Editorial Review Board

Ivar Bjørgen, Norway
James Chapman, New Zealand
Cesare Cornoldi, Italy
Merith Cosden, United States
Gad Elbeheri, Egypt
Linda Elksnin, United States
Paul Gerber, United States
Pol Ghesquière, Belgium
Carol Goldfus, Israel
Lorraine Graham, Australia
Steve Graham, United States
Michael Grosche, Germany
Daniel Hallahan, United States
Ian Hay, Australia
Charles Hughes, United States
Olga Jerman, United States
Asha Jitendra, United States
Sunil Karande, India
Lynda Katz, United States
Barbara Keogh, United States
Youn-Ock Kim, South Korea
Janette Klingner (late), United States
Che Kan Leong, Canada
Joseph Madaus, United States
Lynn Meltzer, United States
Ana Miranda, Spain
Maria Chiara Passolunghi, Italy
Henry Reiff, United States
Georgios Sideridis, Greece
Linda Siegel, Canada
H. Lee Swanson, United States
Rosemary Tannock, Canada
Masayoshi Tsuge, Japan
Annmarie Urso, United States
Delinda van Garderen, United States
Christa van Kraayenoord, Australia
Judith Wiener, Canada

2014 IARLD Executive Board

Christa van Kraayenoord, President, University of Queensland
Linda Siegel, President-Elect, University of British Columbia
Judith Wiener, Immediate Past President, University of Toronto
Margaret Pierce, Treasurer, University of Massachusetts, Amherst
Diane Bryant, Secretary, University of Texas, Austin
Asha Jitendra, Vice President for Fellows, University of Minnesota
Linda Siegel, Vice President for Members, University of British Columbia
Henry B. Reiff, Vice President for Student Members, McDaniel College
Malka Margalit, Vice President for International Development, Tel Aviv University
Lynn Meltzer, Chair of Conference Programs, Research Institute for Learning and Development; Harvard University
Brian Bryant, Academy Historian, University of Texas, Austin
Carol Goldfus, Publications Chair, Levinsky College of Education, Tel Aviv
Delinda van Garderen, By-Laws and Constitution Committee Chairperson, University of Missouri

Members at Large

Rasa Barkauskiene, Vilnius University
James Chapman, Massey University
Marija Kavkler, University of Ljubljana
Table of Contents

Children at Risk of Reading Problems – From Identification to Prevention
By Heikki Lyytinen 2

Predictors of Positive Mood and Negative Mood among Children with Learning Disabilities and Their Peers
By Adi Sharabi and Malka Margalit 18

How Persistent is a Diagnosis of Mathematical Disorder at an Early Age? A Longitudinal Study
By Annemie Desoete, Frauke De Weerdt, Ruth Vanderswalmen and Annemie De Bondt 42

Essay-Writing Interventions for Adolescents with High Incidence Disabilities: A Review of Research
By Lauren Valasa, Linda Mason and Charles Hughes 72

Guest Reviewers 98
William M. Cruickshank Memorial Lecture Delivered at the 2014 Conference of the International Academy for Research in Learning Disabilities

Children at Risk of Reading Problems – From Identification to Prevention

Heikki Lyytinen
University of Jyväskylä, Finland

A fundamental skill for growth and improvement in modern society is the ability to read – a skill still denied to many millions of children across the globe. A number of factors share the responsibility for this dilemma. In developing countries, most potential readers live under conditions where instruction is non-optimal or even unavailable (UNESCO, 2014). However, even in the developed world, some children still face serious problems in learning to read. In this latter context, the most severe bottlenecks result from biological factors that lead to dyslexia. In addition, a large proportion of cases of inadequate literacy result from apathy towards reading. The most important correlate of complete reading skill - which also includes comprehension of the subject matter - is the reading activity itself. After first acquiring the ability to decode written material as a necessary prerequisite, a learner can only achieve full literacy by further prolific reading. My emphasis here is to illustrate the early identification of children who will face such decoding problems and also how such difficulties can be surmounted.

The term ‘dyslexia’ is often used in the description of those children who, despite being in receipt of appropriate instruction, and in the absence of any apparent reason, continue to face severe difficulties in the acquisition of basic reading skill. Dyslexia compromises the positive school learning experiences of fewer than 10% of children during their early school years. As a consequence, however, such a negative start in the quality of their reading acquisition may subsequently result in many such children reading too little in terms of quantity in order to boost the chances of becoming ‘compensated’. Nonetheless, approximately 20% of children who initially demonstrate severe difficulty in learning basic reading skills (decoding accuracy) during the first grades end up as typical readers (Leinonen et al. 2001). This is a promising prognosis, suggesting that individuals with dyslexia (which we now know has a genetic basis) can also overcome their difficulties – as will be described.
below – and sometimes, without recourse to exceptional means. In short, appending initial and appropriate special education support with persistent reading activity can result in ‘compensation’. This compensatory scenario has been demonstrated in both transparent and non-transparent writing environments (see below for an explanation of writing transparency).

The most typical scenario that differentiates ‘compensated’ from other learners who faced similar initial difficulties is that, irrespective of the underlying reason, compensated learners are motivated towards prolific reading and consequently attain full literacy. Our unpublished data reveal that the only discernable ‘dyslexic’ difference in adulthood can be found by measuring auditory memory, and this points saliently to the attenuation of the same core cognitive problem that we define below as characteristic of dyslexia. This finding fuels optimism that dyslexia, independent of its genetic background, can be overcome to such an extent that it does not compromise full literacy, although much more reading effort is required.

In the Jyväskylä Longitudinal study of Dyslexia (JLD) we followed 200 children from birth, half of whom had a familial background of reading problems. Prior to commencing the follow-up of the infants at risk due to their being born to parents with dyslexia, we interviewed a large number of parents who had acknowledged in a questionnaire at the maternity clinic (visited by all Finnish families prior to a birth) that they had had severe problems in learning to read. This defined the risk status of the expected infant. As an acknowledged fact, approximately 20% of the parents showed reading skill that was within normal limits. This transpired despite their having encountered such severe early problems that they repeated a school year (common in Finland at that time for children who did not learn to read during the first grade in line with the very large majority who become accurate readers within the first few months of schooling) and failed to become accurate readers during the early school years. Similar findings have been reported from an English reading environment (Lefly & Pennington, 1991).

The transparency of different writing systems (orthographies) affects the learning burden placed on children during their acquisition of basic reading skill (Seymour et al., 2003). The burden is low in fully transparent and high in non-transparent alphabetic orthographies. Examples of almost transparent orthographies include Finnish and Italian, where each letter represents its own phoneme (or sound) and, in turn, each phoneme is represented consistently by only one letter (or letter combination, grapheme). Thus, the learning burden to acquire basic decoding skill entails little more than to learn the sounds of < 30 letters/graphemes and to realize how to assemble these sounds in the order of letters. Mastery of these basic steps results in the learner’s ability to sound out any pronounceable sequence of letters.

The situation is quite different in non-transparent writing such as in English, where none of its letters consistently represent the same sound, independent of the context of where, in the word, the letter occurs. Furthermore, the meaning of the word may also affect the sound of its written form, which means that the learner first has to know the spoken form of
the word being read. Thus, the differences between various alphabetic orthographies are already substantial, before even considering those writing systems that are non-alphabetic. As previously mentioned, readers of transparent orthographies tend to learn basic reading skills during the first months of school, or even earlier. More than 1/3 of Finnish children are able to read before entering school at 7 years of age. However, children learning to read English require two more years of school instruction in order to reach the mean level of reading of 6-to 7-year-old Finnish children, even when they have not yet been influenced by formal teaching (Seymour et al., 2003). It is, however, prudent to acknowledge that UK children start school around age 5.

The difference between the learning burdens can be illustrated by counting the number of connections that the learner has to store before becoming a relatively accurate reader. In Finnish, this number approximates 30 while in English, it numbers more than a thousand. This is because English has, for example, a number of written words that defy sounding out before all the letters are seen (e.g. ‘view’). In general, due to the aforementioned inconsistency, the size of the written unit that the English reader must learn in order to connect to the spoken language item is much larger than the letter. Larger units (typically >2 letters) can be connected to a spoken segment of English when learning the consistently-behaving connections (i.e., those which are true in all contexts of written English (Ziegler & Goswami, 2005).

This introductory issue should be considered when interpreting the summary of results that I focus on next. These results are based on the previously mentioned JLD follow-up study of Finnish children who have been assessed continuously from birth to puberty. Finnish children who face problems in learning basic reading skills have specific difficulties in a) learning the less than 30 connections between spoken and written units of Finnish and/or b) becoming fluent decoders of Finnish writing. These difficulties were observed in the JLD among learners who have no general cognitive problems (IQ >80).

**Bottlenecks Compromising Reading among Learners with Dyslexia**

A typical child with dyslexia who is learning to read Finnish can be helped towards accuracy simply by sufficient drilling of the sounds of the letters and by motivating him/her to automatize the assembly of letter sequences through repeated practice. Such practice helps them to become fluent readers. In the learning of Finnish spelling, however, the most challenging initial difficulties tend to persevere, despite substantial training, in relation to mastery of the most difficult aspects of such connection building. In Finnish, such a bottleneck is the differentiation between short and long phonemes (i.e. ‘phonemic duration’ whereby a long phoneme is clearly marked by repeating the letter). A further bottleneck often concerns the automatization of reading. That is, to attain sufficient reading fluency/speed in order to still recall the beginning of a sentence when the end is reached and thus facilitate comprehension of long sentences.
The initial focus of learning to read transparent writing is to connect the sounds of single phonemes to the representative letter/grapheme. Therefore, any difficulty with the differentiation of such small speech units may manifest as a substantial bottleneck. This is, in fact, the case, as we have documented in the Jyväskylä Longitudinal study of Dyslexia (JLD; for a review of findings, see Lyytinen et al., 2008). The case is further supported by our findings of very early differences shown by newborns (at familial risk) in their responses to speech stimuli. Brain event-related potentials at age 3-5 days can predict reading acquisition (Guttorm, Alinäveri, Richardson & Lyytinen, 2011; Hämäläinen et al., 2013) and perception of phonemic length is, finally, the most concrete expression of compromised speech perception that can be observed from a very early age and further, at reading age. This has been shown by the JLD results wherein school age children’s perception of the occurrence of long vs. short phonemes still predicts later reading acquisition after the effect of other known predictors has been controlled (Pennala et al., 2013). This finding uncovers a central specific factor within the domain of phonemic awareness that manifests as the most serious bottleneck to delay a learner’s progression towards accurate spelling. It may be important to understand that, in a transparent orthography, ‘phonemic awareness’ per se is not dissimilar to letter sound knowledge. As such, it is achieved no later than the first exposure to initial reading instruction during the first few months of school, which is the first and most necessary step that the learner needs for acquisition of basic reading skill. In short, phonemic awareness is acquired reliably and relatively easily via appropriate instruction of the sounds of the letters in the context of transparent writing. Instruction that includes assembly – by introducing consonants after acquisition of the easy-to-store vowel sounds – together with vowels (i.e., in v-c or c-v syllables) facilitates easy learning of the basic steps of reading skill.

Such an early focusing of the child's attention towards important small units may, however, affect the emergence of later bottlenecks - fluency problems. In transparent orthographies, where only a few fail to become accurate readers, excessively slow reading is the most serious and common difficulty among children with dyslexia. Overly fixating the attention on such small units may therefore culminate in dysfluency. Consequently, the initial instruction must be organized in such a way that it can alert the teacher to any such potential danger and, as such, can be avoided by instigating the transition towards the reading of larger units as early as possible in transparent orthographies. This same process occurs naturally in non-transparent environments where reading cannot start successfully without the use of larger units from the outset.

**Early Identification of Children in Order to Provide Preventive Practice**

The Jyväskylä Longitudinal study of Dyslexia (JLD) has provided a detailed perspective on the obstacles faced by children with a genetic risk for dyslexia. In the JLD, the development of language from an early age was assessed in detail, starting with the earliest imaginable indications. These comprised the brain responses to speech sounds soon after birth. Next, there was a follow-up of the early indications of expressive and receptive speech,
the development of morphological skills, phonological sensitivity, letter knowledge and then naming fluency. Figure 1 summarizes the significant predictors of compromised reading acquisition in second grade, by which time the vast majority of Finnish pupils are accurate readers. The most significant predictors are listed from birth to reading age. The table also shows those predictive measures on which the groups with and without familial risk differed significantly.

<table>
<thead>
<tr>
<th>Age</th>
<th>Variable</th>
<th>P</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 yrs</td>
<td>Reading accuracy &amp; speed</td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>5 yrs</td>
<td>Naming speed</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>4 - 6 yrs</td>
<td>Phonological manipulation</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>5 - 6 yrs</td>
<td>Letter knowledge</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>5 yrs</td>
<td>Verbal memory</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>3 - 6 yrs</td>
<td>Phonological sensitivity</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>3 - 5 yrs</td>
<td>Inflectional skills</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>2 - 3 yrs</td>
<td>Articulation accuracy</td>
<td>P</td>
<td></td>
</tr>
<tr>
<td>2 yrs</td>
<td>Maximum sentence length</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>6 mth</td>
<td>Speech perception</td>
<td>P &amp; D</td>
<td></td>
</tr>
<tr>
<td>Birth</td>
<td>ERP to speech sound</td>
<td>P &amp; D</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1. Early measures with significant predictive correlation (P) to reading acquisition or group difference (D) between children with and without familial risk for dyslexia (D). For more, see Reid et al., Sage Handbook of Dyslexia, 2008.

As previously mentioned, the first indications of the realization of risk can be observed at the age of a few days. The brains of infants with and without risk respond differently to syllabic stimuli (ba, da, ga) and this difference has a significant positive correlation with reading at second grade. The differential features of the waveform of the brain event related responses (ERP; between these groups) reveal correlations to indices that also reflect the earlier steps towards reading skill (Guttor et al., 2005; Guttorm, Leppänen, Hämäläinen, Eklund, & Lyytinen, 2010; Lyytinen et al., 2005). The next observable difference, at age 6 months, reveals compromised speech perception of phonemic duration. Apparently, this is the most difficult-to-differentiate feature of spoken Finnish among children with reading problems, as mentioned above. The first indications were present in categorical perception performance of 6-month-olds (Richardson, Leppänen, Leiwo, & Lyytinen, 2003). Also, as early as 6 months of age, the ERP to change in the phonemic duration feature in the auditory stream revealed a differential response between the groups with and without dyslexia (Leppänen, Pihko, Eklund & Lyytinen, 1999; Leppänen et al., 2002) and also predicted letter knowledge and naming fluency (Leppänen et al., 2012).
Furthermore, those at-risk newborns who ended up facing reading difficulties at age 8 had atypical ERPs to sound frequency changes and showed, before reading age, compromised perceptual differentiation of phonemic duration (Leppänen et al., 2010).

The next important aspect of language development that deserves attention among children who are at familial risk for dyslexia is the delay in expressive language. Late talking was the earliest easy-to-observe indicator to cue the need for help along the developmental route towards acquisition of good reading skills. Children with familial risk and whose speech is delayed at 2 years of age have an elevated risk of facing difficulties in reading acquisition, as can be gleaned from Figure 2. If receptive language (comprehension of spoken messages) is similarly delayed, early support in language development is especially important. Otherwise, the risk of facing difficulties in the attainment of sufficient literacy to comprehend written material in line with the expected developmental milestones is quite high.

![Figure 2. Pisa reading scores reflecting the degree of full literacy including reading comprehension from children who had delayed language development at 2.5 years of age.](image)

The developmental routes that precede the acquisition of reading skill vary. The three most apparent routes, all of which are connected to somewhat compromised reading during school age, are illustrated in Figure 3, together with a typically developing group with no difficulties in reading acquisition. To illustrate this summary, the entire battery of language measures collected in the JLD were fed into the same latent profile analysis (via MPLUS) to highlight those children with sufficient familiarity to form a group sharing a relatively similar developmental route and also to uncover developmentally different groups.
Figure 3. Average profiles of developmental language characteristics for different subgroups across ages 1 to 6.5 years, their average performance in reading and writing composite score during the second grade, and in PISA reading composite at 15 years of age. (Modified from Figure 1 in Lyytinen et al., Merrill-Palmer Quarterly, 2006).

The classic phonological difficulty can be observed therein as a relative decline of phonological development (among 35/199) during the critical pre-school years. The reading level of this group at the end of the second grade is one standard deviation below norm in comparison to other children.

A similarly compromised level of reading skill characterizes the second developmental route, which reveals somewhat delayed early language skills in nearly all measures, with special difficulties in the readiness to rapidly name familiar objects at five and half years of age (Dysfluent group, n=12). The most surprising group, which we call “Unexpected” (n=67), had atypically good early language development. This group’s mean level of reading, however, reached a mean composite score of -.5 (resulting from low scores among a small portion of members of this subgroup) of reading only, during the second grade. The surprising end result consequently revealed by this group is, however, that its achievement in the PISA-assessment at age 15 reached a score +.5 above the mean of all the
participating children – higher than those of the typically developing final group, as shown in Figure 3 (for more, see Lyytinen et al., 2006).

There is one very easy-to-assess measure that seems to be accurate in the identification of the children who will face problems in their reading acquisition. This is “letter knowledge.” Almost in the order that the JLD-children were able to store letter names during the years before they entered school, they learned to read, as revealed by Figure 4. This figure illustrates the individual profiles of all children who faced problems in their reading acquisition in the JLD study. Almost all Finnish children are exposed to letters portrayed on the walls of the kindergarten (obligatory environment for Finnish children), although teachers are advised not to begin reading instruction before children enter school at age 7. This means that children have had the opportunity to become interested in letters and to learn the names. Ultimately, some do not learn the letters in line with the majority. As revealed by Figure 4, from amongst all of the most well-known early age predictors of reading acquisition shown in the profiles, those children whose reading development still continues to be atypical through the early grades are differentiated, without exception (with scores below the level of typical children), by their letter knowledge. No other measure is perfect in the same sense.

We always find that some children who are failing to learn to read in a typical way have a score above the mean on another early predictor included in the figure.

Letter knowledge – especially learning the letter sounds – plays a pivotal role in reading acquisition of transparent orthographies such as Finnish, German, Italian and Spanish. Thus, it is not surprising to find that if we measure letter knowledge and phonological awareness at different ages and model their relationship, we can find what is observed in Figure 5. In transparent orthographies, where letters have a consistent connection to reading, the necessary phonemic awareness is most efficiently supported by the visible letters that the children are interested in naming. Thus, they become motivated to store important cues associated with their sounds. This helps children to gain sufficient impetus from the sounds of the letters and thus prepares them to take the first step towards basic reading skill with efficiency. Those children who need help are therefore easy to identify by using only one simple assessment. Following on from this realisation that early identification can be achieved with such relative ease, naturally, the next step is to know how to help.

Before going on to illustrate the interventions that can help, even when commenced only shortly before school entry, it is important first to consider the implications of the earlier-mentioned results associated with delay in receptive language and its predictive relation to reading comprehension.
Predicting reading fluency

We now know (as illustrated earlier, see Figure 2) that from among the half of children with a familial background of reading problems who face difficulties in learning to read, those whose spoken language development is delayed need early help to support their vocabulary development. Receptive skills have to be strengthened from early on to help them become sufficiently prepared to approach full literacy following reading instruction.

Children with or without familial risk and without any substantial delay in the development of early language, which can be observed in their use of spoken language, can be helped to acquire basic reading skills by starting effective preventive training that focuses specifically on the learning of written language skills just before, or at least at the point of, school entry. At least this is the case if they can use a training regimen which is comparable to the ‘Graphogame’, a learning environment that we have developed for children with a risk factor associated with reading acquisition. Graphogame is used in such a way that those children for whom reading acquisition is predicted to be more difficult and consequently more time-consuming in comparison to the requirements of their peers, can start to use it before such a difference can be observed. If Graphogame is implemented no later than school entry and used in an optimal way, potential obvious disparities between the child in difficulty and his or her peers are unlikely to be noticeable throughout the process of reading skill acquisition. ¹

¹ Footnote 1. What follows informs how Figure 5 can be understood. It reveals an example from a situation where the /N/ sound (in the centre) has been repeatedly heard by the leaner in more than 100 trials when this
Graphogame – an enjoyable mobile or computer game for learning to read: How it helps at risk children to overcome the fuzziness of the phonemic representations with letters

Figure 5. The classical Graphogame display (left, not available any more) illustrating the still applied principle of how GG works. The learner is choosing from the falling balls, the written item corresponding to the one s/he hears from headphones. The right side illustration informs how the "phonemic space" of the learner is becoming differentiated. For an explanation, see footnote 1.

Prevention of Reading Difficulties

We all know that there are different reasons why reading acquisition is not equally easy for all children. Nonetheless, one could argue that the same core content must be learned by all concerned. Ultimately, the initial ‘diagnosis’ of why a child faces difficulty in acquiring initial reading skills may have little effect on what the learner has to do in order to
actually learn basic reading skills. One has to acquire the connections between spoken and written language. This is relatively straightforward under two conditions.

The first condition is that the learners must master the spoken form of the language that is to be read. At least some kind of implicit phonological sensitivity associated with the speech sounds must be available before anyone can begin to connect sounds to letters. As we now know, this condition has not always been met. This has been the case in Africa where attempts have been made to instruct the reading of English to children who have never been exposed to spoken English. Also, in the developed world, many children are instructed to read a language that is not the most familiar language to the learner. Such situations compromise the opportunity to achieve the goal of reading before implicit awareness of the sounds of the intended-to-be-read language have been formulated in the mind of the learner.

The second condition that affects the establishment of connection-building between spoken and written language depends on the consistency of the connections. In transparent orthographies, consistency can be perfect (e.g. Finnish) or partly compromised (e.g. German, where the spelling direction is not fully consistent). If there are inconsistent connections, a way in which to help the learner avoid being confounded by inconsistencies must be found when organizing the instruction. This may be a challenge for English, where the appropriate segmentation of the speech that helps to isolate the relevant sound units that are always connected to the same written unit is a challenging computational problem. Instruction must begin with large units which behave consistently, independent of the context in which they occur, and then inform those contexts where the less consistently behaving connections are valid.

The instruction of initial reading skill can be organized simply and effectively in the aforementioned Graphogame (GG) learning environment. To illustrate its operations, see Figure 6 (left part). When training basic reading within a transparent orthography that is consistent at the letter-sound level, the computer display shows a variable number of falling balls, each containing a different single letter. Simultaneously, the learner hears the sound of one of the visible letters. The task is to choose the letter that represents the sound. Children quickly learn to choose the correct letter which represents the sound in each trial. This makes the connection-building very efficient because new sets of letters are presented all the time and the child receives sufficient repetition to store the connections. The GG program selects items for the subsequent trials by taking into account the knowledge demonstrated by the learner during the preceding trials and this dictates the new content. Thus, it is possible to ensure that the practice is very rewarding because, in most cases, the learner can make a successful choice (e.g. 80% of the time, which we have found to be the best proportion of ‘easy’ trials) and thus, motivation is sustained. This also guarantees sufficient repetition to ensure storage.

