Literacy difficulties and emotional and behaviour disorders: causes and consequences

Jane Hurry

Eirini Flouri

(Department of Psychology and Human Development, UCL Institute of Education, University College London)

Kathy Sylva

(Department of Education, University of Oxford)

Abstract

Background: There is a well-established association between literacy difficulties and emotional and behaviour disorders (EBD). However, the scarcity of experimental studies means directions of causality are not clear.

Aims: This study investigates causal relationships between literacy and EBD, and implications of co-morbidity for intervention.

Sample and method: In the first year of a quasi-experimental study of 258 six-year-olds with reading difficulties: 87 received Phonological Training (PT), 81 received Reading Recovery (RR), and 90 in the control group received standard tuition. Children were followed up immediately post intervention and four years later.

Results: Immediately post intervention, RR significantly improved children's literacy (Cohen's d = .89). Four years later both RR and PT had small effects on literacy (Cohen's d = .25 and .26 respectively). These effects provided the opportunity to test the hypothesis that literacy difficulties cause or exacerbate EBD. This hypothesis was not supported as neither intervention reduced EBD. There was an interaction between hyperactivity symptoms at baseline and the effectiveness of PT, with PT being effective for children with few or no symptoms but ineffective for those with symptoms. EBD did not moderate the effectiveness of RR. Conduct disorder and hyperactivity at baseline had negative effects on literacy progress.

Conclusions: Literacy difficulties do not appear to be a cause of EBD but conduct disorder and hyperactivity exacerbate literacy difficulties. This may be due to EBD interfering with the effectiveness of instruction. If children have more than one problem they probably need more or different support.

Keywords: Comorbidity, reading difficulties, hyperactivity, conduct disorder, emotional disorder

Introduction

In common with other special educational needs, co-occurrence of literacy difficulties and emotional and behaviour disorders (EBD) is a reality for a significant minority of children. EBD is used here as a general overarching term which encompasses the externalising of Attention Deficit Hyperactivity Disorder (ADHD) and Conduct Disorders (also referred to as anti-social behaviour disorders) and the internalising disorders of depression and anxiety (also referred to as emotional disorders). The association between reading difficulties and ADHD (Boyes, Tebbut, Preece & Badcock, 2018; Gilger, Pennington, & DeFries, 1992; McGee, Prior, Williams, Smart, & Sanson, 2002; Medford & McGeown, 2016; Romano, Babchiskin, Pagani, & Kohen, 2010; Shaywitz, Fletcher, & Shaywitz, 1995) and Conduct Disorder (CD) (Boyes et al., 2018; Maughan, Pickles, Hagell, Rutter, & Yule, 1996; Medford & McGeown, 2016; Mundy et al., 2017; Rutter, Tizard, & Whitmore, 1970; Terras, Thompson, & Minnis, 2009), has been reported consistently for a century (Hinshaw, 1992). There is also some evidence of a relationship between reading difficulties and the internalising disorders of depression and anxiety (Boyes et al, 2018, Dahle, Knivsberg, & Andreassen, 2011; Maughan, Rowe, Loeber, & Stouhamer-Loeber, 2003; Siperstein, Wiley, & Forness, 2011; Terras et al., 2009; Willcutt & Pennington, 2000), though associations tend to be weaker than those reported for externalising disorders (Feehan, McGee, Williams, & Nada-Raja, 1995; Hinshaw, 1992; Rutter et al., 1970), and the research evidence is less consistent (Carroll, Maughan, Goodman, & Meltzer, 2005; Fleming & Offord, 1998; Medford & McGeown, 2016). The relationship between literacy difficulties and emotional and behaviour disorders (EBD) also extends to children with spelling difficulties (Anderson,

Williams, McGee, & Silva, 1989; Stevenson & Graham, 1993) and language difficulties (Benner, Nelson, & Epstein, 2002; Dockrell & Hurry, 2018; Riccio & Jemison, 1998). This investigation reports on a quasi-experimental study of literacy interventions for six year olds with literacy difficulties, both of which significantly improved aspects of literacy, and which also measured children's EBD pre and post intervention (Hurry & Sylva, 2007). The implications for meeting the literacy needs of children with co-occuring literacy and emotional and behavioural difficulties are explored, both through examining causal pathways between the two special needs and through investigating the implications of co-occurrence for literacy intervention.

For almost as long as the association between literacy difficulties and EBD has been reported there has been speculation regarding its explanation. Understanding the causal pathways of the relationship informs intervention. If emotional or behaviour disorders increase subsequent risk of reading failure it suggests that addressing children's behaviour needs to be a core part of school business. If reading failure can lead to emotional or behaviour problems this too has implications for both health and education. Finally, co-morbidity itself may have implications for how teachers approach reading remediation or health clinics behaviour management, what works for a child with only one need may not be suitable for a child with additional problems.

Causes of co-morbidity

There are four broad possible types of causal connections behind this co-morbidity: literacy difficulties cause EBD; EBD cause literacy difficulties; literacy difficulties and EBD stem from a common cause or causes, and; the relationship is bi-directional where each

difficulty exacerbates the other. All have some support from research and due to the diversity of both literacy difficulties and EBD multiple explanations may apply. Increasingly the nature of the emotional or behavioural difficulty is seen as significant in determining the nature of the causal connection.

Literacy difficulties and externalising disorders

The evidence of literacy difficulties causing externalising behaviour is equivocal. Hinshaw's review (1992) tentatively concluded that early literacy difficulties could predispose to subsequent externalising disorder, both CD and ADHD (Jorm, Share, Matthews, & Maclean, 1986; McGee, Williams, Share, Andersen, & Silva, 1986). Consistent with this, Bennett, Brown, Boyle, Racine and Offord (2003) found reading difficulties at school entry increased the risk of later conduct problems. However, Fergusson and Lynskey (1997) found that the association between early literacy difficulties and later conduct problems was explained by features present before the onset of literacy difficulties (notably early-onset conduct problems; Fergusson & Horwood, 1995; Maughan, Pickles, Hagell, Rutter & Yule, 1996: Williams & McGee, 1994). Even where literacy difficulties predicted subsequent behaviour problems, effects were weak at best. A rare study of middle-school struggling readers experimentally tested the role of literacy difficulties in subsequent behaviour problems. Ten to eleven year olds were given intensive, response-based reading intervention over three years and this directly improved reading achievement which in turn (indirectly) influenced attention (Roberts et al, 2015).

There is more support for the proposition that externalising problems cause literacy difficulties. Huesmann, Eron and Yarmel (1987, cited in Hinshaw, 1992) found aggression at

eight years old predicted low achievement in adulthood, controlling for IQ at eight. Fergusson and Lynskey (1997) found that CD at age six predicted reading at age eight. Medford and McGeown (2016) reported that hyperactive symptoms on the Strengths and Difficulties Questionnaire at the start of school predicted reading abilities one and a half years later, controlling for pre-reading skills at baseline. However, none of these studies rules out the possibility that their findings may be the result of other common factors, such as, social, family, cognitive and genetic. There is no evidence that intervention to alleviate behaviour problems has affected reading achievement (Maughan & Carroll, 2006).

