., Lim, T., Blazevic, P., & Ritchie, J. M. (2019). Technology-based reading intervention programs for elementary grades: An analytical review. *Computers & Education, 128*, 427–451. https://doi.org/10.1016/j.compedu.2018.10.003

Karmiloff-Smith, A. (2013). 'Microgenetics': No single method can elucidate human learning. *Human Development, 56*(1), 47–51. https://doi.org/10.1159/000345541

Kearns, D. M., & Fuchs, D. (2013). Does cognitively focused instruction improve the academic performance of low-achieving students? *Exceptional Children, 79*(3), 263–290. https://doi.org/10.1177/001440291307900200

Kjeldsen, A., Kärnä, A., Niemi, P., Olofsson, Å., & Witting, K. (2014). Gains from training in phonological awareness in kindergarten predict reading comprehension in grade 9. *Scientific Studies of Reading, 18*(6), 452–467. https://doi.org/10.1080/10888438.2014.940080

Lyytinen, H., Erskine, J., Kujala, J., Ojanen, E., & Richardson, U. (2009). In search of a science-based application: A learning tool for reading acquisition. *Scandinavian Journal of Psychology, 50*(6), 668–675. https://doi.org/10.1111/j.1467-9450.2009.00791.x

Lyytinen, H., Lohvansuu, K., Hämäläinen, J., Pitkänen, M., Leppänen, P., & Richardson, U. (2021). Early identification and prevention of reading problems. In P. Graf & D. J. Dozois (Eds.), *Handbook on the State of the Art in Applied Psychology* (pp. 293–314). Wiley-Blackwell.

Lyytinen, H., Erskine, J., Tolvanen, A., Torppa, M., Poikkeus, A. M. & Lyytinen, P. (2006). Trajectories of reading development: A follow-up from birth to school age of children with and without risk for dyslexia. *Merrill-Palmer Quarterly, 52*(3), 514–546.

Lyytinen, H., & Richardson, U. (2013, July). *A digital environment, GraphoGame, for supporting early reading skills in different orthographies* [Conference presentation]. 20th Annual Meeting of the Society for the Scientific Studies of Reading (SSSR), Hong Kong, China.

Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A-M., & Taanila, M. (2007). Early identification of dyslexia and the usage of computer game-based practice to support reading acquisition. *Nordic Psychology, 59*(2), 109–126. https://doi.org/10.1027/1901-2276.59.2.109

Manolitsis, G., & Tafa, E. (2011). Letter-name letter-sound and phonological awareness: Evidence from Greek-speaking kindergarten children. *Reading and Writing, 24*, 27–53. https://doi.org/10.1007/s11145-009-9200-z

Mayer, R. E., & Moreno, R. (2002). Animation as an aid to multimedia learning. *Educational Psychology Review, 14*, 87–99. https://doi.org/10.1023/A:1013184611077

McTigue, E. M., Solheim, O. J., Zimmer, W. K., & Uppstad, P. H. (2019). Critically reviewing GraphoGame across the world: Recommendations and cautions for research and implementation of computer-assisted instruction for word-reading acquisition. *Reading Research Quarterly, 55*(1), 45–73. https://doi.org/10.1002/rrq.256

Miller, P. H., & Coyle, T. R. (1999). Developmental change: Lessons from microgenesis. In E. K. Scholnick, K. Nelson, S. A. Gelman, & P. H. Miller (Eds.), *Conceptual development: Piaget's legacy* (pp. 209–239). Erlbaum.

Naglieri, J. A. & Das, J. P. (1997). *Das-Naglieri Cognitive Assessment System*. Riverside Publishing (Standardization in Greek by Papadopoulos, T. C., Georgiou, G. K., Kendeou, P., & Spanoudis, G. 2007. Department of Psychology, University of Cyprus).

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel – Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Publication No. 00-4754). U.S. Government Printing Office.

Papadopoulos, T. C. (2019, December). *Reading Difficulties Intervention: Scientific, Technological, and New Community Efforts (ReaDI-STANCE). Research and Innovation Foundation and the Framework Program "Excellence Hubs"* (Grant Agreement No. EXCELLENCE/1216/0508).