A very simple procedure is run in the initial stages with a content of sound-wise and visually easy-to-differentiate items whereby the number of falling alternative written units is few. Next, the number of alternatives from which the learner must choose is increased before
the more difficult-to-differentiate phonemes/letters are included. This may culminate, if necessary, in drilling of minimal pairs to differentiate the phonemic space where their perception is not yet accurate and reliable. After the sounds of the single letters have been reliably learned, the unit size increases to 2 or more letters in each item. This motivates the learner to invent the principle of assembly, how the sounds merge together following the sequence of letters. Thus, syllables, words and even sentences can be used to take steps towards accurate decoding skills.

The present GG developments also include fluency training and we are moving now to support reading comprehension. Further information is available from the info.graphogame.com pages. The way in which related support services are organized for all Finnish children (due to public procurement from us by the Ministry of Education) can be seen at www.lukimat.fi (for an English description push the flag at the top right corner and see the further English description in the next page). Training in basic math skills is also provided within this service. The pages contain substantial information for teachers and parents concerning these basic scholastic skills, ways to learn them and ways to overcome related learning problems.

The provision of very detailed feedback makes such a preventive training environment invaluable to teachers and special education experts. An example is shown on the right side of Figure 5, where the illustration models a hypothetical reflection of the differentiation of phonemic space in the learner’s mind. It illustrates an individual’s learning status, which can be computed and drawn to represent a summary of a group. Any phase of the learning process can be selected in order to identify any as yet unopened bottlenecks, such as which sounds/letters are not sufficiently differentiated. This is revealed by distributions between the target sound in the center and the chosen letters printed at the ends of each distribution. If the distribution has its widest part in the most distant position, the wrong choice has ceased. In the case of severe dyslexia, the learner may require individual tailoring of the content and sequences introduced. Almost all children can reach the goal without further tailoring because of the automatic adaptation that is implemented to organize the steps in the learning process, beginning with easy and small items, moving towards the most demanding-to-differentiate items, and then to larger items by introducing the need to assemble sounds together. The adaptation procedure efficiently accommodates the bottlenecks, which are systematically opened by using the approximately 20% of difficult items to keep the child happy and motivated but by not repeating these too often.

Interested readers can learn more about the use of Graphogame via the published documentation on its efficiency. Saine, Lerkkanen, Ahonen, Tolvanen and Lyytinen (2011) illustrate how a special education teacher can achieve good results by implementing Graphogame as part of face-to-face teaching in small groups. When given the opportunity to play Graphogame as part of their remedial teaching, struggling children, irrespective of the severity or type of their initial risk factors, reached the level of the mainstream children before the end of the third grade. Very comparable results were observed for the elevation of
reading fluency (Heikkilä, Aro, Närhi, Westerholm, & Ahonen, 2013; Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2010).

Today, the illustrated Graphogame is utilized in a research capacity in more than 30 countries, some of which are close to achieving a similar level of implementation to that in Finland where all children can receive help from Graphogame. At best, more than 20,000 children use Graphogame in a single school day. Such a large number of users from amongst Finland’s 60,000 age cohort is understandable because the support is appropriate not only for Grade 1, but also for second graders, whose automatization of decoding skills can also benefit.

Conclusion

My experience has made me optimistic to the extent that we believe that, with optimal use, such an efficient training environment can help most, if not all, children who face difficulties in learning to read. Optimal use entails short (<15 min) sessions, repeated at least three times per day over several subsequent days until the goal of accurate reading skill has been reached. The later fluency training may be conducted with slightly longer sessions when children are older. Overly lengthy sessions increase the risk of boredom, which is why we do not recommend acceptance of long playing sessions. We are also trying to motivate parents not to introduce children to Graphogame too early because if it is not helping at the time that the child starts using it, it may not be of interest to her/him at the time that it can help. We know that the probability of achieving good gains from its use are substantially larger if started close to school entry when the brain of even delayed children is sufficiently mature. However, it must be added that this training is complementary to contact and support from a human teacher. Teachers’ and parents’ supportive advice and their genuine demonstrations of happiness gleaned from witnessing the child’s training is important and should be directed as positive reinforcement towards the child. It is also necessary that, soon after the child has learned the basic reading skill which is trained in the GG-environment to a sufficiently fluent level, the learner is introduced to a lot of reading in order to achieve full literacy. For this purpose, exciting and sufficiently easy-to-begin-with reading material should be made readily available, as soon as the skill is sufficiently prepared. As mentioned at the beginning of this story, learners who become strongly motivated to read a great deal will overcome their problem and attain the status of those ‘compensated’ readers whose competence cannot be differentiated from that of other readers.

In the high income countries, Graphogame can help children with a risk of facing difficulty in learning to read and, most importantly, children who are resistant to the typical instruction available. This means a relatively low (<10) percentage. However, in Africa, the conditions to ensure sufficient instruction are often lacking, as shown in the most recent UNESCO summary (UNESCO, 2014). This is why the percentage of children who can benefit is very high. One principle in the declaration signed with our collaborators from
universities such as Cambridge, Oxford, Harvard, Stanford, Yale and Zurich (see http://info.graphogame.com/wpuploads/2011/04/GraphoWORLDDeclaration110216.pdf) is that the efficiency of the Graphogame implemented in a given language/writing system/educational culture has to be empirically validated before starting any distribution of the support service.

References


Predictors of Positive Mood and Negative Mood among Children with Learning Disabilities and Their Peers

Adi Sharabi
Kibbutzim College and Tel Aviv University, Tel Aviv, Israel

Malka Margalit
Tel Aviv University, Tel Aviv and Peres Academic Center, Rechovot, Israel

Abstract

Positive and negative mood have often been considered indicators of wellbeing, affecting behavior and adjustment. This study evaluated the personal and familial predictors of positive and negative mood among 1,024 Israeli students (children with learning disabilities [LD]: 302 boys, 198 girls; children without LD: 308 boys, 216 girls). They were students in Grades 4-6 (ages 9-12). The authors aimed to identify the direct and indirect contributions of personal and familial factors to positive and negative mood. The participants’ socio-emotional characteristics were assessed with questionnaires measuring mood, hope, effort, family climate, loneliness and the use of the Internet for communication with friends and virtual friends. Structural equation modeling (SEM) results confirmed the conceptual model and showed that the indirect effect of LD on positive and negative mood was mediated by effort, hope, family climate and loneliness. Gender predicted positive mood, loneliness and family cohesion. The differential contribution of types of e-communication extended the understanding of the model, and underscores the role that the children’s perceptions and behavior play in their emotional outcomes.

Positive and negative mood have often been considered indicators of adjustment and psychological wellbeing (Forgas, 2013). Moods are fundamental psychological states that can arise endogenously or in response to an event. They influence how individuals interpret the world around them and are mediated by individuals’ perceptions of their social connections and directing behavior (Kok et al., 2013). Their importance for children with learning disabilities (LD) has been documented, although these children are initially
identified by their academic challenges. Research has acknowledged their experiences of social and emotional distress (Bauminger & Kimhi-Kind, 2008; Bryan, Burstein, & Ergul, 2004; Sharabi & Margalit, 2011b), reporting increased loneliness, tendencies to depression and more social challenges when compared to their typically developing peers (Maag & Reid, 2006). Studies have also focused attention on the diminishment of their hopeful thinking and lower levels of effort investment (Heath & Wiener, 1996; Margalit, 2010). The recent increase in the use of e-communication for different types of social connections (with friends and with strangers) amplifies the distinctions between the experiences of loneliness among adolescents with LD and their peers (Sharabi & Margalit, 2011b). Differences in family climate between children and adolescents with LD have also been linked to the children’s adjustment (Al-Yagon, 2012; Sullivan, 1993). However, a comprehensive model of the predictors of children’s positive and negative mood, including personal and familial perceptions, has not yet been constructed.

The goals of the current study were to propose an integrated conceptual framework for predicting positive and negative mood among children with LD and their peers, using their participation in two types of e-communication and through personal and familial perceptions. We proposed three exogenous variables: LD, two types of e-communication and gender. The contribution of hope, effort, loneliness and family climate factors (cohesion and adaptability) were examined as mediators of mood. Figure 1 illustrates the proposed model, which is grounded in a theoretical rationale for integrating the predictors of mood.

Positive and Negative Mood: The Dependent Variables

Positive and negative mood are generalized, pervasive, affective states that are less intense than emotions, but have profound effects on thought processes, behaviors and wellbeing (Bono, McCullough, & Root, 2008; George, 2011). Mood may serve as an immediate and direct source of information about people’s internal state of affairs and the available resources they have for meeting environmental challenges, preparing them to react by approaching or withdrawing from their goals (Larsen, 2000; Tillema, Cervone, & Scott, 2001). Research has demonstrated that positive and negative mood are two separate activation systems that operate independently, but also interact with each other (Carver & Scheier, 1990; Watson, Wiese, Vaidya, & Tellegen, 1999). Their impact on cognitive processing (such as attention and memory), academic performance and behavior have been widely documented in adult research (Hamann, 2009; Osaka, Yaoi, Minamoto, & Osaka, 2013). Several studies have considered negative mood as stronger than positive mood (Baumeister, 1999; Forgas, 2013), while others have emphasized the strength of positive affect in promoting wellbeing (Fredrickson & Joiner, 2002) and buffering the effects of negative affect (Riskind, Kleiman, & Schafer, 2013). Thus, both types of mood have important roles as indicators of wellbeing.

However, only a few studies have investigated moods reported by children with LD. For example, Sideridis (2007) documented the relationship between academic achievement...
in math and affect. Several studies have focused attention on adolescents with LD, who reported higher levels of negative mood and lower levels of positive mood than their peers (Heath & Wiener, 1996; Lackaye, Margalit, Ziv, & Ziman, 2006; Sharabi & Margalit, 2011b). Given that children with LD participate in the current youth culture of using various types of e-communication, we must also consider their relationship with positive and negative mood.

**E-communication**

The Internet has grown rapidly as a major communication medium. More than half of all American youngsters have used online social networking sites. E-communication enables children and adolescents to expand their circle of friends, create social ties and share information about themselves (Madden et al., 2013). Adolescents report that social networking sites help them manage their relationships with their friends by supporting connections with existing friends and creating new friendships with individuals that they sometimes never meet (Lenhart & Madden, 2007). Among adolescents with LD, e-communication that supports existing friendship predicted reduced loneliness, while virtual friendships predicted higher levels of loneliness (Sharabi & Margalit, 2011a). The current study focused on a younger group of children, in order to explore the contribution of these two types of e-communication to their positive and negative mood.

**Gender and LD**

Research on the social and emotional experiences of children and adolescents with LD has focused on their cognitive and affective challenges (Al-Yagon & Margalit, 2013), showing the heterogeneity of their challenges, as well as gender and age differences. Inconsistent results were reported for gender comparisons among children with LD and their peers. In several studies, girls reported higher levels of negative mood than boys, lower levels of positive mood (Sharabi & Margalit, 2011b) and decreased levels of loneliness (Idan & Margalit, 2014). In other studies, no significant group or gender differences were found (Lackaye et al., 2006), implying the need to consider the roles of personal factors (i.e., hope) and interpersonal dynamics (i.e., family climate) as mediators.

**Hope and Effort**

Snyder (2002, p.249) defined hope as “the perceived capability to derive pathways to desired goals, and motivate oneself via agency thinking to use those pathways.” This definition suggests that hope is not an emotion, but rather a dynamic, motivational, personality characteristic (Phan, 2013). Accordingly, the guiding assumption is that human actions are goal directed, and goals provide the targets for the sequences of mental actions. Agentic thinking refers to people’s ability to motivate themselves to pursue goals. Pathways are the cognitive routes or strategies to achieving a goal (Snyder, 2002). High hope children can identify pathways and strategies for attaining these goals and master the mental energy needed to pursue these pathways. Although they may have confidence in their pathways, they can also identify alternative pathways when they encounter obstacles in achieving their goals (Hellman, Pittman, & Munoz, 2013).
Hope is related not only to a positive future outlook, but also to effort and academic engagement (Adelabu, 2008). Students who exhibit heightened feelings of hope are more likely to invest time and effort in school activities, and to achieve academic success (Valle, Huebner, & Suldo, 2006). Studies have revealed that adolescents with LD often reported lower levels of hope and less investment of effort than their peers without LD (Lackaye et al., 2006). A longitudinal study that examined the developmental trajectory of hope over four years among adolescents revealed that girls reported higher levels of hope than boys in Grade 7, but lower levels of hope by Grade 10 (Heaven & Ciarrochi, 2008).

Loneliness

Social relations play an important role in children’s wellbeing. Therefore, it is not surprising that loneliness has been considered a distressful emotional experience that affects children’s quality of life and is considered a major developmental risk for their future adjustment (Margalit, 2010). Studies have also documented the contribution of loneliness to additional psychological problems such as social phobia, depression, and anxiety (Lasgaard, Goossens, Bramsen, Trillingsgaard, & Elklit, 2011). In addition, increased feelings of loneliness were found to be related to poorer academic achievement (Cheng & Furnham, 2002). However, studies have found inconsistent results with regard to loneliness among children with LD. Several studies documented their social distress. Children with LD reported higher levels of loneliness from an early age (Al-Yagon, 2003; Yu, Zhang, & Yan, 2005). In contrast, several other studies (Wiener, 2003) that sensitized awareness to environmental factors and focused attention on the important contribution of parental roles and family climate to feelings of social alienation failed to find increased levels of loneliness among children with LD.

Family Climate

The family system plays a major role in children's adaptation and development (Al-Yagon, 2012). Olson identified two main dimensions of the family system: adaptability and cohesion (Olson & Gorall, 2003). Adaptability refers to the ability of the family system to be flexible in their adaptation to environmental changes such as recognizing the special and extended needs of children with LD. Cohesion reflects the emotional bonding and closeness between family members, and their ability to provide social support for the struggling child.

In an earlier study (Idan & Margalit, 2014), adolescents with LD reported lower levels of family adaptability and decreased family cohesion. They considered their family more rigid and less able to adapt to the new reality presented by the adolescent with LD. While family cohesion was negatively related to loneliness in both groups (with and without LD), family adaptability was negatively related to loneliness only among adolescents without LD (Idan & Margalit, 2014).
Theoretical Integration and the Purpose of the Present Study

In summary, research has demonstrated that the mood of children with LD reflects their struggle with challenges and is related to lower levels of hope and effort, high levels of loneliness and differential use of e-communication. In addition, the studies have also confirmed the value of families as a source of support. Utilizing a comprehensive theoretical framework may facilitate the identification of mood predictors and ultimately advance the approaches to supporting the students’ wellbeing. Based on prior studies that examined the relationships between these measures separately, the current study proposed an integrated conceptual framework through which we may better understand how e-communication, LD and gender interact with individual perceptions (hope and effort) and interpersonal perceptions (loneliness and family climate) as well as predict positive and negative mood. Thus, as presented in Figure 1, our aim is to present and examine the complex and interactive nature of several predictors of children’s wellbeing. We hypothesize that positive and negative mood will be explained, not only directly by the children’s characteristics as exogenous variables (LD and gender), as well as their technology supported interpersonal forms (e-communication, and virtual friendship) but also indirectly by personality and familial characteristics. Acknowledging earlier research that emphasized the significant role of family climate variables as well as children’s hopeful thinking, effort investment and loneliness distress, we expected that these variables would mediate the impact of the personal and interpersonal exogenic variables. Especially, we expected that the children’s LD would predict, directly and indirectly, the emotional experiences and wellbeing. Yet, LD is just one of the four exogenous variables, focusing attention on the prediction of the gender, and technology-supported interpersonal relations mediated by the family variables (cohesion and adaptability), as well as of hopeful thinking, social distress (loneliness) and effort investment.

Method

Participants

The sample consisted of 1,024 children with and without LD: students in Grades 4 to 6 (ages 9-12) from 28 large elementary schools in urban areas of central Israel. The participants were divided into two groups: 500 children with LD (302 boys, 198 girls) and 524 students without LD (308 boys, 216 girls). The non-LD comparison group was matched by gender and class to the children with LD in their classes, and in every class they were randomly selected. There were no significant differences between the two groups of children in the proportion of the sample by gender ($\chi^2$(1df) =0.28, p=.60) or by class ($\chi^2$(2df) =0.19, p=.91).

Students in the LD group were previously diagnosed with LD as their primary disability using the Israeli Ministry of Education’s criteria and the Israeli Special Education Law. These criteria included the presence of a verbal and/or performance IQ score in the average range or above (ranging from 85 to 120), achievement scores at least one standard deviation below the expected score in one or more areas of academic functioning, and
evidence of a processing deficit in one or more cognitive or linguistic domains. Each child was diagnosed in reading, writing, and/or mathematics. The most common difficulties in reading and writing were slow reading and spelling inaccuracies.

The local municipal psycho-educational agencies and each school's psycho-educational team conducted the diagnostic evaluations. Assessments included the Wechsler Intelligence Scale for Children-Third Edition (Wechsler, 1991) and/or the Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983), the Bender-Gestalt Test (Koppitz, 1975), and the Hebrew adaptation of the Rey Auditory Verbal Learning Test (Vakil & Blachstein, 1993), as well as evaluations of reading and writing levels in Hebrew. Due to confidentiality directives, group data were available for this study, rather than specific information about individual children's disabilities. Students with special difficulties other than LD were not included in the sample.

Instruments

Mood. The Hebrew adaptation (Margalit & Ankonina, 1991) of the Mood Scale (Moos, Cronkite, Billings, & Finney, 1987) is a 20-item measure of students' perceived mood. The scale is divided into two subscales: 10 positive mood items (e.g., “happy,” “in control,” “friendly”) and 10 negative mood items (e.g., “sad,” “tired,” “worried”) rated on a 5-point Likert-type scale, ranging from 1 (not at all appropriate) to 5 (very appropriate). In the current study, a Cronbach’s alpha of .81 was obtained for the positive mood items, and .80 for the negative mood items.

Use of the Internet for communication. Two items from the Internet Self-Report Scale (Sharabi & Margalit, 2011a) were used in this study to assess two types of e-communication on a 5-point Likert-type scale ranging from 1 (not at all) to 5 (a lot). The first item assessed the extent of e-communication with friends: “I keep in touch with my friends from everyday life on the Internet.” The second item assessed e-communication with virtual friends: “I have friends who I know only through the Internet.”

Hope. The Hebrew adaptation (Sharabi, Levi, & Margalit, 2012) of the Children’s Hope Scale (Snyder, 2002) taps the belief in one’s ability to pursue desired goals and use strategies to achieve them. The scale consists of six statements to which children respond on a 5-point Likert-type scale, ranging from 1 (almost never) to 5 (almost always). There are three agency items (e.g., “I think I am doing pretty well”) and three pathways items (e.g., “I can think of many ways to get things in life”). The hope score is the average of the six items on the scale. Thus, a higher score reflects a higher level of hope. In the current study, internal consistency (Cronbach’s alpha) for the scale was .74.

Effort. The Effort Scale (Sharabi et al., 2012) consists of four items to which children respond on a 5-point Likert-type scale, ranging from 1 (almost never) to 5 (almost always), with items such as “I don’t give up even when it is difficult for me.” In the current study the internal consistency for the measure (Cronbach’s alpha) was .70.

The Family Climate Scale (FACES III). The Family Adaptability and Cohesion Evaluation Scale (Olson, 1986) assesses the degree of flexibility and cohesion within the
family. The Hebrew adaptation (Teichman & Navon, 1990) of the scale consists of 20 items with a 5-point Likert-type scale, ranging from 1 (almost never) to 5 (almost always), comprising two subscales of 10 items each. The adaptability subscale refers to the flexibility within the family, focusing on how the family system balances stability versus change (e.g., “We shift household responsibilities from person to person”). The cohesion subscale refers to emotional bonding, family boundaries, and time spent together (e.g., “Family members feel closer to other family members than to people outside the family”). The scores are the averages of the items on each one of the subscales. Thus, a higher score reflects a higher level of adaptability and cohesion within the family. In the current study, a Cronbach’s alpha of .69 was obtained for adaptability and .80 for cohesion.

Loneliness. The Hebrew adaptation (Margalit, 1991) of the Loneliness Scale (Asher, Parkhurst, Hymel, & Williams, 1990) consists of 16 primary items rated on a 5-point Likert-type scale, ranging from 1 (never) to 5 (always), for example, “I have many friends in my class”, “I have nobody to talk to in my class”, “I am lonely.” The additional eight filler items cover various activity areas and were not included in the analysis. Higher scores reflect higher levels of loneliness. In the current study, a Cronbach’s alpha of .92 was obtained.

Procedure
Prior to data collection, approval was obtained from the Israeli Ministry of Education, the schools, and the parents. Teachers and school counselors identified the fourth to sixth grade students with LD in their schools. Students with and without LD from the same classes completed the set of questionnaires about mood, hope, effort, family climate, loneliness, and Internet use as a group in their classrooms during school hours. Students were informed that participating in the study was voluntary and that the information would remain confidential. The examiner read some of the items aloud for those students with LD who were entitled to such accommodations. Completion of the questionnaires took from 35 to 45 minutes. Afterwards, the students with LD were matched with students without LD by class and gender, and randomly selected.

Data Analysis
In the first stage, the Statistical Package for the Social Sciences (SPSS 21) was used for computing the descriptive statistics, Cronbach’s alphas, and bivariate correlations to examine the relationships between the measures. In addition, to test group differences (children with LD vs. non-LD/boys vs. girls/learning), analyses of variance were performed including partial η² as an estimate of effect size (Cohen, 1988).

The conceptual model illustrated in Figure 1 was tested and analyzed using structural equation modeling (SEM), carried out using AMOS 21.0 (Arbuckle, 2012). We used SEM, since Path analysis is a special case of SEM, and has a more restrictive set of assumptions than SEM (McDonald, 1996; Xue, 2007). It corrects for measurement error and can estimate both direct and indirect (mediated) effects simultaneously. The SEM analyses provided information about the direct and indirect associations between mood (positive and negative),
hope, effort, family climate (adaptability and cohesion), loneliness, two types of e-communication, and belonging to the LD/non-LD groups and gender. An additional advantage of using SEM is that it produces estimates of total effects (i.e. the sum of the overall direct effects and overall indirect effects). In evaluating the fit of models, multiple indices of fit, corresponding to different types of fit evaluation (Hoyle & Panter, 1995), were reported. Constructs were represented with item parcels (i.e. sums of items), with parcels created randomly (Bandalos & Finney, 2001).

The model was tested using several indices (Hu & Bentler, 1999). Specifically, the fit was assessed using Hu and Bentler’s (1999) root mean square error of approximation (RMSEA) and Bollen’s goodness-of-fit index (GFI). According to Hu and Bentler (1999), values of 0.06 and lower on the RMSEA, and 0.95 and higher on the GFI indicate a good fit between the model and the data. In addition to reporting the chi-square test statistic and the chi-square by degrees-of-freedom value as measures of absolute fit, a normed-fit index (NFI) and comparative fit index (CFI) were reported as measures of incremental fit (Bentler & Bonett, 1980).

