On the cusp of predictability: Disruption in the typical association between letter and word identification at critical thresholds of RAN and phonological skills

Ola Ozernov-Palchik a,b,*,1, Georgios D. Sideridis c,1, Elizabeth S. Norton d, Sara D. Beach a,c, Maryanne Wolf e, John D.E. Gabrieli a, Nadine Gaab b,c

a McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA
b Harvard Graduate School of Education, Harvard University, Cambridge, MA, USA
c Harvard Medical School, Boston, MA, USA
d Department of Communication Sciences and Disorders, Department of Medical Social Sciences, Northwestern University, Evanston, IL, USA
e School of Education & Information Studies, University of California, Los Angeles, California, USA

ARTICLE INFO

Keywords:
Reading
Rapid automatized naming
Phonological processing
Catastrophe theory
Early identification

ABSTRACT

Early identification is important for improving reading outcomes for children at risk for reading difficulties, but current methods tend to perform sub-optimally in identifying true risk. One possible reason is that whereas most prediction models assume linear relationships among risk and reading outcome measures, qualitatively different patterns of association may exist among the measures at different skill levels. We implemented dynamic non-linear modeling to test this possibility in two distinct samples of children: 1) 225 pre-kindergarten and kindergarten with concurrent data and 2) 104 children with longitudinal (pre-)kindergarten and second-grade reading scores. Cusp catastrophe modeling was used to evaluate the moderating effects of rapid automatized naming (RAN) and phonological processing, two foundational pre-reading skills, on the concurrent and longitudinal relationship between letter identification and word reading. We further tested whether RAN and phonological processing have independent non-linear effects on the letter-word reading relationship above and beyond that of a single skill. Deficits in RAN and phonological processing beyond a critical level were associated with non-linear changes in the prediction of word reading from letter knowledge, both concurrently and longitudinally, fully supporting the cusp model over the competing models. These findings demonstrate the importance of implementing non-linear models for predicting risk for reading difficulties. There was no evidence for the interactive effects of RAN and phonological processing on reading. Instead, in accordance with the basic tenets of the double-deficit hypothesis, current results suggest that the constructs represent two salient but separable causes of reading impairment, even at the earliest stages of reading ability. These findings suggest that models predicting which at-risk children will develop poor reading must diverge from assumptions of relationships observed in typical reading.

1. Introduction

Early identification is important for improving intervention outcomes for children at risk for developing reading difficulties (Catts et al., 2015; Lovett et al., 2017; Wanzek & Vaughn, 2007). There are significant challenges, however, to identifying these at-risk children reliably (Catts & Petscher, 2018; Poulsen, 2018). Although there are multiple widely accepted pre-reading indicators of later reading skills, existing models tend to have relatively poor accuracy in predicting poor readers, despite their high accuracy in predicting good readers (Badian, 1993; McCardle et al., 2001). For example, across studies aiming to predict reading outcomes from kindergarten and first grade, high sensitivity values yielded high false-positive rates (e.g., Catts et al., 2009; Fletcher et al., 2020; Petscher & Koon, 2020; Satz & Friel, 1978; Thompson et al., 2015). Here, we examined whether the challenges of predicting poor reading outcomes stem from the linear assumptions most studies make...
about the relations among early literacy constructs across skill levels. We tested these assumptions of linearity by applying a cusp catastrophe model, a dynamic non-linear approach, to concurrent and longitudinal associations among early literacy measures. We examined whether a deficit in two subcomponents of reading, namely phonological processing and rapid automatized naming (RAN), might moderate the relation between letter knowledge and single-word reading skills in a large longitudinal sample of English-speaking children in pre-kindergarten and kindergarten.

1.1. Early reading development

Early reading relies on a series of developmental milestones, from acquisition of oral language and identification of native phonemes, to letter knowledge, to word identification (Blaiklock, 2004; Leppanen et al., 2008; Walsh et al., 1988). Phonological processing – the ability to remember, identify, and manipulate the sounds of oral language – is foundational to successful reading acquisition (Schatschneider et al., 2004; Wagner & Torgesen, 1987; Wagner et al., 1997). Letter-sound knowledge encompasses the ability to map these sounds onto their visual representations and to combine these sounds into words. Here, we use the general term “letter knowledge” to refer to letter-name and letter-sound knowledge, recognizing that the two show similar correlations with word reading in kindergarten (Blaiklock, 2004; Kim et al., 2010) and beyond (Georgiou et al., 2012; Leppanen et al., 2008; Pennington & Lefly, 2001; Share et al., 1984). Differences in word reading include both accuracy and fluency. Word identification accuracy is central across stages of reading development ( Ehri, 1995). Fluency becomes an important contributor to explaining differences in reading only after a certain level of accuracy has been achieved (Juul et al., 2014; Kameenui & Simmons, 2001). Therefore, because the focus of this paper is on the earliest stages of reading development in kindergarten, the outcome of interest is word identification accuracy rather than fluency.

Another key skill for reading acquisition is rapid automatized naming (RAN), the ability to quickly and accurately name an array of familiar items (Dencelka & Cutting, 1999; Norton & Wolf, 2012). RAN is thought to reflect the automaticity and efficiency with which an individual can retrieve information (e.g., high frequency object names or highly familiar letter names) and integrate across the multiple processes necessary for reading and naming aloud (Norton & Wolf, 2012). The multi-componential nature of RAN has been demonstrated across orthographies, with studies finding both a direct association between RAN and reading and an indirect association through processing speed, attention, working memory, and orthographic skills (e.g., Cutting & Dencelka, 2001; Juul et al., 2014; Papadopoulos et al., 2016; Sunseth & Greig Bowers, 2002).