Papadopoulos, T. C., Csépe, V., Aro, M., Caravolas, M., Diakidoy, I. A., & Olive, T. (2021). Methodological issues in literacy research across languages: Evidence from alphabetic orthographies. *Reading Research Quarterly, 56*(S1), S351–S370. https://doi.org/10.1002/rrq.407

Papadopoulos, T. C., Das, J. P., Parrila, R. K., & Kirby, J. R. (2003). Children at-risk for developing reading difficulties: A remediation study. *School Psychology International, 24*(3), 340–366. https://doi.org/10.1177/01430343030243006

Papadopoulos, T. C., Georgiou, G. K., & Apostolou, T. (2020). The role of distal and proximal cognitive processes in word reading fluency in Greek. In R. Alves, T. Limpo, & M. Joshi (Eds.), *Reading-writing connections: Towards integrative literacy science* (pp. 171–184). Springer.

Papadopoulos, T. C., Georgiou, G. K., & Kendeou, P. (2009). Investigating the double-deficit hypothesis in Greek: Findings from a longitudinal study. *Journal of Learning Disabilities, 42*(6), 528-547. https://doi.org/10.1177/0022219409338745

Papadopoulos, T. C., & Kendeou, P. (2010). Is there a remedy for reading difficulties? A comparison of two theory-driven programs. *Psychological Science*, *33*(6), 1299–1306.

Papadopoulos, T. C., Georgiou, G. K., Kendeou, P., & Spanoudis, G. (2009). *Standardization in Greek of the Das-Naglieri Cognitive Assessment System.* Department of Psychology, University of Cyprus. (Original work published 1997).

Papadopoulos, T. C., Kendeou, P., & Spanoudis, G. (2012). Investigating the factor structure and measurement invariance of phonological abilities in a sufficiently transparent language. *Journal of Educational Psychology, 104*(2), 321–336. https://doi.org/10.1037/a0026446

Papadopoulos, T. C., Ktisti, C., Christoforou, C., & Loizou, M. (2015). Cognitive and linguistic dynamics of reading remediation. In T. C. Papadopoulos, R. K. Parrila, & J. R. Kirby (Eds.), *Cognition, intelligence, and achievement* (pp. 311–343). Academic Press.

Papadopoulos T. C., Spanoudis G., & Kendeou, P. (2009). *Early Reading Skills-Assessment Battery (ERS-AB).* Department of Psychology: University of Cyprus.

Perry, C., Zorzi, M. & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science, 30*(3), 386–395. https://doi.org/10.1177/0956797618823540

Protopapas, A., & Vlahou, E. L. (2009). A comparative quantitative analysis of Greek orthographic transparency. *Behavior Research Methods, 41*, 991–1008. https://doi.org/10.3758/BRM.41.4.991

Richardson, U., & Lyytinen, H. (2014). The GraphoGame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments, 10*(1), 39–60. https://doi.org/10.17011/ht/urn.201405281859

Ronimus, M., Eklund, K., Westerholm, J., Ketonen, R., & Lyytinen, H. (2020). A mobile game as a support tool for children with severe difficulties in reading and spelling. *Journal of Computer Assisted Learning, 36*(6), 1011–1025. https://doi.org/10.1111/jcal.12456

Saine, N. L., Lerkkanen, M. K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at-risk for reading disability. *Child Development, 82*(3), 1013–1028. https://doi.org/10.1111/j.1467-8624.2011.01580.x

Savage, R. S., Abrami, P. C., Hipps, G., & Deault, L. (2009). A randomized controlled trial study of the ABRACADABRA reading intervention program in Grade 1. *Journal of Educational Psychology, 101*(3), 590–604. https://doi.org/10.1037/a0014700

Scammacca, N. K., Roberts, G. J., Cho, E., Williams, K. J., Roberts, G., Vaughn, S. R., & Carroll, M. (2016). A century of progress: Reading interventions for students in grades 4–12, 1914–2014. *Review of Educational Research, 86*(3), 756–800. https://doi.org/10.3102/0034654316652942

Schneider, W., Ennemoser, M., Roth, E., & Küspert, P. (1999). Kindergarten prevention of dyslexia: Does training in phonological awareness work for everybody? *Journal of Learning Disabilities, 32*(5), 429–436. https://doi.org/10.1177/002221949903200508

Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programs. *Journal of Educational Psychology, 92*(2), 284–295. https://doi.org/10.1037/0022-0663.92.2.284

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Shapiro, L. G., & Stockman, G. C. (2001). *Computer vision.* Prentice-Hall Inc.