Before turning to the results, it is important to add a word of caution about the language we use in reporting the findings of our study. The results are explicated in terms of associations rather than effects, as causal effects cannot be established in the absence of experimental data. However, an important feature of SEM analysis is that it allows for distinctions between direct, indirect, and total effects. Thus, in discussing the SEM results, we have chosen to use the word “effect”, but its use does not imply that causality has been established with any certainty.
Results

Descriptive Analysis

To decrease the chance of Type 1 errors, a multivariate analysis of variance (MANOVA) was conducted with groups (LD/non-LD) and gender (boys/girls) as the independent variables, and positive and negative mood, hope, effort, loneliness, family adaptability, family cohesion and two internet e-communication types as the dependent variables. The MANOVA yielded a main effect for group, $F(9, 1012) = 6.86, p< .001$, partial $\eta^2 = .057$; and a main effect for gender, $F(9, 1012) = 5.49, p< .001$, partial $\eta^2 = .047$; but the interaction was not significant. The follow-up univariate analysis (see Table 1) revealed a main effect for children with LD on the following variables: e-communication with friends and virtual friends, loneliness, hope, effort, and family adaptability. The follow-up univariate analysis for gender revealed significant differences in the following measures: negative mood, loneliness, e-communication with friends, effort and family cohesion. The correlation analysis revealed that for both groups of children, e-communication with friends was associated with positive mood, while e-communication with virtual friends was related to the negative mood.

Students with LD expressed lower levels of hope and effort as well as fewer e-connections with everyday friends through the Internet. They reported more e-connections with strangers, considered themselves lonelier than their peers, and assessed their family system as more adaptable to changes. No significant differences were found in mood or family cohesion. With regard to gender differences, boys indicated that they had higher levels of negative mood and loneliness. They connected with virtual friends more than girls did. They invested less effort in their studies and reported lower levels of cohesiveness in their families. No significant differences were found in the remaining measures. Table 2 shows the correlations among the research variables.

Estimations of the Model

SEM was used to identify the factors that explain positive and negative mood, and their relationship to learning disabilities, e-communication with friends and virtual friends, gender, family cohesion and adaptability, loneliness, effort, and hopeful thinking. This section describes the estimations of the hypothesized base-model via SEM (AMOS 21.0), which measured the interactions between the model’s components.

According to the conceptual model (see Figure 1), learning disabilities, gender and the two e-communication measures (the exogenous variables) were expected to contribute (i.e. have direct paths) to effort, hope, the two family measures (cohesion and adaptability) and loneliness (the mediating endogenous variables), as well as to positive and negative mood (the endogenous variables). Direct paths were also expected between the mediating endogenous variables – family cohesion and adaptability, loneliness, effort and hope – and positive and negative mood. This model also assumed relationships among all of the exogenous variables themselves. Specifically, it was also expected that the mediating family
measures of cohesion and adaptability would have direct paths with hope, effort and loneliness. Given that gender differences were identified in several of the research measures, they were also included in the model. The measurement model established the connections between the constructs in the model and the underlying data that defined them.

Testing of the current model was conducted in two steps of estimation: (1) the base model and (2) the modified model composed only of the pathways that emerged as significant in the first step.

**Table 1**

*Group and Gender Means, SDs and F Scores*

<table>
<thead>
<tr>
<th>Variable</th>
<th>SLD</th>
<th>Non-SLD</th>
<th>F(1,1024)</th>
<th>Partial Eta^2</th>
<th>Boys</th>
<th>Girls</th>
<th>F(1,1024)</th>
<th>Partial Eta^2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive mood</td>
<td>4.14(0.65)</td>
<td>4.15 (0.65)</td>
<td>.42</td>
<td>.000</td>
<td>4.16 (0.66)</td>
<td>4.14 (0.64)</td>
<td>34</td>
<td>.000</td>
</tr>
<tr>
<td>Negative mood</td>
<td>2.43(0.81)</td>
<td>2.34 (0.82)</td>
<td>3.34</td>
<td>.003</td>
<td>2.46 (0.80)</td>
<td>2.29 (0.82)</td>
<td>10.61**</td>
<td>.010</td>
</tr>
<tr>
<td>E-communication</td>
<td>3.25 (1.63)</td>
<td>3.52 (1.55)</td>
<td>10.00**</td>
<td>.010</td>
<td>3.45 (1.55)</td>
<td>3.30 (1.65)</td>
<td>2.40</td>
<td>.002</td>
</tr>
<tr>
<td>Virtual friend</td>
<td>2.04 (1.45)</td>
<td>1.87 (1.38)</td>
<td>3.63</td>
<td>.004</td>
<td>2.14 (1.51)</td>
<td>1.69 (1.22)</td>
<td>25.53**</td>
<td>.024</td>
</tr>
<tr>
<td>Hope</td>
<td>3.98 (0.68)</td>
<td>4.18 (0.64)</td>
<td>27.37**</td>
<td>.026</td>
<td>4.06 (0.68)</td>
<td>4.11 (0.64)</td>
<td>1.14</td>
<td>.001</td>
</tr>
<tr>
<td>Effort</td>
<td>4.16 (0.73)</td>
<td>4.31 (0.65)</td>
<td>12.83**</td>
<td>.012</td>
<td>4.20 (0.71)</td>
<td>4.30 (0.67)</td>
<td>4.88*</td>
<td>.005</td>
</tr>
<tr>
<td>Adaptability</td>
<td>3.33 (0.68)</td>
<td>3.22 (0.69)</td>
<td>5.62*</td>
<td>.005</td>
<td>3.27 (0.70)</td>
<td>3.27 (0.66)</td>
<td>.00</td>
<td>.000</td>
</tr>
<tr>
<td>Cohesion</td>
<td>4.19 (0.63)</td>
<td>4.22 (0.63)</td>
<td>.29</td>
<td>.000</td>
<td>4.17 (0.66)</td>
<td>4.26 (0.59)</td>
<td>4.79*</td>
<td>.005</td>
</tr>
<tr>
<td>Loneliness</td>
<td>1.89 (0.81)</td>
<td>1.74 (0.77)</td>
<td>9.30**</td>
<td>.009</td>
<td>1.87 (0.82)</td>
<td>1.73 (0.75)</td>
<td>7.78**</td>
<td>.008</td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses are standard deviations.

* p<.05; ** p<.01

**The Base Model**

The first step in the analysis was to test the measurement model of the scales. This involved a confirmatory factor analysis procedure that tested assumptions about the factor structures of the various scales. Parceling, a measurement practice used with latent-variable analysis techniques, was used to reduce the number of items (Little, Cunningham, Shahar, & Widaman, 2002). In this approach, items from the same scale are aggregated into multiple parcels (i.e., mini scales) and then used as indicators of the latent variable. This approach is commonly employed to obtain more consistently distributed variables and to reduce the number of parameters in the structural equation model, thereby creating a more optimal
variable-to-sample-size ratio. A parcel can be defined as an aggregate-level indicator that comprises the average of two or more items. In this model, two parcels were created for every latent variable. The random assignment of items to parcels led to the creation of parcels that contained roughly equal common factor variance (Little et al., 2002).

Table 2

*Pearson’s Correlations among Variables for Students with and without LD*

<table>
<thead>
<tr>
<th>LD Group (n = 500)</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Positive Mood</td>
<td></td>
<td>-</td>
<td>.34**</td>
<td>.53**</td>
<td>.45**</td>
<td>.31**</td>
<td>.48**</td>
<td>-.45**</td>
<td>.23**</td>
</tr>
<tr>
<td>Negative Mood</td>
<td>-.48**</td>
<td>-</td>
<td>-.27**</td>
<td>-.25**</td>
<td>.11*</td>
<td>-.17**</td>
<td>.51**</td>
<td>-.03</td>
<td>.14**</td>
</tr>
<tr>
<td>Hope</td>
<td>.55**</td>
<td>-.44**</td>
<td>.62**</td>
<td>.31**</td>
<td>.42**</td>
<td>-.35**</td>
<td>.09*</td>
<td>.01</td>
<td></td>
</tr>
<tr>
<td>Effort</td>
<td>.46**</td>
<td>-.40**</td>
<td>.65**</td>
<td>.231**</td>
<td>.44**</td>
<td>-.28**</td>
<td>.06</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td>Family Adaptability</td>
<td>.19**</td>
<td>.02</td>
<td>.10*</td>
<td>.05</td>
<td>-</td>
<td>.49**</td>
<td>-.09*</td>
<td>13**</td>
<td>.26**</td>
</tr>
<tr>
<td>Family Cohesion</td>
<td>.47**</td>
<td>-.30**</td>
<td>.42**</td>
<td>.38**</td>
<td>.42**</td>
<td>-</td>
<td>-.28**</td>
<td>.07</td>
<td>.06</td>
</tr>
<tr>
<td>Loneliness</td>
<td>-.52**</td>
<td>.54**</td>
<td>-.49**</td>
<td>-.34**</td>
<td>-.10*</td>
<td>-.34**</td>
<td>-</td>
<td>-.22**</td>
<td>-.02</td>
</tr>
<tr>
<td>E-communication</td>
<td>.16**</td>
<td>-.08</td>
<td>.08</td>
<td>.01</td>
<td>.04</td>
<td>.02</td>
<td>-.23**</td>
<td>-</td>
<td>.32**</td>
</tr>
<tr>
<td>Virtual Friends</td>
<td>.05</td>
<td>.14**</td>
<td>-.09*</td>
<td>-.14**</td>
<td>.12**</td>
<td>.01</td>
<td>.05</td>
<td>.33**</td>
<td>-</td>
</tr>
</tbody>
</table>

*p<.05; **p<.01

Non-LD Group (n = 524)

SEM analysis demonstrated a good fit between the conceptual model (see Figure 1) and the studied measures: the $\chi^2$ test was significant, $\chi^2$ (df=70, N=1,024) = 108.499, $p$ = .002, and indices values were high: NFI = .985, CFI = .995, GFI = .988, and RMSEA = .023. In sum, step 1 revealed a good fit between the conceptual model and the empirical data. However, given that several paths were not significant, we expected that a modified model would result in a better fit.

**The Modified Model**

The second step in the analysis tested a modified and more concise model of students’
positive and negative mood (see Figure 2) by considering only the paths that emerged as significant in the base model (Byrne, 2001). This analysis revealed a very strong fit between the modified model and the empirical findings: the $\chi^2$ test was non-significant, $\chi^2 (88 \text{df}, N=1024) = 98.041, p = .218$, and the indices-of-fit values were high: NFI = .986, CFI = .999, GFI = .990, and RMSEA = .011.

In the modified model, the exogenous variable LD had a direct path to hope ($\beta = -.07, p < .01$), effort ($\beta = -.11, p < .01$) and family adaptability ($\beta = .08, p = .014$). Children with LD rated their families as more flexible, and reported lower levels of hope and effort in engagement. The exogenous variable of e-communication with friends had a direct path to positive mood ($\beta = .10, p < .01$) and negative mood ($\beta = .13, p < .01$), loneliness ($\beta = -.22, p < .01$) and effort ($\beta = .08, p = .019$). The exogenous variable of e-communication with virtual friends (strangers) had a direct path to effort ($\beta = -.11, p < .01$) and adaptability ($\beta = .19, p < .01$). Students with higher levels of e-communication with friends had higher levels of positive mood as well as increased negative mood. They reported lower levels of loneliness and more engagement in academic effort. However, children who e-communicated with strangers considered their families more flexible and reported less academic effort. The exogenous variable of gender had a direct path to the endogenous variables of positive mood ($\beta = -.10, p < .01$), loneliness ($\beta = -.09, p < .01$) and family cohesion ($\beta = .08, p = .014$). Boys
rated their families as more cohesive, and reported higher levels of loneliness and lower levels of positive mood than girls.

The following mediating variables had a direct path to positive mood: hope ($\beta = .29$, $p < .01$), family cohesion ($\beta = .29$, $p < .01$) and family adaptability ($\beta = .15$, $p < .01$). Students with higher levels of hope reported higher levels of positive mood, as did students with higher family cohesion and adaptability. In addition, two mediating variables had a direct path to negative mood: family adaptability ($\beta = .10$, $p < .01$) and loneliness ($\beta = .75$, $p < .01$). Students who reported higher levels of family flexibility and increased loneliness also reported higher levels of negative mood.

Hope had a direct path to loneliness ($\beta = -.41$, $p < .01$) and positive mood ($\beta = .30$, $p < .01$). In addition, there was a direct path from effort to hope ($\beta = .77$, $p < .01$). Direct paths were found from cohesion to adaptability ($\beta = .53$, $p < .01$), effort ($\beta = .54$, $p < .01$), loneliness ($\beta = -.38$, $p < .01$) and positive mood ($\beta = .29$, $p < .01$), emphasizing the important positive role of family cohesion. Adaptability had a direct path to loneliness ($\beta = .20$, $p < .01$), negative mood ($\beta = .10$, $p < .01$) and positive mood ($\beta = .15$, $p < .01$). The endogenous variable of negative mood had a direct path to positive mood ($\beta = -.41$, $p < .01$), meaning that students who reported higher levels of negative mood had lower levels of positive mood.

The exogenous variables were also related. LD was negatively associated with e-communication with friends ($\beta = -.08$, $p < .01$) and positively related to virtual friends ($\beta = .06$, $p = .056$). E-communication with virtual friends was also related to e-communication with friends ($\beta = .31$, $p < .01$) and gender ($\beta = -.14$, $p < .01$).

Summary of the Model

The modified model can be summarized as follows (see Table 3). The measures that contributed directly to positive mood were hope, negative mood, family cohesion and adaptability, gender, and e-communication with friends. LD contributed to positive mood only indirectly through family adaptability, as well as through effort and hope.

Negative mood was predicted directly by loneliness, e-communication with friends, and family adaptability. Indirectly, LD contributed to negative mood through loneliness, effort through hope, and family adaptability. Hope was predicted directly and indirectly by LD, directly by effort, and indirectly by family cohesion, e-communication with friends (a positive relationship) and with strangers (a negative relationship). Family variables (adaptability and cohesion) contributed differently to mood (positive and negative), loneliness, hope, and effort. LD and virtual friends contributed only indirectly to positive and negative mood.
Table 3
Direct, Indirect and Total Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Direct</th>
<th>Indirect</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>On Mood (positive)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>.00</td>
<td>-.06**</td>
<td>-.06**</td>
</tr>
<tr>
<td>Gender</td>
<td>-.10**</td>
<td>.08**</td>
<td>-.02</td>
</tr>
<tr>
<td>E-communication</td>
<td>.10**</td>
<td>.04*</td>
<td>.14**</td>
</tr>
<tr>
<td>Virtual friends</td>
<td>.00</td>
<td>-.03</td>
<td>-.03</td>
</tr>
<tr>
<td>Hope</td>
<td>.29**</td>
<td>.13**</td>
<td>.42**</td>
</tr>
<tr>
<td>Effort</td>
<td>.00</td>
<td>.32**</td>
<td>.32**</td>
</tr>
<tr>
<td>Loneliness</td>
<td>.00</td>
<td>-.31**</td>
<td>-.31**</td>
</tr>
<tr>
<td>Family adaptability</td>
<td>.15**</td>
<td>-.11**</td>
<td>.04</td>
</tr>
<tr>
<td>Family cohesion</td>
<td>.29**</td>
<td>.31**</td>
<td>.61**</td>
</tr>
<tr>
<td>Negative mood</td>
<td>-.41**</td>
<td>.00</td>
<td>-.41**</td>
</tr>
<tr>
<td><strong>On Mood (negative)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>.00</td>
<td>.07**</td>
<td>.07**</td>
</tr>
<tr>
<td>Gender</td>
<td>.00</td>
<td>-.09**</td>
<td>-.09**</td>
</tr>
<tr>
<td>E-communication</td>
<td>.13**</td>
<td>-.18**</td>
<td>-.06*</td>
</tr>
<tr>
<td>Virtual friends</td>
<td>.00</td>
<td>.08**</td>
<td>.08**</td>
</tr>
<tr>
<td>Hope</td>
<td>.00</td>
<td>-.31**</td>
<td>-.31**</td>
</tr>
<tr>
<td>Effort</td>
<td>.00</td>
<td>-.24**</td>
<td>-.24**</td>
</tr>
<tr>
<td>Loneliness</td>
<td>.75**</td>
<td>.00</td>
<td>.75**</td>
</tr>
<tr>
<td>Family adaptability</td>
<td>.10**</td>
<td>.15**</td>
<td>.25**</td>
</tr>
<tr>
<td>Family cohesion</td>
<td>.00</td>
<td>-.28**</td>
<td>-.28**</td>
</tr>
<tr>
<td><strong>On Hope</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>-.07*</td>
<td>-.09**</td>
<td>-.16**</td>
</tr>
<tr>
<td>Gender</td>
<td>.00</td>
<td>-.03*</td>
<td>.03*</td>
</tr>
<tr>
<td>E-communication</td>
<td>.00</td>
<td>.06*</td>
<td>.06*</td>
</tr>
<tr>
<td>Virtual friends</td>
<td>.00</td>
<td>-.09**</td>
<td>-.09**</td>
</tr>
<tr>
<td>Effort</td>
<td>.77**</td>
<td>.00</td>
<td>.77**</td>
</tr>
<tr>
<td>Family cohesion</td>
<td>.00</td>
<td>.41**</td>
<td>.41**</td>
</tr>
<tr>
<td><strong>On Loneliness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LD</td>
<td>-.00</td>
<td>.08**</td>
<td>.08**</td>
</tr>
<tr>
<td>Gender</td>
<td>-.09*</td>
<td>-.04*</td>
<td>-.13**</td>
</tr>
<tr>
<td>E-communication</td>
<td>-.22**</td>
<td>-.02*</td>
<td>-.24**</td>
</tr>
<tr>
<td>Virtual friends</td>
<td>.00</td>
<td>.07**</td>
<td>.07**</td>
</tr>
</tbody>
</table>
Negative mood was predicted directly by loneliness, e-communication with friends, and family adaptability. Indirectly, LD contributed to negative mood through loneliness, effort through hope, and family adaptability. Hope was predicted directly and indirectly by LD, directly by effort, and indirectly by family cohesion, e-communication with friends (a positive relationship) and with strangers (a negative relationship). Family variables (adaptability and cohesion) contributed differently to mood (positive and negative), loneliness, hope, and effort. LD and virtual friends contributed only indirectly to positive and negative mood.

**Discussion**

The objective of the present study was to examine the personal and familial predictors of positive and negative mood among children with LD and their typically developing peers. Although negative emotional experiences are often dramatic and impactful, and have captured researchers’ interest more than positive experiences, the current study investigated both (positive and negative) emotional experiences and showed their relationships to children with LD. The study demonstrated a strong fit between the conceptual model and the empirical findings as well as the pattern of relationships between the model’s components. The results underscored the complexity of LD related challenges, but pointed to the need to include personal and familial qualities in predicting outcomes. The model provided important information about the different direct and indirect effects of gender, e-communication, hopeful thinking, loneliness and perceptions of family characteristics on children’s wellbeing as reflected in their mood. Before addressing the results of the SEM analyses, the following section briefly discusses the findings of the preliminary analysis.

**Preliminary Analysis**

As a preliminary analysis, the two groups of children were compared with regard to the tested variables. Children with LD reported lower levels of hope and effort, and higher levels of loneliness than their typically developing peers. They considered their family more
flexible than that of their peers. In part, these results supported early studies that focused on adolescents with LD who reported lower levels of hope and effort, and increased loneliness (Lackaye et al., 2006). The family climate results also supported in part the results of family climate differences among adolescents with and without LD (Idan & Margalit, 2014). Similarly, younger children with LD in the current study reported that the adaptability and flexibility of their family climate were higher than that of their typically developing peers. However, the expected differences in family cohesion were not found in the current study. With regards to gender differences, boys reported higher levels of negative mood and increased loneliness compared to girls, and they had more virtual friends. Girls reported higher levels of effort investment than boys, and they considered their families as more cohesive and supportive.

However, the reported differences should be treated with caution (as further explained in the limitation section), considering the large sample, medium effect size of the general MANOVA, and especially the small effect sizes of the univariate analyses. No significant interaction was found, and both variables were considered in the model.

**Estimation of the Model**

The results supported the hypotheses in part, emphasizing the important role of personal and familial perceptions in predicting the children’s wellbeing. In line with the proposed conceptual model, the results showed that positive and negative mood were predicted by personality and familial characteristics, which mediated the exogenous variables: the children’s characteristics (LD and gender), as well as their technology, supported interpersonal behavior (e-communication, and virtual friends). Acknowledging the significant roles of family climate variables as well as children’s hopeful thinking, effort investment and loneliness distress that have been identified by earlier research, the current model confirmed that these variables mediated the impact of the personal and interpersonal exogenous variables. Thus, the roles of family cohesion and adaptability, as well as of hopeful thinking, loneliness and effort investment were confirmed in the model together with the direct contribution of exogenous variables.

In the current study, learning disabilities contributed to positive and negative mood only through hope, effort and family adaptability. An earlier study showed that children with LD were less hopeful and invested less effort in their studies (Al-Yagon, 2011; Lackaye et al., 2006). Regarding the first mediator variable – children’s hope – the present study provided additional support for the relevance of hope theory in explaining the association between LD and mood. Hope was positively associated with positive mood and negatively associated with the experience of loneliness that contributed to negative mood.

With regard to the mediator of loneliness, LD contributed to it only indirectly, together with additional factors that will be further clarified. The outcomes of experiencing loneliness support its importance as a risk to wellbeing. While loneliness contributed to negative mood, an examination of the factors that contributed to loneliness revealed a complex picture. Loneliness was predicted negatively by gender (boys more lonely than
girls) and by the two family measures: negatively by family cohesion, and positively by family adaptability. It was also negatively predicted by e-communication with friends and by hope. Thus, these results extended our understanding of the research inconsistencies with regard to the relationship between loneliness and LD (Pavri & Monda-Amaya, 2000; Wiener, 2003; Yu et al., 2005).

Several personal and familial factors including LD and gender contributed to experiences of loneliness and accentuated the advantages of the comprehensive conceptual model. Gender also predicted family cohesion (with girls in more cohesive families), and positive mood. Yet, these connections have to be treated with caution, since although they were significant in fact they were low. The current study also pinpointed the comprehensive aspects of loneliness. As expected, it was related to social connections with children who used e-communication for supporting existing social relationships and reported lower levels of loneliness. Social support from the cohesive family also contributed to reducing loneliness. The family’s flexibility was negatively related to loneliness, perhaps reflecting the impact of reduced stability and the psychological price of making changes within the family system (Olson & Gorall, 2003). Hopeful thinking also contributed to reducing loneliness, underscoring the personality aspects in the construct.

The family’s adaptability and cohesion provided an index of the family context. The learning disabilities in the current study contributed directly to the family’s flexibility, but not to its cohesion. These families have to adapt to the changes dictated by the extended needs of their child’s challenges. The increased flexibility takes an emotional toll, reflected in higher levels of loneliness and higher levels of negative mood. However, the family’s flexibility also contributed to positive mood when it was combined with family cohesion and hope, reflecting the importance of different relationships. For example, regardless of their academic challenges, children with LD who were engaged in their academic work and made a focused effort in school felt more hopeful, and experienced higher levels of positive mood and lower levels of loneliness. Indeed, the family’s increased flexibility may result in increased distress expressed in greater loneliness and higher levels of negative mood. However, the increased openness to change may also predict positive mood, when combined with the family’s social support (i.e., cohesion) and hope. While prior studies have examined the impact of LD on separate aspects of children’s perceptions, the proposed conceptual model in the current study demonstrates that LD should be treated only as a risk factor. When it is mediated by personal strengths such as hope, social initiatives using e-communication, and combined familial strengths such as cohesion and flexibility, different outcomes may be expected.