Externalising behaviour and literacy difficulties may share common causes; genetic or environmental. For genetic risk, the evidence is most persuasive for ADHD (Carroll et al., 2005). Using the Environmental Risk Longitudinal Twin Study, Trzesniewski, Moffitt, Caspi, Taylor and Maughan (2006) conclude that the association between literacy difficulties and ADHD is best explained by common genetic influences. This is consistent with a number of studies reporting a common genetic aetiology between literacy difficulties and ADHD (Chadwick, Taylor, Taylor, Heptinstall & Danckaerts, 1999; Cheung et al., 2012; Cornish, Savage, Hocking & Hollis, 2011; Stevenson et al., 2005). There is little support for a genetic link between literacy difficulties and CD (Carroll et al., 2005; Stevenson & Graham, 1993; Trzesniewski et al., 2006). This relationship is more frequently attributed to shared environmental risks, in particular, socioeconomic status (Carroll et al., 2005). However, these environmental factors only very partially account for the association between literacy difficulties and CD (Hinshaw, 1992; Trzesniewski et al, 2006). Finally, there is consensus that at least part of the relationship between CD and literacy difficulties is accounted for by co-

morbid ADHD, particularly inattention (Hinshaw, 1992; Carroll et al., 2005; Willcutt & Pennington, 2000; Rabiner, Malone & the Conduct Problems Prevention Research Group, 2004). It is indeed intuitively plausible that failure to attend interferes with learning to read (or anything else for that matter) and supported by theoretical understandings of ADHD (Barkley, 1997; Tsang, Sam, Wong, Cheng, Sin and Ho, 2016). An experimental study, showing that the benefits of reading tutoring in first grade were reduced where there were inattention problems, provides support for this position (Rabiner et al., 2004).

Finally, there is also some evidence for a bi-directional relationship between externalising disorder and literacy difficulties (Maughan & Carroll, 2006). This is the model preferred by Trzesniewski and colleagues (2006) to explain the association between CD and literacy difficulties for the boys in their sample. Spira and Fischel (2005) also argue that evidence tends to support complex multidirectional links across development.

Literacy difficulties and internalising disorders

Exploring the link between literacy difficulties and internalising disorders in children and adolescents is more difficult than for externalising disorders. Third parties, who have historically been the informants on children's behaviour, are less reliable reporters of internalising than externalising disorders (Puura et al., 1998). There is certainly less consistent evidence of association. However, in their review of the relationship between literacy difficulties and mental disorders, Maughan and Carroll (2006) concluded that reading failure is a direct risk factor for depressed mood. Maughan et al. (2003) had earlier shown that associations between literacy difficulties and depressed mood were not simply a by-product of co-morbidity with disruptive disorders but suggested a direct causal link with an increased

risk of depression amongst poor readers in primary school. Maughan and colleagues (2003, 2006) conceptualise internalising disorders as a consequence of literacy difficulties.

Implications of co-morbidity for reading remediation

Understanding the cause of literacy difficulties or EBD would inform teachers' approach to children with difficulties but in practice this information is rarely available in a clear form and some element of a bi-directional relationship is likely. Irrespective, teachers must support their students to acquire the central skill of reading and a significant minority of the children needing reading remediation will have co-occurring EBD. This has potential implications for appropriate teaching approaches. EBD, particularly attentional problems, may moderate the impact of literacy intervention. Rabiner et al., (2004) found that whilst a phonics intervention led to significant improvement for first graders with reading problems it was not effective for those who also had attention problems. Al Otaiba and Fuchs (2006) also reported problem behaviours were associated with non-responsiveness to reading intervention in kindergarten and first grade. For slightly older children with ADHD, medication improved their responsiveness to intervention (Tannock et al. 2018).

Conclusion

The evidence regarding the causal nature of the connections between literacy difficulties and internalising and externalising problems suggests that the mechanisms differ depending on the nature of the problem. There is some evidence that literacy difficulties can exacerbate CD, depression and anxiety. Inattention problems may interfere with learning to read. Models which include testing bi-directional relationships between literacy difficulties and EBD are the most plausible. Behavioural problems, particularly relating to hyperactivity,

reduce the effectiveness of reading remediation. The paucity of any experimental studies is striking.

Rationale for the present study

Virtually all our knowledge on the nature of the causal link between literacy difficulties and EBD comes from longitudinal studies. Such studies are informative as they explore whether one condition temporally precedes another, one of three important criteria for inferring cause advocated by Cook and Campbell (1979): "1) covariation between presumed cause and effect; 2) the temporal precedence of the cause; and 3) the need to use the 'control' concept…to rule out alternative interpretations for a possible cause and effect connection" (1979, p. 31). Longitudinal studies do not however address the third of Cook and Campbell's criteria, the need to actively manipulate presumed causal agents.

More than two decades ago Hinshaw (1992) proposed an experimental design that would satisfy this requirement in clarifying the nature of the link between behaviour and reading problems. This design would require: (a) clinical samples of either children with reading or behaviour problems, (b) treatment conditions directed toward either reading or behaviour, and (c) assessment of both reading and behaviour outcomes over an extended time period. To date few such studies have been conducted (Spira & Fischel, 2005; Tannock et al., 2018). Hurry and Sylva (2007) conducted a study which satisfied all these requirements. Children were assessed on literacy and emotional and behavioural outcomes at baseline (age six years), immediately post-intervention and four years later at the end of elementary school. The measurement of emotional and behavioural outcomes was a screening measures for internalising symptoms of depression and anxiety and for externalising symptoms of

hyperactivity and conduct disorder (see Methods for details). In a mixed quasi-experimental and experimental design children received either Reading Recovery (RR) or Phonological Training (PT), two very different interventions aimed at children experiencing reading difficulties and representing the major categories of early literacy intervention, both of which have been found to be effective (Suggate, 2016), whilst the third group received standard school provision. The effects of the interventions on children's literacy, reported by Hurry and Sylva (2007), were mixed. Medium to large effect sizes on a range of literacy outcomes immediately post intervention were reported for RR (Cohen's d = .63 to .87). PT had immediate post intervention effects on phonological awareness only (Cohen's d = .30 to .72). At long-term follow up the effects of RR were only evident for non-readers at six years (about half the sample; Cohen's d = .48). PT had significant effects on an overall measure of reading and spelling (Cohen's d .26) in the quasi-experimental design, though not in the experimental design.

Research Question 1

In the present paper, the large effect of RR on literacy progress at immediately post intervention and some effects of both RR and PT four years later, demonstrated by Hurry and Sylva (2007), provided the opportunity to test the hypothesis: do literacy difficulties cause or exacerbate EBD using a quasi-experimental design?

Research Question 2

Does EBD predict literacy difficulties, using the longitudinal data?

Research Question 3

What are the implication of co-morbidity for literacy intervention? In the context of two interventions, varying in both intensity and content, we tested the hypothesis: does the presence of EBD, particularly hyperactivity, reduce the effectiveness of reading intervention?

Method

Sampling

Schools

Twenty-two RR schools participated in the study, virtually all the schools offering RR in England at the time. For each RR school, two neighbouring schools with similar intake were identified by the local authority and then randomly assigned to control or phonological training conditions. The final sample was: Control schools $N = 18$, Phonological Training schools $N = 23$. Schools were selected from seven local authorities in south-east England, six of which were in Greater London.