Siegler, R. S. (2000). The rebirth of children's learning. *Child Development, 71*(1), 26–35. https://doi.org/10.1111/1467-8624.00115

Siegler, R. S., & Svetina, M. (2002). A microgenetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development, 73*(3), 793–809. https://doi.org/10.1111/1467-8624.00439

Stafford, T., & Vaci, N. (2022). Maximizing the potential of digital games for understanding skill acquisition. *Current Directions in Psychological Science, 31*(1), 49–55. https://doi.org/10.1177/09637214211057841

Suggate, S. P. (2010). Why "what" we teach depends on "when": Age as a moderator of reading intervention. *Developmental Psychology, 46*(6), 1556–1579. https://doi.org/10.1037/a0020612

Torppa, M., Parrila, R. K., Niemi, P., Lerkkanen, M. K., Poikkeus, A. M., & Nurmi, J. E., (2013). The double deficit hypothesis in the transparent Finnish orthography: A longitudinal study from kindergarten to Grade 2. *Reading and Writing, 26*, 1353–1380. https://doi.org/10.1007/s11145-012-9423-2

Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology, 105*(3), 213–231. https://doi.org/10.1016/j.jecp.2009.11.005

Vanden Bempt, F., Economou, M., Van Herck, S., Vanderauwera, J., Glatz, T., Vandermosten, M., Wouters, J., & Ghesquière, P. (2021). Digital game-based phonics instruction promotes print knowledge in pre-readers at cognitive risk for dyslexia. *Frontiers in Psychology, 12*, 720548. https://doi.org/10.3389/fpsyg.2021.720548

Van Garderen, D., Stormont, M., & Goel, N. (2012). Collaboration between general and special educators and student outcomes: A need for more research. *Psychology in the Schools, 49*(5), 483–497. https://doi.org/10.1002/pits.21610

Verwimp, C., Snellings, P., Wiers, R. W., & Tijms, J. (2023). Goal-directedness enhances letter-speech sound learning and consolidation in an unknown orthography. *Child Development, 94*(4), 836–852. https://doi.org/10.1111/cdev.13901

Verwimp, C., Tijms, J., Snellings, P., Haslbeck, J. M., & Wiers, R. W. (2021). A network approach to dyslexia: Mapping the reading network. *Development and Psychopathology, 35*(3), 1011–1025. https://doi.org/10.1017/S0954579421000365

Virinkoski, R., Lerkkanen, M. K., Holopainen, L., Eklund, K., & Aro, M. (2018). Teachers' ability to identify children at early risk for reading difficulties in grade 1. *Early Childhood Education Journal, 46*, 497–509. https://doi.org/10.1007/s10643-017-0883-5

Wechsler, D. (1992). *Intelligence Scale for Children, WISC-III-R* (Standardization in Greek by Georgas, D. D., Paraskevopoulos, I. N., Bezevegis, E. G., & Giannitsas, N. D., 1997. Ellinika Grammata).

Werner, H. (1957). *Comparative psychology of mental development* (Rev. Ed.). International Universities Press.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology, 105*(2), 249–265. https://doi.org/10.1037/a0031311

Ziegler, J. C., Castel, C., Pech-Georgel, C., George, F., Alario, F., X., & Perry, C. (2008). Developmental dyslexia and the dual-route model of reading: Simulating individual differences and subtypes. *Cognition, 107*(1), 151–178. https://doi.org/10.1016/j.cognition.2007.09.004

Zigler, E. (1969). Developmental versus difference theories of mental retardation and the problem of motivation. *American Journal of Mental Deficiency, 73*(4), 536–556.

Zigler, E., & Balla, D. (1982). Selecting outcome variables in evaluations of early childhood special education programs. *Topics in Early Childhood Special Education, 1*(4), 11–22. https://doi.org/10.1177/027112148200100406

Zigler, E., & Hodapp, R. (1986). *Understanding mental retardation.* Cambridge University Press.