The use of e-communication with friends and strangers was also associated with LD, and contributed to the mood outcomes directly and indirectly. Prior studies of e-communication reported that adolescents used it to preserve and support existing social relationships with actual friends from everyday life, and predicted lower levels of loneliness (Sharabi & Margalit, 2011a, 2011b). In the current study, e-communication with friends predicted both types of mood among children with and without LD. These findings provided
further support to the consideration of positive and negative mood as two separate systems that operate independently but also interact with one another (Watson et al., 1999). E-communication with strangers predicted family adaptability, negatively predicted effort, and was associated indirectly with both types of mood.

These results underscore the relationship between positive and negative mood in the present study. They provide support to the construct of two separate yet interacting activating systems and to the research that has documented the power of positive mood to buffer the effects of negative mood (Riskind et al., 2013). In the current study, girls expressed higher levels of positive mood than boys. Gender comparison has yielded inconsistent results regarding socio-emotional aspects within different age groups (Idan & Margalit, 2014; Lackaye et al., 2006; Sharabi & Margalit, 2011b). In a study that compared gender differences in adolescents with LD, boys reported lower levels of negative mood and higher levels of positive mood than girls (Sharabi & Margalit, 2011b), but in another study, the gender differences were not significant (Lackaye et al., 2006).

Conclusions, Implications, Limitations and Future Directions

In conclusion, the results of this study supported the proposed conceptual model of the relationships between predictors of students’ positive and negative mood. The results emphasized the heterogeneous characteristics of students with LD, showing that personal characteristics, types of e-communication and family qualities mediate their effect on adjustment and wellbeing. The importance of positive mood and negative mood as two separate activation systems should be considered in future studies. The results demonstrated that children who are hopeful have a more positive mood, and family cohesion positively predicts effort in school, hope, and positive mood, and negatively predicts loneliness and negative mood. This applies to children with and without LD.

Implications. The study has important practical implications for education and family intervention. It is essential to sensitize teachers not only to the academic challenges and needs of students with LD, but also to the children’s personal and familial characteristics. The introduction of hope intervention approaches to educational systems (Davidson, Feldman, & Margalit, 2012), and the supporting curriculum and strategies that help to promote a classroom climate of hopeful thinking may enhance the students’ positive mood and resiliency. The study also calls for the development of an educational awareness of the possible complex contribution of e-communication among children. The planned responsiveness of families and educators to the virtual connections among children, and to the different uses of e-communication and their emotional outcomes, may lead to planning future interventions.

The study also focused attention on the important role of family variables, especially family cohesion. The significance of cohesion and support within the family were emphasized. They predicted not only positive mood, but also the adaptability of the family to changes, children’s investment in learning, and decreased loneliness. These results call for
focused effort at empowering families through family therapy and parental counseling in order to enhance their closeness, support and cohesion, as well as their abilities to adapt to the changing demands in their environment. Future programs have to sensitize parents to their important role in enhancing their children’s wellbeing through parental counseling that emphasizes the critical role of family climate in children’s school engagement and well-being.

Limitations. Several limitations in this study call for further research. First, the study focused on the heterogeneous, self-reported characteristics of students with LD. Information from additional sources such as teachers and parents may provide an extended perspective for understanding the wellbeing of children with LD. In addition, the correlational nature of the study requires caution in interpreting its conclusions, emphasizing the need for longitudinal studies that follow the direction of developmental paths and further clarify the impact of the various predictors. More studies are needed to clarify how e-communication and friendship are related to face-to-face companionship with peers, as well as to the loneliness experience.

Future directions. Different research methodologies such as in-depth interviews and daily diary methods will enrich our understanding of the factors that contribute to the wellbeing of children with and without LD, promote an understanding of their struggles and happiness, and lead to the development of effective contextual support.

Considering the indirect impact of the LD, and the small effect sizes of the comparisons, this comprehensive model may refer to children in general, with and without LD. Future studies are needed to further examine the distinct aspects as well as shared characteristics among children with and without LD, as well as between different levels and subgroups of LD.

While it is important to focus attention on the social and emotional aspects of children with LD, there is also the danger of denying their academic and cognitive challenges. We can suspect that they have many negative experiences with academic and social failure during their development. Still, we also hope that many of them experience some success and fulfillment when they overcome those difficulties. In order to determine the mechanisms that promote children’s resilience and success, we need to acknowledge as well as put forward a coherent and comprehensive analysis of their challenges and strengths within various familial and educational systems in broader cultural contexts.

References

Children with LD: Positive and Negative Mood by Adi Sharabi and Malka Margalit


---

**How Persistent is a Diagnosis of Mathematical Disorder at an Early Age?**

**A Longitudinal Study**

Annemmie Desoete  
Ghent University, Association Research Platform “Typical and Atypical Learning” and Artevelde University College, Learn+

Frauke De Weerdt  
Formerly of Ghent University, currently in Clinical Practice

Ruth Vanderswalmen  
Artevelde University College and Ghent University Association Research Platform

Anнемie De Bond  
Artevelde University College and Ghent University Association Research Platform

**Abstract**

The study was conducted to look at differences between children who outgrew and did not outgrow an early diagnosis of mathematical learning disorder (MD; *n* =13), and peers without MD (*n* =13). Children were tested at 5, 6, 7 and 10 years of age. About 54% of the children with an early diagnosis of MD still experienced mathematical difficulties at the age of 10. All 10-year-olds with MD still had more difficulties than peers without MD on fact retrieval. Seriation in kindergarten and spelling and reading pseudo words in elementary school, but not gender and intelligence, predicted whether MD was outgrown. Spatial span best predicted children outgrowing MD. Digit recall was a good predictor of persistent MD. Results emphasize the dynamic aspect of MD and the importance of assessing the numerical and central executive domain of working memory, as well as seriation, reading and spelling, in children at risk for MD.
Children with a specific learning disorder with impairment in mathematics, also called mathematical learning disorder (MD), show a significant degree of impairment in mathematics. Their mathematical abilities are situated substantially and quantifiably below those expected for the individual’s chronological age, causing interference with academic performance. In addition, the MD-related problems cannot be better accounted for by intellectual disabilities or external factors (such as inadequate educational instruction) that could provide sufficient cause for scholastic failure. Finally, the symptoms persist for at least 6 months despite the provision of interventions that target the specific difficulties (APA, 2013; Fletcher, Francis, Morris, & Lyon, 2005; Landerl, Bevan, & Butterworth, 2004; Mazzocco, Devlin, & McKenney, 2008; Mazzocco & Myers, 2003; Passolunghi, Cargnelutti, & Pastore, 2014).

The operationalization and cut-off scores used to define MD have varied substantially (Bartelet, Ansari, Vaessen, & Blomert, 2014; Moeller, Fischer, Cress, & Nuerk, 2012), with measuring mathematics performance multiple times and documenting symptoms persisting for at least 6 months as two of the important criteria to consider in assessing MD (APA, 2013). Although they might contribute to our understanding of the development of mathematical ability, studies whose conclusions do not including stability as a criterion (e.g., Bartelet et al., 2014) might not be generalizable to children who experience math impairments across grades (e.g., De Weerdt, Desoete, & Roeyers, 2013; Mazzocco & Myers, 2003; Pieters, Roeyers, Rosseel, Van Waelvelde, & Desoete, 2013; Toll, Van der Ven, Kroesbergen, & Van Luit, 2011). The latter definition is more restricted than the first one and might indicate a more stable, chronic pattern of MD. However, even the empirical picture provided by the existing studies that include a ‘persistence criterion’ is still far from clear-cut (Geary, 2011a, 2011b; Murphy, Mazzocco, Hanich, & Early, 2007; Vukovic & Siegel, 2010). Although stability is one of the criteria for MD (APA, 2013), some children with MD do appear to outgrow their math problems to some extent (Jordan, Mulhern, & Wylie, 2009), with MD only persistent in between 40% and 63% of cases (Mazzocco & Myers, 2003; Shalev, Manor, & Gross-Tsur, 2005; Silver, Pennett, Black, Fair, & Balise, 1999). A major drawback of previous research is that only a few studies have compared in detail children outgrowing and not outgrowing MD with a control group of children of the same age but experiencing no learning disabilities at the start of elementary school (at the ages of 6 and 7). It also remains unclear whether different general (gender, intelligence, working memory) or more specific (preparatory math skills) factors influence the outgrowing or chronicity of MD during elementary school.

Since mathematics is a complex ability composed of a variety of skills (Bartelet et al., 2014; Dowker, 2005b; Shalev, 2004), the development of mathematics learning in elementary education might not be a linear process, with timed fact retrieval skills, performances on untimed mental arithmetic and number knowledge having different rates of progress. In most studies, researchers hypothesized that fact retrieval difficulties in MD would be rather persistent and show little improvement over time (Jordan, Hanich, & Kaplan, 2003a).
whereas an impairment in procedural skills (e.g., in mental arithmetic or in number knowledge) might reflect a developmental delay with slow improvement over grades (Geary, 2011a, 2011b). Thus, the evidence on stability of problems might depend on the arithmetic skill that is tested. Since several studies in the field of MD emphasize the importance of incorporating a multicomponent approach (Hart, Petrill, & Thompson, 2010; Jordan et al., 2009; Mazzocco, 2009; Simms, Cragg, Gilmore, Marlow, & Johnson, 2013) and because arithmetic is a complex ability composed of a variety of skills (Bartelet et al., 2014; Dowker, 2005b), it will be important to include several timed and untimed math skills simultaneously to address the gap in the previous studies.

The current study was intended to better refine MD assessment by investigating whether all children with an early diagnosis of MD based on the acquisition of arithmetic skills at the ages of 6 and 7 still have problems with timed fact retrieval and untimed mental calculation and number knowledge at the age of 10.

The Contribution of Gender, Socio-economic Status and Intelligence to Mathematics Achievement at Age 10

In MD, a number of studies have provided evidence in favour of balanced gender ratios (Lachance & Mazzocco, 2006; Shalev et al., 2005) with a gender distribution across subtypes of MD that is not significantly different (Bartelet et al., 2014). However, not all studies agree on this topic. Indeed, Landerl and Moll (2010) have lent support for a preponderance of girls with MD and Judge and Watson’s longitudinal study (2011) reported an association between gender and mathematics growth, with girls experiencing smaller growth than boys.

Mixed results were also found for the influence of socio-economic status (SES) on MD. Whereas some longitudinal studies did not find any influence of educational status and profession of parents (Barnes et al., 2014; Krajewski & Schneider, 2009; Navarro et al., 2012; Shalev et al., 2005); others did (e.g., Aunio, Hautamaki, Heiskari, & Van Luit, 2006).

Finally, several studies have demonstrated that intelligence is a strong predictor for mathematics achievement (Geary, Hoard, Nugent, & Bailey, 2012; Passolunghi et al., 2014; Passolunghi & Lanfranchi, 2012; Praet & Desoete, 2014). However, only a few studies have investigated the contribution of intelligence to the persistence of an early diagnosis of MD. Those that did find that this factor had a significant impact: children not outgrowing MD had a lower IQ than children outgrowing MD or control children (Shalev et al., 2005; Stock Desoete, & Roeyers, 2010).

Thus, previous studies have generated a mixed empirical picture of gender, SES and intelligence as contributors to mathematics achievement. One purpose of the current study, therefore, was to compare children outgrowing and not outgrowing MD and peers without MD on these components.
The Contribution of Preparatory Mathematical Abilities to Mathematics Achievement at the Age 10

Research has demonstrated how preparatory mathematical abilities, such as the Piagetian logical abilities seriation and classification and post-Piagetian conceptual and procedural counting knowledge, are able to predict math achievement in primary school (e.g., Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Desoete, 2014; Praet, Titeca, Ceulemans, & Desoete, 2013; Stock, Desoete, & Roeyers, 2009). However, the empirical picture provided by existing studies on the value of preparatory abilities for MD is less clear-cut. Stock et al. (2010) revealed how 7-year-old children with MD already performed worse than 5-year-old (kindergarten) control children on seriation, classification and magnitude comparison. Further, Toll et al. (2011) showed how preparatory mathematical skills, measured at the end of kindergarten, predicted MD in second grade to a lesser extent than did working memory. Attout, Noel and Majerus (2014) confirmed the importance of working memory for order in early calculation acquisition, as well as in recognizing ordinal and magnitude representations.

Thus, previous studies have provided evidence for the need to explore the combined effect of predictors, such as working memory and preparatory mathematical skills, over a longer period of time. This study addressed this need by investigating preparatory abilities in addition to working memory in children outgrowing and not outgrowing MD at age 10.

Reading and Spelling Skills in the Prediction of MD

Several studies have revealed genetic factors as overlapping predictors contributing to typical reading and mathematics abilities (e.g., Davis et al., 2014; Desoete, 2008; Hart, Petrill, & Dush, 2010; Hart, Petrill, Thompson, & Plomin, 2009; Kovas & Plomin, 2006; Kovas, Harlaar, Petrill, & Plomin, 2005, Kovas, Haworth, Petrill, & Plomin, 2007), as well as to atypical development or learning disorders (Alarcon, DeFries, Light, & Pennington, 1997; Gross-Tsur, Manor, Kerem, Friedlander, & Shalev, 1998; Shalev et al., 2001). Recently, using twin and genome-wide analysis, Davis and colleagues (2014) revealed that around one half of the observed correlation between reading and mathematics ability at age twelve is due to shared genetic effects (so-called “generalist genes”). Attout, Fias, Salmon and Majerus (2014) demonstrated shared neural correlates in the intraparietal cortex, suggesting the existence of domain-general, potentially ordinal comparison processes, supported by the left intraparietal sulcus. Strong relationships were also found on a behavioural level among mathematics, reading and spelling (e.g., Davis et al., 2014; Landerl & Moll, 2010). Bull and Scerif (2001) demonstrated a correlation of .61 between reading and mathematics and a longitudinal study by Jordan et al. (2003a) revealed that reading influenced math achievement.

It is estimated that between 3.4% (Badian, 1999) and 7.6% (Dirks, Spyer, van Lieshout, & de Sonneville, 2008) of children with MD also have a reading disorder. Moreover, children with combined reading and mathematical learning disorders experienced more generalized and persistent problems than children with isolated MD (Dirks et al., 2008;
Peng & Fuchs, 2014). Strong relationships were also found among behavioural-level deficiencies in mathematics, reading and spelling skills (e.g., Davis et al., 2014; Landerl & Moll, 2010). Several longitudinal studies have demonstrated a relationship between MD and spelling (Shalev et al., 2005) and between MD and a lower level of reading (e.g., Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Mazzocco & Myers, 2003). For instance, results of a study by Murphy et al. (2007) revealed that performance of children with MD on a pseudo word reading test was worse than performance of control children on the same test, from the age of 5 (kindergarten) to the age of 8 (third grade). A longitudinal study with elementary school children from 6 to 9 years of age showed that children with MD were more likely than control children to experience reading problems (Vukovic & Siegel, 2010). However, other longitudinal studies in elementary school children found no association between reading and MD (e.g., Shalev et al., 2005). Moreover, research examining spelling in children with MD is scarce (Dirks et al., 2008).

In summary, although inconsistent, findings from the literature highlight the importance of attending to reading and spelling skills in order to understand mathematical performance over time. This study aimed to investigate the contribution of reading and spelling proficiency in children outgrowing and not outgrowing MD at the age of 10 years.

The Importance of Working Memory in the Prediction of Achievement in Mathematics

Baddeley (1986, 2000) defines working memory as the active system that regulates complex cognitive behavior and consists of a central executive (CE) attentional control system, answering for the processing aspect of a task and strongly interacting with two domain-specific storage systems, and a multidimensional capacity store. The phonological loop (PL) is responsible for the storage and maintenance of verbal information (Raghubar, Barnes, & Hecht, 2010). The visuospatial sketchpad (VSSP) has similar responsibilities for visual and spatial information (Baddeley, 1986; Barnes & Raghubar, 2014). The episodic buffer was added to the model at a later stage (Baddeley, 2000) and conceptualized as a multidimensional but essentially passive store that can be fed from the other working memory components, from long term memory or through perception (Baddeley, Allen, & Hitch, 2010). To our knowledge, no research on working memory in children with learning disorders has taken this component into account, probably because the episodic buffer has to be seen as a vague, shadowy concept, research on which – in spite of its high importance – is still in its infancy (Baddeley et al., 2010). Since Baddeley’s 1986 model is without doubt the most empirically verified (Miyake et al., 2000), the focus of our study lies in this model.

Several studies have demonstrated that working memory is involved in learning mathematics (Andersson & Lyxell, 2007; Attout, Noel, et al., 2014; Barnes et al., 2014; Bull, Espy, & Wiebe, 2008; De Smedt et al., 2009; Desoete & De Weerdt, 2013; De Weerdt et al., 2013; Gathercole, Alloway, Willis, & Adams, 2006; Geary et al., 2007; Meyer, Salimpoor, Wu, Geary, & Menon, 2010; Passolunghi et al., 2014; Passolunghi & Siegel, 2004; Passolunghi, Vercelloni, & Schadee, 2007; Raghubar et al., 2010; Vukovic & Siegel, 2010).
Mathematical Disorder by Annemie Desoete, Frauke De Weerdt, Ruth Vanderswalmen and Annemie De Bond

However, a recent data-driven cluster analysis by Bartelet et al. (2014) strengthened the notion that MD is a heterogeneous disorder, with their data providing support against the notion that MD is strongly underpinned by domain-general factors, such as working memory. In addition, Peng and Fuchs (2014) revealed in their review that children with reading and mathematical disabilities show more severe working memory deficits than peers with isolated learning disorders. Moreover, they noted that MD and RD children showed comparable verbal working memory deficits, but MD children had more severe numerical WM deficits than RD children.

The present study aimed to add some nuance to the literature by highlighting differences between children outgrowing and not outgrowing MD. Research focusing on this aspect might explain some of the inconsistencies across studies and reveal potential risk factors and strengths that can be used as protective factors.

Research Objectives

The purpose of this longitudinal study was to examine the differences between children outgrowing and not outgrowing their math difficulties during elementary school (by the age of 10 years), although they had been previously diagnosed as children with MD (at ages 6 and 7).

This research objective resulted in the following research questions:

1. Do all children diagnosed with MD at the age of 7 years still have mathematical difficulties at the age of 10 years, or is there a group of children who outgrow MD during elementary school?
2. What is the differential contribution of gender, SES, intelligence, and preparatory mathematical abilities (at age 7) to the development of fact retrieval, mental arithmetic and number knowledge at age 10?
3. What are the differential contributions of working memory, reading and spelling to outgrowing math difficulties?

Method

Participants

A three-year longitudinal study was conducted with a large sample \((n=471)\). Children were followed from kindergarten to elementary school and tested at the ages of 5, 6, and 7 (Stock et al., 2010). At the age of 7, 43 children were retrospectively assigned to the MD group. To belong to this group, one had to score below the 25\(^{th}\) percentile on at least one mathematics test both at age 6 and 7, have a clinical diagnosis of MD, and be non-responsive to remediation (measured by the stability of the below average performance on encoding arithmetic facts into long-term memory and the use of effortful procedures to solve arithmetic problems such as finger counting). We refer to Stock et al. (2010) for an overview of these results.
Fifteen 10-year-old children from the MD group of the large sample were randomly selected for the present follow-up study. For practical reasons it was not possible to follow up more children. In order to control for possible confounding variables, this MD group was matched individually on age, gender and intelligence with 15 children without learning problems who scored above the 25\textsuperscript{th} percentile on all mathematics measures at both ages 6 and 7. The children in the control group all had arithmetical scores above the 50\textsuperscript{th} percentile (although this was not an a priori criterion, it strengthened the fact that control children were not at risk for MD). After the first of two test sessions at age 10, two children with MD dropped out because of their parents’ lack of time. For this reason, the two control children who were individually matched with these two MD children were eliminated from the study as well. Hence, the final sample consisted of 13 children with MD and 13 control children \((n = 26)\). Participant characteristics are presented in Table 1.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No MD Control ((n=13))</th>
<th>MD ((n=13))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in months</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>First test session</td>
<td>69.46 (1.41)</td>
<td>69.92 (1.41)</td>
</tr>
<tr>
<td>Second test session</td>
<td>81.23 (3.83)</td>
<td>81.46 (4.27)</td>
</tr>
<tr>
<td>Third test session</td>
<td>92.62 (4.13)</td>
<td>92.08 (5.30)</td>
</tr>
<tr>
<td>Fourth test session</td>
<td>123.08 (8.69)</td>
<td>121.15 (8.57)</td>
</tr>
<tr>
<td>Male : female</td>
<td>6:7</td>
<td>6:7</td>
</tr>
<tr>
<td>IQ</td>
<td>104.54 (10.63)</td>
<td>104.54 (10.63)</td>
</tr>
<tr>
<td>SES Mother</td>
<td>15.46 (1.61)</td>
<td>15.33 (3.11)</td>
</tr>
<tr>
<td>SES Father</td>
<td>16.17 (2.33)</td>
<td>15.00 (3.42)</td>
</tr>
</tbody>
</table>

\textit{Note.} MD = mathematical disorder group; SES Mother = socio-economic status of the mother, as measured by years of education; SES Father = socio-economic status of the father, as measured by years of education.

### Measures

At all ages, parents were asked to complete a questionnaire concerning familial history and SES, as measured by years of education of both the father and mother (e.g., Shalev et al., 2005). At the age of 5, preparatory mathematical abilities were tested. At the ages of 6, 7 and 10, math performance was assessed. In addition, a shortened intelligence test was administered at the age of 7, while at the age of 10 these children were tested on reading, spelling and working memory as well. For a chronology of the tests, see Table 2.
Intelligence. At the age of 7, an estimated IQ was calculated, using an abbreviated version of the Dutch WISC-III (Wechsler et al., 2005). This shortened version is recommended by Grégoire (2000) and has a high correlation ($r = .93$) with a Full Scale IQ (Kaufman, Kaufman, Balgopal, & McLean, 1996). It consists of four subtests: Vocabulary, Similarities, Picture Arrangement and Block Design.

Table 2

*Tests Used in This Study.*

<table>
<thead>
<tr>
<th>Tests</th>
<th>5 years</th>
<th>6 years</th>
<th>7 years</th>
<th>10 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparatory mathematical abilities</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Intelligence</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
</tr>
<tr>
<td>Reading (words, pseudowords)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Spelling</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
<tr>
<td>Mental arithmetic/number knowledge</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Fact retrieval</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Working memory</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
</tr>
</tbody>
</table>

Mathematics. In kindergarten, at the age of 5, the preparatory mathematical abilities (namely procedural and conceptual knowledge of counting, seriation, classification and magnitude comparison) were individually tested with the Test for the Diagnosis of Mathematical Competencies (TEDI-MATH; Grégoire, Noel, & Van Nieuwenhoven, 2004).

Procedural knowledge of counting was assessed with subtest 1 of TEDI-MATH (Grégoire et al., 2004), using accuracy in counting numbers, counting forward to an upper bound (e.g., ‘count up to 6’), counting forward from a lower bound (e.g., ‘count from 3’), and counting forward with an upper and lower bound (e.g., ‘count from 5 up to 9’) as indications of procedural counting knowledge.

Conceptual knowledge of counting was assessed with subtest 2 of TEDI-MATH (Grégoire et al., 2004). Children were asked: “How many objects are there in total?” or “How many objects are there if you start counting with the leftmost object in the array?” When children had to count again to answer, they did not gain any points, as this was considered to represent good procedural knowledge, but a lack of understanding of the counting principles.
Seriation was tested with subtest 4 of TEDI-MATH. Children had to seriate numbers (e.g., “Sort the cards from the one with the fewest trees to the one with the most trees.”) Classification was also tested with subtest 4, where children had to make groups of cards in order to assess the classification of numbers (e.g., “Make groups with the cards that go together.”).

