Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models

Johannes C Ziegler, Conrad Perry, Marco Zorzi

Current Directions in Psychological Science
1–8

DOI: 10.1177/0963721420915873
www.psychologicalscience.org/CDPS

Learning to Read and Dyslexia: From Theory to Intervention Through Personalized Computational Models

Johannes C. Ziegler[1], **Conrad Perry**[2], and **Marco Zorzi**[3,4]

[1]Laboratoire de Psychologie Cognitive, Centre National de la Recherche Scientifique (CNRS), Aix-Marseille University; [2]Faculty of Health, Arts and Design, Swinburne University of Technology; [3]Department of General Psychology, University of Padova; and [4]IRRCS San Camillo Hospital, Venice-Lido, Italy

Abstract

How do children learn to read? How do deficits in various components of the reading network affect learning outcomes? How does remediating one or several components change reading performance? In this article, we summarize what is known about learning to read and how this can be formalized in a developmentally plausible computational model of reading acquisition. The model is used to understand normal and impaired reading development (dyslexia). In particular, we show that it is possible to simulate individual learning trajectories and intervention outcomes on the basis of three component skills: orthography, phonology, and vocabulary. We therefore advocate a multifactorial computational approach to understanding reading that has practical implications for dyslexia and intervention.

Keywords

computational modeling, dyslexia, literacy, reading, reading development

Literacy is one of the greatest achievements of both human civilization and the human mind. Through reading and writing, "we can shape events in each other's brains . . . bridging gaps of time, space, and acquaintanceship" (Pinker, 1994, pp. 1–2). One of the biggest challenges in cognitive and developmental sciences is to understand the complex machinery that is behind this extraordinary ability. As stated by Edmund Burke Huey (1908/1968), who was one of the pioneers of experimental psychology, "to completely analyse what we do when we read, would almost be the acme of the psychologist's achievements, for it would be to describe many of the most intricate workings of the human mind" (p. 6).

Ever since the first connectionist model of letter and word perception (McClelland & Rumelhart, 1981), computational models of reading have played an important role in our understanding of the "intricate workings" that make it possible to read and comprehend written words. Computational models of reading are computer programs that specify the ingredients of the reading process and implement the units and computations that are necessary to transform visual information into linguistic information (phonemes, stress, words, meaning). Once a model is implemented, it can be used to simulate real reading performance in terms of reading latencies (how long it takes to compute the pronunciation of a word or nonword) and reading accuracy (whether the output of the model is correct or, with nonwords, the output is the same as what people produce).

Computational models offer far more than black-box predictions of reading behavior. They allow us to better understand reading impairments, such as developmental dyslexia, a neurodevelopmental disorder that affects between 5% and 10% of the population and is characterized by a failure to automatize word-recognition skills despite normal intelligence and appropriate schooling (Peterson & Pennington, 2015). Indeed, model components can be "impaired" in very specific and focal ways, and the consequences of these impairments can

Corresponding Author:
Johannes C. Ziegler, Aix-Marseille University, Laboratoire de Psychologie Cognitive, 3 Place Victor Hugo, 13003 Marseille, France
E-mail: Johannes.Ziegler@univ-amu.fr

be analyzed through computer simulations (Harm & Seidenberg, 1999; Perry, Zorzi, & Ziegler, 2019; Woollams, 2014; Ziegler et al., 2008; Ziegler, Perry, & Zorzi, 2014). This allows one to establish causal relations between deficits in components of the reading network and reading outcomes.

The first computational model of reading, the interactive activation model (McClelland & Rumelhart, 1981), did not have any phonological components and did not learn—all connections were hardwired. Subsequent models implemented phonological processes in order to tackle reading aloud. Although some models used connectionist learning mechanisms to acquire the mapping between letter strings and their corresponding sounds (Harm & Seidenberg, 1999; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989; Zorzi, Houghton, & Butterworth, 1998), it is probably fair to say that none of these models simulated reading development in a developmentally plausible way (for a discussion, see Ziegler et al., 2014). For example, the very influential model of reading development by Harm and Seidenberg (1999) required 10 million supervised learning trials to learn a few thousand monosyllabic words. Such a training regimen is akin to learning all words through direct instruction, which is very different from the way children learn to read (see below; Share, 1995).