RECEIVED 20.10.2022 | ACCEPTED 28.07.2023

SUPPLEMENTARY MATERIAL SM1: GROUP COMPARISONS ON READING ACHIEVEMENT

The group of 56 Greek-speaking children with RD assigned through an RCT design to a 5-week intervention focusing on phonological (GraphoGame; GG, n = 14) cognitive (PREP, n = 14) training or the two combined (PREP-to-GG or GG-to-PREP; n = 14 in each group) were compared to a chronological-age matched control (CA-C) group (n = 17) of typically developing readers. Outcomes were assessed in multiple cognitive, linguistic, reading, and orthographic measures, before (Time 1), during (Time 2), and after treatment (Time 3), as well as at a follow-up a year later (Time 4). Here, we report only the preliminary findings on the groups' performance in reading fluency. Remediation consisted of daily 30-min sessions, administered individually, during school hours by certified special education teachers or trained graduate psychology students.

Results on Word Reading Fluency Performance

Two 5 (group) x 3 (time) between-subjects analysis of covariance was performed for word reading and phonemic decoding fluency to determine the effects of the four training programs on post-intervention and follow-up reading performance. Pre-intervention reading performance was used as a covariate. In both analyses, results of the evaluation of the assumptions of normality of sampling distributions, linearity, and homogeneity of covariance were satisfactory. Tables 1 and 2 present the unadjusted and adjusted intervention means for mid-intervention, post-intervention, and follow-up reading performance with pre-intervention scores as a covariate for word reading fluency and phonemic decoding fluency.

Word reading fluency: Results showed that after adjustment for pre-intervention performance (Time 1) on word reading fluency, no significant differences of the type of treatment were found, $F(4, 67) = .43$, $p > .05$, $\eta^2 = .03$, nor was there a significant interaction between group and time ($p > .05$). However, statistically significant changes in word reading fluency were revealed over time, $F(1,67) = 35.94$, $p < .001$, $\eta^2 = .35$. Subsequent analyses revealed significant differences from mid- to post-intervention, $F(1, 67) = 27.15$, $p < .001$, $\eta^2 = .29$, from post-intervention to follow-up, $F(1, 67) = 21.32$, $p < .001$, $\eta^2 = .24$, and from mid-intervention to follow up, $F(1, 67) = 53.03$, $p < .001$, $\eta^2 = .44$. In other words, children participating in this study were learning to read irrespective of the group they belonged to.

Phonemic Decoding Fluency: Results showed that after adjustment for pre-intervention performance (Time 1) on phonemic decoding fluency, no significant differences among the groups were found, $F(4, 67) = 1.39$, $p > .05$, $\eta^2 = .08$, nor was there a significant interaction between group and time ($p > .05$). However, statistically significant changes in phonemic decoding fluency were found over time, $F(1, 67) = 46.08$, $p < .001$, $\eta^2 = .41$. Subsequent analyses showed significant differences between mid- and post-intervention scores, $F(1, 67) = 27.14$, $p < .001$, $\eta^2 = .29$, between post-intervention and follow-up scores, $F(1, 67) = 51.92$, $p < .001$, $\eta^2 = .44$, and between mid-intervention and follow up scores, $F(1, 67) = 66.60$, $p < .001$, $\eta^2 = .50$. These results indicate that all treatments groups developed decoding skills enabling them to reliably identify words that are unfamiliar to them in print.

Overall, these findings show that the development in reading ability seen in all treatment groups was comparable to the development seen in the CA-C group, after controlling for their initial score, which was far faster than expected over participants' school careers.

TABLE A1.
Unadjusted and Adjusted Group Means for Mid-Intervention, Post-Intervention, and Follow Up Reading Performance with Pre-Intervention Scores as a Covariate for Word Reading Fluency

		Mid-intervention				Post-intervention				Follow-up			
		Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
Groups	*N*	*M*	*SD*	*M*	*SE*	*M*	*SD*	*M*	*SE*	*M*	*SD*	*M*	*SE*
CA-C	17	18.71	5.50	14.44	0.86	19.82	5.95	16.53	0.95	30.35	6.59	25.13	1.65
PREP	14	13.93	3.12	13.94	0.82	15.36	3.00	15.36	0.91	27.71	9.57	27.73	1.58
GG	14	11.50	5.36	13.47	0.85	15.71	3.17	17.24	0.94	26.29	7.29	28.70	1.63
PREP+GG	14	12.36	4.22	14.27	0.84	16.00	3.53	17.48	0.94	26.21	5.91	28.56	1.62
GG+PREP	14	12.64	5.33	13.92	0.83	15.76	5.37	16.78	0.92	26.57	7.54	28.14	1.60

Note. SE = Standard Error; CA-C = chronological-age control group; PREP = PREP cognitive intervention group; GG = GraphoGame phonological intervention group; PREP+GG and GG+PREP = combined treatments sharing equal time of both interventions; reading performance was measured in fluency scores.