Individual differences in phonological processing, RAN, and letter knowledge strongly and independently affect reading development (Cardoso-Martins & Pennington, 2004; Christopher et al., 2015; Deacon, 2012; Fletcher et al., 2020; Ozernov-Palchik et al., 2017; Schatschneider et al., 2002; Shapiro et al., 2013; Thompson et al., 2015; Wagner et al., 1997), but it remains to be investigated whether these associations vary qualitatively along skill continua during early reading acquisition. Specifically, little is known about how early differences in phonological processing skills (which encompass the constructs of phonological awareness and phonological memory) and RAN influence the relationship between letter knowledge and single-word reading, a relationship that is at the core of reading development. This lack of understanding may be due in part to the limitation of many methods of analysis that assume linear associations among phonological processing, RAN, letter knowledge, and reading outcomes. There are important implications, however, for gaining insight into the distinct roles of RAN and phonological processing for predicting reading ability.

1.2. Double-deficit hypothesis of developmental dyslexia

Phonological processing and RAN abilities are moderately correlated (e.g., Swanson et al., 2003) and are sometimes theorized to fall under the same larger construct of phonological processes (Torgesen et al., 1997; Ramus & Szenkovi, 2008), but some evidence suggests they should be considered separately. Many studies across languages find that phonological processing and RAN each account for unique variance in reading ability (e.g., Boets et al., 2010; Clayton et al., 2020; Cutting & Dencelka, 2001; Georgiou et al., 2008; Ho & Lai, 1999; Jiménez et al., 2008; Katzir et al., 2006; Kirby et al., 2003, 2010; Landerl & Wimmer, 2008; Lervåg & Hulme, 2009; Manis et al., 2000; De Jong & Van der Leij, 1999; Vaessen & Blomert, 2015; Vaessen & Blomert, 2010; Parrilla et al., 2004; Papadopoulos et al., 2009; Torppa et al., 2013; Shechter et al., 2018).

For example, a longitudinal study of English-speaking children from kindergarten to fifth grade reported unique contributions of phonological processing and RAN latent factors to explaining variance in reading across grades ( Kirby et al., 2005). In another study of English-speaking children, RAN, phonological processing, and letter knowledge were tested four times in kindergarten. All three measures independently accounted for the growth in word reading (Clayton et al., 2020). In non-English shallower orthographies RAN has been shown to explain greater variance in reading than in deeper orthographies (Vaessen et al., 2010).

Further, phonological processing and RAN have both been proposed as independent core deficits in dyslexia, a neurobiological condition that makes it difficult for children to acquire accurate word reading skills (Lyon et al., 2003; Peterson & Pennington, 2015). According to the Double-Deficit Hypothesis of developmental dyslexia (Wolf & Bowers, 1999), a deficit in either phonological processing or RAN can cause dyslexia, and children with both deficits are often the most impaired readers. The existence of phonological, RAN, and double-deficit subtypes has been validated across languages (e.g., Escribano, 2007; Jiménez et al., 2008; Lovett et al., 2000; Ozernov-Palchik et al., 2017; Shany & Share, 2011), but some studies have failed to identify these distinct subtypes (e.g., Vukovic & Siegel, 2006). In a longitudinal study of 1215 English-speaking kindergartners, a data-driven latent profile analysis identified distinct profiles of early literacy (Ozernov-Palchik et al., 2017). Three deficit profiles emerged in this study: Phonological deficit, RAN deficit, and double deficit. Two years later, there was 100% stability in group membership in these profiles. These findings suggest that RAN and phonological processing represent distinct and stable deficits.

There remains some debate as to whether having both phonological and RAN deficits affects one’s reading “more than the sum of the parts,” or whether it is simply the case that two deficits are worse than one. Phonological processing and RAN scores are not fully independent, and so grouping children based on deficit cutoff scores creates a statistical artifact of poor reading in the double-deficit group (Schatschneider et al., 2002). Creating arbitrary deficit cutoff points may also not capture meaningful differences in continuous variables (Compton et al., 2001). Thus, an important lingering question is whether RAN and phonological processing abilities exert unique and interactive effects on reading development and whether these effects are linear or non-linear (Kirby et al., 2010; Kruk et al., 2014).

1.3. Reading ability and the assumption of linearity

There is some evidence that the effects of RAN and phonological skills on reading development may not be simple and linear (Jones et al., 2016; Lervåg & Hulme, 2009; Juul et al., 2014; Protopapas et al., 2013; McBride-Chang & Manis, 1996; Johnston & Kirby, 2006). Poor phonological processing is thought to be most critical for children with poor word reading skills and earlier in development, with increased influence of RAN with age (de Groot et al., 2015; Kirby et al., 2003; Lervåg et al.,...
One study, for example, used word-reading cutoff scores to label children as poor, typical, and good readers and compared the strength of the predictive relationships from phonological processing to word reading across the three groups (de Groot et al., 2015). Stronger associations between phonological processing and word reading were found in poor readers than in good readers. Other studies, however, reported equal importance of phonological processing for reading across the distribution of reading ability (McIlraith & Language and Reading Research Consortium, 2018; Savage et al., 2005).

Similar inconsistencies have been demonstrated in developmental studies of RAN. RAN has been reported to be a stronger predictor of reading and spelling abilities among poor readers than among typical readers (Johnston & Kirby, 2006; McBride-Chang & Manis, 1996; Meyer et al., 1998; Pennington & Lefly, 2001). Other studies, however, reported a weaker (Kirby et al., 2010; McIlraith & Language and Reading Research Consortium, 2018) or equal (de Groot et al., 2015) predictive role of RAN in poor readers and increased contribution of RAN to explaining variance in reading with age (Kirby et al., 2003). This suggests that the RAN-reading relationship is not consistent across levels of reading ability.