Children

In each of these 63 schools, teachers identified their six poorest Year 2 readers in the age range six to six years six months (approximately the bottom 12% of readers) and these children were assessed on a range of literacy measures described below, reduced to one overall score. On the basis of this overall score the four lowest scoring children in each school were selected for the analyses reported here, as follows: in the 22 RR schools, the bottom four scorers (usually) were offered intervention (RR). In each of the 18 Control schools, the bottom four scorers were selected for the control group. In each of the 23 Phonological Training schools, the six poorest readers were randomly assigned in the ratio 2:1 to PT or

control group The bottom four scorers per Phonological school assigned to control (N = 26) were added to the control group from the Control schools (N = 64) to make up the Control group.

Table 1 shows the numbers of children in each group at the three measurement points for whom both reading and behaviour data was available. Attrition at first follow-up was less than 1% and at second follow-up 12.0%.

Table 1: Number of children at each measurement point

Intervention Condition	Baseline, beginning of school year	First follow-up, 9 months later	Second follow-up, 4 years later
Control	90	89	77
RR	81	80	74
PT	87	87	76
Total	258	256	227

Table 2 shows children's demographics. Boys were overrepresented at % of the sample (class average 52% boys); 43% of the sample were receiving free school meals (class average 32%, national average 16%); 15% (class average 17%) spoke English as an additional language (EAL). The groups (RR, PT, Control) were well matched on these demographic factors, with no significant differences.

Table 2: Student demographics

Intervention Condition	% boys	% FSM	% White British	% EAL	Age year:month	IQ *M* (*SD*)
Control	58.9%	45.6%	73.3%	11.1%	6:3	94.6 (12.6)
RR	58.0%	40.0%	64.2%	21.0%	6:3	92.3 (12.5)
PT	63.2%	42.5%	70.1%	14.9%	6:3	92.1 (12.2)
Total	60.1%	42.9%	69.4%	15.5%	6.3	93.0 (12.4)

FSM = Free school meals at baseline; EAL = English as an additional language; IQ = British Ability Scales short form IQ at baseline.

Research design

Children receiving RR and children receiving PT were compared with similar Control children receiving their school's standard provision.

Children were tested before intervention on a battery of literacy tests and a behaviour measure, at the beginning of the school year in September/October (baseline). Short-term gains were assessed at the end of the school year in June/July, after the interventions were completed (first follow-up). Long-term effects were assessed four years after pre-test in September/December, when children were in their final year of elementary school (second follow-up).

Measures

Literacy

Baseline and first follow-up.

Children were assessed on a battery of standardised tests sensitive to early reading and writing skills, as follows:

1) The British Ability Scales (BAS) Word Reading test (Elliot, Murray & Pearson, 1984) assesses children's ability to read a list of 90 single words of increasing difficulty. Raw scores are a count of the number of words read correctly. There is reported test retest reliability = .98 and validity = .71 with Wechsler Objective Reading Dimensions (WORD) (Elliott et al., 1997). At first follow-up in the present study raw scores correlated .85 with Book Level (Hurry & Sylva, 2007).

2) Neale Analysis of Reading (Neale, 1989) was the most widely used test of its kind in the UK, consisting of six graded passages which were used to assess reading accuracy and comprehension, producing standardised scores on each.. It had good reliability (Cronbach's alpha for parallel forms: Accuracy = .81 and .82; Comprehension = .91 and .93) and validity (correlations with the Schonell Graded Word Reading Test for parallel forms: Accuracy *r* = .96 and .95; Comprehension *r* = ,88 and .88: Neale, Christophers & Whetton, 1989). At first follow-up in the present study accuracy scores correlated .85 with Book Level (Hurry & Sylva, 2007).

4) The Observation Survey of Early Literacy Achievement (OSELA; Clay, 1985, 2002, 2013; D'Agostino, Rodgers & Mauck, 2018) which includes: Book Level (text level of a series of graded texts 0-26 read with 90% accuracy), Letter Identification, Concepts about Print (children's knowledge of the conventions of print), a word reading test, Written Vocabulary (the number of correctly spelt words written in 10 minutes) and Hearing Sounds in Words

(marked for phonetically plausible spelling). Raw scores from each subtest were transformed to z scores and summed.

Each individual sub test has good reliability and validity (Clay, 2013). At first follow-up, summed z-scores for the OSELA correlated significantly with the BAS Word Reading test (*r* = .80), the Neale (*r* = .79).

5) The Oddities Test (Kirtley, Bryant, Maclean & Bradley, 1989), which measures awareness of rhyme and of initial and final phonemes. Bryant, MacLean, Bradley and Crossland (1990) report a Spearman-Brown reliability coefficient of 0.78 in their sample of sixty-four children aged five years seven months; in the present study, internal consistency of the Oddities Test was .83 (Cronbach's alpha). However, scores on the Oddities Test were only modestly correlated with the Dictation task (OS) at first follow-up (Spearman's rho = .44), suggesting that these tests measure different sub-skills.

An overall literacy score was calculated by summing z scores for each of these assessments.

Second follow-up.

At second follow-up, children were assessed on:

1) The NFER-Nelson Group Reading Test 6-12, (NFER-Nelson, 1985), was a standardised test of reading comprehension at the sentence level, using cloze procedures and consisting of 48 multiple-choice items. It had internal consistency *r* = .92 and .96 (Cronbach's alpha for parallel forms) and criterion validity: *r* coefficients ranging from .65 to .86 compared with six other tests of reading (NFER-Nelson, 1985).

2) The Parallel Spelling Test (Young, 1983), a standardised test of single word spelling, with internal consistency r = .93 (Cronbach's alpha), and correlation with the Vernon Spelling Test r = .92.

An overall literacy score was calculated by summing the z scores for the reading and spelling tests.

Inter-correlations were examined for children in the control group who were assessed on all three occasions (N = 153). Literacy at baseline correlated with literacy at first follow-up (identical measures) r = .790, and with literacy at second follow-up (different measures) r = .629. Literacy at first follow-up correlated with literacy at second follow-up (different measures) r = .754. The fact that the correlations between first and second follow-up is only marginally lower than the correlation between baseline and first follow-up suggests that the change in measurement is not unduly problematic.

EBD

Baseline and first follow-up.

Class teachers completed the Child Behaviour Questionnaire (CBQ, Rutter, 1967) at baseline and at the end of the school year (first follow-up). The CBQ was the principle screening measure for childhood psychological disorders in the UK at that time, used n major epidemiological studies and birth cohort studies. The teachers' scale has 26 items scored, 0, 1 or 2 ('does not apply', 'applies somewhat' and 'certainly applies'). The scale offers an overall score of behaviour difficulties, with a score of nine or more deemed to identify probable disorders. Subscales are available for conduct disorders (six items: destroys own or others belongings, fights, disobedient, lies, stolen on more than one occasion, bullies), emotional disorders (four items: often worries, miserable or tearful, fearful of new situations, tears on

arrival at school or school refusal) and hyperactivity (three items: restless hardly ever still, squirmy or fidgety, poor concentration or short attention span; Elander & Rutter, 1996). The CBQ teachers' scale has a test-retest reliability r = .89, inter-rater reliability (ratings made by two different teachers) r = .72. The validity of the score of nine or more being indicative of probable disorder was tested comparing children attending psychiatric clinics with children in the general population. In the general population, 11% of boys and 3.5% of girls scored nine or more compared with 80% of boys and 60% of girls in a clinical sample.