In the present article, we first introduce the major theory of how children learn to read. We then present a developmentally plausible computational reading model that is an implementation of this theory (Ziegler et al., 2014). We finally present our recent attempt to personalize computational models to simulate normal and impaired reading development for individual children (Perry et al., 2019). This approach has important implications for our understanding of dyslexia and allows us to predict intervention outcomes for individual children.

How Do Children Learn to Read?

Although the ultimate goal of learning to read is to comprehend what we read (Castles, Rastle, & Nation, 2018), the initial stages of learning to read are all about cracking the orthographic code. That is, writing systems code spoken language, and to some extent meaning, through morphology and etymology. Children have to understand how this code works in their language. In alphabetic writing systems, children have to learn how letters or groups of letters (graphemes) map onto their corresponding phonemes. In some alphabetic writing systems, such as English, there is a trade-off between the extent to which spellings prioritize the consistent spelling of morphemes over the consistent spelling of phonemes (see Bowers & Bowers, 2018). This creates inconsistencies at the phoneme level, which are a major hurdle for learning to read (Ziegler & Goswami, 2006). Yet in most alphabetic writing systems, including English, learning instruction starts through the explicit teaching of letter–sound or grapheme–phoneme rules. Children can then use these rules or associations to decode words they have heard but never seen before. This process is referred to as *phonological decoding* (Share, 1995). Phonological decoding is at the heart of reading acquisition in all alphabetic writing systems because it provides an extremely parsimonious and straightforward way to retrieve the spoken form and therefore the meaning of the thousands of words children have stored in their phonological lexicon (Ziegler & Goswami, 2005).

In fact, such a theory predicts that inconsistency in the mapping between letters and sounds (i.e., when the same letter has multiple pronunciations) should slow down the initial stages of reading acquisition (decoding, reading aloud, word identification, spelling). As can be seen in Figure 1, comparisons of the rate of single-word reading aloud across different languages show that this is indeed the case. The more inconsistent a writing system is, the longer it takes children to acquire basic reading skills. Thus, the difficulty with which basic grapheme–phoneme correspondences can be taught and learned predicts the speed of reading acquisition in different languages.

Once children have learned basic decoding skills, explicit teaching is largely replaced by self-teaching (Share, 1995). That is, children start to decode words autonomously. If they find a word in the phonological lexicon that fits the context, they create an orthographic representation for the decoded and retrieved word. Every successfully decoded word provides children with an opportunity to acquire the word-specific orthographic information that is the foundation of skilled word recognition. Thus, phonological decoding provides a powerful self-teaching device because the explicit learning of a small set of spelling-sound correspondences allows children to decode an increasingly large number of words or, as Share (1995) puts it, "minimum number of rules, maximum generative power" (p. 156). We refer to this learning loop as the *phonological-decoding self-teaching theory* (Ziegler et al., 2014).

A Computational Model of Reading Development

We have developed a computational model that implements the core principles of the phonological decoding and self-teaching theory (Ziegler et al., 2014) within the processing architecture of the connectionist dual-process