Based on these findings of varied associations, it has been proposed that the shape of the predictive relationship between RAN and reading is curvilinear in that it is steep at lower levels of ability and flat at higher levels of ability (Kirby et al., 2010). Slow naming speed was thought to represent a salient problem for low-ability individuals, although fast naming speed was of little benefit to high-achieving individuals (Cutting & Denckla, 2001). In contrast, an “accuracy-before-speed” pattern has also been suggested, where RAN starts to explain variance in reading once a minimal level of decoding accuracy has been achieved (Juul et al., 2014). Another study found quadratic-type relationships in the prediction of word decoding from RAN, giving merit to the hypothesis of non-linearity (Kruk et al., 2014). In that study, growth in word-decoding skills from Grade 1 to Grade 3 slowed over time when critical levels of RAN time were exceeded. Yet, a limitation of previous studies is that the shape of the predictive relationship between RAN and reading is not consistent across levels of reading ability.

1.4. Using the non-linear cusp catastrophe approach to study reading ability

We employed a cusp catastrophe model, which is based on the idea that continuous changes in independent variables are oftentimes associated with abrupt, sudden, and discontinuous changes in dependent variables (Thom, 1975). According to the cusp model (Guastello, 2001, 2002), a change in behavior (the dependent variable) is a function of two (or more\(^2\)) control variables, termed asymmetry and bifurcation (Fig. 1), within a three-dimensional surface. When bifurcation has low values, the change in behavior as a result of incremental changes in the independent variable will be continuous and predictable. When bifurcation takes on high values, the change in observed behavior is sudden and discontinuous, with the direction of change depending on the level of the asymmetry factor, which exerts a linear effect on the dependent variable. That is, below a critical threshold of the independent variable, behavior becomes non-linear, discontinuous, and unpredictable; when this happens, a specific value of the independent variable can take on different values of the dependent variable, leading to unpredictability. Thus, according to cusp catastrophe theory, when phonological or RAN abilities are below certain critically low levels, relations between letter knowledge and word reading may become non-linear and chaotic. Previous findings point to the need to evaluate the complex letter

---

\(^2\) In the sense that more than one source of asymmetry or bifurcation can be tested simultaneously.
part of a larger study of reading development (Ozernov-Palchik et al., 2017). This study was approved by the authors’ institutional review board. Parents gave informed written consent and children gave verbal assent to participate. The original sample was from diverse backgrounds (Table 1) and included all children in the participating KG classrooms whose parents consented. The analysis was conducted in two independent samples recruited from one larger study using identical inclusion criteria: a concurrent sample and a longitudinal sample. The original sample included 1433 children. Participants were excluded from analysis if they did not complete a measure of word reading or were determined by the testers (trained post-baccalaureate research assistants or masters-level speech-language pathology students) to be non-native speakers of English or to have a language delay or disorder. Participants were also excluded if their age-based standard score on a measure of nonverbal cognitive ability (the Matrices subtest of the Kaufman Brief Intelligence Test (KBIT-2); Kaufman & Kaufman, 2004) was below 80.

The concurrent sample included 225 children who completed all assessments in KG. Because the research questions focused on relations between letter knowledge and word reading, children who read zero words on the Woodcock Reading Mastery Test, 3rd edition (WRMT-3) Word Identification subtest were excluded from the analysis.

The longitudinal sample was recruited from the original larger sample with the goal of oversampling at-risk readers and included 104 participants. Although the complete longitudinal sample included 181 participants in KG and 166 participants in second grade, because of a mid-study switch from an older version of Woodcock Reading Mastery Tests, Revised/Normative Update (WRMT-R/NU) to WRMT-3, complete data was available for 104 participants only. Additionally, the concurrent sample excluded children who read zero words, but a greater number of at-risk children were recruited for the longitudinal sample. Therefore, only 15 participants overlapped between the concurrent and longitudinal samples. All measures were collected in KG and again at the end of second grade. Second-grade word reading was the longitudinal outcome measure.

Score and demographic information for both samples is provided in Table 1. Mean household income was determined from 2010 census data using participants’ home zip code. Race and ethnicity were based on the researcher’s best estimate during the first testing session. For the longitudinal sample, additional information about race, income, and parental education was collected via parent questionnaire (Supplemental Table 1).

2.2. Behavioral assessment

Trained research testers administered behavioral assessments to children on an individual basis. Testing occurred in schools in either the spring preceding kindergarten or the fall of the kindergarten year. Longitudinal reading scores were collected during the spring or summer preceding third grade. All assessment sessions were audio-recorded and test administration and scoring were checked for validity and accuracy by a second scorer.

Behavioral assessments included nonverbal cognitive ability (KBIT-2 Matrices), phonological processing (Comprehensive Test of Phonological Processing (CTOPP) Elision, Blending Words, and Nonword Repetition subtests; Wagner et al., 1999), rapid automatized naming (Rapid Automatized Naming and Rapid Alternating Stimulus Tests (RAN-RAS) Objects, Colors, and Letters subtests; Wolf & Denckla, 2005), letter name knowledge and single-word reading (WRMT-3 Letter Identification (ID) and Word ID subtests, respectively; Woodcock, 2011; the Word ID subtest was administered in both KG and second grade). Standard scores based on age were determined for each measure and used for all modeling. For tests that did not include score norms for 4-year-olds (CTOPP, RAN-RAS, and WRMT-R/NU), score norms for age 5;0 (years; months) were used for children aged 4;9–4;11 (n = 40). Administration and reliability information for each measure and Pearson’s correlations among measures is reported in Supplemental Materials.

2.3. Data analysis using non-linear dynamical systems theory (NLDST)

We first estimated a cusp catastrophe model evaluating the effects of each of the RAN subtests (Objects, Colors, Letters) and phonological processing subtests (Elision, Blending Words, Nonword Repetition) on the letter knowledge-word reading relationship (similar to a moderation analysis). In the cusp catastrophe model, as values of an independent variable (RAN and/or phonological processing) decrease, the “response surface” bifurcates. In other words, moving below some critical level of
δy = y0 + by – α

in which a continuous response variable ‘y’ (word reading) is predicted by the bifurcation variable(s) ‘b’ (estimated for each independent variable model, RAN and/or phonological processing) and the asymmetry variable ‘a’ (letter knowledge).