Second follow-up.

Class teachers completed the Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) in November of children's last year of primary school. The SDQ had replaced the CBQ as the primary screening measure for childhood psychological disorders in the UK and is widely used internationally. Correlations between the SDQ and the CBQ teacher scale are around or above .9 for both total and subscale scores (Goodman, 1997). The SDQ's 25 items are scored on a 3-point Likert scale, similar to the CBQ on which it was based, and offers an overall score of behaviour difficulties, with a score of 16 or more being deemed to identify probable disorders (classified borderline or abnormal). Subscales are available for conduct problems (five items), emotional symptoms (five items) and hyperactivity-inattention (five items). In a large sample of British children (Goodman, 2001), good levels of reliability are indicated with internal consistency = .73 (Cronbach's alpha) and retest stability after 4 to 6 months r = .62. In terms of validity, SDQ scores above the 90th percentile predicted a substantially raised probability of independently diagnosed psychiatric disorders (odds ratio 15.2 for teacher scales).

Demographics and IQ

Background information was collected on age, sex, English as an Additional Language (EAL: 0 = native speaker, 1 = EAL) and free school meals status at baseline (0 = no free meals, 1 = free meals), at the end of the next school year and at second follow-up. A free school meals index was created summing free school meals status at each measurement point and provides a proxy measure for social disadvantage.

IQ was measured using the British Ability Scales (BAS) short form ((Elliot et al., 1984) consisting of the following sub-tests: Naming Vocabulary, Recall of Digits, Similarities and Matrices.

Procedures

All children were pre-tested by researchers who had been trained over several days in correct test administration procedures. Researchers were either experienced psychologists and/or experienced primary teachers. At each follow-up, researchers tested the children blind to their intervention condition. Informed consent was obtained from both parents and teachers.

Interventions

Reading Recovery (RR)

Reading Recovery, developed by Marie Clay in collaboration with experienced primary educators on the basis of close observation, research and classroom development (Clay, 2005; Watson & Askew, 2009), conceptualises reading as involving a rapid processing of a range of available information, including alphabetic decoding, language, orthographic, semantic and syntactic information, with comprehension and fluency as target outcomes. The intervention was delivered by experienced primary teachers, trained in Reading Recovery by an accredited trainer, though some teachers were in their training year. Each session included reading of several graded

texts, word-level phonics work and writing. Children were withdrawn from class for individual tuition daily for half an hour, until either they were reading at the average level for their class, sufficient to manage independently in the classroom (a 'discontinued' programme), or they were classified as 'referred' back to school and in need of ongoing support (for full programme details, see Clay, 1993; Douëtil, Hobsbaum & Maidment, 2013). In the present study children received on average 21 weeks intervention, with an average of 77 sessions. Eighty-nine percent of the children made sufficient progress to be 'discontinued'. All children receiving Reading Recovery were included in the analyses reported below, irrespective of their discontinued status.

Fidelity of instruction is managed within the Reading Recovery programme in a 'three-tiered system' (Hurry, 1996; Schmitt, Askew, Fountas, Lyons & Pinnell, 2005), with national leaders based in universities running master's degrees and continuing professional development for teacher leaders, who work in local regions to provide professional development for school-based teachers. Teacher leaders undertake one year of in-depth training where theory is critiqued and their practice is repeatedly observed. They continue to be supported by their national leader and network with an annual visit to observe the teaching of children and teaching of teachers, and five days of professional learning alongside other Reading Recovery teacher leaders. School based teachers receive an initial one year training in their local context, involving fortnightly 2.5 hour professional learning and development sessions, ongoing individual support in school and monitoring. In subsequent years they attend six sessions and receive at least one support visit annually with a focus on deepening and refining decision-making and maintaining fidelity to the teaching procedures and key features of implementation. Data on the selection process, the number of sessions delivered

per child; individual child progress and outcomes, is gathered annually to provide information and to monitor the efficiency and efficacy of implementation at local and national levels (Amott, Hindmarsh & Morris, 2013). Around the time this intervention was conducted Clay evaluated the implementation of Reading Recovery in England and confirmed that it was on the whole being faithfully implemented, with the proviso that teachers were sometimes prevented from providing lessons daily due to other demands being made on their time (Douëtil et al.,, 2013), a criticism supported by the figures above that an average 77 sessions were delivered over 21 weeks intervention (ie. 3.67 sessions weekly rather than 5).

Phonological Training (PT)

Following Bradley and Bryant (1985), this involved sound awareness training plus word building with plastic letters. The training focused initially on alliteration and rhyme but also included work on boundary sounds, vowels and digraphs in response to the child's progress. Children also matched sounds with plastic letters and constructed words. Each child received forty, ten-minute, individual sessions, spread over seven months.

The five teachers who delivered the Phonological Training were all highly experienced primary teachers, but, unlike the Reading Recovery teachers, they were part of the research team and did not share details of the intervention with classroom teachers. They were given a one-day training session in the required techniques by Kirtley and MacLean, researchers who had taught the phonological programme in the original Bryant and Bradley studies (Bradley & Bryant, 1985; Bryant & Bradley, 1985; Kirtley et al., 1989), together with a training manual, and one week's practice delivering the programme to non-participating children. Problems encountered were discussed with Kirtley and MacLean, who also gave feedback. Further training sessions were held monthly for the duration of the intervention

period. Programme fidelity was monitored by the senior research officer who observed each member of the team teaching and listened to audio tapes of five sessions. The researchers recorded the content of every lesson, for every child, to facilitate monitoring. Lessons delivered were faithful to the intervention protocols though content varied in response to the individual child's needs and progress. At the end of the intervention, the performances of children by phonological tutor were compared. Regression analysis (controlling for pre-test scores on the Diagnostic Survey and the BAS Word Reading test) established that there were no significant differences on any of the first follow-up measures which were due to the tutor delivering PT

Provision for the Control Group

Children in both within and between school control groups received the standard provision available in their school. As weak readers, they often received extra, specialized help with reading, on average 21 minutes weekly. Classroom teachers of all participating children in the intervention year were asked to complete a questionnaire describing their practice (closely based on one devised by Ireson, Joscelyne, & Blatchford, 1994). One hundred and ten of the 127 teachers involved returned the questionnaire (86% response rate). There were no statistically significant differences between teachers from the different types of school (Reading Recovery, Phonological Training and Control) on the basis of average years teaching experience or the frequency with which they used most types of reading activities in class.

Results

Student demographics

Baseline assessments of literacy difficulties and EBD, and their association

The mean chronological age of the children at baseline was six years three months but they were performing roughly one year below age norms on the BAS Word Reading test (*M* = 5 years, 2 months, *SD* = 3 months). At baseline there were no significant group differences on the overall literacy z score or on the measure of emotional and behavioural difficulties, the CBQ or any of its sub scales (Table 3). A quarter to a third of the children in the study manifested emotional or behavioural problems at baseline and a third to nearly a half were deemed 'borderline' to 'abnormal' at age ten.