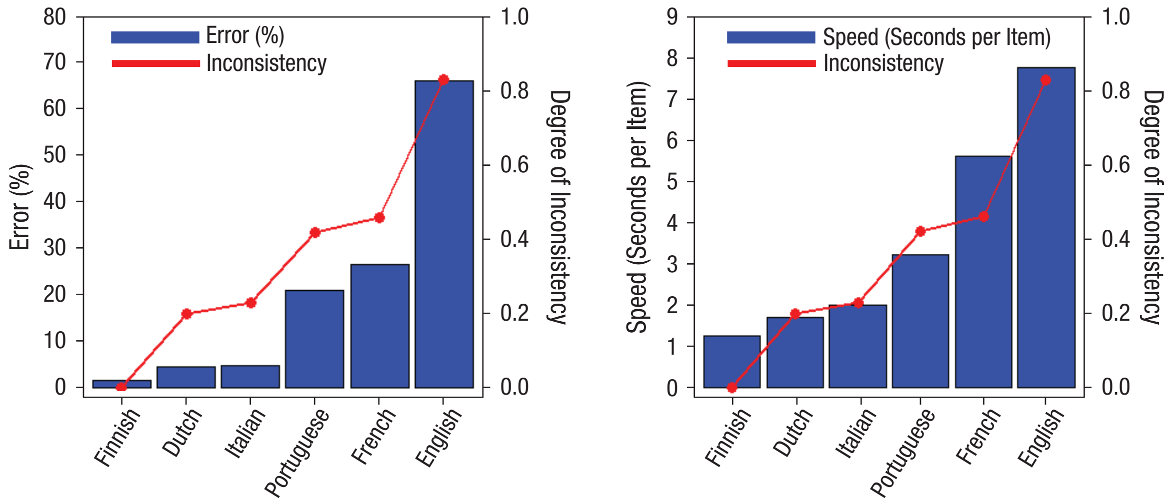

Fig. 1. Reading performance at the end of Grade 1 in six European countries and a measure of the inconsistency of a writing system in terms of letter–sound correspondences (i.e., onset entropy; see Ziegler et al., 2010). Reading performance is indexed by both the percentage of errors (left) and speed of reading (right). Reading data were taken from Seymour, Aro, and Erskine (2003).

model, which is a leading computational model of skilled reading aloud (Perry, Ziegler, & Zorzi, 2007, 2010). Figure 2 illustrates the developmental connectionist dual-process model. In a nutshell, we first implemented a decoding network that was pretrained on a small set of grapheme–phoneme correspondences (e.g., b → /b/, p → /p/). We chose to implement this process in a simple two-layer associative network that takes graphemes (letters or simple letter combinations, e.g., TH, OO, EA) as input and uses phonemes as output. During this stage, learning is supervised. We believe that this process mirrors, to a large extent, the explicit teaching of grapheme–phoneme correspondences (Department for Education, 2014; Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012). Note that the model would also acquire grapheme–phoneme correspondences even if the teaching were based on syllables, morphemes, or whole words, although at the expense of a slower learning rate (see simulations in Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004).

From there on, the model enters the self-teaching phase. Thus, the model is presented with several thousand written words to be learned (i.e., a real-sized child lexicon). The initially rudimentary decoding network generates a phoneme sequence that potentially activates entries in the phonological lexicon (i.e., phonological

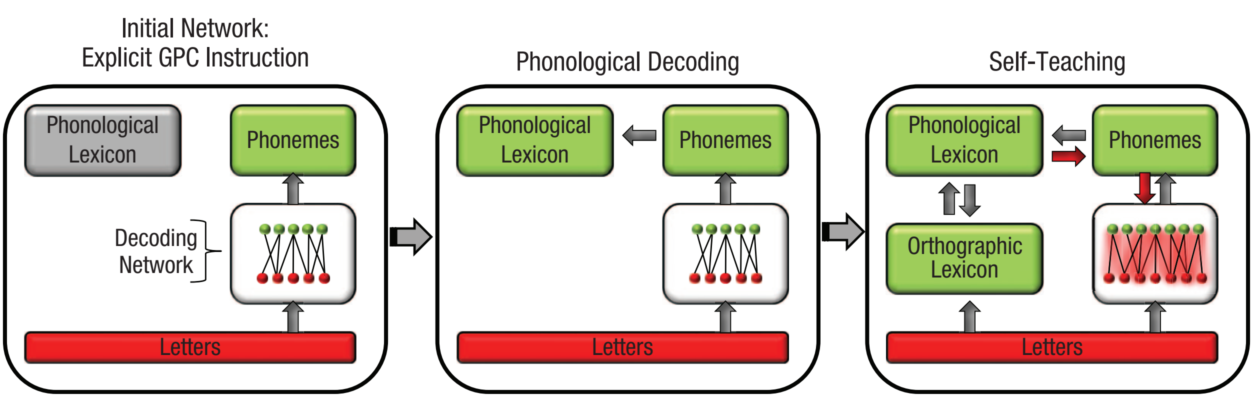

Fig. 2. Implementation of the phonological-decoding self-teaching hypothesis in the developmental connectionist dual-process model. After initial explicit teaching on a small set of grapheme–phoneme correspondences (GPCs), the decoding network is able to decode words that have a preexisting representation in the phonological lexicon but no orthographic representation. If the decoding mechanism activates a word in the phonological lexicon, an orthographic entry is created, and the phonology is used as an internally generated teaching signal (red arrows) to refine and strengthen letter–sound connections, thereby improving the efficiency of the decoding network. Figure was adapted from Ziegler, Perry, and Zorzi (2014).