Magnitude comparison was assessed by comparison of dot sets. Children were asked where they saw the most dots. One point was given for a correct answer.

Raw scores were the number of correct items and were converted into percentile- and z-scores. All z-score conversions were based on the entire sample of a longitudinal mathematical study \((n = 471, \text{see Stock et al., 2010})\). TEDI-MATH is an individual assessment battery that was constructed to detect MD. For a more detailed description of this test, we refer to Stock et al. (2010).

In elementary school, at 6, 7 and 10 years, all children were tested on mental arithmetic, number knowledge and on fact retrieval skills (see Table 2). Mental arithmetic and number knowledge were tested with the Kortrijk Arithmetic Test Revision (KRT-R; Baudonck et al., 2006) at the age of 6, 7 and 10 years. The KRT-R is frequently used in Flemish education as a measure of math achievement (e.g., Stock et al., 2010). Raw scores were the numbers of correct items and were converted into percentile- and z-scores. At the ages of 6 and 7, all z-score conversions were based on the entire sample of a longitudinal mathematical study \((n = 471, \text{see Stock et al., 2010})\). At the age of 10, raw scores were converted to z-scores based on the entire sample \((n = 204)\) of a study of working memory (De Weerdt et al., 2013).

Fact retrieval was tested with the Arithmetic Number Facts Test (TTR; De Vos, 1992) at the ages of 6, 7 and 10. The TTR is a numerical facility test consisting of five subtests with arithmetic number fact problems: addition, subtraction, multiplication, division and mixed exercises. Children have to solve as many items as possible in five minutes; they can work for one minute on every column. The TTR is a standardized test that is frequently used in Flemish education as a measure of number fact retrieval (e.g., Stock et al., 2010). Raw scores were the numbers of correct items and were converted into z-scores.

Reading and spelling. In elementary school, children were also tested on reading and spelling proficiency. At 10 years of age, all children were tested with standardized Dutch reading and spelling measures. Word reading accuracy was assessed using the One Minute Reading Test (EMT; Brus & Voeten, 1999) and pseudo word reading using the Klepel test (Van den Bos, Spelberg, Scheepstra, & de Vries, 1994). Both reading tests consist of lists of 116 unrelated words. Children are instructed to read as many words as possible in one (EMT) or two minutes (Klepel) without making errors. On both tests, the raw scores were the numbers of words read correctly. These raw scores were then converted into standard scores (SS) (mean: 10, SD: 3) and z-scores based on the entire sample \((n = 204)\) of a working memory study (De Weerdt et al., 2013).
Spelling was assessed using Paedological Institute-dictation (PI-dictation; Geelhoed & Reitsma, 2000), a Dutch standardized test in which children have to write down the repeated word from each sentence. The test consists of nine blocks of 15 words. Each block has a higher difficulty level and testing is stopped once a child makes seven or more errors in a block. The raw score was the number of words spelled correctly and was converted into a z-score based on the entire sample (n = 204) of a working memory study.

Working memory. Working memory was tested in elementary school at age 10. Besides the backward digit, word list, listening, and block recall subtests of the Working Memory Test Battery for Children (WMTB-C; Gathercole & Pickering, 2001; Gathercole, Pickering, & Braams, 2002), backward word list recall and backward block recall (e.g., Passolunghi & Mammarella, 2010) were used. In addition, in line with St. Clair-Thompson and Gathercole (2006), all children were tested on spatial span, an adapted version of the Automated Working Memory Assessment (AWMA, Alloway, 2007). Each block consisted of six trials. The task was discontinued if three errors or more were made in one block. Span score was calculated by counting each correct trial as one sixth and adding the total number of sixths – except for backward digit, backward word list, and backward block recall, where the total number of sixths was counted and incremented by one (Imbo, Szmalec, & Vandierendonck, 2009; Smyth & Scholey, 1992).

All tasks were programmed in Affect 4.0 (Hermans, Clarysse, Baeyens, & Spruyt, 2005) and presented on a desktop computer. The main task was not started until the child thoroughly understood the task instructions. For spatial span and backward block recall, a mouse was used as the response device. For backward digit recall, backward word list recall and listening recall, a voice key was used. In all tasks, the experimenter pressed a key in order to trigger the next item (Landerl et al., 2004). As the subtest order of the WMTB-C was followed, all tasks were presented in a fixed order.

Phonological loop. Digit- and word-list recall are measures of the verbal recall of sequences. Children have to repeat sequences of digits or high frequency words (see Figure 1 for a trial representation). Digit sequences were random lists of digits ranging from 1 to 9.

Visuospatial sketchpad. Block recall measures visuospatial recall of sequences of cubes. Places of the nine cubes stayed fixed during the whole task. Cubes that were part of a to-be-recalled sequence were highlighted in orange. Cubes that were not, remained blue. After the sequence ended, a screen with nine blue cubes was shown. Children were asked to repeat the sequence of the orange cubes by clicking on the different blue cubes (see Figure 1).

Central executive. In backward digit recall and backward word list recall, children are required to recall sequences of digits or words in the reverse order (see Figure 1). In listening recall, children are presented with a sequence of spoken sentences (e.g., ‘Lions have four legs’). In the processing task, they have to verify the sentence by stating ‘true’ or ‘false’. In the memorization task, the final word of each sentence has to be recalled in sequence (see Figure 1). In the spatial span, a picture of two identical shapes in which the
shape on the right side has a red dot is shown to the children. In the processing task, the child has to identify whether the shape on the right side is the same as or opposite to the shape on the left. In the recall task, the child has to show the location of each red dot on the shape in the correct sequence (see Figure 1). Backward block recall measures visuospatial recall of sequences of cubes in the reverse order (see Figure 1).

**Figure 1. Visualization of the working memory tasks**

- **Backward digit recall and backward word list recall**
  - 500 ms
  - 1000 ms/
  - max. 15000 ms
  - number or word

- **Backward block recall**
  - 500 ms
  - 1000 ms/
  - max. 15000 ms
  - block

- **Listening recall**
  - 500 ms
  - 2000 ms/
  - max. 15000 ms
  - sentence

- **Spatial Span**
  - 500 ms
  - max. 15000 ms
  - 1000 ms
  - max. 15000 ms
Procedure

**Data collection.** All children were tested (refer to Table 2) by a trained researcher. At 5, 6 and 7 years of age, preparatory math tests were conducted during school hours in a separate and quiet room at school. This lasted for a maximum of 50 minutes. At age 7, a short version of the WISC-III was administered individually, and the duration was approximately 45 minutes. At age 10, all children were tested in a quiet room at home for two different sessions, each session lasting up to 70 minutes. During the first session, tests were used to tap mathematics, reading and spelling. In the second session, working memory tasks were administered. To maximize vigilance and persistence in completing tasks, breaks were included.

**Outlier analysis.** Outlier analysis was done for working memory data. At the sample level, accuracy measures exceeding the group mean by 3 SDs were replaced by values 3 SDs from the group mean (Friedman & Miyake, 2004). Outliers represented less than 1% of responses.

**Statistical analyses.** A 3 (children not outgrowing MD, children outgrowing MD and control) x 3 (z-scores on the TTR at 6, 7 or 10 years) factorial repeated measures analysis was performed to examine fact retrieval performances over time. Since assumptions of normality and homogeneity were not met for mental arithmetic and number knowledge, non-parametric tests were performed using z-scores on the KRT-R Mental Arithmetic and the KRT-R Number Knowledge results. The Kruskal-Wallis test was used to compare the three groups at the age of 6, 7 and 10 years. Performance over time was analyzed with the Wilcoxon signed-rank test. In addition, multinomial logistic regression analyses were carried out to clarify to what extent SES, intelligence, preparatory mathematical abilities, working memory, reading and spelling predicted the probability of outgrowing MD.

Results

**Persistence of Mathematical Disorders**

In line with Vikovic and Siegel (2010), the 25th percentile was used as a cutoff to define mathematical abilities that situate themselves below those expected for the individual’s chronological age, one of the criteria of MD (APA, 2013). Three girls and three boys with an early diagnosis of MD at the age of 6 and 7 years (46% of the MD group) achieved math scores above the 30th percentile on all math tests at the age of 10 years. In addition, there was no longer severe interference with academic mathematics performance, so these children were classified in the group of children who had outgrown MD during elementary school. One of these children, however, revealed a severe and persisting reading disorder at 10 years of age.
In seven children (four girls and three boys) with an early diagnosis of MD at the age of 6 and 7 years, there was still evidence of scholastic failure. Math abilities were situated below the 25th percentile on at least one math test at 10 years of age. Three of the children still had scores below the 10th percentile. One boy scored at the 11th percentile on mental arithmetic and number knowledge. One girl and one boy scored at the 23 percentile, with additional clinical scores for spelling and/or reading (due to a comorbid reading disorder).

Except for one child, all control children (children without learning disorders at the age of 7) achieved math scores above the 25th percentile at 10 years of age. For fact retrieval, scores were even above the 75th percentile. We decided not to exclude the child with math scores below the 25th percentile from the control group, since the focus of our study was on the profiles of the children who did not outgrow MD (MD for 3 years) and the children who did outgrow MD (MD for two consecutive years), and the child’s parents reported no interference with academic performance (APA, 2013) and at least average math performance at school.

Math Performance over Time

A significant main effect on the factorial repeated measures analysis was found for the within factor, “fact retrieval” \( (F(2, 22) = 12.36, p < .001, \eta^2 = 0.53) \). Contrasts revealed that fact retrieval skills at the age of 6 \( (p = .002) \) and 7 \( (p < .001) \) were significantly worse than at the age of 10. Moreover, results revealed a main effect for group \( (F(2, 23) = 27.85, p < .001, \eta^2 = 0.71) \). Both the children not outgrowing MD \( (p < .001) \) and the children outgrowing MD \( (p < .001) \) achieved lower fact retrieval scores than the control children. There was no interaction effect of group and time measurement \( (F(4, 44) = 0.81, p = .525) \).

Kruskal-Wallis tests showed significant group differences for mental arithmetic \( (\chi^2(2) = 8.56, p = .014) \) and a trend for number knowledge \( (\chi^2(2) = 8.56, p = .065) \), as measured at the age of 6. The children who did not outgrow MD differed significantly from the control group on both mental arithmetic \( (p = .016) \) and number knowledge \( (p = .046) \). The children who did outgrow MD differed significantly from the control group on mental arithmetic \( (p = .017) \), with only a trend for number knowledge \( (p = .072) \). Performances of both groups of children with MD did not differ significantly from each other. No significant results were found for mental arithmetic at the ages of 7 \( (\chi^2(2) = 4.54, p = .103) \) and 10 \( (\chi^2(2) = .91, p = .634) \), nor for number knowledge at the ages of 7 \( (\chi^2(2) = 3.58, p = .167) \) and 10 \( (\chi^2(2) = 3.25, p = .197) \). Wilcoxon Rank tests revealed a significant difference only between mental arithmetic performance at the age of 10 and at the age of 6 for the children who did not outgrow MD \( (Z = -2.03, p = .043) \).

Significant differences in performance of the children who outgrew MD over time were found between number knowledge scores at the ages of 6 and 7 \( (Z = -1.99, p = .046) \), at the ages of 6 and 10 \( (Z = -2.20, p = .028) \), and at the ages of 7 and 10 \( (Z = -1.99, p = .046) \), as well as between mental arithmetic scores at the ages of 6 and 10 \( (Z = -2.00, p = .046) \). No
significant differences over time were found for the control group. Math performance over time is plotted in Figure 2.

**Influential Factors**

To detect the influence of intelligence, SES, etc. on children who outgrew MD or not, in comparison with control children, multinomial logistic regression analyses were carried out (see Table 3 for an overview of means and standard deviations). Children with an early diagnosis of MD who did not outgrow their mathematic learning difficulties at the age of 10 performed more poorly on digit recall ($M=3.95; SD=0.14$) than did control children ($M=4.89; SD=0.14$) of the same age. In addition, our data demonstrated poorer spatial span performance in children who outgrew MD ($M=2.11; SD=0.24$) as compared to the control children ($M=3.60; SD=0.27$). The differences between children who not outgrow MD ($M=2.60; SD=0.38$) and control children ($M=3.60; SD=0.27$) were not significant on spatial span tasks.
Logistic regression analyses were conducted with the predictors, producing odds ratios. Odds ratios represent the ratio change in the odds of the event (e.g., belonging to the control group) for a one unit change in the predictor variable (e.g., accuracy of the phonological loop) and may vary from 0 to infinity. An odds ratio can be seen as an estimate of effect size. An odds ratio below 1 indicates a higher risk not to be in the reference group and as such reflects problems if the control group functions as the reference category. In contrast, an odds ratio higher than 1 suggests a higher chance of belonging to the reference category and thus suggests a protective factor if the control group is the reference group. When the odds ratio is 1 (or close to it), no effect is found. Not only their nearness to 1, but also the significance of odds ratios - indicated by the \(p\) value of the Wald statistic - plays an important role in the decision process about the strength of a model.

### Table 3

**Means and Standard Deviations on Predictors of Achievement**

<table>
<thead>
<tr>
<th></th>
<th>No MD Control (n=13)</th>
<th>Not outgrowing MD (n=7)</th>
<th>Outgrowing MD (n=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M (SD))</td>
<td>(M (SD))</td>
<td>(M (SD))</td>
</tr>
<tr>
<td><strong>Preparatory Math Abilities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conceptual counting</td>
<td>0.80 (0.23)</td>
<td>0.33 (0.32)</td>
<td>0.84 (0.19)</td>
</tr>
<tr>
<td>Procedural counting</td>
<td>0.53 (0.10)</td>
<td>-0.20 (0.42)</td>
<td>0.22 (0.19)</td>
</tr>
<tr>
<td>Seriation</td>
<td>0.64 (0.38)(^a)</td>
<td>-0.16 (1.08)(^b)</td>
<td>-0.01 (1.00)(^ab)**</td>
</tr>
<tr>
<td>Classification</td>
<td>-0.06 (0.70)</td>
<td>-0.62 (0.70)</td>
<td>-0.10 (1.01)</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>0.04 (0.23)</td>
<td>-0.19 (0.41)</td>
<td>-0.31 (0.47)</td>
</tr>
<tr>
<td><strong>Reading and spelling</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading existing words</td>
<td>1.00 (1.86)(^a)</td>
<td>-0.26 (0.15)(^h)</td>
<td>0.18 (0.22)(^ab)**</td>
</tr>
<tr>
<td>Reading pseudo words</td>
<td>1.05 (0.15)(^a)</td>
<td>-0.33 (0.18)(^h)</td>
<td>0.36 (0.26)(^ab)**</td>
</tr>
<tr>
<td>Spelling</td>
<td>1.05 (0.10)(^a)</td>
<td>-0.03 (0.20)(^h)</td>
<td>0.25 (0.26)(^ab)**</td>
</tr>
<tr>
<td><strong>Working memory</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Phonological loop</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit recall</td>
<td>4.89 (0.14)(^a)</td>
<td>3.95 (0.14)(^b)</td>
<td>4.39 (0.25)(^ab)**</td>
</tr>
<tr>
<td>Word list recall</td>
<td>3.92 (0.11)</td>
<td>3.67 (0.15)</td>
<td>4.06 (0.14)</td>
</tr>
<tr>
<td><strong>Visuo-spatial sketchpad</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block recall</td>
<td>4.82 (0.15)</td>
<td>4.41 (0.08)</td>
<td>4.36 (0.08)</td>
</tr>
<tr>
<td><strong>Central executive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backward digit recall</td>
<td>3.59 (0.17)</td>
<td>3.12 (0.18)</td>
<td>3.22 (0.07)</td>
</tr>
<tr>
<td>Backward word recall</td>
<td>3.03 (0.12)</td>
<td>2.79 (0.14)</td>
<td>2.92 (0.10)</td>
</tr>
<tr>
<td>Listening recall</td>
<td>2.23 (0.13)</td>
<td>1.93 (0.13)</td>
<td>2.11 (0.13)</td>
</tr>
<tr>
<td>Backward block recall</td>
<td>4.40 (0.10)</td>
<td>4.16 (0.06)</td>
<td>4.11 (0.16)</td>
</tr>
<tr>
<td>Spatial span</td>
<td>3.60 (0.27)(^a)</td>
<td>2.60 (0.38)(^ab)</td>
<td>2.11 (0.24)(^ab)**</td>
</tr>
</tbody>
</table>

*Note. SES = socio-economic status, as measured by years of parental education; \*\(p < .05\); **\(p \leq .01\); ***\(p \leq .001\). \(^a,b\) posthoc indices at \(p < .05\).*
Besides, model fitting results provide information about the significance of the model and log likelihood ratio tests show us to what extent the model changes if we omit a particular predictor. Nagelkerke’s $R^2$ is used to express the explanation power; it ranges from 0 to 1. Finally, the model predicts group membership by trying to classify participants correctly. Due to the small sample size, no more than two predictors at a time were entered in one model. Only predictors with a good fit were combined together until the model could not be improved anymore and, hence, was maximized.

For each influential factor, the best logistic regression model is presented in Table 4.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Prediction of Outgrowing MD or Not, or Having No Learning Difficulties, at 10 Years of Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group comparison</td>
<td>Model</td>
</tr>
<tr>
<td><strong>SES</strong></td>
<td></td>
</tr>
<tr>
<td>Outgrowing MD vs control$^a$</td>
<td>SES Father</td>
</tr>
<tr>
<td>Not outgrowing vs control$^a$</td>
<td>SES Father</td>
</tr>
<tr>
<td>Outgrowing or not MD$^b$</td>
<td>SES Father</td>
</tr>
<tr>
<td><strong>Intelligence</strong></td>
<td></td>
</tr>
<tr>
<td>Outgrowing or not MD$^b$</td>
<td>Intelligence</td>
</tr>
<tr>
<td><strong>Preparatory Math</strong></td>
<td></td>
</tr>
<tr>
<td>Not outgrowing vs control$^a$</td>
<td>Seriation</td>
</tr>
<tr>
<td>Outgrowing MD vs control$^a$</td>
<td>Seriation</td>
</tr>
<tr>
<td>Outgrowing or not MD$^b$</td>
<td>Seriation</td>
</tr>
<tr>
<td><strong>Reading and spelling</strong></td>
<td></td>
</tr>
<tr>
<td>Outgrowing MD vs control$^a$</td>
<td>Klepel</td>
</tr>
<tr>
<td>Not outgrowing vs control$^a$</td>
<td>PI-dictation</td>
</tr>
<tr>
<td>Outgrowing or not MD$^b$</td>
<td>Klepel</td>
</tr>
<tr>
<td><strong>Working memory</strong></td>
<td></td>
</tr>
<tr>
<td>Not outgrowing vs control$^a$</td>
<td>Digit recall</td>
</tr>
<tr>
<td>Outgrowing MD vs control$^a$</td>
<td>Spatial span</td>
</tr>
<tr>
<td>Outgrowing or not MD$^b$</td>
<td>Digit recall</td>
</tr>
<tr>
<td><strong>Note.</strong></td>
<td></td>
</tr>
<tr>
<td>OR = odds ratio; CI = confidence interval; SES = socio-economic status, as measured by years of parental education.$^a$ control group as reference category; $^b$ outgrowing MD as reference category.</td>
<td></td>
</tr>
<tr>
<td>$^*$ $p \leq .05$.</td>
<td></td>
</tr>
</tbody>
</table>
**Socio-economic status.** For SES, only the father’s was a significant predictor, $\chi^2 (2, n = 26) = 6.80, p = .033$, Nagelkerke’s $R^2 = 0.26$. The model could classify 53.8% of the children correctly: 57.1% of the children who had not outgrown their early diagnosis of MD, 0% of the children who had outgrown their MD and 76.9% of the control group without learning difficulties at age 7 were classified correctly.

**Intelligence.** The MD group was carefully matched with the control group on intelligence, in order to rule out the possible influence of this factor. For this reason, only the children who had and had not outgrown the early diagnosis of MD were compared with each other. There was no significant result for the model, $\chi^2 (2, n = 26) = 2.69, p = .101$, Nagelkerke’s $R^2 = 0.25$; 71.4% of the children who had not outgrown MD and 83.3% of the children who had outgrown MD were classified correctly.

**Preparatory mathematical abilities.** Seriation appeared to be the only significant preparatory mathematical abilities predictor, $\chi^2 (2, n = 26) = 6.37, p = .041$, Nagelkerke’s $R^2 = 0.25$. Overall, 57.7% of the children were classified correctly: 57.1% of the children who had not outgrown an early diagnosis of MD and 84.6% of the control group without learning difficulties at age 7 were classified correctly.

**Reading and spelling.** The best model consisted of pseudo word reading and spelling, $\chi^2 (4, n = 26) = 27.21, p < .001$, Nagelkerke’s $R^2 = 0.74$. Log-likelihood tests showed significant results for both pseudo word reading ($\chi^2 (2, n = 26) = 7.32, p = .026$) and spelling ($\chi^2 (2, n = 26) = 7.57, p = .023$). About 71.4% of the children who had not outgrown an early diagnosis of MD, 50% of the children who had outgrown an early diagnosis of MD and 92.3% of the control group without learning difficulties at age 7 years were classified correctly. Overall, the model classified 76.9% of the children correctly.

**Working memory.** The best logistic regression model consisted of numerical and visuospatial central executive working memory components. Digit recall (numerical component) and spatial span (central executive component) appeared to be the best working memory predictors. Relative to the control group, the odds ratios of the children who outgrew or did not outgrow an early diagnosis of MD showed a significant decrease in both domains: the higher the numerical (phonological) and visuospatial central executive working memory scores, the higher the chance of belonging to the control group of children without mathematic learning difficulties at the ages of 6 and 7. The model was significant, $\chi^2 (4, n = 26) = 20.77, p < .001$, Nagelkerke’s $R^2 = 0.63$. Moreover, log-likelihood tests showed significant results for digit recall ($\chi^2 (2, N = 26) = 9.51, p = .009$) and spatial span ($\chi^2 (2, N = 26) = 6.61, p = .037$). The model was able to classify 76.9% of the children correctly: 66.7% were correctly categorized as the group of children who outgrew an early diagnosis of MD they had received at ages 6 or 7.
Discussion

A first aim of this study was to investigate whether children with an early diagnosis of MD, including the criterion of consistency in performance over time (at the ages of 6 and 7) and receiving interventions that target their difficulties, can outgrow their mathematical difficulties by 10 years of age. Fact retrieval, mental arithmetic and number knowledge were measured at the ages of 6, 7 and 10.

The results demonstrated that about 46% of the children with an early diagnosis of MD no longer experienced severe mathematical difficulties in fact retrieval, number knowledge and mental arithmetic at 10 years of age. These children no longer situated themselves substantially and quantifiably below non-diagnosed children at the chronological age of 10. However 54% (7 out of 13) of the children diagnosed as ‘children with MD’ at the age of 7 still experienced severe mathematical difficulties at the age of 10. This was the case for four girls, and hence a balanced gender ratio was found (Shalev et al., 2005). This percentage is in line with previous findings (Mazzocco & Myers, 2003;; Shalev et al., 2005;; Silver et al., 1999), pointing to the fact that although some children seem to be non-responsive to remediation at the beginning of mathematical instruction (ages 6 and 7), we cannot speak of stability of the mathematical difficulties in all children with an early diagnosis of MD.