Table 3: Means and standard deviations of literacy and EBD scores by group

	Intervention Condition		
Baseline	Control Mean (*sd*)	RR Mean (*sd*)	PT Mean (*sd*)
BAS WR raw	3.21 (3.97)	2.43 (3.73)	3.16 (4.28)
Literacy z score	-0.12 (0.82)	-0.29 (0.51)	-0.24 (.70)
Total CBQ	6.77 (6.05)	6.77 (6.67)	7.343 (6.20)
Emotional	1.39 (1.62)	1.42 (1.86)	1.51 (1.58)
Conduct	1.39 (2.32)	1.12 (2.06)	1.85 (2.54)
Hyperactive	2.09 (1.93)	2.04 (2.09)	2.30 (1.83)

% CBQ score ≥ 9 (probable disorder)	27.8% (n = 25)	30.9% (n = 25)	33.3% (n = 29)
First follow-up			
Literacy z score	-0.38 (0.91)	0.40 (0.69) ***	-0.35 (.90)
Total CBQ	6.00 (5.50)	6.21 (5.76)	6.41 (5.33)
Emotional	1.29 (1.48)	1.54 (1.72)	1.11 (1.48)
Conduct	1.26 (2.18)	0.94 (1.61)	1.45 (2.21)
Hyperactive	1.74 (1.60)	1.58 (1.75)	2.00 (1.77)
% CBQ score ≥ 9 (probable disorder)	23.1% (n = 18)	26.3% (n = 21)	26.3% (n = 20)
Second follow-up			
Literacy z score	-0.21 (0.94)	-0.04 (0.83)	-0.09 (1.06)
Total SDQ	10.11 (6.63)	11.39 (7.04)	11.11 (6.56)
Emotional	2.22 (2.14)	2.97 (2.64)	2.40 (2.37)
Conduct	1.66 (2.26)	1.70 (2.29)	1.77 (2.32)
Hyperactive	4.56 (3.01)	4.49 (3.15)	4.64 (3.04)
% SDQ score ≥ 16 (borderline/abnormal)	35.6% (n = 26)	47.2% (n = 34)	42.9% (n = 33)

*** $p < .001$ between RR and Control and RR and Phono

CBQ = Child Behaviour Questionnaire; SDQ = Strengths and Difficulties Questionnaire.

The impact of EBD on literacy

The following analyses lay the ground for addressing Research Question 1 by establishing the effectiveness of the interventions at improving children's literacy, and directly address Research Questions 2 and 3, EBD as a predictor of literacy difficulties and as a moderator of the effectiveness of intervention. Predictive factors of children's overall literacy scores at first and second follow-up were examined through fixed entry multiple regression (SPSS Statistics 24) with control variables for poverty, English as an additional language, IQ, literacy and CBQ at baseline entered in Model 1. The intervention conditions (RR and Phono) were entered in the form of dummy variables (0=control; 1=intervention) in Model 2. Interaction terms between RR and CBQ and PT and CBQ were entered in Model 3 in separate regression analyses. The results of the models using total CBQ scores and interaction terms are presented in Table 4. Separate regression models were run for each CBQ subscale and its interaction with the literacy interventions and these are reported in the text where they differ significantly from the results for total CBQ. The results of the regression analyses were re-run in STATA 15 at school level and child level. The amount of variance explained at school level once the child level variables were accounted for was less than 5% both at first and second follow-up therefore multi-level analysis was deemed unnecessary.

Table 4: Factors predicting literacy at first and second follow-up

	Model 1				Model 2				Model 3 Interaction between RR and EDB				Model 3 Interaction between Phono and EBD			
	B	*SE*	*p*	β	*B*	*SE*	*p*	β	*B*	*SE*	*p*	β	*B*	*SE*	*p*	β
First follow-up (N = 249)																
Poverty (base)	-0.14	.09	.132	-0.08	-0.09	.08	.251	-0.05	-0.08	.08	.318	-0.04	-0.09	.08	.254	-0.05
EAL	0.45	.13	.001	0.18	0.37	.11	.001	0.15	0.37	.11	.001	0.15	0.36	.11	.001	0.15
BAS IQ	0.01	.01	.788	0.02	0.01	.01	.343	0.05	0.01	.01	.343	0.05	0.01	.01	.363	0.04
Literacy (base)	0.75	.07	.001	0.57	0.79	.06	.001	0.60	0.78	.06	.001	0.60	0.79	.06	.001	0.61
CBQ (base)	-0.02	.01	.025	-0.12	-0.02	.01	.015	-0.11	-0.02	.01	.015	-0.11	-0.01	.01	.185	-0.07
RR					0.88	.10	.001	0.44	0.74	.13	.001	0.44	0.88	.10	.001	0.44
PT					0.14	.609	.137	0.07	0.14	.09	.137	0.07	0.25	.13	.058	0.13
Intervention x CBQ (base)									0.02	.01	.129	0.10	0.02	.02	.231	0.09
Adjusted R^2			.40				.57				.57				.57	
Second follow-up (N = 226)																
Poverty (base)	-0.12	.05	.010	-0.16	-0.12	.05	.013	-0.15	-0.11	.05	.017	-0.15	-0.11	.05	.014	-0.15
EAL	0.34	.16	.032	0.13	0.32	.16	.042	0.12	0.33	.16	.038	0.12	0.32	.16	.039	0.12
BAS IQ	0.01	.01	.146	0.09	0.01	.01	.118	0.10	0.01	.01	.102	0.10	0.01	.01	.134	0.10

Literacy (base)	0.49	.08	.001	0.37	0.51	.08	.001	0.38	0.50	.08	.001	0.37	0.51	.08	.001	0.38
CBQ (base)	-0.02	.01	.021	-0.14	-0.02	.01	.018	-0.15	-0.03	.01	.003	-0.23	-0.01	.01	.264	-0.08
RR					0.24	.13	.072	0.12	0.03	.18	.866	0.02	0.24	.13	.073	0.12
PT					0.25	.13	.056	0.13	0.26	.13	.047	0.13	0.47	.18	.010	0.24
Intervention x EBD (base)									0.03	.02	.072	0.16	-0.30	.02	.083	-0.17
Adjusted R^2			.28				.29				.30				.30	

Poverty (base) = Free school meals at baseline; EAL = English as an additional language; BAS IQ = British Ability Scales short form IQ; Literacy (base) = literacy z score at baseline; CBQ (base) = Child Behaviour Questionnaire at baseline; Intervention x CBQ (base) = RR or PT x CBQ at baseline.

First follow-up

At first follow-up, literacy score at baseline was a powerful predictor of literacy 10 months later. RR significantly improved literacy progress over the 10 months (Cohen's *d* = .95) but PT did not. Regarding Research Question 2, CBQ independently predicted slower literacy progress at first follow-up. For the subscales, conduct symptoms significantly predicted slower literacy progress (conduct: *B* = -0.05, *SE* = .02, *p* = .029, β = -0.11) but the emotional and hyperactive subscales did not (hyperactivity showed a trend: *B* = -0.05, *SE* = .02, *p* = .056, β = -0.09). Regarding Research Question 3, there were no significant interaction effects between CBQ and RR, either for the total score or the sub-scales at baseline. A significant interaction effect was observed between PT and hyperactivity: it worked better at improving literacy for with children with fewer hyperactive symptoms than for those with more hyperactive symptoms (interaction term PT x Hyperactivity: *B* = -0.09, *SE* = .04, *p* = .027, β = 0.16). The selective effect of PT was further tested by dichotomising children into those with a maximum score of one hyperactive symptom at baseline (non-hyperactive, n = 122) and those with scores of two to six (hyperactive, n = 136). PT had a significant effect at first follow-up for non-hyperactive children (*B* = 0.33, *SE* = .15, *p* = .031, β = 0.16; Cohen's *d* = .35) but not for hyperactive children (*B* = -0.01, *SE* = .11, *p* = .933, β = -0.01). RR had large effects for both groups (non-hyperactive, Cohen's *d* = .89; hyperactive, Cohen's *d* = 1.03).