TABLE A2.
Unadjusted and Adjusted Group Means for Mid-Intervention, Post-Intervention, and Follow Up Reading Performance with Pre-Intervention Scores as a Covariate For Phonemic Decoding Fluency

		Mid-intervention				Post-intervention				Follow-up			
		Unadjusted		Adjusted		Unadjusted		Adjusted		Unadjusted		Adjusted	
Groups	*N*	*M*	*SD*	*M*	*SE*	*M*	*SD*	*M*	*SE*	*M*	*SD*	*M*	*SE*
CA-C	17	14.41	4.09	10.38	0.87	17.29	4.52	13.62	1.04	19.00	4.99	16.09	1.23
PREP	14	10.57	2.74	11.05	0.83	13.00	3.19	13.43	0.98	18.14	5.55	18.49	1.16
GG	14	9.14	5.45	10.15	0.83	10.57	6.03	11.49	0.99	16.71	4.12	17.44	1.17
PREP+GG	14	9.64	4.40	11.98	0.87	13.57	3.94	15.70	1.03	17.43	3.55	19.12	1.21
GG+PREP	14	9.86	5.45	10.93	0.83	12.86	5.50	13.84	0.99	17.43	5.96	18.20	1.17

Note. SE = Standard Error; CA-C = chronological-age control group; PREP = PREP cognitive intervention group; GG = GraphoGame phonological intervention group; PREP+GG and GG+PREP = combined treatments sharing equal time of both interventions; reading performance was measured in fluency scores.

SUPPLEMENTARY MATERIAL SM2: THE PASS READING ENHANCEMENT PROGRAM (PREP)

PREP cognitive remediation focuses on training proximal (e.g., phonological skills) and distal (e.g., working memory) processes related to reading. It was designed to improve selected aspects of children's information-processing skills and increase their word reading and decoding abilities (Papadopoulos et al., 2003). PREP is based on the assumption that principles transfer can be facilitated through inductive rather than deductive inference (Carlson & Das, 1997). Accordingly, the training is structured to allow inductive inference spontaneously with the internalization of principles and strategies rather than deductive rule learning. Remedial training of this kind is more likely to ensure the transfer of learned principles and produce strategies for novel situations with higher success rates (Das et al., 1995).

PREP was originally designed for Grades 3 and 4 (Das et al., 1995). Parrila et al. (2000) and Papadopoulos et al. (2003) expanded on that work by developing and implementing a version suitable for Grade 1 readers. Each task includes a global training component and a curriculum-related bridging component. The global components require applying simultaneous or successive strategies, based on the PASS theory of intelligence (e.g., Das et al., 1994), and include structured non-reading tasks. These tasks also facilitate transfer by providing the opportunity for children to internalize strategies in their way (Papadopoulos et al., 2004). The bridging tasks also include simultaneous and successive processing, which are practiced with reading-related materials (letters, syllables, and words). Each task is designed to facilitate the development of strategies such as the rehearsal, categorization, monitoring of performance, prediction, revision of prediction, sounding and sound blending, and children develop their ability to use these strategies through experience with the tasks (Papadopoulos et al., 2003).

The global and bridging components are further divided into three levels of difficulty. In addition, a system of prompts is an integral part of each global and bridging component. The prompts create a scaffolding network that supports and guides the child to ensure that tasks are completed with a minimal amount of assistance and a maximal amount of success. A record of these prompts was used as a monitoring system to determine when the material was too difficult for a child or when a child could progress to a more difficult level successfully. A criterion of 80 percent correct responses was required in this study before a child could proceed to the next difficulty level. If the criterion was not met, alternate tasks at the same difficulty level were used to provide the additional training required. The following eight tasks were selected for use with Grade 1 participants and presented to the children in the order listed: Window Sequencing, Connecting Letters, Joining Shapes, Matrices, Related Memory Set, Transportation Matrices, Tracking, and Shape Design. For a detailed description of the program, see Papadopoulos et al. (2003).