The fit of the non-linear cusp model was then compared to both a logistic model, similar in conceptualization to the cusp, and to a competing linear model (as in Cancer & Antonietti, 2018). Superiority of the cusp model would be supported by (a) the presence of bimodality or multimodality in the responses that fall within the bifurcation area, (b) low values of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) in the cusp model as compared to both the linear and logistic models, and (c) significance of all coefficients linking the asymmetry and splitting or bifurcation variables to reading achievement. All models were run using the R software package cusp, which employs a modified model based on Cobb’s conceptualization of a cusp catastrophe (Cobb, 1981).

2.4. Power analysis

We performed a power analysis involving a non-linear regression model with five independent variables for which a medium effect size (squared multiple correlation $R^2 = 0.35$) is expected. We found that 43 participants were sufficient for obtaining this effect size (Cohen, 1992). Thus, our modeling approach had ample power to decipher small-to-medium effect sizes and detect significant effects.

2.5. Summary of model testing approach

Our research questions regarding the non-linear (splitting) role of RAN and phonological processing, alone and in combination, on the relation between letter knowledge and word reading, were tested with eight unique cusp models (Table 2 and Summary of cusp models, below). All models but Model 8 included nonverbal IQ as a regressor (asymmetry variable) in order to test whether the observed effects are evident over and above the effects of general cognitive ability.

2.5.1. Summary of cusp models

First, models were tested with individual single variables as sources that could independently lead to a “catastrophe” in word-reading behavior (i.e., an association with letter knowledge that cannot be explained using linear terms). The variables tested included each of the three RAN tasks (Letters, Objects, Colors; Models 1–3, respectively) and the three phonological processing tasks (Elision, Blending Words, Nonword Repetition; Models 4–6). By testing each variable separately, we could assess their individual effects and determine whether they show the same patterns as predictors of reading given that, despite their high correlation, their independence has been verified (Badian et al., 1991; Felton et al., 1990; Ho & Lai, 1999; Parrila et al., 2007; Vukovic et al., 2004).

Next, we tested the combined and interaction effects of RAN and phonological processing as potential mediators of the relation between letter knowledge and word reading (Models 7 and 8). In Model 7, the simultaneously tested direct effects of RAN and phonological processing predicted by the double-deficit hypothesis were tested with one RAN task (Letters) and one phonological processing task (Elision). RAN Letters and Elision were selected as they showed the strongest individual effects (indicated by the greatest $B$ values in the individual models).

Finally, Model 8 tested the prediction of the double-deficit hypothesis that RAN and phonological processing both exert direct and interactive effects on reading. That is, poor reading performance is expected.

Table 1

Behavioral scores on reading-related measures and demographics.

<table>
<thead>
<tr>
<th>Race</th>
<th>Mean/count</th>
<th>SD/percent</th>
<th>Range</th>
<th>Longitudinal (N = 104)</th>
<th>Mean/count</th>
<th>SD/percent</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>21</td>
<td>9%</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>1%</td>
<td>–</td>
</tr>
<tr>
<td>Black/African American</td>
<td>52</td>
<td>23%</td>
<td>–</td>
<td>18</td>
<td>18</td>
<td>17%</td>
<td>–</td>
</tr>
<tr>
<td>Hispanic</td>
<td>16</td>
<td>7%</td>
<td>–</td>
<td>3</td>
<td>3</td>
<td>3%</td>
<td>–</td>
</tr>
<tr>
<td>N/A</td>
<td>11</td>
<td>5%</td>
<td>–</td>
<td>18</td>
<td>18</td>
<td>17%</td>
<td>–</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>126</td>
<td>56%</td>
<td>–</td>
<td>4</td>
<td>4</td>
<td>4%</td>
<td>–</td>
</tr>
<tr>
<td>White</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>60</td>
<td>60</td>
<td>58%</td>
<td>–</td>
</tr>
<tr>
<td>Gender</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Female</td>
<td>125</td>
<td>56%</td>
<td>–</td>
<td>51</td>
<td>51</td>
<td>49%</td>
<td>–</td>
</tr>
<tr>
<td>Male</td>
<td>100</td>
<td>44%</td>
<td>–</td>
<td>53</td>
<td>53</td>
<td>51%</td>
<td>–</td>
</tr>
<tr>
<td>Median family income</td>
<td>84,612</td>
<td>29,018</td>
<td>29,638-22,0441</td>
<td>81,791.69</td>
<td>26,042</td>
<td>26,092-20,8194</td>
<td>34,931-20,8194</td>
</tr>
<tr>
<td>Age</td>
<td>66.66</td>
<td>4.08</td>
<td>–</td>
<td>57.00-78.00</td>
<td>99.19</td>
<td>4.11</td>
<td>92.00-113.00</td>
</tr>
<tr>
<td>KBIT-2 Matrices SS</td>
<td>100.47</td>
<td>11.14</td>
<td>–</td>
<td>83.00-154.00</td>
<td>98.76</td>
<td>9.26</td>
<td>80.00-124.00</td>
</tr>
<tr>
<td>CTOPP Elision SS</td>
<td>10.51</td>
<td>2.43</td>
<td>5.00-18.00</td>
<td>9.59</td>
<td>2.23</td>
<td>5.00-16.00</td>
<td>5.00-16.00</td>
</tr>
<tr>
<td>CTOPP Blending Words SS</td>
<td>11.05</td>
<td>1.97</td>
<td>5.00-17.00</td>
<td>9.93</td>
<td>2.4</td>
<td>5.00-17.00</td>
<td>5.00-17.00</td>
</tr>
<tr>
<td>CTOPP Nonword Repetition SS</td>
<td>8.68</td>
<td>2.64</td>
<td>4.00-17.00</td>
<td>8.65</td>
<td>2.68</td>
<td>4.00-17.00</td>
<td></td>
</tr>
<tr>
<td>RAN Letters SS</td>
<td>107.3</td>
<td>13.27</td>
<td>61.00-134.00</td>
<td>99.2</td>
<td>14.67</td>
<td>55.00-125.00</td>
<td></td>
</tr>
<tr>
<td>RAN Objects SS</td>
<td>101.78</td>
<td>15.96</td>
<td>57.00-144.00</td>
<td>95.89</td>
<td>15.02</td>
<td>62.00-130.00</td>
<td></td>
</tr>
<tr>
<td>RAN Colors SS</td>
<td>99.11</td>
<td>18</td>
<td>54.00-144.00</td>
<td>93.33</td>
<td>17.2</td>
<td>54.00-132.00</td>
<td></td>
</tr>
<tr>
<td>WRMT-R/NU Letter ID SS</td>
<td>109.63</td>
<td>7.98</td>
<td>84.00-134.00</td>
<td>106.55</td>
<td>8.44</td>
<td>85.00-130.00</td>
<td></td>
</tr>
<tr>
<td>WRMT-R/NU Word ID SS</td>
<td>121.81</td>
<td>25.48</td>
<td>83.00-175.00</td>
<td>105.95</td>
<td>23.39</td>
<td>80.00-175.00</td>
<td></td>
</tr>
<tr>
<td>2nd Grade WRMT-R/NU Word ID SS</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