Second follow-up

At second follow-up (Table 4), literacy at baseline was still strongly predictive of literacy attainment and, with regard to Research Question 2, baseline CBQ still independently predicted literacy at 10 years old, also evident in the subscales of conduct disorder and

hyperactivity (conduct: B = -0.06, SE = .03, p = .034, β = -0.14; hyperactivity: B = -0.06, SE = .03, p = .049, β = -0.12). However, RR no longer significantly predicted higher literacy scores but showed a non-significant trend (p = .072; Cohen's d = .25). The effect of PT on literacy progress also showed a non-significant trend (p = .056; Cohen's d = .26) to improve literacy progress. Regarding Research Question 3, there were no significant interaction effects between CBQ or its subscales and RR. For the less intensive PT, a significant interaction effect was observed between intervention and hyperactivity in the opposite direction with children with fewer hyperactive symptoms benefitting more from PT (interaction term PT x Hyperactivity: B = -0.16, SE = .04, p = .008, β = 0.26). The selective effect of PT was further tested by dichotomising children into those with a maximum score of one hyperactive symptom at baseline (non-hyperactive, N = 122) and those with scores of two to six (hyperactive, N = 136). PT had a significant effect at second follow-up for non-hyperactive children (B = 0.53, SE = .19, p = .006, β = 0.26; Cohen's d = .59) but not for hyperactive children (B = 0.12, SE = .18, p = .666, β = 0.06).

The impact of literacy intervention on EBD

The large effect of RR on literacy progress at first follow-up provides the opportunity to test the hypothesis that literacy difficulties cause or exacerbate EBD using a quasi-experimental design (Research Question 1).

Separate regression analyses were run for total CBQ score (first follow-up) and total SDQ (second follow-up), and for each of their associated subscales; emotional sub-score; conduct sub-score, and hyperactive sub-score. In each regression analysis, the explanatory variables entered were: free school meal status, EAL, IQ, literacy ability at pre-test and the

pre-test score on the CBQ corresponding to the response variable (so pre-test score on emotional sub-scale where emotional sub-score was the response variable, and so on), the intervention conditions (in the form of dummy variables).

First follow-up

Score on the appropriate CBQ sub-scale at pre-test significantly predicted the score at first follow-up (Table 5). Neither reading intervention significantly predicted total CBQ at first follow-up, or any of its subscales, thus the hypothesis that literacy difficulties cause or exacerbate externalising and internalising problems was not supported.

Table 5: Factors predicting emotional and behavioural disorders at first and second follow-up

First follow-up (N=233)																
	Total CBQ				Emotional sub-scale				Conduct sub-scale				Hyperactive sub-scale			
	B	*SE*	*p*	β	*B*	*SE*	*p*	β	*B*	*SE*	*p*	β	*B*	*SE*	*p*	β
Poverty (base)	0.91	.53	.089	0.08	0.34	.18	.056	0.11	0.23	.19	.215	0.06	0.23	.17	.175	0.07
EAL	-0.91	.73	.217	-0.06	-0.25	.25	.305	-0.06	-0.45	.26	.081	-0.08	0.28	.24	.905	0.01
BAS IQ	-0.01	.02	.627	-0.03	0.06	.01	.473	-0.04	-0.01	.01	.237	-0.06	-0.01	.01	.280	-0.06
Literacy (base)	0.50	.41	.224	0.06	-0.10	.14	.477	-0,04	0.28	.14	.052	0.10	0.22	.13	.096	0.09
CBQ (base)	0.62	.04	.001	0.70	0.52	.05	.001	0.55	0.64	.04	.001	0.74	0.61	.04	.001	0.68
RR	0.58	.64	.359	0.05	0.40	.22	.064	0.12	-0.12	.22	.598	-0.03	-0.11	.21	.586	-0.03
PT	-0.01	.61	.994	0.00	-0.17	.21	.418	-0.05	-0.16	.22	.451	0.04	0.13	.20	.502	0.04
Adjusted R^2	.51				.33				.55				.48			
Second follow-up (N = 219)																
	Total SDQ				Emotional sub-scale				Conduct sub-scale				Hyperactive sub-scale			
Poverty (base)	0.56	.34	.103	0.11	0.10	.13	.475	0.05	0.29	.12	.017	0.16	0.24	.14	.096	0.10
EAL	-1.92	1.15	.095	-0.10	-0.91	.46	.050	-0.09	-0.08	.41	.852	-0.01	-0.63	.49	.206	-0.07
BAS IQ	0.02	.04	.538	-0.04	-0.01	.01	.365	-0.07	0.01	.01	.478	-0.05	-0.01	.02	.379	-0.06
Literacy (base)	-1.27	.62	.042	-0.14	-0.14	.25	.565	-0.04	-0.27	.22	.223	0.05	-0.68	.26	.011	-0.16
matched CBQ sub-	0.39	.07	.001	0.31	0.19	.01	.045	0.36	0.37	.06	.001	0.08	0.70	.09	.001	0.46

scale (base)																
RR	1.23	1.00	.216	0.09	0.77	.40	.057	0.15	0.20	.35	.573	0.04	-0.02	.43	.955	-0.01
PT	0.30	.97	.755	0.02	0.06	.39	.884	0.01	0.12	.35	.726	0.03	-0.07	.42	.875	-0.01
Adjusted R^2			.24				.04				.21				.31	

Poverty (base) = Free school meals at baseline; EAL = English as an additional language; BAS IQ = British Ability Scales short form IQ; Literacy (base) = literacy z score at baseline; CBQ (base) = Child Behaviour Questionnaire at baseline; Intervention x EBD (base) = RR or PT x the relevant sub-scale of the CBQ at baseline.

Second follow-up

At second follow-up a generally similar pattern of results emerged (Table 5), with score on the appropriate CBQ sub-scale at pre-test significantly predicting the score on SDQ at second follow-up. For the emotional sub-scale this relationship only just reached statistical significance and the model only explained four percent of the variance overall, compared with 24% for total SDQ and 31% for hyperactivity. Neither reading intervention significantly predicted total SDQ or any of its subscales. Lower literacy scores at baseline predicted higher scores on the conduct and hyperactive sub-scales.

Discussion

As has been consistently reported in the past there is co-morbidity between literacy difficulties and EBD. This is evident both at baseline, when a quarter to a third of the children, all with early literacy difficulties, are above the caseness threshold, and at age ten, when a third to a half are borderline or abnormal, 2.6 times that of a large normative British sample using the same measure of EBD, the SDQ (aged 5-15; Meltzer, Gatward, Goodman, & Ford, 2000).