Math Performance over Time

An objective of the current study was to gain insight into children outgrowing or not outgrowing an early diagnosis of MD. Our results concerning fact retrieval are to some degree in congruence with literature which states that children with MD experience persistent fact retrieval deficits (Jordan, Hanich, & Kaplan, 2003b;; Rousselle & Noel, 2007). In accordance with the cumulative growth model (Aunola et al., 2004; Morgan, Farkas, & Wu, 2009), all children with MD (outgrowing their mathematical difficulties or not) still experience more difficulties in the retrieval of arithmetic facts from long-term memory compared to age matched peers without learning problems (control children) at 10 years of age. These results point to the fact that encoding facts into long term memory is a problem in MD during the whole of elementary school and not only at the beginning of mathematical instruction. Therefore, it will be important to attune instructional strategies to this persisting memory-related difficulty in order to optimize the mathematical learning, and hence, performance, of children with MD.

The findings of the present study are also suggestive of a developmental pathway, with children with MD catching up on number knowledge, especially those who outgrow their early diagnosis of MD. This catching up shows some resemblance to the theoretical lag model of mathematical development. This lag model (Aunola et al., 2004; Morgan et al., 2009) suggests that children with less mathematical knowledge can catch up with their higher skilled peers due to the provision of systematic instruction. In this study, children with MD received extra interventions that targeted their difficulties due to their diagnosis of MD at the
ages of 6 and 7. However, the theoretical lag model seems to be in contrast with the cumulative growth model (Aunola et al., 2004; Morgan et al., 2009) that could serve as an explanation for the remaining severe difficulties of all children with MD diagnosed at an early age, and in fact for problems with retrieval at 10 years of age.

Outgrowing MD or Not and Possible Influential Factors

In congruence with previous MD studies (e.g., Shalev et al., 2005), no gender differences were found. There were an equal number of boys and girls with MD who both outgrew and did not outgrow their mathematical difficulties.

In this study, SES was indicated by the years of parental education. Results revealed that the less educated the father was, the higher the child’s chance of belonging to the group of children who did not outgrow MD. However, the model could not correctly classify the children who outgrew MD. Except for one child, they were all categorized as control children. This might be explained by the longitudinal research of Krajewski and Schneider (2009), revealing that parental SES, as measured by educational status and trained and current professions, only became important at 10 years of age (and not at ages 6 to 8). Intelligence was not significant predictor of outgrowing MD. This result contradicts other findings, for instance those of Stock et al. (2010), and may be due to the small sample size and power of the study.

As to preparatory mathematical abilities, in contrast with previous research (e.g., Praet & Desoete, 2013; Stock et al., 2010), only seriation was a significant predictor in kindergarten. We found no significant results for classification, counting and magnitude comparison tested in kindergarten in the prediction of mathematical performance at the age of 10. However, in line with the importance of ordinal processing in the numerical domains (Attout, Fias, et al., 2014), seriation predicted the failure to outgrow mathematical difficulties in children with an early diagnosis of MD.

This finding might encourage clinicians to select preparatory mathematical subtests in kindergarten with caution when assessing and aiming to define MD at an early age. The fact that none of the other preparatory mathematical abilities was of importance might be clarified by considering the dynamic aspect of mathematical abilities (Shalev et al., 2005). Most preparatory mathematical abilities research is restricted to the first years of elementary school. Hence, one can expect a declining influence of preparatory mathematical abilities such as classification, counting and magnitude comparison on MD over time (e.g., by the age of 10 years). Indeed, Toll et al. (2011) could classify 76.9% of the children correctly if persistence was not taken into account, against 57.1% when controlling for chronicity.

This study also aimed to test whether reading- and spelling-related abilities contributed to the prediction of children with an early diagnosis of MD outgrowing or not outgrowing their mathematical difficulties. We demonstrated that five children who did not outgrow MD and one child who did outgrow MD achieved a score below the 25th percentile on at least one reading or spelling test, and that all reading and spelling accuracy scores were significantly
lower in children with MD than in control children. In addition, children with an early diagnosis of MD still having mathematical difficulties at the age of 10 showed lower spelling and pseudo word reading accuracy scores than peers without learning difficulties. These results are in line with the findings of Geary et al. (2007), Mazzocco and Myers (2003) and Murphy et al. (2007), but contradict the findings of Shalev et al. (2005), who found that spelling, but not reading, was associated with the persistence of MD. However, our longitudinal study provided evidence for the importance of assessing the accuracy of reading pseudo words and spelling in mathematics achievement and in the prediction of whether MD can be outgrown or not.

**Working Memory in the Prediction of MD**

In line with previous studies (Attout, Noel, et al., 2014; Raghubar et al., 2010; Toll et al., 2011) that have demonstrated the importance of working memory for order in early calculation acquisition, working memory was the strongest predictor of whether MD was outgrown or not, even stronger than the children’s reading and spelling scores. Children not outgrowing an early diagnosis of MD performed more poorly on digit recall (phonological loop, numerical working memory) than control children. These results are in line with the meta-analysis of Peng and Fuchs (2014) revealing, among other differences, poorer numerical working memory in children with MD in comparison with control children. In addition, research has shown that the phonological loop plays an important role in mathematical abilities, in particular in mental calculation, mature addition strategies, verbal counting strategies and fact retrieval, especially later on in elementary school (De Smedt et al., 2009; Raghubar et al., 2008). Thus, clinicians might be encouraged to test the numerical working memory component when aiming to predict mathematical proficiency or persistence of MD.

Our findings also revealed poorer spatial span as a central executive component of working memory in children who outgrew MD, compared to the control children. The central executive seems involved in controlling and monitoring complex operations, including the spatial aspects of calculations and the use of concrete representations (De Smedt et al., 2009; McLean & Hitch, 1999). This evidence is in line with the empirical study of Passolunghi and Cornoldi (2008), who also found that a visuospatial central executive task was one of the most important predictors of differences between children with MD and control children at the ages of 8 and 10. However, it remains unclear why this component did not differ significantly between children who did outgrow MD and the control children in this study; this is perhaps due to the small sample size.

Taken together, clinicians might think critically about the selection and interpretation of working memory tests, while being aware of the difference between numerical and visuospatial central executive working memory in the assessment of children at risk. Our results suggest that word list recall, block recall, backward digit and word recall, listening recall and backward block recall may not provide additional prediction of MD, and a shorter
test (with a digit recall and a spatial span task) may be given, reducing administration costs and improving scores because participants are less fatigued.

**Limitations**

We recognize that there are severe limitations to the present study that should be mentioned. This longitudinal study occurred with an individually matched, though small, sample of children with and without an early diagnosis of MD. As a consequence, some results might not be detected due to a lack of power. This might, for example, have been the case for intelligence and spatial span in the prediction of children who outgrew MD. More profound conclusions can only be made by increasing the number of participants with and without MD and by looking at differences between children with isolated MD and children with a combined reading and mathematical disorder (MD+RD). In that way, we would have a more detailed understanding of mathematical development and possible influencing factors in outgrowing or not outgrowing MD. The small number of participants in the current study certainly limited the generalization and implication possibilities of its results.

Moreover, although our sample was carefully matched on gender, age, and intelligence, for practical reasons we were not able to match each child with MD with a control child from the same classroom. This might have additionally affected results in that instruction level, motivation of the teacher, etc., might have influenced the performance of the children, making our findings between children marginally interpretable. In addition, it might be worthwhile to investigate in more depth the evolution in numerical and central executive working memory, and to add language proficiency and ordinal representation and comparison as predictors. Such studies with a larger sample of children with MD and MD+RD are needed before more accurate conclusions about outgrowing MD can be made.

**Implications and Future Research**

The finding that mathematical abilities strongly develop over time, even in children diagnosed at an early age with MD, may have some implications for both education and clinical diagnostic practice.

For education, it seems important that children with an early diagnosis of MD get enough time with a daily re-looping and explicit rehearsal of number facts, since they seem to continue to experience severe difficulties with fact retrieval, even at 10 years of age. Number fact retrieval appeals to rote memory and is taught systematically and straightforwardly in regular education until age 7 in the Belgian curriculum. Although additional research is needed, it seems that children with MD at the age of 10 persist in non-retrieval or effortful procedures such as finger counting or breaking problems into multiple steps, to solve these kinds of problems (Bartelet et al., 2014), since they remain less accurate in encoding arithmetic facts into long-term memory. Therefore, children with MD might benefit from an adjusted speed and adequate support to solve these kinds of arithmetic problems. In addition, since our findings indicated that number knowledge improved over the
years, it might be important to support this number knowledge during educational instruction. Further investigation is needed to determine whether children with an early diagnosis of MD can outgrow difficulties with number knowledge more quickly if help that targets number knowledge difficulties is provided.

In diagnostic practice our findings seem to indicate that one cannot assume that all children diagnosed with MD at an early age (6 or 7) still experience severe learning difficulties at the end of elementary school, even if math problems were present during two consecutive years at the beginning of academic instruction. It is recommended that these children be retested at the end of elementary school to determine whether they outgrow their mathematical difficulties or to confirm the stability of MD based on continuous low arithmetic performance (Fletcher et al., 2005; Murphy et al., 2007). In addition, to get a full picture of children with MD, one should not only assess mathematical performance, but also test seriation skills in kindergarten and the accuracy of reading pseudo words and spelling in elementary school, as well as determine the proficiency of the numerical and central executive working memory components at 10 years of age.

Conclusion

Overall, this longitudinal study revealed that clinicians and researchers should not neglect the results of seriation tasks in kindergarten as a predictor for MD. In addition, in elementary school numerical and visuospatial working memory tasks, reading pseudo words and a spelling test should be included in an assessment battery for children at risk for MD. Finally, nearly half of the children with an early diagnosis of MD seem to outgrow their mathematical difficulties by the age of 10.

References


Mathematical Disorder by Annemie Desoete, Frauke De Weerdt, Ruth Vanderswalmen and Annemie De Bond


Mathematical Disorder by Annemie Desoete, Frauke De Weerdt, Ruth Vanderswalmen and Annemie De Bond


Mathematical Disorder by Annemie Desoete, Frauke De Weerdt, Ruth Vanderswalmen and Annemie De Bond


Essay-Writing Interventions for Adolescents with High Incidence Disabilities: A Review of Research

Lauren L. Valasa
The Pennsylvania State University

Linda H. Mason
The University of North Carolina

Charles Hughes
The Pennsylvania State University

Abstract

Many students with disabilities have written language production deficits. As a result, these students are failing to meet the demands of government-initiated standards, higher education, and employment. In this review, quantitative experimental intervention studies for improving persuasive, narrative, and expository compositions for adolescents, 6th – 12th grade, with high incidence disabilities in the United States are evaluated. The review focuses on standard-based essay writing, specifically U.S. Common Core State Standards (CCSS, 2013). Twenty-six single subject and group studies were reviewed. Effect sizes for group designs and percentage of non-overlapping data (PND) for single-case designs were provided to enable standardized assessment of intervention strength. Quality indicators were used to evaluate strength of designs. Results indicate Self-Regulated Strategy Development and the Strategic Instruction Model as promising intervention approaches for facilitating essay-writing skills for adolescents with high-incidence disabilities. Further intervention research is needed, specifically in relation to CCSS, and to identify methods for supporting maintenance and generalization skills across content-area curricula.

Academic intervention research often reflects classroom curriculum demands. Initiatives, such as the Common Core State Standards Initiative (CCSS, 2013) in the United States (U.S.) and Common European Framework of Reference for Languages (Council of Europe, 2001), mandate that individuals master narrative, persuasive, and expository essay writing across writing tasks. In these contexts, students are also expected to demonstrate skilled keyboarding for the production and publication of written compositions. In the U.S.,
specifically, students are required to begin generalizing essay-writing skills across all genres in all academic content areas by the onset of sixth grade. Thus, it is crucial that teachers understand the skills students need for essay writing and have instructional methods for addressing standards and benchmarks.

Several benefits of the initiatives for writing standards have been identified. Graham and Harris (2013), for example, note that CCSS (2013) writing reforms may result in increased emphasis for writing performance improvement. Increased emphasis for writing performance could positively influence reading comprehension, subject-matter knowledge, and network-communication skills such as email, text, blogs, and social media. In addition, CCSS benchmarks provide a clear understanding of what is needed for writing proficiency and help educators better identify appropriate grade-level skills. Benchmarks also help teachers identify students who may not be meeting standards and who may need remediation. CCSS emphasizes writing instruction as a school-wide responsibility, rather than a skill taught during Language Arts instruction, providing increased opportunities for students to practice using writing skills across settings using content-knowledge.

However, despite CCSS (2013) mandates and advantages, most typically achieving students in the U.S. are not proficient writers. According to results from the National Assessment of Educational Progress report (NAEP, 2011), 80% of students in 8th grade and 70% of students in 12th grade are not writing proficient persuasive (argumentative essays intended to support claims with clear reasons and relevant evidence), narrative (essays intended to develop real or imagined experiences or events), and expository (essays intended to examine a topic and convey ideas, concepts, and information) essays. An NAEP proficient writing level indicates competent writing skills; any level below proficient is indicative of partial mastery. Adolescents with high-incidence disabilities (i.e., learning disability [LD], emotional behavior disorder [EBD], attention deficit hyperactivity disorder [ADHD], speech or language impairment [SLI], and developmental delay [DD]) have even greater difficulty writing the coherent essays required. Students with disabilities often struggle with idea generation, have difficulty using genre-specific text, and generally produce shorter, less structured essays with more errors than their peers without disabilities (Taft & Mason, 2011). Furthermore, written compositions of adolescents with high-incidence disabilities are well below the high standards required for higher education and employment (Mason & Graham, 2008).

Unfortunately, essay-writing intervention approaches for adolescents with disabilities is understudied, leaving teachers without clearly defined evidence-based practices that are directly reflective of CCSS (2013) demands (Graham & Harris, 2013; Graham, Harris, & McKeown, 2013; Mason & Graham, 2008; Taft & Mason, 2011). In a synthesis of writing intervention techniques, Graham and Perin (2007a) reviewed empirical studies (i.e. true and quasi-experimental) conducted with adolescents in fourth through 12th grade. Elements such as providing strategies for planning, writing, and revising, goal-setting, and explicit models were recognized as effective, research-based writing intervention techniques. However, these
techniques were selected as effective for typically achieving students without an analysis of effective techniques for adolescents with disabilities. Moreover, omission of single-case studies may have excluded effective writing intervention research for struggling writers and students with disabilities, as single-case experimentation is a commonly used methodology in special education research (Horner et al., 2005).

Mason and Graham (2008) also reviewed writing intervention programs of study for adolescents in fourth through 12th grade. Although the review was expanded to include single-case designs, only intervention studies for students with LD were included. Recommendations for effective instruction components for students with LD included imbedding self-regulation practices (goal-setting, self-monitoring, self-reinforcements, and self-instructions) throughout instruction, allowing opportunities for teacher-student conferencing, and providing scaffolded, guided practice to foster generalization and maintenance. In a subsequent review, Taft and Mason (2011) synthesized writing research for students with disabilities other than LD by providing an analysis of all types of writing approaches and interventions across all grade levels. Self-Regulated Strategy Development (SRSD), an instructional program used to facilitate strategy use and develop self-regulation skills, was noted to be an effective program for students with disabilities other than LD.

More recently, Graham et al. (2013) conducted a meta-analysis to examine the effects of SRSD writing instruction for all students with LD. Results indicated a large average effect size (ES) of 2.37 for both quasi-experimental and experimental designs. The authors indicated that SRSD was an effective instructional procedure for teaching writing to students with LD and recommended future researchers investigate new writing strategies to address a wider range of tasks across genres and contents.

The reviews of writing research for adolescents with disabilities have noted effective procedures used to improve writing performance; however, reviews have not evaluated intervention with a focus on specific CCSS (2013) writing standards. CCSS calls for student achievement in three writing genres - persuasive, narrative, and expository. Table 1 displays key standards for mastery across genres. With increased demands on students with disabilities to write and generalize persuasive, narrative, and expository essays across all content areas by sixth grade, a review of effective essay interventions with regard to government-initiated standards is timely.

**Current Study**

The purpose of this literature review is to synthesize the research on essay-writing interventions for adolescents with high incidence disabilities and to discuss the extent to which the research base is reflective of high CCSS (2013) initiatives. Quantitative experimental studies focused on interventions for improving narrative, persuasive, and expository essays of students in sixth through 12th grade with high-incidence disabilities are reviewed. Results are organized around the following research questions:
1) What were outcomes of essay-writing interventions for adolescents with high-incidence disabilities across classroom setting, participant characteristics, and narrative, persuasive, and expository writing genres?

2) What was the quality of the research for each writing genre?

3) How well is the research base within each writing genre addressing CCSS (2013)?

Method

Studies included in this review met all of the following criteria: (a) used quantitative intervention methods (i.e., true experiment [randomized control trial], quasi-experimental group [non-randomized group studies, single group pretest/posttest], or single-case) to analyze the effects of an essay-writing intervention on narrative, persuasive, or expository compositions, (b) targeted writing performance of students in sixth through 12th grade in secondary (i.e., middle or high school) settings who were diagnosed with EBD, LD, ADHD, SLI, or DD, and (c) published in peer-reviewed journals. Only interventions for instruction in persuasive, narrative, or expository essay writing were considered acceptable independent variables. Thus, interventions that did not focus on constructing an essay (e.g., editing strategies, fluency interventions, interactive journaling, and note-taking methods) were excluded from this review. Quick writes (i.e., an organized, short constructed response to teacher prompts) and summary writing, while a valuable form of written expression, were also excluded from the review, as the review is intended to focus on essay compositions (Graham & Perin, 2007a; 2007b; Mason, Benedek-Wood, & Valasa, 2009).

To locate studies for review, searches of ERIC, ProQuest Education Journals, and PsychINFO databases were conducted using the keywords writing, writing intervention, middle school, high school, essay intervention, narrative essay, persuasive essay, expository essay, compare-contrast essay, special education, and disability. Keywords and keyword combinations were entered into title, abstract, and descriptor fields. Next, a hand search of reference lists of all identified articles was conducted to avoid potential omissions. Twenty-six studies met inclusion criteria.

Effect sizes (ES; standardized mean difference) based on researcher reported essay strategy-specific elements and quality scores for group studies and percentage of all non-overlapping data (PND: the percentage of data in the treatment phase that exceeds the most positive result documented in the baseline phase) for single-case studies were obtained (Thompson, 2007). Effect sizes, if not reported, were calculated by subtracting the posttest mean of the control group from the posttest mean of the treatment group and dividing by the pooled standard deviation. For pre and posttest or repeated measures quasi-experimental designs, effect sizes were calculated using gain scores. Cohen’s benchmarks for group studies were classified as small (.20), medium (.50), or large (.80 or greater; Cohen, 1988). It is important to note effect sizes from designs with a control group (randomized control trials and quasi-experimental group designs) are not comparable to designs without a control group (repeated measures and pre and posttest designs), as randomized control designs and quasi-
experimental designs could yield smaller effects due to more rigorous design methodology. Effect sizes are calculated to provide standardized information about overall treatment impact on behavior.

PND is a widely implemented and recommended method for quantitatively synthesizing single-case treatment effects and was therefore chosen to evaluate effect sizes in single-case studies (Scruggs & Mastropieri, 2001). As noted in the recent meta-analysis of writing interventions conducted by Graham et al. (2013), PND is the most commonly used method of evaluating effects across writing reviews and meta-analyses. In other words, use of PND makes it easy to situate findings within the current research base and compare findings across other reviews and meta-analyses of writing. Moreover, PND has established benchmarks within writing research for small, medium, and large effects, which is useful for the purpose of this review. PND is presented simply to provide standardized information about overall treatment impact on behavior.

If not researcher reported, PND was calculated by dividing the number of data points exceeding the most positive result in baseline by the total number of data points in the intervention phase and multiplying by 100. PND for single-case studies were also classified as small (50-70%), medium (70-90%), or large (90% and above). While effects sizes and PND are not comparable, each provides meaningful, standardized information about overall treatment impact on behavior and is, therefore, reported to enable standardized assessment of intervention strength (Mason & Graham, 2008). Unless reported by the researcher, all ES and PND calculations were computed using the data provided by study authors.

Quality indicators were also utilized to assess quality of research. Nine indicators adapted from Gersten et al., (2005) by Graham and Perin (2007a), each holding a value of 1 point, were used to evaluate group studies: (1) random assignment of subjects, (2) mortality equivalence between conditions, (3) no ceiling or floor effects for the primary measure, (4) pretest equivalence across conditions, (5) instructor training described, (6) type of control condition described, (7) the Hawthorne effect controlled, (8) treatment fidelity established, and (9) teacher effects controlled (e.g. instructors blind to research questions). For pre and posttest and repeated measures designs, mortality equivalence between conditions, pretest equivalence across conditions, and type of control condition described were considered unmet, as these designs do not involve a control or comparison condition. As a result, randomized control trials and quasi-experimental group designs that included a comparison condition were able to meet a higher number of indicators, signifying a stronger design with greater internal validity. Additionally, if no confound for the Hawthorne effect was documented, the Hawthorne effect was assumed controlled.

Quality indicators for single subject studies (Horner et al., 2005) were adapted by Mason and Graham (2008) and were reported for groups of studies based on persuasive, narrative, and expository genre. One point was assigned for each of the 11 indicators: (1) participants adequately described, (2) participant selection adequately described, (3) description of instructional setting adequately described, (4) dependent measures quantified,
(5) dependent measures reliable, (6) multiple baseline data points collected, (7) multiple intervention points collected, (8) treatment fully described, (9) treatment fidelity established, (10) testing procedures adequately described, and (11) social validity established. Indicator criteria were based on procedures described by Horner et al. (2005). All quality indicators were calculated for group and single-case designs in this way by the first author.

Results

Findings are organized into three main sections to address research questions: participant and setting characteristics, intervention outcomes, and CCSS (2013) standards (see Table 1). A total of 26 studies were reviewed. Table 2 displays study design characteristics, participants, and setting. Eight studies were randomized control trials (RTC), six used quasi-experimental group designs, three utilized a pre and posttest single group method, and nine used single-case designs.

Fifteen studies implemented a persuasive essay-writing intervention, three studies implemented a narrative essay-writing intervention, and eight studies implemented an expository essay-writing intervention as the independent variable.

Participants and Setting

A total of 417 students with disabilities were participants in 26 studies. Disability, gender, setting, and instructional delivery information for individual studies is noted in Table 1. Eighty-five percent (n=354) of students were diagnosed with LD, 12% (n=51) with EBD, 2% (n=9) with ADHD, less than 1% (n=1) with ADHD/SLI, less than 1% (n=1) with ADHD/DD, and less than 1% (n=1) with SLI. In the 21 studies that specified participant gender, 74% (n=226) were male and 26% (n=80) were female.

Twenty-seven percent (n=7) of studies took place in an inclusion classroom, 23% (n=6) in the resource room, 8% (n=2) in an alternative program for students with emotional/behavioral difficulties, and 42% (n=11) administered treatment individually or outside of the regular daily setting. Instruction was delivered by the classroom teacher in 27% (n=7) of studies, trained graduate students in 23% (n=6), trained undergraduate students in 12% (n=3), and the researcher or a research assistant in 35% (n=9). One study did not specify who provided instruction (Ferretti, MacArthur, & Dowdy, 2000).