candidates). Critically, we assume that context, meaning, and morphosyntactic information are used to select the correct candidate among the activated phonological competitors (for a review, see Tunmer & Chapman, 2004). Although the specification of these mechanisms was beyond the scope of our model, there is substantial evidence to suggest that children can use their oral vocabulary to correct a partial decoding attempt (Tunmer & Chapman, 2012) and can correct imperfect decoding attempts by reference to the known pronunciation and meaning of a word (Dyson, Best, Solity, & Hulme, 2017). Importantly, the internally generated phonological representation is then used as a teaching signal (i.e., self-teaching) to improve the decoding network. This leads to the learning of a richer, more complex, and context-sensitive set of spelling–sound associations.

In a first set of simulations (Ziegler et al., 2014), we assessed whether such a simple but developmentally plausible learning mechanism can work for a language such as English, which is known for its rather difficult letter–sound mapping (see Fig. 1). We pretrained the two-layer associative network with 65 correspondences and then let it run through cycles of self-teaching with several thousands of words. The results showed that it successfully learned more than 80% of the words through decoding and self-teaching.[1] In addition, very rapidly in the course of learning, the most active item in the phonological lexicon tended to be the correct word. This is the reason why self-teaching works so well. This simulation is a proof of concept for the claim that phonological decoding and self-teaching provide a powerful bootstrapping mechanism (Share, 1995) that allows the beginning reader to start small (i.e., with a small set of explicitly taught letter–sound correspondences) and to build on this knowledge to self-learn the majority of words through a simple decoding mechanism that gets more efficient with every successfully decoded word (see also Hutzler et al., 2004; Pritchard, Coltheart, Marinus, & Castles, 2018).

How do we learn the remaining 20%, which are too irregular to be learned through decoding, such as *yacht*, *aisle*, or *choir*? To simulate irregular word learning and reading, we had to add a mechanism that gets irregular words into the orthographic lexicon (Perry et al., 2019). The basic idea is that children use a variety of strategies to do this. These include direct instruction (Department for Education, 2014), the teaching of a small number of sight words (Shapiro & Solity, 2016), the use of context and partial decoding (Share, 1995; Tunmer & Chapman, 2004), and morphological information and etymology (Bowers & Bowers, 2018). It was not possible to faithfully implement these processes, so each time a phonological decoding attempt was unsuccessful, we allowed for the possibility that a word might enter the orthographic lexicon. To do this, we used a computational shortcut in which high-frequency words had a greater chance of being learned by other means than low-frequency words. The results showed that the combination of these strategies, which can vary from one child to another, allowed the model to learn as well as children do.

Modeling Dyslexia: The Multideficit-Component Approach

Our model has five critical components: letters, phonemes, a phonological lexicon, a decoding network, and an orthographic lexicon (see Fig. 2). We have previously shown that one can impair these components and investigate the consequences of such impairments for the learning-to-read process (Ziegler et al., 2014). For example, it is well known that many children with dyslexia in different countries have poor phoneme-awareness skills (Landerl et al., 2013; Ziegler et al., 2010). We can assume that children with poor phoneme awareness have problems mapping letters onto phonemes, a process that is modeled by the two-layer associative decoding network. One can simulate such a deficit through the switching of phonetically similar phonemes during learning, which is a reasonable assumption because children with dyslexia tend to confuse phonetically similar phonemes (Ziegler, Pech-Georgel, George, & Lorenzi, 2009). Thus, the core idea of the multideficit-component approach was to estimate the efficiency of the component processes through component tasks and then create personalized models for each child to simulate his or her learning trajectory.