Note: SS = Standard score, using age-based norms. CTOPP standard scores have a mean of 10; all other standard scores have a mean of 100.
### Table 2 (continued)

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>B</th>
<th>S.E.</th>
<th>Z-value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>b(RAN Letters)</td>
<td>-0.757</td>
<td>0.137</td>
<td>-5.529</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>b(Ellision)</td>
<td>-0.275</td>
<td>0.125</td>
<td>-2.196</td>
<td>0.028*</td>
</tr>
<tr>
<td>b(RAN Letters * Ellision)</td>
<td>-0.105</td>
<td>0.113</td>
<td>-0.925</td>
<td>0.354</td>
</tr>
<tr>
<td>w(Intercept)</td>
<td>-1.020</td>
<td>0.055</td>
<td>-18.590</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>w(Word ID)</td>
<td>1.198</td>
<td>0.043</td>
<td>27.626</td>
<td>&lt;0.001***</td>
</tr>
</tbody>
</table>

### 3. Results

#### 3.1. Behavioral scores and the letter knowledge-word reading relationship

Scores for all behavioral measures are reported in Table 1. We first tested whether the predicted strong association between letter knowledge and word reading was present in our sample. A Pearson’s correlation revealed that WRMRT-R/NU Letter ID and Word ID standard scores were correlated at \( r = 0.75 \) (\( p < 0.001 \)); Fig. 2 provides a scatterplot of these scores. This correlation coefficient indicates a large effect size (Cohen, 1992). The longitudinal correlation was at \( r = 0.48 \) (\( p < 0.001 \)), indicating a medium effect size (Fig. 3).

![Figure 2](image)

**Fig. 2.** Scatterplot of letter knowledge (WRMT-R/NU Letter ID subtest) and word reading (WRMT-R/NU Word ID subtest) standard scores in the sample (n = 225) showing their strong linear relationship (\( r = 0.75 \)) but also their distinct nature (49% common variance, 51% unexplained), using simple linear regression.
Fig. 3. Scatterplot of pre-kindergarten/kindergarten (KG) letter knowledge (WRMT-R/NU Letter ID subtest) and second-grade word reading (WRMT-R/NU Word ID subtest) standard scores in the sample (n = 104) showing their strong linear relationship (r = 0.48) but also their distinct nature (49% common variance, 51% unexplained), using simple linear regression.

Additionally, we verified a lack of outlying cases across models. A univariate analysis of outliers for all terms (i.e., dependent, asymmetry, and bifurcation) indicated that all participants’ scores were less than three standard deviations from the mean, and given that the distributions were normal, these deviations were within acceptable standards. To evaluate the presence of multivariate outliers, the dependent variable of word identification was regressed on three independent variables for all models in Table 2. Mahalanobis distances were transformed to probabilities and were evaluated at a p = 0.001 level of significance, again revealing an absence of outliers.

3.2. Cusp-catastrophe concurrent models

3.2.1. Single-variable models

Information for each of the cusp models is given in Table 2. A negative B value indicates a pattern consistent with a cusp catastrophe, such that better performance on the bifurcation variable (RAN or phonological processing) is associated with linearity and lower performance with non-linearity. Model 1 considered RAN Letters as the bifurcation variable. This cusp model was well supported by the data: at low levels of RAN Letters performance, Letter ID performance no longer had a linear relationship with word reading and the reading scores deviate markedly from the predictions of the linear model. The bifurcating role of the RAN task was also observed with RAN Objects (Model 2). RAN Colors, on the other hand, did not exceed conventional levels of significance as the bifurcation variable, although it approached such levels (Model 3). Similarly, when investigating the role of phonological processing, the cusp model was supported with the Elision task (Model 4) and partially with the Blending Words task (Model 5); however, the cusp model incorporating Nonword Repetition (Model 6) did not show a good fit to the data.