We have sought to clarify the nature of the causal relationships, as this has implications for how co-morbidity is managed. The most persuasive evidence regarding the nature of the causal connection between literacy difficulties and EBD comes from the quasi-experimental design. Although RR has a large short term effect on literacy, this does not alter children's EBD, either internalising or externalising. Small effects of both RR and PT on literacy four years later (Cohen's d = .25 and .26 respectively) similarly do not influence

children's EBD. This does not support the hypothesis that literacy difficulties cause or exacerbate EBD (Research Question 1). There is also little evidence of literacy difficulties at baseline predicting the exacerbation of EBD at later follow-ups. This is consistent with previous research reviewed in the Introduction and is welcome news to the extent that literacy difficulties do not appear to worsen children's behaviour to any substantial degree. However, the present study does not offer a good test of relationships in this direction because all of the children were poor readers at baseline, substantially reducing the variability.

In the test of Research Question 2, that EBD predict literacy difficulties, children who have behaviour problems at six make less progress in literacy, both in the short term (nine months on) and in the longer term (four years on). This is true for the externalising symptoms of conduct disorder and hyperactivity but not for the internalising, emotional symptoms. This is consistent with Maughan et al. (2003), who found that effects of depression on literacy did not persist, and with the low stability of emotional disorders over time reported by Hinshaw (1992). This difference in the stability of internalising and externalising disorders is evident in the data reported here, where the regression models for conduct disorder and hyperactivity at baseline predicts 21% and 31% of the variance for the same symptoms four years on, but the regression model for emotional symptoms only predicts 4% of the variance four years on. The implication for teachers here is that they should notice and try to address their students' externalising behaviours to reduce their risk of later problems with literacy.

Regarding Research Question 3, that EBD may moderate the impact of literacy intervention, the concern raised by Rabiner et al., (2004) is that high levels of inattention may interfere with the effectiveness of a literacy intervention. We replicate these results for the

less intensive phonological intervention. The more intensive RR (on average 90 hours, compared to the 40 hour tutoring programme used in the Rabiner et al study) was equally effective for children with and without hyperactive symptoms. It is notable that an intensive intervention for older readers, responsive to need and available over a three year period was effective for those with attention problems (Roberts et al., 2015). It is plausible that children with attentional problems in addition to reading problems may need a larger dose of intervention than children with reading problems alone. We cannot state with certainty the precise nature of intervention suitable for children with co-occurring hyperactivity and reading difficulties, but our results support the notion that teachers should observe the effectiveness of intervention for individual pupils and be prepared to be respond with alternative evidence-based approaches where children fail to learn, perhaps using a Responsiveness to Intervention model (Velluntino, Scanlon, Zhang, & Schatschneider, 2008).

Limitations

There are two major limitations in the present study, the lack of random assignment to experimental conditions and the crudity of the measurement of behaviour disorders.

Quasi-experimental groups were well-matched at baseline but non-random assignment is not ideal, for well-rehearsed reasons (Hurry, 2017). However, quasi-experimental designs are a valuable alternative to Randomised Controlled Trials in addressing causal questions. Attaining true randomisation and true control group designs is often highly problematic (Plewis & Hurry, 1998). In a review of literacy and numeracy interventions, Seethaler and Fuchs (2005) found only 34 of 806 relevant articles (4.22%) used random allocation.

The measures of behaviour disorder are fairly crude screening measures, in particular for identifying specific disorders, although stability over the four and a half year study was reassuring, especially for hyperactivity ($\beta = 0.46$). The measure of emotional (internalising) disorder, which relies solely on teacher reports, is likely to be particularly weak as teachers may not be aware of children's internal states (Boetsch, Green, & Pennington, 1996; Puura et al., 1998).

Conclusions

Co-morbidity between literacy and behaviour difficulties has long been recognised as a concern and in the present study we estimated that around one third of children with literacy difficulties manifested a behaviour disorder at least at borderline level. The long-term consequences of such co-morbidity are likely to be serious (Maughan, Gray, & Rutter, 1985; Sylva, Scott, Totsika, Ereky-Stevens, & Crook, 2008). Unfortunately we found that an effective reading intervention, on the whole, failed to have any effect on behaviour, and our findings support Maughan and Carroll's (2006) conclusion that each disorder needs separate treatment.. The replication of Rabiner et al.'s (2004) finding that inattention/hyperactivity interferes with the effectiveness of less intensive reading programmes suggests that such children with attention problems need more intensive interventions. The simple message we can draw from this is that if a child has more than one problem they probably need more attention to thrive than a child with only one problem. The implication is that teachers and school programme designers need to take co-morbidity seriously, part of a recognition that effective programmes may be effective for some children but not for others.

References

Al Otaiba, S., & Fuchs, D (2006). Who are the young children for who best practices in reading are ineffective?: An experimental and longitudinal study. *Journal of Learning Disabilities, 39,* 414-431. doi.:10.1177/00222194060390050401

Amott, P., Hindmarsj, V., & Morris, H., (2013). Creatively responding to the imperative of scaling up every Child a Reader. In S. Burroughs-Lange & A. Ince (Eds.) *Reading Recovery and Every Child a Reader: History, policy, and practice.* London: IOE Press.

Anderson, J., Williams, S., McGee, R., & Silva, P. (1989). Cognitive and social correlates of DSM-III disorders in preadolescent children. *Journal of the American Academy of Child* & Adolescent Psychiatry, 28, 842-846. doi.:10.1097/00004583-198911000-00005

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94. doi.:10.1037/0033-2909.121.1.65

Benner, G., Nelson, J., & Epstein, M. (2002). Language skills of children with EBD: A literature review. *Journal of Emotional and Behavioral Disorders, 10,* 43-59. doi.:10.1177/106342660201000105

Bennett, K., Brown, K., Boyle, M., Racine, Y., & Offord, D. (2003). Does low reading achievement at school entry cause conduct problems? *School Science & Medicine, 56,* 2443-2448. doi.:10.1016/S0277-9536(02)00247-2

Boetsch, E., Green, P., & Pennington, B. (1996). Psychosocial correlates of dyslexia across the life span. *Development and Psychopathology, 8,* 539-562. doi.:10.1017/S0954579400007264

Boyes, M., Tebbutt, B., Preece, K., & Badcock, N. (2018). Relationships between Reading Ability and Child Mental Health: Moderating Effects of Self-Esteem, *Australian Psychologist*, *53*, (125-133). doi.:10.1111/ap.12281

Bradley, L., & Bryant, P. (1985). *Rhyme and Reason in Reading and Spelling*. University of Michigan Press, Ann Arbor.

Bryant, P., & Bradley, L. (1985). *Children's Reading Problems*. Oxford: Blackwell.

Bryant, P., MacLean, M., Bradley, L., & Crossland, J. (1990) Rhyme and Alliteration, phonemic detection, and learning to read. *Developmental Psychology, 26,* 429-438. doi.:10.1037/0012-1649.26.3.429

Carroll, J., Maughan, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders: Evidence for comorbidity. *Journal of Child Psychology and Psychiatry, 46,* 524-532. doi.:10.1111/**j**.1469-7610.2004.00366

Chadwick, O., Taylor, E., Taylor, A., Heptinstall, E., & Danckaerts, M. (1999). Hyperactivity and reading disability: A longitudinal study of the nature of the association. *Journal of Child Psychology and Psychiatry, 40,* 1039-1050. doi.:10.1111/1469-7610.00522

Cheung, C. H. M., Wood, A. C., Paloyelis, Y., Arias-Vasquez, A., Buitelaar, J. K., Franke, B., & Kuntsi, J. (2012). Aetiology for the covariation between combined type ADHD and

reading difficulties in a family study: the role of IQ. *Journal of Child Psychology and Psychiatry, 53,* 864–873. doi.:10.1111/j.1469-7610.2012.02527.x

Clay, M. (1985). *The early detection of reading difficulties: diagnostic survey with recovery procedures* (3rd ed.). Auckland: Heinemann.