In the simulations reported by Perry et al. (2019), we selected three component tasks from one of the biggest dyslexia samples, which contained reading-aloud data (on regular words, irregular words, and nonwords) as well as performance measures in other nonreading tasks for 622 English-speaking children, including 388 children with dyslexia (Peterson, Pennington, & Olson, 2012). We specified how performance on these component tasks could map onto the component processes in the model. Orthographic choice was taken as a measure for processing efficiency in the orthographic lexicon, phoneme deletion was taken as a measure for the efficiency of activating phonemes correctly, and vocabulary score was taken as a measure of the size of a child's phonological lexicon. We used performance on these three tasks to create individual models, one for each child, in which the parameterization of the models' components and processes was changed using a simple linear function based on the child's performance on the three component tasks.

A full learning simulation was performed for each individual model, and its performance after learning was assessed by presenting the same words and nonwords used by Peterson et al. (2012). This allowed a

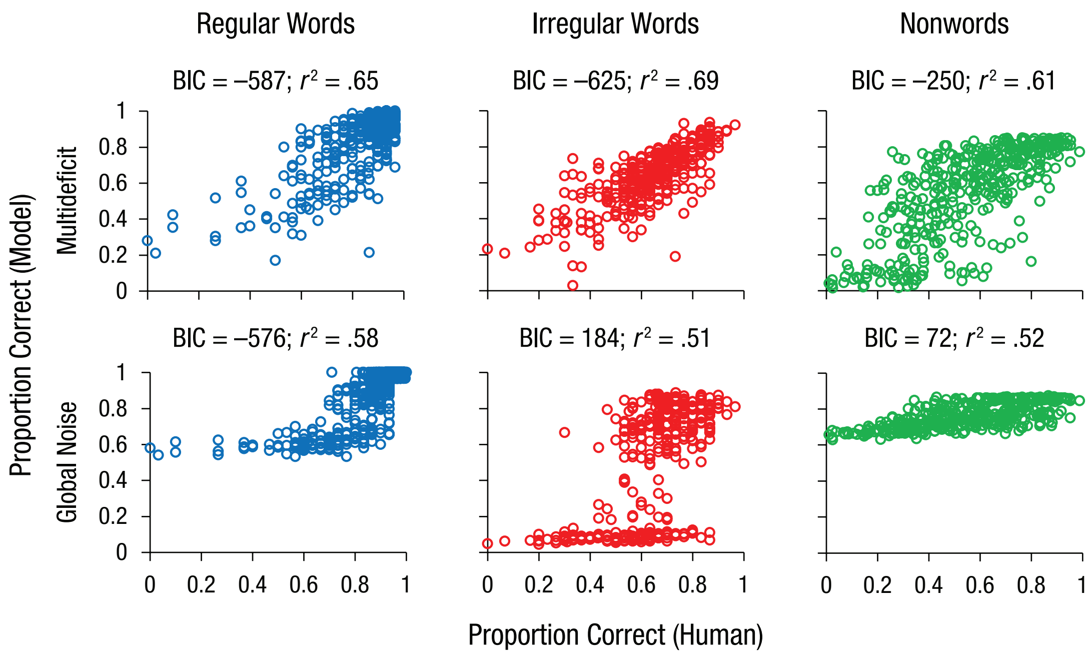

Fig. 3. Correlation between predicted reading performance and actual reading performance of 388 children with dyslexia, separately for regular words, irregular words, and nonwords. For predicted reading, the top row shows results from the multideficit model, and the bottom row shows results from a noisy computation model. A larger negative Bayesian information criterion (BIC) is an index of better fit, and a BIC difference of 10 corresponds to a posterior odds ratio of approximately 150:1. Children's reading data were taken from Peterson, Pennington, and Olson (2012), and model results were taken from Perry, Zorzi, and Ziegler (2019).

direct comparison between learning outcomes in the simulation and actual reading performance of the child that the simulation was meant to capture. Note that we investigated only accuracy (whether a word or nonword has been pronounced correctly) and not fluency. This was done because fluency for individual words was not available in Peterson et al.'s data. Also, because of the high inconsistency of English (see Fig. 1), accuracy rather than fluency tends to be a more sensitive measure in beginning and dyslexic readers. We obtained a striking correlation between the model predictions and the real learning outcomes (see Fig. 3, top row). Note that the model read exactly the same words as the 622 children from the Peterson et al. (2012) study.