3.2.2. Combined RAN-phonological processing model

Model 7 assessed the roles of RAN and phonological processing simultaneously as bifurcation variables, since according to the double-deficit hypothesis, they represent two salient but distinct sources of reading impairment. As planned, we used the RAN and phonological processing subtests that showed the strongest relations with Word ID in the single-deficit models: RAN Letters and Elision. Model 7 showed significant linear effects of letter knowledge and IQ on word reading (b_{LK} = 0.172, p < 0.001; b_{IQ} = 0.029, p < 0.001) and significant non-linear contributions of both RAN Letters and Elision (b_{RAN\_L} = −0.055, p < 0.001; b_{Elision} = −0.087, p < 0.05). The cusp model was superior to both the linear and logistic models (AIC\_Linear = 1887.735, AIC\_Logistic = 1843.075, AIC\_Cusp = 286.831). Fig. 4 displays the pattern of responding that is observed within the response surfaces (lower and upper) of the cusp model. If the data supported the proposed model, we would see observations within the bifurcation (folded) area manifest in bimodal or multimodal distributions, suggesting various levels of performance within those regions. Furthermore, observations to the left and to the right of the bifurcation area would show positive and negative skew, respectively. Indeed, these premises of a well-fitted cusp model were observed with the current data (Fig. 4), with more than 10% of the observations falling within the bifurcation area (Fig. 5). Last, Fig. 6 shows the observations transitioning from the upper to the lower surface, entering the shaded “inaccessibility” area for which observations take on various values, before reaching equilibrium (lower surface). An ancillary plot (Fig. 7) displays how the cusp catastrophe works, using the RAN Objects task as an example. RAN Objects acts as a bifurcation variable relative to word reading and letter knowledge in a threedimensional scatterplot. The circled observations show that at moderate levels of letter knowledge (standard scores approximately between 90 and 110) and across the range of scores on RAN Objects, reading performance is rather low and does not show an identifiable visual pattern (linear or non-linear). This relationship is estimated to be around...
zero, suggesting that RAN does not relate to word reading when letter knowledge scores are in this range. However, at high values of letter knowledge, low RAN standard scores are associated with high word reading scores, and as RAN time increases, Word ID scores decrease.

3.2.3. The interaction model for testing the double-deficit hypothesis

We also tested the hypothesis that RAN and phonological processing exert both independent and combined effects on reading (Model 8). The model with nonverbal IQ as an asymmetry term did not converge properly as the error terms of nonverbal IQ could not be estimated; consequently, the model was run without it. The model included both the linear and interaction terms. Results indicated that uncertainty in reading performance was introduced by the independent RAN and phonological processing terms, but not by their interaction. In other words, there was no additional bifurcation effect due to the combined influence of RAN and phonological processing.
### Table 3 (continued)

<table>
<thead>
<tr>
<th>Parameter estimates</th>
<th>B</th>
<th>S.E.</th>
<th>Z-value</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>a[Letter ID]</td>
<td>0.161</td>
<td>0.056</td>
<td>2.888</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>b[RAN Letters]</td>
<td>0.272</td>
<td>0.049</td>
<td>5.586</td>
<td>&lt;0.001***</td>
</tr>
<tr>
<td>a[Blending]</td>
<td>2.226</td>
<td>0.026*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w[2nd Grade Word ID]</td>
<td>0.027</td>
<td>0.005</td>
<td>-5.868</td>
<td>&lt;0.001***</td>
</tr>
</tbody>
</table>

*p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001 marginal.

### 3.3. Cusp-catastrophe longitudinal models

#### 3.3.1. Single-variable longitudinal models

We tested the influence of RAN and phonological processing on the longitudinal association between KG Letter Knowledge and second-grade Word ID Table 3. The bifurcation role of RAN Objects (Model 10), Blending Words (Model 13), and Elision (Model 12) was supported by the results. RAN Letters (Model 9) was not significant. Consistent with the concurrent KG findings, cusp models that included RAN Colors (Model 11) and Nonword Repetition (Model 14) did not show a good fit to the data.

#### 3.3.2. Combined RAN-phonological processing model

We included RAN Letters and Elision (Model 17) simultaneously as bifurcation variables, both independently and as an interaction term. There was a significant non-linear contribution of RAN Letters (b_{RAN.L}=0.331, p < 0.05) and of Elision (b_{Elision}=0.380, p < 0.05) over the longitudinal relationship with Word ID. The fit of the cusp model was superior to that of the linear model (AIC_{Linear}=270.433, AIC_{Cusp}=253.764, with significant non-linear and independent contributions to the longitudinal models (b_{RAN.L}=0.331, p < 0.05; b_{Elision}=0.526, p < 0.05). Model 17 showed a better fit to data than the linear models (AIC_{Linear}=802.731, AIC_{Cusp}=265.591). The cusp model that included RAN Letters and Elision (Model 16) performed better than the linear models (AIC_{Linear}=334.052, AIC_{Cusp}=124.068) and showed a significant bifurcation. The individual RAN Letters (b_{RAN.L}=0.331, p = 0.052) and Elision (b_{Elision}=0.380, p = 0.033) terms were significant and approaching significance, but the interaction term was not (b_{Interaction}=0.032, p = 0.812). The model with RAN Objects and Blending Words (Model 19) also showed a better fit to the data than the linear models (AIC_{Linear}=253.764, AIC_{Cusp}=-154.698), with significant individual terms for RAN Objects (b_{RAN.O}=0.393, p = 0.023) and Blending Words (b_{Blending}=0.590, p = 0.001), but not for their interaction (b_{Interaction}=0.029, p = 0.874).

In summary, the results in the longitudinal models remained stable with regard to the roles of RAN, phonological processing, and their interaction, but not for IQ. Specifically, IQ was no longer a significant positive predictor of word reading. RAN was a significant bifurcation factor with Letters and Objects subtests, but to a lesser extent with Colors; phonological variables remained significant bifurcation factors with regard to Elision and Blending Words but not Nonword Repetition. Both RAN and phonological processing remained significant bifurcation factors in the combined models but their interaction term did not provide additional predictive value in the explanation of word reading. All requirements of the cusp model were met in Model 17, suggesting that the longitudinal relationship between letter knowledge and word reading was positive and linear until a point at which RAN or phonological processing ability was below a certain critical level, after which point word reading became unpredictable. This finding of a combined phonological processing and RAN model supporting the double-deficit hypothesis was replicated with other combinations of the bifurcation variables – that is, significant effects were also observed when the bifurcation variables were Elision and RAN Objects. These findings further support the notion that both naming speed and phonological...
processing are critical precursors of reading development (Table 3).