In the 23 studies that documented school district classification, 57% (n=13) of schools were considered suburban school districts, 17% (n=4) urban, 17% (n=4) rural, and 9% (n=2) urban/suburban.

Intervention Outcomes

The following section is a review of outcomes by essay strategy specific elements and quality for each study across persuasive, narrative and expository genres. Outcomes of studies within each writing genre are organized by intervention program (e.g., SRSD or Strategic Instruction Model [SIM]) or instructional methodology (e.g., goal-setting, concept mapping, or strategy instruction). Key characteristics of intervention programs and
Methodologies are outlined in Table 3. It is very important to note many components of various instructional programs/methodologies overlap. The original authors’ terms for their interventions are used to categorize instructional programs/methodologies. Quality indicators are used to review quality of research by genre.

Table 1

CCSS Writing Standards

<table>
<thead>
<tr>
<th>Genre</th>
<th>Key Standards for Essay Writing</th>
<th>Addressed in Existing Research</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Persuasive</td>
<td>Write organized arguments that introduce claims, provide reasons and explanations to support claims, and offer concluding remarks to support the stated argument</td>
<td>Yes</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Write discipline-specific persuasive essays that acknowledge counter arguments and present accurate data using credible sources</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Utilize technology to produce typed essays</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Narrative</td>
<td>Write compositions to describe real or imagined experiences through description of context and characters, appropriately sequenced events, and development of a logical conclusion</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Include narrative techniques (i.e. transition words, dialog, descriptive details) within essays</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Utilize technology to produce typed essays</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td>Expository</td>
<td>Introduce a topic, organize ideas using strategies, and provide a conclusion</td>
<td>Yes</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Include formatting (e.g., headings), graphics (e.g., charts, tables), and multimedia to enhance reader comprehension</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Write domain-specific expository essays using domain-specific vocabulary and an objective tone</td>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Utilize technology to produce typed essays</td>
<td>Yes</td>
<td>2</td>
</tr>
</tbody>
</table>

Note. Standards were coded as “yes” for “addressed in existing research” if the standard was all or partly addressed in studies reviewed.
# Table 2

## Participants, Settings and Procedures

<table>
<thead>
<tr>
<th>Author</th>
<th>Sample Size (n)</th>
<th>Gender</th>
<th>Disability</th>
<th>Grade</th>
<th>Setting</th>
<th>Design</th>
<th>Instructional Procedure</th>
<th>Genre</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barry &amp; Moore, 2004</td>
<td>20</td>
<td>15m, 5f</td>
<td>LD</td>
<td>9</td>
<td>Rural/Resource/Teacher</td>
<td>Pre and Post Test</td>
<td>SI</td>
<td>Expository/Persuasive</td>
</tr>
<tr>
<td>Chalk et al., 2005</td>
<td>15</td>
<td>11m, 4f</td>
<td>LD</td>
<td>10</td>
<td>Suburban/Resource/Researcher</td>
<td>Quasi-Experimental</td>
<td>SRSD</td>
<td>Expository</td>
</tr>
<tr>
<td>Cihak &amp; Castle, 2011</td>
<td>19</td>
<td>15m, 4f</td>
<td>LD, ADHD, EBD</td>
<td>8</td>
<td>Rural/Inclusion/Teacher</td>
<td>Quasi-Experimental</td>
<td>SI</td>
<td>Expository</td>
</tr>
<tr>
<td>Cuenca-Sánchez et al., 2012</td>
<td>21</td>
<td>20m, 1f</td>
<td>EBD</td>
<td>7</td>
<td>Alternative Program/Teacher</td>
<td></td>
<td>RCT</td>
<td>SRSD</td>
</tr>
<tr>
<td>De La Paz, 1999</td>
<td>6</td>
<td>5m, 1f</td>
<td>LD</td>
<td>7-8</td>
<td>Suburban/Inclusion/Teacher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Expository</td>
</tr>
<tr>
<td>De La Paz, 2001</td>
<td>3</td>
<td>1m, 2f</td>
<td>ADD, ADHD, SLI</td>
<td>7-8</td>
<td>Suburban/Inclusion/Teacher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Expository</td>
</tr>
<tr>
<td>De La Paz, 2005</td>
<td>11</td>
<td>8m, 3f</td>
<td>LD</td>
<td>8</td>
<td>Suburban/Inclusion/Teacher</td>
<td>Quasi-Experimental</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>De La Paz &amp; Graham, 1997</td>
<td>42</td>
<td>33m, 9f</td>
<td>LD</td>
<td>5-7</td>
<td>Suburban/Pull-out/Graduate Students</td>
<td>RCT</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Ferretti et al., 2000</td>
<td>32</td>
<td>NS</td>
<td>LD</td>
<td>6</td>
<td>Urban/Suburban/Inclusion/Not specified</td>
<td>RCT</td>
<td>Elaborated goal setting, Goal setting</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Ferretti et al., 2009</td>
<td>24</td>
<td>NS</td>
<td>LD</td>
<td>6</td>
<td>Urban/Suburban/Inclusion/Undergraduate Students</td>
<td>RCT</td>
<td>Elaborated goal setting, Goal setting</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Graham &amp; Harris, 1989a</td>
<td>11</td>
<td>NS</td>
<td>LD</td>
<td>6</td>
<td>Suburban/Pull-out/Undergraduate Students</td>
<td>RCT</td>
<td>SRSD/SI</td>
<td>Narrative</td>
</tr>
<tr>
<td>Graham &amp; Harris, 1989b</td>
<td>3</td>
<td>1m, 2f</td>
<td>LD</td>
<td>6</td>
<td>Suburban/Pull-out/Graduate Students</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Author</td>
<td>Sample Size (n)</td>
<td>Gender</td>
<td>Disability</td>
<td>Grade</td>
<td>School/ Setting/ Instructor</td>
<td>Design</td>
<td>Instructional Procedure</td>
<td>Genre</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>------------</td>
<td>-------</td>
<td>------------------------------</td>
<td>--------------------------</td>
<td>-------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Jacobson &amp; Reid, 2010</td>
<td>3</td>
<td>3m</td>
<td>ADHD</td>
<td>11-12</td>
<td>Rural/ Pull-out/ Researcher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Kiuhara et al., 2012</td>
<td>6</td>
<td>4m, 2f</td>
<td>EBD, LD, ADHD</td>
<td>10</td>
<td>Suburban/ Pull-out/ Researcher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>MacArthur &amp; Philippakos, 2012</td>
<td>6</td>
<td>5m, 1f</td>
<td>LD</td>
<td>6-8</td>
<td>Private School/ Pull-out/ Researcher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Compare-Contrast</td>
</tr>
<tr>
<td>Mastropieri et al., 2009</td>
<td>15</td>
<td>14m, 1f</td>
<td>EBD</td>
<td>8</td>
<td>Alternative Program/ Graduate Students</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Mastropieri et al., 2012</td>
<td>12</td>
<td>12m</td>
<td>EBD</td>
<td>7-8</td>
<td>Suburban/ Pull-out/ Graduate Students</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Monroe &amp; Troia, 2006</td>
<td>6</td>
<td>5m, 1f</td>
<td>LD</td>
<td>6-8</td>
<td>Urban/ Pull-out/ Researcher</td>
<td>Quasi-Experimental</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Page-Voth &amp; Graham, 1999</td>
<td>30</td>
<td>NS</td>
<td>LD</td>
<td>7-8</td>
<td>Suburban/ Pull-out/ Graduate Students</td>
<td>RCT</td>
<td>Goal setting, Goal setting and SI</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Patel &amp; Laud, 2009</td>
<td>3</td>
<td>1m, 2f</td>
<td>LD</td>
<td>7</td>
<td>Urban/ Resource/ Researcher</td>
<td>Pre and Post Test</td>
<td>Combined SRSD and VV</td>
<td>Narrative</td>
</tr>
<tr>
<td>Sawyer et al., 1992</td>
<td>43</td>
<td>25m, 18f</td>
<td>LD</td>
<td>5-6</td>
<td>Suburban/ Pull-out/ Undergraduate Students</td>
<td>RCT</td>
<td>SRSD, SI, Direct teaching</td>
<td>Narrative</td>
</tr>
<tr>
<td>Sexton et al., 1998</td>
<td>6</td>
<td>4m, 2f</td>
<td>LD</td>
<td>5-6</td>
<td>Suburban/ Pull-out/ Researcher</td>
<td>Single Case</td>
<td>SRSD</td>
<td>Persuasive</td>
</tr>
<tr>
<td>Sturm &amp; Rankin-Erickson, 2002</td>
<td>12</td>
<td>8m, 4f</td>
<td>LD</td>
<td>8</td>
<td>Urban/ Inclusion/ Teacher</td>
<td>Quasi-Experimental</td>
<td>Concept Mapping</td>
<td>Expository</td>
</tr>
</tbody>
</table>
**Persuasive essay-writing.** SRSD was utilized as an instructional approach in ten persuasive essay intervention studies (Cuenc-Sanchez, Mastropieri, Scruggs, & Kidd, 2012; De La Paz, 2005; De La Paz & Graham, 1997; Graham & Harris, 1989b; Jacobson & Reid, 2010; Kiuhara, O'Neill, Hawken, & Graham, 2012; Mastropieri et al., 2009, 2012; Monroe & Troia, 2006; Sexton, Harris, & Graham, 1998). Strategy instruction was implemented in two studies (Barry & Moore, 2004; Wong, Butler, Ficzere, & Kuperis, 1996).

**SRSD.** Effects of SRSD in the reviewed persuasive writing studies were evaluated by counting the number of strategy specific functional persuasive essay elements (premise, reason, conclusion, elaboration, and nonfunctional elements) and calculating quality according to a traditional holistic rating scale (i.e. a point scale designed to enable scorers to assign a numerical value to represent overall essay organization, sentence structure, vocabulary, ideas, and coherence).

SRSD was used to teach the TREE (Topic sentence, Reasons, Examine reasons, Ending) strategy to adolescents with disabilities in two single-case studies (Graham & Harris, 1989b; Sexton et al., 1998). Both studies yielded an increased number of essay strategy specific elements and essay quality during the intervention phase. PND was calculated using data provided by the study authors: PND for Sexton et al. (1998) was medium, 88%, for quality and small, 58%, for essay strategy specific elements. PND for Graham and Harris (1989b) was large, 100% for essay strategy specific elements. PND for quality was not computable, as mean holistic quality ratings for each student were provided rather than individual quality scores. Graham and Harris (1989b) found students were not able to generalize writing gains across genres. Sexton et al. (1998), however, noted generalization through students’ use of TREE when administered a prompt by a teacher other than the study instructor. Both studies documented slightly decreased essay strategy specific element and quality scores during maintenance phase in treatment phase comparison to; however, gains maintained above baseline (PND = 100%).
Table 3

**Instructional Programs and Methodologies**

<table>
<thead>
<tr>
<th>Instructional Methodology</th>
<th>Description</th>
<th>Key Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRSD</td>
<td>Instructional program used to facilitate strategy use and develop self-regulation skills (Harris, Graham, Mason, &amp; Friedlander, 2008).</td>
<td>• Stages for acquisition: (a) develop necessary prerequisite skills for strategy mastery (b) discuss the strategy and explain how the strategy will improve writing skills (c) memorize the strategy steps (d) provide an explicit model that includes verbalization of the thought process used to apply the strategy, (e) provide guided practice and self-regulation procedures (f) provide ample independent practice and opportunities for generalization. • Four self-regulation procedures (goal-setting, self-monitoring, self-instructions, and self-reinforcement)</td>
</tr>
<tr>
<td>Strategy Instruction</td>
<td>Strategy instruction is a teaching methodology used to assist students in appropriately executing and maintaining strategy-use (Wong, 1998).</td>
<td>• May look different across studies • Modeling, collaborative planning, scaffolding, drafting/revising, and collaborative revising (Swanson, 1999)</td>
</tr>
<tr>
<td>Goal-setting</td>
<td>Goal setting studies aimed to display the impact of goals on written compositions and self-regulatory skills of adolescents with LD (Ferretti et al., 2000; Ferretti et al., 2009; Page-Voth &amp; Graham, 1999).</td>
<td>• Attainable, proximal, genre-specific goals guide the writing process</td>
</tr>
<tr>
<td>SIM</td>
<td>SIM utilizes explicit/direct instruction of writing strategies (Therrien et al., 2009)</td>
<td>• Components of SIM instruction: (a) establishing the purpose of the strategy (b) teaching how, when, and why to use the strategy (c) different ways to remember the strategy (d) developing goals for learning the strategy (e) modeling the strategy (f) guided practice (Lenz &amp; Deshler, 2004)</td>
</tr>
<tr>
<td>Concept-Mapping</td>
<td>Instructional approach used to enable students to create a visual representation of text-structure and ideas.</td>
<td>• Also known as semantic mapping and graphic organizing</td>
</tr>
</tbody>
</table>

Mastropieri et al. (2009) and Mastropieri et al. (2012) used single-case methodology to implement SRSD for the POW (P = Pick my idea, O = Organize my notes, W = Write and Say More) + TREE strategies for students with EBD, again documenting increased number of essay strategy specific elements and essay quality with PND at 100% for both measures in both studies. As in the Graham and Harris (1989b) and Sexton et al. (1998) studies, scores in
the Mastropieri et al. (2009, 2012) studies decreased during the maintenance phase in comparison to the treatment phase but were above those noted at baseline (PND = 100%). In an RTC study, Cuenca-Sanchez et al. (2012) also implemented SRSD for POW + TREE strategies. A large effect size for number of essay strategy specific elements (ES = 3.19) and essay quality (ES = 3.43) was calculated using the data reported by study authors, who had reported significant results on both measures. Students in the treatment condition outperformed students in the control condition on a maintenance probe given 2 weeks after instruction (ES = 1.36). To facilitate student ability to generalize the strategy across content areas, instructors encouraged students to think of other contexts where they could use the strategy and modeled how to write a persuasive essay using science and social studies content. Students in the treatment condition wrote qualitatively better essays than students in the control condition on a generalization probe (ES = 1.8).

A second line of SRSD persuasive writing research stems from an RTC conducted by De La Paz and Graham (1997) during which students were taught the STOP and DARE (Suspend judgment, Take a side, Organize ideas, and Plan more as you write; Develop your topic sentence, Add supporting ideas, Reject at least one argument for the other side, and End with a conclusion) strategy. A large, significant effect for dictation and advanced planning strategy use in quality (ES = .90) and number of essay strategy specific elements (ES = 1.15) was calculated using the data reported by study authors. Gains were maintained according to a maintenance probe administered 2 weeks after instruction. Jacobson and Reid’s (2010) single case study used the STOP and DARE strategy and documented a large effect (PND = 100%) for both essay strategy specific elements and quality. However, Jacobson and Reid (2010) reported decreased quality scores from post-instruction to maintenance phases (PND = 67%). To meet high school writing task demands, Kiuhara et al. (2012) added the AIMS (Attract the reader’s attention, Identify the problem of the topic, Map the context of the problem, and State the thesis) strategy to SRSD for STOP and DARE. The intervention produced a moderate effect (PND = 74%) for quality and moderate-large effect (PND = 90%) for essay strategy specific elements. PND was calculated using study data reported by the authors.

In a quasi-experimental group study, De La Paz (2005) used SRSD for the STOP and DARE strategy combined with strategy instruction in historical reasoning. Students were taught to reconcile conflicting historical information and to display understanding of content in a persuasive essay. To meet standards for students at the secondary level, De La Paz (2005) increased expectations for length, elaborations, and number of supporting reasons. Using pre and posttest data reported by the study author, a large effect (ES = 1.09) for quality and moderate effect (ES = .70) for essay strategy specific elements was calculated. De La Paz (2005) also reported large, significant effects at post-test for quality (ES = 1.19) and essay strategy specific elements (ES = 1.17) compared to a control group. No students with disabilities were in the control group. Small-medium effects were documented on a historical
accuracy measure for the pre and posttest comparison (ES = .42) and control group (ES = .57) comparison. Maintenance and generalization data were not collected.

In the tenth SRSD persuasive writing study, Monroe and Troia (2006) used SRSD for the CDO (Compose, Diagnose, and Operate) and SEARCH (Set goals, Examine paper to see if it makes sense, Ask if you said what you meant, Reveal picky errors, Copy over neatly, Have a last look for errors) strategies. Although number of essay strategy specific elements and quality increased, a standardized mean difference could not be calculated because neither standard deviations nor individual scores were provided.

Results of SRSD persuasive writing studies can be examined with confidence. Five single-case designs met 11 out of 11 quality indicators (Graham & Harris, 1989b; Kiuhara et al., 2012; Mastropieri et al., 2009; 2012; Sexton et al., 1998). Jacobson and Reid (2010) met 10 out of 11 indicators; social validity was not established. The RTC studies conducted by De La Paz & Graham (1997) and Cuenca-Sanchez et al. (2012) met 9 out of 9 quality indicators. The quasi-experimental study conducted by De La Paz (2005) met 8 out of 9 quality indicators; students were not randomly assigned to treatment and control conditions. Monroe and Troia’s (2006) pre and posttest design met 2 out of 9 indicators. Missing indicators included lack of random assignment of subjects, mortality equivalence between conditions, pretest equivalence across conditions, teacher training description, control condition description, methods to control for teacher effects, and establishment of treatment fidelity.

**Goal setting.** Effects of goal setting in the reviewed persuasive writing studies were evaluated by counting the number of strategy specific functional persuasive essay elements (premise, reason, conclusion, elaboration, and nonfunctional elements) and calculating quality according to a traditional holistic rating scale (i.e. a point scale designed to enable scorers to assign a numerical value to represent overall essay organization, sentence structure, vocabulary, ideas, and coherence). In an RTC study, Page-Voth and Graham (1999) demonstrated student ability to create higher quality persuasive essays with a greater number of essay strategy specific elements when provided with genre-specific goals (e.g., a goal to increase the number of supporting reasons, a goal to increase the refutation of counterarguments). As reported by the study authors, students in both a goal setting and goal setting plus a strategy conditions significantly outperformed students in the no treatment control condition on quality (ES = 1.18) and essay strategy specific element (ES = 1.53) measures.

Two studies conducted by Ferretti and colleagues analyzed the effects of general goal setting (i.e., students were instructed to write a persuasive response and support their position) versus elaborated goal setting (i.e., students were provided with the general goal as well as genre-specific sub-goals [e.g., statement of belief, two or three reasons for belief, examples or supporting information for each reason, two or three reasons why others might disagree, why those reasons are wrong] based on elements of argumentation; Ferretti et al., 2000; Ferretti, Lewis, & Andrews-Weckerly, 2009). As reported by Ferretti and colleagues, students in both studies, in the elaborated goal conditions, produced qualitatively stronger
Essay Interventions for Adolescents with Disabilities by Lauren Valasa, Linda Mason and Charles Hughes

persuasive essays (average ES = .63) according to a holistic rating scale with more essay strategy specific elements. Feretti and colleagues reported significant effects in both investigations. Effect sizes could not be calculated for number of essay strategy specific elements because mean and standard deviation scores for essay strategy specific elements were not reported.

Page-Voth and Graham’s (1999) RTC met 9 out of 9 quality indicators; results should be examined with great confidence. Studies by Feretti and colleagues met 6 out of 9 indicators: instructor training, methods of controlling for teacher effects, and methods of establishing treatment fidelity were not addressed.

**Strategy instruction.** Following common principles of strategy instruction outlined in Table 3, Wong et al. (1996) taught students to write persuasive essays with a planning, writing, and revising strategy. Students worked collaboratively to revise each other’s compositions. Following instruction, Wong et al. (1996) reported students wrote with greater clarity (i.e., degree of absence of ambiguities in essays) and cogency (i.e., degree of persuasiveness of arguments presented in essays) according to a holistic scale (ES for clarity = 2.17; ES for cogency = 2.74). Wong et al. (1996) also reported a large, significant effect compared to the control condition (ES = 2.55). Effects were reported to maintain one week after instruction. Wong et al. (1996) did not measure essay strategy specific elements or collect generalization data. This quasi-experimental design met 5 of the 9 quality indicators used to measure strength of group designs. Elements lacking included random assignment of subjects, instructor fidelity described, and establishment of treatment fidelity. The authors also reported a floor effect.

Strategy instruction was also utilized in a quasi-experimental study aimed to increase state competency exam persuasive writing scores (Barry & Moore, 2004). Stages of strategy instruction in the study conducted by Barry and Moore (2004) involved explaining the purpose of the strategy, modeling, providing opportunities for student practice, providing corrective feedback, and holding a peer review session. Students were taught to use their fingers as an iconic memory stimulus for identifying the paragraphs (introduction, 3 body paragraphs containing supporting reasons, and conclusion) of a persuasive composition. While a large, significant effect (ES = .92) on state testing scores was computed using the data reported by study authors, the quasi-experimental study met only 3 (control condition described, no ceiling effects or floor effects for the primary measure, and Hawthorne effect controlled) of the 9 quality indicators.

**Narrative essay-writing.** All three narrative intervention studies explored the effects of variations of SRSD on the narrative essays of students with LD (Graham & Harris, 1989a; Patel & Laud, 2009; Sawyer, Graham, & Harris, 1992). Effects of SRSD in the reviewed narrative writing studies were evaluated by counting the number of essay strategy specific story-grammar elements (main character, locale, time, starter event, goal, action, ending, and reaction) and calculating quality according to a traditional holistic rating scale (i.e. a point scale designed to enable scorers to assign a numerical value to represent overall essay
organization, sentence structure, vocabulary, ideas, and coherence). Graham and Harris (1989a) taught the W-W-W, What = 2, How = 2 strategy to help students remember the parts of a narrative essay by answering the following questions: (a) Who is the main character? Who else is in the story? (b) When does the story take place? (c) Where does the story take place? (d) What does the main character want to do? (e) What happens when he or she tries to do it? (f) How does the story end? (g) How does the main character feel? Graham and Harris (1989a) implemented W-W-W, What = 2, How = 2 across two treatment conditions: (1) instruction including self-regulation procedures, and (2) instruction without explicit instruction in self-regulation. Combined pre and posttest means and standard deviations of intervention conditions were provided by study authors. Both intervention conditions produced significant increased essay strategy specific elements (ES = 2.2) and essay quality (ES = .61) in comparison to pretest scores. Graham and Harris (1989a) found students maintained writing skills two weeks after instruction and were able to independently generalize skills to the general education setting.

In a replication study, Sawyer et al. (1992) compared SRSD for W-W-W, What = 2, How = 2 across full SRSD instruction, strategy instruction without explicit instruction for self-regulation, and direct teaching. A control condition was added in order to strengthen internal validity of results. Students in the “full SRSD condition” outperformed students in the “instruction without explicit instruction for self-regulation condition” and “direct teaching condition”. A large effect on number of essay strategy specific elements (ES = 3.67) and overall story quality (ES =1.85) in comparison to the control condition was calculated using data reported by study authors (although study authors reported significant results for number of strategy specific element measures, no significant differences were found for quality measures). Students in instruction without explicit instruction for self-regulation (ES = .52) and direct teaching (ES = .97) also improved quality of narrative compositions compared to the control condition. Strategy use across all intervention conditions slightly decreased on maintenance probes administered four weeks after instruction. However, researchers noted results suggested maintenance was highest in the full SRSD condition.