We then compared the model with a number of alternative models: a phonological-deficit model, which assumed deficits in activating correct phonemes (i.e., deficits in phonological awareness, phoneme discrimination, and categorical perception of phonemes); a visual-deficit model, which assumed impoverished orthographic processing due to poor letter-position coding (e.g., letter reversals); and a global-noise model, which assumed general processing inefficiency (set as a function of the child's overall level of performance) due to noisy computations (Hancock, Pugh, & Hoeft, 2017). The phonological-deficit model was best, followed by the visual-deficit and the noisy-computation models, but none of them reached the performance of the multideficit model. For illustrative purposes, the results of the noisy computation model are presented in the bottom row of Figure 3. As can be seen, the model fits are much poorer, which suggests that one needs to take into account the specific deficits on the three component tasks. Just adding noise to these components to fit the overall reading level does not work.

Predicting Intervention Outcomes

The strong correlations between predicted and actual reading performance on different types of words make it possible to use the model as a tool to predict how remediating one component would change reading performance on the different types of words. To do this, we set up a three-dimensional deficit space in which each component task had one dimension (see Perry et al., 2019, for details). This allowed each child to be represented as a point in this space (see Fig. 4). We then obtained simulations that sampled the entire space, which is akin to moving each point along the three dimensions. This makes it possible to predict how reading performance on words and nonwords changes as a function of improving component skills through intervention.

Such simulations of potential intervention outcomes are of great theoretical interest because in a model, in contrast to real life, one can test all possible changes and