4. Discussion

The purpose of the present study was to test the hypothesis that the contribution of RAN and phonological processing to single-word reading is well characterized by the non-linear cusp catastrophe model. We tested this hypothesis in two independent samples: a concurrent sample that included 225 pre-kindergarten and kindergarten children and a longitudinal sample of 104 children followed from pre-kindergarten and kindergarten to second grade. Results clearly supported this hypothesis, revealing non-linear associations between letter knowledge and word reading in cases of low phonological skills or RAN ability. The cusp model fit the data significantly better than a linear model for both RAN and phonological processing. By demonstrating that the assumptions of linearity are violated at the lowest levels of RAN and phonological skills, these findings provide insight into why previous models perform sub-optimally when predicting subsequent reading difficulties in at-risk children.

Although these are the first findings in emerging readers in pre-kindergarten and kindergarten, the ideal time frame for risk identification, they are consistent with a previous study that applied cusp modeling to investigate the association between RAN (but not phonological processing) and reading in second and third graders (Sideridis et al., 2019). The results from that study support the role of RAN in disrupting the association among reading variables. Our results also extend previous literature that has demonstrated that the predictive nature of RAN varies across reading skill levels (Compton, 2003; Kirby et al., 2003; Kruk et al., 2014; Lervåg et al., 2009; McBride-Chang & Manis, 1996), again confirming the non-linear nature of the associations among various reading-related constructs. Importantly, these previous studies applied a linear framework to modeling, such as testing for a moderating role of RAN on bivariate associations between two variables. In some cases, more complex (quadratic) models were applied, but such analyses assume that although the strength of the association may vary, the relationship remains predictive at all skill levels. Here we demonstrate that at critically low levels of RAN and phonological processing, the relationship becomes qualitatively different and the assumption of predictability is violated.

The disruption of the letter knowledge-word knowledge association in children with poor phonological and RAN skills could be due to qualitative differences in how reading develops in typical and poor readers. The developmental transition from mastering letters to using them to put the letter sounds together to form words (i.e., phonemic awareness). A child with poor phonological skills, therefore, may know all their letters, but may not be able to put the letter sounds together to form words (Byrne & Fielding-Barnsley, 1989). Similarly, reading words presents additional cognitive and perceptual challenges beyond those of naming letters (Wolf & Katzir-Cohen, 2001). A child with poor RAN, therefore, may struggle to develop automaticity in connecting letters into words and may demonstrate a discrepancy between their letter-knowledge and word-identification skills. Indeed, there is considerable evidence that children with poor reading development rely on different strategies for identifying words than typically developing children. Whereas typical readers utilize knowledge of letter-sound correspondences to decode words, poor readers rely to a greater extent on sight memory or context-based prediction (Nation & Snowling, 1998; Perfetti & Roth, 1980; Stanovich et al., 1981; Siegelman et al., 2020).

4.2. Effects of RAN and phonological processing are independent

Overall, the phonological processing and RAN measures significantly explained word reading, evidenced by the testing of two independent models and a combined model with both measures. Including the two predictors as bifurcation variables (Models 7 and 15) suggested strong independent effects for both of them. The independence of RAN and phonological processing observed here is consistent with a consensus in the literature that the two processes exert unique effects on reading ability (see Norton & Wolf, 2012 for a review). Crucially, all models tested for these independent variables accounted for nonverbal cognitive ability, which has been suggested as an important factor early in reading development (van Bergen et al., 2013). Thus, the disruptive effects of RAN and phonological processing in combination were not due to an overall pattern of low cognitive skills often associated with a multi-deficit profile (e.g., Ozernov-Palchik et al., 2017). The contribution of RAN to reading is considered multi-componential (Norton & Wolf, 2012) and has been partially explained by contributions of processing speed, orthographic processing, attention, articulation, and working memory (Cutting & Denckla, 2001; Juul et al., 2014; Papadopoulos et al., 2016; Sunseth & Greig Bowers, 2002; Holland et al., 2004; Neuhaus & Swank, 2002).

We tested whether the interaction of RAN and phonological processing would be associated with additional effects compared to those of the individual terms on their own (Models 8 and 16). There was no specific effect of the interaction of the two variables on the letter knowledge-word reading relation. Although ample evidence exists with regard to their differentiation, the present analysis demonstrated that their combined effect (e.g., in low-low patterns or high-high patterns) was not associated with concomitantly non-linear changes in reading behavior. This finding is also consistent with other studies that have reported independent effects of phonological processing and RAN (Compton et al., 2001; Sunseth & Greig Bowers, 2002). The significance of the RAN and phonological direct effects support their conceptual independence and provide important counterevidence for the notion that a RAN deficit represents a mere extension of the phonological retrieval deficit (Ramus & Szenkovits, 2008).

The findings that RAN and phonological processing on their own and in combination influence reading performance is consistent with the double-deficit hypothesis. The double-deficit hypothesis proposes that those children who evince deficits in both phonological processing and RAN will be the most severely impaired word readers; however, the authors do not directly predict an interactive statistical relationship among phonological, RAN, and word reading measures. Similar results have emerged in previous studies, in that both variables together predict word reading, but that there is not an interaction between them that improves prediction (Compton et al., 2001; Sideridis et al., 2019). Indeed, the impetus for the double-deficit hypothesis itself was to open the door to considering multiple, heterogeneous causes of reading difficulty (Wolf & Bowers, 1999).

4.3. Differences across RAN and phonological processing subtests

We examined multiple subtests tapping phonological and RAN skills.
and observed that the parameters of the cusp model were significant for most of the models that tested single-variable effects of phonological processing or RAN on reading behavior. The exceptions were the models with RAN Colors (Models 3 and 11) and Nonword Repetition (Models 6 and 14).