Reviews on the efficacy of PREP can be found in several papers (e.g., Das et al., 2008; Kearns & Fuchs, 2013; Mahapatra et al., 2010; Papadopoulos et al., 2003; Papadopoulos, 2013). Generally, PREP has produced positive results in terms of cognitive performance and reading ability, in non-transparent (e.g., Carlson & Das, 1997; Das et al., 1995; Papadopoulos et al., 2003; Parrila et al., 2000) and transparent orthographies (e.g., Papadopoulos et al., 2004; Papadopoulos & Kendeou, 2010); with children at-risk for reading difficulties in Kindergarten (e.g., Papadopoulos et al., 2004), poor readers in Grades 1 and 2 (e.g., Papadopoulos et al., 2003; Parrila et al., 2000), Grades 3 and 4 (e.g., Das et al., 1995; Das et al., 2008) or Grades 5 and 6 (Boden & Kirby, 1995); with First-Nations children in Canada (e.g., Das et al., 2008; Hayward et al., 2007) or poor readers learning English as a second language (Mahapatra et al., 2010); in small groups (Carlson, 1996; Carlson & Das, 1997; Papadopoulos et al., 2003) or on an intensive one-to-one basis (Papadopoulos et al., 2004; Papadopoulos & Kendeou, 2010); in comparison with other experimental groups receiving different treatment programs, such as phonics-based (e.g., Das et al., 2008), meaning-based (Papadopoulos et al., 2003) or neuropsychologically-based programs (Papadopoulos & Kendeou, 2010); and with designs including a follow-up component allowing examination of the long-term efficacy of PREP (Papadopoulos et al., 2003, 2004; Papadopoulos & Kendeou, 2010).

REFERENCES

Boden, C., & Kirby, J. R. (1995). Successive processing, phonological coding, and the remediation of reading. *Journal of Cognitive Education, 4*, 19–32.

Carlson, J., & Das, J. P. (1997). A process approach to remediating word-decoding deficiencies in chapter 1 children. *Learning Disability Quarterly, 20*(2), 93–102. https://doi.org/10.2307/1511217

Das, J. P., Hayward, D., Georgiou, G., Janzen, T., & Boora, N. (2008). Comparing the effectiveness of two reading intervention programs for children with reading disabilities. *Journal of Cognitive Education and Psychology, 7*(2), 199–222. https://doi.org/10.1891/194589508787381836

Das, J. P., Mishra, R.K., & Pool, J.E. (1995). An experiment on cognitive remediation or word-reading difficulty. *Journal of Learning Disabilities, 28*(2), 66–79. https://doi.org/10.1177/002221949502800201

Das, J. P., Naglieri, J. A., & Kirby, J. R. (1994). *Assessment of cognitive processes: The PASS theory of intelligence*. Allyn & Bacon.

Hayward, D., Das, J. P., & Janzen, T. (2007). Innovative programs for improvement in reading through cognitive enhancement: A remediation study of Canadian First Nations children. *Journal of Learning Disabilities, 40*(5), 443–457. https://doi.org/10.1177/00222194070400050801

Mahapatra, S., Das, J. P., Stack-Cutler, H., & Parrila, R. (2010). Remediating reading comprehension difficulties: A cognitive processing approach. *Reading Psychology, 31*(5), 428–453. https://doi.org/10.1080/02702710903054915

Papadopoulos, T. C. (2013). PASS theory of intelligence in Greek: A review. *Preschool and Primary Education, 1*(1), 41-66.

Papadopoulos, T. C., Charalambous, A., Kanari, A., & Loizou, M. (2004). Kindergarten intervention for dyslexia: The PREP remediation in Greek. *European Journal of Psychology of Education, 19*, 79–105. https://doi.org/10.1007/BF03173238

Papadopoulos, T. C., Das, J. P., Parrila, R. K., & Kirby, J. R. (2003). Children at-risk for developing reading difficulties: A remediation study. *School Psychology International, 24*(3), 340–366. https://doi.org/10.1177/01430343030243006

Papadopoulos, T. C., & Kendeou, P. (2010). Is there a remedy for reading difficulties? A comparison of two theory-driven programs. *Psychological Science, 33*(6), 1299–1306.

Parrila, R. K., Das, J. P., Kendrick, M. E., Papadopoulos, T. C., & Kirby, J. R. (2000). Cognitive reading remediation for at-risk children in grade 1. *Journal of Cognitive Education and Psychology, 1*, 114–139.