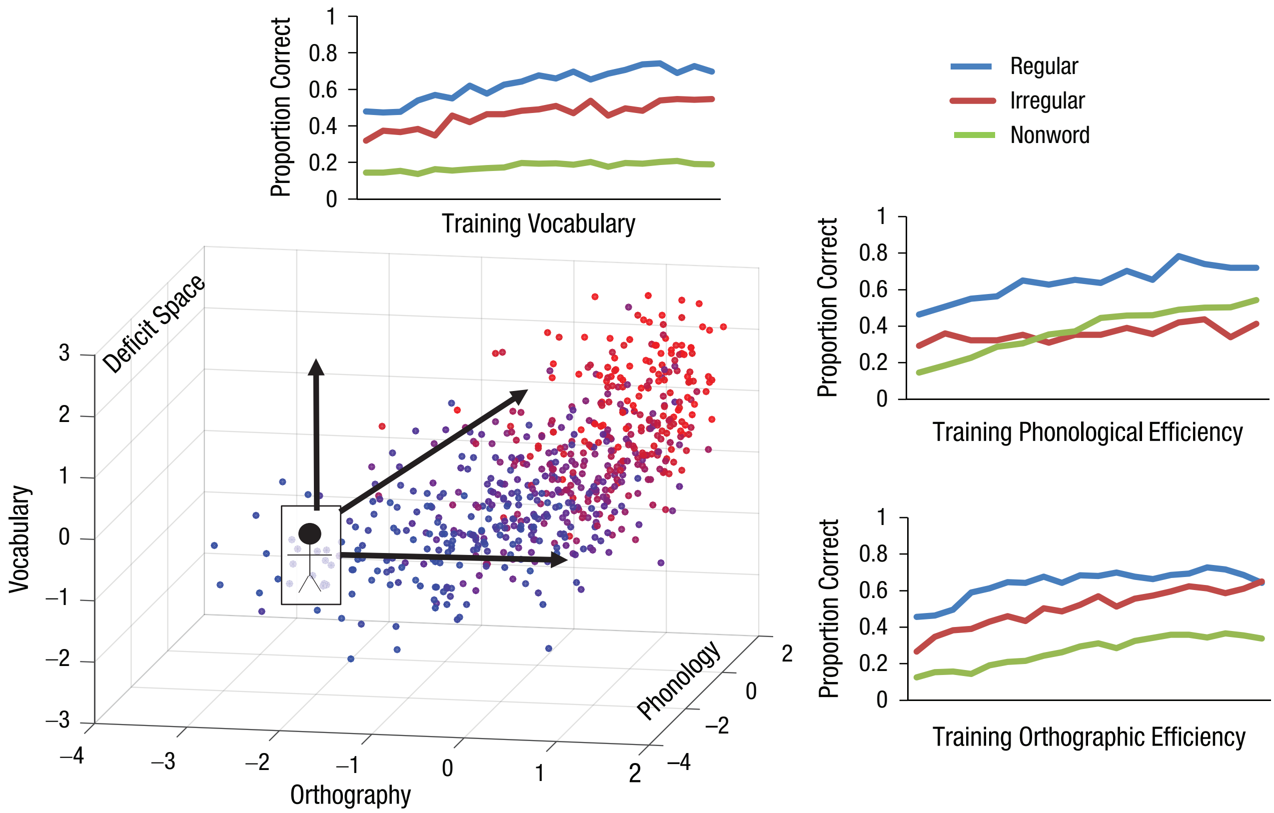

Fig. 4. Predicting intervention outcomes through personalized computer simulations. Each child can be represented as a point in a three-dimensional deficit/performance space (bottom left). The color of the dots represents the average reading-aloud level of a child on a blue-red scale, where dark blue indicates high performance and bright red low performance. Depending on the child's position in that space, the model predicts the individual learning outcomes (proportion of correct responses) for regular-word, irregular-word, and nonword scores. By systematically moving the point along all possible dimensions (i.e., training the components), the model can predict intervention outcomes as a function of training orthographic efficiency, phonological efficiency, and vocabulary (shown separately in each graph). The data are for an individual child with a mixed dyslexia profile.

their direct effects on reading. Further, the link between changes and outcomes is necessarily causal. However, we do not want to give the false impression that these components are independent and should be trained independently. We know from an accumulating number of studies on evidence-based approaches to early reading intervention that the highly structured, early, and intense teaching of decoding skills, metaphonological awareness, and fluency are key ingredients of effective intervention programs (Savage & Carless, 2005; Vellutino et al., 1996). From a developmental standpoint, it is actually highly questionable whether one can train orthographic efficiency without improving phonological decoding skills first. At the same time, there is an interplay between decoding and oral language because vocabulary and knowledge of word meanings are important to correct incorrect or partial decoding attempts, and these skills can be trained efficiently (Dyson et al., 2017). Indeed, in a longitudinal study, Nation and Snowling (2004) showed that three measures of nonphonological oral language tapping vocabulary knowledge and listening comprehension predicted individual differences in reading comprehension, word recognition, and irregular word reading a few years later, which suggests that the developing reading system is, of course, part of a wider language system.

Conclusions

Children come to the task of learning to read with large interindividual differences in vocabulary, phonology, and orthographic skills. Our work shows that taking into account the starting point of each child in this multidimensional space allows one to predict individual learning outcomes through large-scale personalized computational models. This is highly relevant for our understanding of dyslexia. It has become increasingly clear over the years that there is not a single developmental trajectory (Pennington, 2006), and the idea that children with dyslexia have deficits on either the orthographic route (surface dyslexia) or the phonological route (phonological dyslexia) has been shown to be false (Sprenger-Charolles, Siegel, Jimenez, & Ziegler, 2011; Ziegler et al., 2008). Instead, we show that knowing the efficiency of the component skills of the reading network is helpful for our understanding of reading difficulties. Alternative models that affect only one component or randomly affect all components (noisy computation) do not fare as well and cannot predict interindividual differences. Importantly, although the human data are only correlational, the relation between deficits and outcomes in the model is a causal one and can be used to derive empirical predictions. Finally, such personalized models can be used to explore how changing the efficiency of one component through intervention is likely to change reading performance for an individual child. We believe that personalized computational modeling will play an important role not only in the early detection of dyslexia but also in the context of evidence-based interventions.

Recommended Reading

Castles, A., Rastle, K., & Nation, K. (2018). (See References). An excellent and timely review on how children learn to read and how reading skills should be taught.

Harm, M. W., & Seidenberg, M. S. (1999). (See References). A historical classic on a computational model of reading acquisition and dyslexia.

Pennington, B. F. (2006). (See References). An important article convincingly demonstrating that developmental disorders, such as dyslexia, are probabilistic and multifactorial, whereas the prevailing cognitive models have often been deterministic and focused on single cognitive causes.

Share, D. L. (1995). (See References). The foundational article on the importance of phonological decoding and self-teaching for reading acquisition.