We observed that phonological tasks of Elision and Blending Words showed a different effect on the cusp model than did Nonword Repetition. Whereas all three processes are often subclassed under the larger construct of phonological processing, many studies show that phonological memory, as indexed by Nonword Repetition, is only correlated at about \( r = 0.55 \) with phonological awareness at this age (Alloway et al., 2005; Nation & Hulme, 2011). Furthermore, unlike Elision and Blending skills, which are characteristically low in individuals with reading impairment, a deficit in phonological memory has emerged as a robust indicator of risk for both language and reading impairment and has demonstrated an association with language skills other than phonological processing, such as grammar ability (Bishop et al., 1996; Botting & Conti-Ramsden, 2001; Gathercole & Baddeley, 1996; Gray et al., 2019; Marton et al., 2016). Furthermore, children with specific language impairments were more likely to have dyslexia if they had deficits in nonword repetition (Cutts & Adlof, 2011; Kelso et al., 2007; McArthur & Castles, 2013). In a meta-analysis that dissociated different types of working memory (Peng et al., 2018), verbal working memory represented a unitary domain-general construct of executive functioning in relation to reading in younger children before fourth grade, but in older children there was domain-specificity in this relationship, with verbal working memory showing a stronger association with reading. This suggests that phonological working memory is partially distinct from the overall phonological processing construct and could therefore exert qualitatively different effects on the association between letter knowledge and word reading (Duke & Cartwright, 2021).

Among the three RAN measures used in the current study, Letters is the closest to early reading since it is associated with processes related to retrieval of letter knowledge (versus, for example, semantic knowledge). This finding is consistent with a previous paper in first- and second-graders that shows a weaker association of RAN Colors with word reading (\( r = 0.31 \)) as compared to RAN Objects (\( r = 0.38 \)) and RAN Letters (\( r = 0.43 \); Papadopoulos et al., 2016). Similarly, while some studies found that RAN is a unitary construct early in reading development (Kirby et al., 2003; Van den Bos et al., 2002), in other studies, alphabetic and non-alphabetic RAN stimuli followed different patterns of relationships to other variables (e.g., Denckla & Cutting, 1999; Lervåg et al., 2009). The alphabetic stimuli were better concurrent discriminators of reading skills, while the non-alphabetic stimuli were better predictors of reading outcomes (Denckla & Cutting, 1999). In our study, although RAN Letters and Objects were significant predictors of reading, RAN Colors was not (although it followed the same pattern of relationships as the other subtests). Among non-alphabetic stimuli, it is possible that, early in development, objects exert a stronger influence on the relations among reading constructs than do colors due to differences in automaticity.

### 4.4. Replication of findings across two independent samples

A noteworthy strength of the current study is that we replicated our findings across two largely independent samples (with an overlap of 15 participants), one concurrent and one longitudinal. We confirmed that the disruption of linearity observed between letter knowledge and word reading in the concurrent (kindergarten) sample is present in the relations between letter knowledge and longitudinal reading outcomes. We also confirmed the independent contributions of RAN and phonological processing to word reading and the weaker and non-significant effects of Nonword Repetition and RAN Colors subtests. Therefore, we can interpret our findings with greater confidence, but future studies in larger independent samples and in non-English orthographies are needed to confirm the generalizability of the current findings to other populations.

One important difference between the concurrent and longitudinal models is the effects of kindergarten IQ on reading, which were significant in the concurrent models but not in the longitudinal models. These findings could speak to the increased uncoupling of IQ from reading (Ferrer et al., 2007; Ferrer et al., 2010) and from language (Kievit et al., 2019; Kievit et al., 2017) across development. Such findings have been interpreted as showcasing the importance of cognition for skill mastery in early education, when these skills are systematically taught (Peng & Kievit, 2020). However, because there was little participant overlap in our two samples, these findings could also indicate sample-specific differences. For example, the association between IQ and reading is stronger in good readers than in poor readers (Ferrer et al., 2010). Thus, the lack of IQ effects in the longitudinal sample could reflect the overall lower word-reading performance of this group as compared to the concurrent sample.

### 4.5. Limitations and future directions

A limitation of our study is the exclusion of a subset of children from the concurrent sample because of their word-reading scores of zero. This exclusion has likely resulted in over-representation of good readers in the sample. Although we replicated the findings in another, largely independent, longitudinal sample, additional replication in an independent study is needed to establish the disruptive role of low RAN and phonological skills on reading development. Furthermore, it is important to evaluate whether the current findings of the disruptive effects of both phonological processing and RAN will replicate in different – and shallower – orthographies (Phillips Galloway et al., 2020). Additionally, although the current findings make a theoretical advance about the influence of RAN and phonological processing on reading development, an important future direction is to translate these findings to more precise identification methods for at-risk students (Solari et al., 2020). Finally, we emphasize that causal inferences cannot be made from the current analyses because the cusp model utilizes covariances between the independent and dependent variables.

### 5. Summary

In summary, we reveal an important non-linear effect of low phonological abilities, low RAN abilities, or both, on the otherwise predictable path of reading development. Our findings also support the double-deficit hypothesis of dyslexia, which predicts additive effects of rapid-naming and phonological deficits on reading development, reinforcing the importance of multi-componential conceptualizations of dyslexia (e.g., Pennington et al., 2012) and more multi-componential intervention (Lovett et al., 2017). Further, these combinatorial influences of phonological processing and RAN support the notion that both should be considered in early screening as potential explanatory factors in poor reading ability, even at the earliest stages of reading development. These results also extend recent findings indicating that interactive relations govern the association between phonological processing, RAN, and reading (Landerl et al., 2019).

### Funding

This research was supported by the National Institutes of Health (F32-HD100064 to OO; R01-HD067312 to JDEG and NG; F31-HD10101 to SDB).

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.lindif.2022.102166.
word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. Developmental Psychology, 33, 468–479.