Patel and Laud (2009) evaluated SRSD for POW + W-W-W, What = 2, How = 2; however, during the “W” phase students were encouraged to visualize and verbalize (V&V) structure words- what, size, color, number, shape, where, when, background, movement, mood, and perspective- to enhance story detail. Large, positive gains were computed using study data on the number of essay strategy specific elements (ES = 8.97) and the quality (ES= 1.38) of student compositions. The study was conducted with three students. Significance was not reported.

Studies conducted by Graham and Harris (1989a) and Sawyer et al. (1992) were well-constructed, randomized experimental designs that met 9 out of 9 quality indicators for group studies. Results of these studies can be examined with confidence. However, Patel and Laud (2009) met only 2 (no ceiling or floor effects for the primary measure, Hawthorne effect controlled) of the 9 quality indicators.
Essay Interventions for Adolescents with Disabilities by Lauren Valasa, Linda Mason and Charles Hughes

**Expository essay-writing.** Strategy instruction was utilized as an instructional approach in two expository essay intervention studies (Cihak & Castle, 2011; Wong, Butler, Ficzere, & Kuperis, 1997), SRSD in four studies (Chalk, Hagan-Burke, & Burke, 2005; De La Paz, 1999; 2001; MacArthur & Philippakos, 2010), SIM in one study (Therrien, Hughes, Kapelski, & Mokhari, 2009), and concept mapping in one study (Sturm & Rankin-Erickson, 2002).

**Strategy instruction.** In an early study, students with LD were taught to write compare-contrast type essays on computers with a planning, writing, and revising strategy (Wong et al., 1997). Following principles of effective writing strategy instruction (modeling, collaborative planning, scaffolding, drafting/revising, and collaborative revising), students were taught to effectively collaborate to revise each other’s compositions. Following instruction, students wrote with greater clarity, aptness (i.e., appropriateness of ideas/details in supporting comparisons and contrasts), and organization according to a holistic scale (ES = 1.56). The large, statistically significant effect was reported to maintain one week after instruction. Wong et al. (1997) did not measure expository essay strategy specific elements. This pre and posttest design met 5 of the 9 quality indicators used to measure strength of group designs. Elements lacking included random assignment of subjects, mortality equivalence between conditions, pretest equivalence between conditions, and type of control described.

In a quasi-experimental study, Cihak and Castle (2011) explored effects of strategy instruction for the expository writing program, Step-Up to Writing. Five lessons were given to provide students with instruction in organizing, outlining, structuring, and using transitions and details in expository essays. A large, statistically significant effect (ES = 3.80) was calculated on student writing according to a state test-scoring rubric using data reported by study authors. Number of essay strategy specific elements was not measured. While the intervention yielded a large effect size, caution should be taken in interpreting results. The study met only 2 (Hawthorne effect controlled and no ceiling or floor effects for the primary measure) out of 9 quality indicators.

**SRSD.** Effects of SRSD in the reviewed expository writing studies were evaluated by counting the number of strategy specific essay elements (premise, reason, conclusion, elaboration, and nonfunctional elements) and calculating quality according to a traditional holistic rating scale (i.e. a point scale designed to enable scorers to assign a numerical value to represent overall essay organization, sentence structure, vocabulary, ideas, and coherence). De La Paz (1999, 2001) explored the effects of SRSD on the expository compositions of adolescents in two single-case designs. Students were taught to write expository essays using SRSD for the PLAN (Pay attention to the prompt, List main ideas, Add supporting ideas, Number your ideas) and WRITE (Work from your plan to develop your thesis statement, Remember your goals, Include transition words, Try to use different kinds of sentences, and Exciting, interesting, $100,000 words) strategies. Instruction resulted in increased essay strategy specific elements and quality in both studies. De La Paz (1999, 2001) was also able
to document lasting effects on a maintenance probe administered four weeks following instruction. PND for essay strategy specific elements and quality were calculated using data reported by study authors; essay strategy specific element scores resulted in 89% PND and quality scores resulted in 89% PND for the first study (De La Paz, 1999). For the second study, essay strategy specific element scores resulted in 100% PND. PND could not be calculated for quality using data provided in the second study because individual quality scores were not reported.

Chalk et al. (2005) also examined the effects of SRSD on the expository essays of adolescents with LD. Using a repeated measures group design, Chalk et al. (2005) used SRSD to teach (a) Think, who will read this and why I am writing it, (b) Plan what to say using DARE, and (c) Write and say more. Researchers observed a medium, statistically significant effect (ES = .60) on the quality of expository compositions. Students sustained improved scores on a maintenance measure given two weeks following instruction and during a generalization probe administered in the social studies classroom. Essay strategy specific elements were not measured.

MacArthur’s and Philippakos’ (2010) study taught adolescents with LD to develop and write compare-contrast essays. SRSD for the IBC (Introduction, Body, and Conclusion) and TAP (Topic, Audience, and Purpose), Brainstorm and Organize strategies were taught along with instruction on developing compare-contrast text structure. MacArthur and Philippakos noted increased essay strategy specific elements (PND = 100%) and quality (PND = 85%), indicating a large effect for essay strategy specific elements and a moderate effect for quality. Of the four students probed during the maintenance phase, two were able to maintain gains made immediately following instruction.

The studies by De La Paz (1999) and MacArthur and Philippakos (2012) met 11 out of 11 quality indicators. De La Paz (2001) met 10 out of 11 indicators; social validity was not established. Results of these single-case studies should be analyzed with confidence. The quasi-experimental group study conducted by Chalk et al. (2005), however, met 5 (no ceiling or floor effects for the primary measure, instructor training described, Hawthorne effect controlled, treatment fidelity established, and teacher effects controlled) out of 9 quality indicators.

**SIM.** Therrien et al. (2009) evaluated the Strategic Instruction Model (SIM) for The Essay Test-Taking Strategy, ANSWER: (a) Analyze the action words in the question, (b) Notice the requirements of the question, (c) Set up an outline, (d) Work in detail, (e) Engineer your answer, and (f) Review your answer. Effects of SIM were evaluated by counting the number of strategy specific essay elements (i.e. action words, underlining requirements, setting up an outline, listing details, engineering answer, and reviewing answer) and calculating quality according to a holistic rating scale (i.e., a point scale designed to enable scorers to assign a numerical value to represent overall essay organization, sentence fluency, word choice, voice, ideas and content, and conventions). Therrien and colleagues reported a medium effect (ES = .51-.68) for quality and large, significant effect (ES = 1.69) for number
of strategy-specific elements. While the authors reported significant student gains in quality of student writing according to sections of the rubric that were aligned with the strategy, no significant differences were found on sections of the quality rubric that were not aligned with the strategy. This RTC study met 9 out of 9 quality indicators established by Graham and Perin (2007a), signifying a strong experimental design.

**Concept mapping.** To examine the effects of concept mapping on the expository essays of eighth grade students with LD, Sturm and Rankin-Erickson (2002) used a repeated measures within-subjects design comprising three writing conditions: (1) no-mapping, (2) hand-mapping, and (3) computer-mapping. Students typed all essays on computers. Effect sizes were calculated using data reported by study authors. Results indicated students wrote higher quality essays (according to a holistic rating scale) using hand-mapping or computer mapping in comparison to a baseline probe (ES = .93). However, students also wrote qualitatively better essays in the no-mapping condition compared to the baseline (ES = 1.63). Researchers attributed statistically significant gains in both conditions to writing instruction containing information about expository writing processes and conventions. Number of essay strategy specific elements was not used as a dependent measure. Sturm and Rankin-Erickson’s (2002) concept mapping study met five (no ceiling or floor effects for the primary measure, instructor training described, treatment fidelity established, teacher effects controlled, and Hawthorne effect controlled) out of the nine quality indicators.

**Addressing CCSS**

Table 1 displays writing standards addressed in existing research. Persuasive writing research provided methods for enhancing students’ ability to write organized arguments that introduce claims, provide reasons and explanations to support claims, and offer concluding remarks to support the stated argument, as required by CCSS (2013). However, results reveal several aspects of CCSS persuasive writing demands that have not been adequately addressed. First, only one of 15 studies considered students’ ability in writing discipline-specific persuasive essays that acknowledge counter arguments (De La Paz, 2005). In all other studies, treatment was administered in a pull-out setting, resource room, self-contained classroom, or in an unspecified classroom setting. None of the studies were conducted in a science classroom setting or aimed to encourage student use of science-specific content. However, Cuenca-Sanchez et al. (2012) provided a model of how to generalize the POW + TREE strategy across science and social studies settings and documented successful generalization of the strategy in comparison to the treatment group. Next, studies have not focused on teaching students to present accurate data using credible sources. Only one study facilitated student use of accurate data from credible sources (De La Paz, 2005). Finally, researchers have not developed methods to enhance students’ ability to utilize technology to produce typed compositions within persuasive writing intervention studies.

Analysis of existing narrative essay intervention research shows researchers have developed strategies to address some aspects of CCSS demands. As mentioned, researchers have established instructional practices aimed at students’ development of imagined
experiences through description of context and characters, appropriately sequenced events, and development of a logical conclusion (Graham & Harris, 1989a, 1989b; Patel & Laud, 2009; Sawyer et al., 1992). However, results also reveal several aspects of CCSS narrative writing demands that have not been addressed. First, studies have not focused on the production of narratives based on personal experience. All studies used W-W-W, What = 2, How = 2 to engender imaginative story-writing skills. Furthermore, none of the studies provided personal narrative or story-writing prompts using words. Instead, pictures were used as prompts to develop imagined experiences. Next, researchers have not included strategies to assist students in incorporating narrative techniques (i.e. transition words, dialog, and descriptive details). Finally, researchers have not developed methods to enhance students’ ability to utilize technology to produce typed compositions within narrative writing intervention studies.

Expository essay writing research has addressed some CCSS demands including instruction in strategies to enable students to introduce a topic, organize ideas, and provide a conclusion (Cihak & Castle, 2011; Chalk et al., 2005; De La Paz, 1999, 2001; MacArthur & Philippakos, 2010; Sturm & Rankin-Erickson, 2002; Therrien et al., 2009; Wong et al., 1997) and utilization of technology to produce typed essays (Sturm & Rankin-Erickson, 2002; Wong et al., 1997). However, several CCSS writing standards have not yet been addressed. First, aside from compare-contrast strategies, research has not provided expository text-structure (definition, classification, and cause/effect) specific strategies. It is not known whether existing interventions generalize across all expository essay text-structure types required by CCSS. None of the studies collected generalization data across various expository text-structures. Next, none of the expository writing interventions include instruction for including graphics to enhance reader comprehension. Finally, research has not targeted students’ ability to utilize domain-specific vocabulary. While one study (Chalk et al., 2005) documented successful strategy-use within the social studies classroom, dependent measures (essay length and holistic quality) did include student utilization of domain-specific vocabulary or content.

Discussion

The purpose of the present review was to explore the literature on essay-writing interventions for adolescents with high incidence disabilities. Specifically, essay-writing interventions implemented across persuasive, narrative, and expository genres for adolescents with high-incidence disabilities were reviewed and the extent to which research has addressed CCSS standards within each genre was evaluated. Results reveal several major findings: (1) Certain participant and school setting populations are underrepresented within the literature; (2) Most essay interventions designed to enhance writing skills for adolescent writing skills are persuasive; few interventions aimed at improving narrative and expository essay writing skills; (3) Within each genre, essay interventions that utilized SRSD and strategy instruction methodologies had the most support across the literature; (4) Research
Essay Interventions for Adolescents with Disabilities by Lauren Valasa, Linda Mason and Charles Hughes

has not addressed many key components of CCSS for writing instruction. These major findings are discussed in detail below.

The majority of participants (85%) across studies were students with LD, further emphasizing the need for more writing intervention research aimed to enhance essay-writing skills of students with disabilities other than LD. Instruction was provided by the classroom teacher in only 27% (n = 7) of studies. Without data documenting successful teacher implementation of instruction, it is difficult to ascertain whether teachers can independently implement interventions and engender outcomes similar to those documented in research. Thus, more research is needed to ensure interventions can be feasibly implemented by the classroom teacher. Furthermore, only 17% (n = 4) of studies took place in urban and 17% (n = 4) in rural school districts. In order to ensure results are applicable to students in urban and rural districts, studies should be replicated across urban and rural settings, as students in these settings could possess different characteristics/cultural differences that may impact outcomes.

Examination of results by genre revealed effective instructional approaches and interventions as well as gaps within essay writing intervention research. More than half of the studies that met inclusion criteria targeted persuasive writing performance. Within the persuasive genre, the SRSD and goal-setting (a component of SRSD) studies yielded large effects on writing quality through strong study designs. These results extend and strengthen results of prior reviews noted in the introduction that documented the effectiveness of SRSD for teaching writing (Graham et al., 2013; Graham & Harris, 2003; Graham & Perin, 2007; Mason & Graham, 2008; Taft & Mason, 2011). To address middle and high school CCSS standards, future persuasive writing research should incorporate instruction in writing persuasive essays across content areas using accurate data and credible sources. Researchers should also match study settings to CCSS classroom expectations by incorporating student opportunities to produce typed essays.

Examination of results within the narrative genre reveals somewhat similar findings. However, only 2 studies, both conducted over 20 years ago, were strong experimental designs (Graham & Harris, 1989a; Sawyer et al., 1992). Both found medium effects for strategy instruction without explicit instruction in self-regulation on essay quality measures. Only one of the strong designs produced a large effect using SRSD for the W-W-W, What = 2, How = 2 strategy (Sawyer et al., 1992). These results are surprising considering the well-documented, highly positive effects of SRSD and strategy instruction for the W-W-W, What = 2, How = 2 strategy on narrative writing tasks at the elementary level (Harris et al., 2012). While SRSD and strategy instruction for the W-W-W, What = 2, How = 2 strategy has resulted in large effects on the quality of narrative compositions at the elementary level (Harris et al., 2012), it could be that a different strategy, designed to match writing prompts and demands at the secondary level, would yield consistent, large effects for adolescent students with high-incidence disabilities. More research is needed to conclusively recommend SRSD and strategy instruction without explicit instruction in self-regulation as best-practice narrative essay intervention techniques for adolescents with disabilities. To
better address CCSS standards in secondary settings, future narrative writing intervention research for students with disabilities should include methods for responding to written, grade-level, personal experience and story prompts using appropriate narrative sequence. Moreover, instruction should include methods for facilitating student use of narrative techniques, and technology to produce typed compositions.

Within the expository genre, three strong single-case designs documented medium-large effects on the quality of student essays for SRSD (De La Paz, 1999, 2001 MacArthur & Philippakos, 2010). SIM was also found to have medium-large effects on quality of expository compositions of adolescents in a strong RTC design (Therrien et al., 2009). This result is parallel to previous research noted in the introduction documenting the effectiveness of SRSD and SIM for improving writing performance (Graham & Harris, 2003; Graham & Perin, 2007a, 2007b; Mason & Graham, 2008; Taft & Mason, 2011). Although these strong expository intervention studies for adolescent students with disabilities provide information about effective programs of instruction in the expository genre, CCSS standards require student mastery of compare/contrast, cause/effect, classification, and definition expository text structures across content-areas. Thus, future intervention studies should specify expository text-structure prompt type, document whether or not the intervention can be generalized across text-structures, and should include instruction for including discipline-specific content and vocabulary.

Limitations and Implications

The lack of long-term maintenance and generalization data across genres represents a large gap in the research base. Of the 15 studies that collected maintenance data, 6 reported decreased scores on maintenance probes in comparison to post instruction scores (Graham & Harris, 1989b; Jacobson & Reid, 2010; Mastropieri et al., 2009, 2012; Sawyer et al., 1992; Sexton et al., 1998). This result further emphasizes the need for booster sessions as part of the writing curriculum, previously recommended by Graham and Harris (1989a). Previous reviews have also stressed the need for assessing the effects of writing interventions over an extended period of time (Graham et al., 2013).

As noted in the results, generalization data were reported in 8 studies. In 5 studies, successful generalization was reported because students could write an essay in a different classroom or with a different teacher (Chalk et al., 2005; Graham & Harris, 1989a; Mastropieri et al., 2009, 2012; Sexton et al., 1998). Two studies reported that students were unable to transfer newly acquired writing skills across genres (Graham & Harris, 1989b; Monroe & Troia, 2006). Student inability to simply generalize writing strategies across various genres highlights the need for further analysis and development of writing research within each genre.

Disappointingly, only 8% \((n = 2)\) of studies documented successful generalization across varied content-area classrooms (Cuenca-Sanchez et al., 2012; De La Paz, 2005). In both studies, persuasive essay interventions were implemented. Students with disabilities are expected to write persuasive, narrative, and expository essays across all content areas (CCSS,
2013). For example, a student must write a cause/effect expository essay in science class using domain-specific vocabulary, then must not only recognize a cause/effect expository prompt in social studies class, but also use domain specific vocabulary to respond to the prompt. Clearly, more research in writing across content-area curricula for adolescents with high-incidence disabilities is needed to ensure students can recognize/differentiate between text structures of, and use the appropriate strategy to respond to, prompts of various genres across content areas. Future researchers should incorporate generalization instruction within interventions and generalization measures within designs to enable students at the secondary level to write seamlessly across content areas and avoid over/under-generalization of strategy use.

**Future Research**

To meet CCSS demands, Graham and Harris (2013) have recommended enhancing teacher knowledge of writing development and implementation of evidence-based writing procedures for students with disabilities in general education settings. Thus, identification of specific, effective essay intervention procedures that are grounded in effective, research-based instructional methodology, and address CCSS writing standards, may help teachers provide effective essay writing instruction to students. However, results reveal that U.S. CCSS demands are higher and broader than the research has addressed. Based on the results of this review, it is recommended that future writing research for students at and above middle school level should (a) incorporate stronger methods for facilitating maintenance and generalization, (b) address student ability to utilize technology to produce typed compositions, (c) include methods of instructing students to incorporate domain-specific vocabulary, data, and credible sources, (d) ensure students can identify genres of prompts to allow for appropriate strategy use, (e) further develop strategies within the narrative and expository genres to account for all text-structures within those genres, and (f) utilize SRSD and SIM programs of writing instruction.

Students with disabilities often have severe persuasive, narrative, and expository writing deficiencies (Santangelo, Harris, & Graham, 2008). As a result, students with high-incidence disabilities are failing to meet the demands of the CCSS (2013), higher education, and employment. SRSD and SIM approaches are promising intervention methods for facilitating essay-writing skills for adolescent students with high-incidence disabilities. However, further empirical research is needed to develop the research base to meet CCSS standards and identify methods to help students maintain and generalize skills across curriculum content areas. Such investigations would be valuable to teachers who are in need of effective, evidence-based instructional techniques to enable their students with writing difficulties to achieve high levels of academic success across the curriculum.
References

*Reviewed Studies


Essay Interventions for Adolescents with Disabilities by Lauren Valasa, Linda Mason and Charles Hughes


Essay Interventions for Adolescents with Disabilities by Lauren Valasa, Linda Mason and Charles Hughes


Guest Reviewers

Thank you to the following members of the IARLD who provided their expertise as guest reviewers in the past year:

Wendy Cavendish
University of Miami

Raymond J. Rodriguez
James Madison University

Thank you to the following professional peers who provided their expertise as guest reviewers in the past year:

Salim Abu-Rabia
University of Haifa

Elinor Saiegh-Haddad
Bar-Ilan University
39th Annual IARLD Conference

July 7-9, 2015

University of British Columbia

Vancouver Canada

The conference will take place on the campus of The University of British Columbia.

Please find the information about call for proposals, registration, scientific program, conference venue, accommodation and social program on the website

http://www.iarld.com/?page_id=84

Vancouver, judged to be one of the most beautiful cities in the world, is on the sea and in the shadow of the Coast Mountains.

There will be reasonably priced accommodations on campus.

If you have any questions, please contact

linda.siegel@ubc.ca

We look forward to seeing you in Vancouver in July 2015
**Manuscript Submission Guidelines**

English is used for submissions to the journal, correspondence and publication. All submissions must be formatted consistent with the 6\textsuperscript{th} Edition of the Publication Manual of the American Psychological Association (APA). Manuscripts must include a 100 to 150 word abstract summarizing the contents.

A critical concern in the learning disabilities field is the definition of the population. Therefore, authors are expected to operationally define the study participants in accordance with professional standards (see CLD Research Committee: Rosenberg et al., 1993. Minimum standards for the description of participants in learning disabilities research. Learning Disability Quarterly, 26(4), 210-213). In addition, parameters of the setting in which the research took place are to be clearly delineated. Such descriptions facilitate replication and application of results. Manuscripts that fail to specify participant and setting variables will be rejected or returned to the authors for clarification. Authors of research manuscripts are encouraged to include brief (e.g. one to two sentences) explanations of why specific procedures and analysis methods were employed. Peer reviewers will evaluate the appropriateness of this and all aspects of the reported study.

Manuscripts are not to exceed 35 double spaced pages, consistently employing a 12 point font (including references, tables, figures and appendices). Please limit tables and figures to those essential to conveyance of your content. Please present figures and tables in portrait format and use grey scale rather than colored images.

The manuscript submission and review process will be conducted electronically. To submit your manuscript digitally, you must use one of the following formats: Microsoft Word (.doc or .docx), RTF or PDF. Manuscripts must be saved as “letter” (“U.S. letter”) page length/paper size. Submissions in Word or RTF format will be converted into a PDF before being sent for blind peer review. If you submit a PDF file it is your responsibility to ensure it can be read and printed by others. To avoid delays please embed all fonts and use Adobe PDF Distiller instead of PDF Writer to ensure that others can view the article exactly as intended. No email attachment should exceed 15 MB.

Each manuscript must be accompanied by a cover letter that communicates (a) that the manuscript is an original work, (b) that the manuscript is not under consideration by any other journal, and (c) any other disclosures as required by the APA (e.g., ethical treatment of participants, financial relationship disclosures). A cover page that provides the name, affiliation, mailing address, phone number, fax and email address of each author must also accompany each manuscript. All communications will be with the lead author. No author identifying information should be included directly in manuscripts submitted for review (e.g. explicit reference to previous publications, acknowledgements, author biographies). Please send the required cover letter, separate cover page with author identifying information, and manuscript as three separate attachments to a single email to: ijrd@bc.edu (please do not send tables, figures, appendices or other supporting materials in separate files). Manuscripts may not be submitted by fax or in paper form.

For further information, please refer to:
The IARLD (International Academy for Research in Learning Disabilities) is an international professional organization dedicated to conducting and sharing research about individuals who have learning disabilities.

The IARLD is an elected group of premier scientists, educators and clinicians in the field of learning disabilities throughout the world. The Academy was formed in 1976 by Dr. William Cruickshank (United States of America) and Dr. Jacob Valk (The Netherlands), meeting in Canada with the intention of providing a forum for the exchange of information and the advancement of knowledge regarding learning disabilities.

Since its inception, the Academy has realized its mission of being a professional, international, interdisciplinary consortium of scientists. The Academy currently has a membership of nearly 200 distinguished scholars, representing 26 different countries and thirty disciplines.

IARLD members represent:

- distinguished researchers,
- distinguished practitioner/clinicians,
- young researchers, and
- promising doctoral students.

www.iarld.com

ISSN: 2325-565X (print)
ISSN: 2329-3764 (online)