Transparency

Action Editor: Randall W. Engle
Editor: Randall W. Engle
Declaration of Conflicting Interests
The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.
Funding
This research was supported by grants from the Australian Research Council (DP170101857), the European Research Council (210922-GENMOD), the Agence National de la Recherche (ANR-13-APPR-0003 GraphoGame), the Institute of Convergence on Language, Cognition and the Brain (ANR-16-CONV-0002), the Aix-Marseille University Initiative d'Excellence (A*MIDEX; ANR-11-IDEX-0001-02), and the Cariparo Foundation (Excellence Grants 2017).

ORCID iDs

Johannes C. Ziegler https://orcid.org/0000-0002-2061-5729
Marco Zorzi https://orcid.org/0000-0002-4651-6390

Acknowledgments

This work was performed in part on the swinSTAR supercomputer at Swinburne University of Technology.

Note

1. In later simulations (Perry et al., 2019), orthographic learning was made a probabilistic process. That is, a word has a certain chance of entering the orthographic lexicon depending on an individual parameter reflecting the child's orthographic-learning potential (e.g., a standardized measure of orthographic ability).

References

Bowers, J. S., & Bowers, P. N. (2018). Progress in reading instruction requires a better understanding of the English spelling system. *Current Directions in Psychological Science*, *27*, 407–412.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*, 5–51.

Department for Education. (2014). *Statutory guidance: National curriculum in England: Framework for key stages 1 to 4*. Retrieved from https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4

Dyson, H., Best, W., Solity, J., & Hulme, C. (2017). Training mispronunciation correction and word meanings improves children's ability to learn to read words. *Scientific Studies of Reading*, *21*, 392–407.

Hancock, R., Pugh, K. R., & Hoeft, F. (2017). Neural noise hypothesis of developmental dyslexia. *Trends in Cognitive Sciences*, *21*, 434–448.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review*, *106*, 491–528.

Huey, E. B. (1968). *The psychology and pedagogy of reading*. New York, NY: Macmillan. (Original work published 1908)

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, *23*, 572–577.

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, *91*, 273–296.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppanen, P. H., Lohvansuu, K., . . . Schulte-Korne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, *54*, 686–694.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: 1. An account of basic findings. *Psychological Review*, *88*, 375–407.

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*, 342–356.

Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, *101*, 385–413.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*, 273–315.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the connectionist dual process (CDP++) model. *Cognitive Psychology*, *61*, 106–151.

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological Science*, *30*, 386–395.

Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology*, *11*, 283–307.

Peterson, R. L., Pennington, B. F., & Olson, R. K. (2012). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition*, *126*, 20–38.

Pinker, S. (1994). *The language instinct*. New York, NY: Harper Perennial Modern Classics.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science*, *42*, 722–770.

Savage, R., & Carless, S. (2005). Learning support assistants can deliver effective reading interventions for 'at-risk' children. *Educational Research*, *47*, 45–61.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.

Shapiro, L. R., & Solity, J. (2016). Differing effects of two synthetic phonics programmes on early reading development. *British Journal of Educational Psychology*, *86*, 182–203.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*, 151–218.

Sprenger-Charolles, L., Siegel, L. S., Jimenez, J. E., & Ziegler, J. C. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Studies of Reading*, *15*, 498–521.

Tunmer, W. E., & Chapman, J. W. (2004). The use of context in learning to read. In T. Nunes & P. Bryant (Eds.), *Handbook of children's literacy* (pp. 199–212). Dordrecht, The Netherlands: Springer.

Tunmer, W. E., & Chapman, J. W. (2012). Does set for variability mediate the influence of vocabulary knowledge on the development of word recognition skills? *Scientific Studies of Reading*, *16*, 122–140.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology*, *88*, 601–638.

Woollams, A. M. (2014). Connectionist neuropsychology: Uncovering ultimate causes of acquired dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), Article 20120398. doi:10.1098/rstb.2012.0397

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*, 551–559.

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*, 151–178.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*, 3–29.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, *9*, 429–436.

Ziegler, J. C., Pech-Georgel, C., George, F., & Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Developmental Science*, *12*, 732–745.

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), Article 20120397. doi:10.1098/rstb.2012.0397

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception & Performance*, *24*, 1131–1161.