Clay, M. (1993). *Reading Recovery: A guidebook for teachers in training.* Auckland: Heinemann.

Clay, M.M. (2002). *An observation survey of early literacy achievement (2nd ed.).* Portsmouth, NH: Heinemann.

Clay, M. (2005). *Literacy lessons designed for individuals, Part two: Teaching procedures.* Portsmouth, NH: Heinemann.

Clay, M (2013). *An Observation Survey Of Early Literacy Achievement*. London: Heinemann

Cook, T., & Campbell, D. (1979). *Quasi-Experimentation.* Boston: Houghton Mifflin Company.

Cornish, K. M., Savage, R., Hocking, D. R., & Hollis, C. P. (2011). Association of the DAT1 genotype with inattentive behavior is mediated by reading ability in a general population sample. *Brain and Cognition, 77,* 453–458. doi.:10.1016/j.bandc.2011.08.013

D'Agostino, J., Rodgers, E., & Mauck, S. (2018). Addressing Inadequacies of the Observation Survey of Early Literacy Achievement. *Reading Research Quarterly, 53,* 56-69. doi.:10.1002/rrq.181

Dahle, A., & Knivsberg, A-M., & Brit Andreassen, A. (2011). Coexisting problem behaviour in severe dyslexia. *Journal of Research in Special Educational Needs, 11*, 162-170. doi.:10.1111/j.1471-3802.2010.01190.x

Dockrell, J., & Hurry, J. (2018) The Identification of Speech and Language Problems in Elementary School: Diagnosis and Co-occurring needs. Research in Developmental Disabilities. Advance online publication. doi.:10.1016/j.ridd.2018.04.009

Douëtil, J., Hobsbaum, A., & Maidment, P. 2013). Reading Recovery, and early literacy intervention. In S. Burroughs-Lange & A. Ince (Eds.) *Reading Recovery and Every Child a Reader: History, policy, and practice.* London: IOE Press.

Elander, J., & Rutter, M. (1996). Use and development of the Rutter parents' and teachers' scales. *International Journal of Methods in Psychiatric Research, 6,* 63-78. doi.:10.1002/(SICI)1234-988X(199607)6:2<63::AID-MPR151>3.3.CO;2-M

Elliot, C., Murray, D., & Pearson, L., (1984). *British Ability Scales.* Londonn; NFER-Nelson.

Elliott, C. D., Smith, P. and McCulloch, K. (1997). *British Ability Scales Second Edition (BAS II). Technical Manual*. London: Nelson.

Feehan, M., McGee, R., Williams, S., & Nada-Raja, S. (1995). Models of adolescent psychopathology: Childhood risk and transition to adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry, 34,* 670-679. doi.10.1097/00004583-199505000-00019

Fergusson, D., & Horwood, L. (1995). Early conduct problems and later life opportunities. *Journal of Child Psychology and Psychiatry, 39,* 1097-1108. doi.:10.1111/1469-7610.00414

Fergusson, D., & Lynskey, M. (1997). Early reading difficulties and later conduct problems. *Journal of Child Psychology and Psychiatry, 38,* 899-907. doi.:10.1111/j.1469-7610.1997.tb01609.x

Fleming, J., & Offord, D. (1998). Epidemiology of childhood depressive disorders: A clinical review. *Journal of the American Academy of Child & Adolescent Psychiatry, 29,* 571-580. doi.:10.1097/00004583-199007000-00010

Gilger, J., Pennington, B., & DeFries, J. (1992) A twin study of the etiology of comorbidity: Attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy of Child & Adolescent Psychiatry, 31,* 343-348. doi.:10.1097/00004583-199203000-00024

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry, 38,* 581-586. doi.:10.1111/j.1469-7610.1997.tb01545.x

Goodman R. (2001). Psychometric properties of the Strengths and Difficulties Questionnaire (SDQ) *Journal of the American Academy of Child and Adolescent Psychiatry. 40,* 1337–1345. doi.:10.1097/00004583-200111000-00015

Huesmann, L. R., Eron, L. D., & Yarmel, P. W. (1987). Intellectual functioning and aggression. Journal of Personality and Social Psychology, 52(1), 232-240. doi.:10.1037/0022-3514.52.1.232

Hinshaw, S. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin, 111,* 127-155. doi.:10.1037/0033-2909.111.1.127

Hurry, J. (1996). What's so Special about Reading Recovery. *The Curriculum Journal, 7,* 93-108. https://doi.:10.1080/0958517960070107

Hurry, J (2017) Experimental design. In J. Swain (ed) *Designing Research in Education: Concepts and Methodologies.* London: Sage

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early reading intervention. *Journal of Reading Research, 30,* 227-248. doi.:10.1111/j.1467-9817.2007.00338.x

Ireson, J., & Blatchford, P. (1995) What Do Teachers Do? Classroom Activities in the Initial Teaching of Reading, *Educational Psychology*, 15, 245-56. doi.:10.1080/0144341950150302

Kirtley, C., Bryant, P., Maclean, M., & Bradley, L. (1989). Rhyme, rime and the onset of reading. *Journal of Experimental Child Psychology, 48,* 224-245. doi.:10.1016/0022-0965(89)90004-0

Jorm, A., Share, D., Matthews, R., & Maclean, R. (1986). Behavioural problems in specific reading retarded and general backward children: A longitudinal study. *Journal of Child Psychology and Psychiatry, 27,* 33-43. doi.:10.1111/j.1469-7610.1986.tb00619.x

Maughan, B., & Carroll, J. (2006). Literacy and mental disorders. *Current Opinion in Psychiatry*, *19*, 350-354. doi.:10.1097/01.yco.0000228752.79990.41

Maughan, B., Gray, G., & Rutter, M. (1985). Reading retardation and antisocial behaviour: a follow-up into employment. *Journal of Child Psychology and Psychiatry, 26,* 741-758. doi.:10.1111/j.1469-7610.1985.tb00588.x

Maughan, B., Pickles, A., Hagell, A., Rutter, M., & Yule, W. (1996). Reading problems and antisocial behaviour: Developmental trends in co-morbidity. *Journal of Child Psychology and Psychiatry, 37,* 405-418. doi.:10.1111/j.1469-7610.1996.tb01421.x

Maughan, B., Rowe, R., Loeber, R., & Stouthamer-Loeber, M. (2003). Reading problems and depressed mood. *Journal of Abnormal Child* Psychology, *31,* 219-229. doi.:10.1023/A:1022534527021

McGee, R., Prior, M., Williams, S., Share, D., & Sanson, A. (2002). The long-term significance of teacher-rated hyperactivity and reading ability in childhood: Findings from two longitudinal studies. *Journal of Child Psychology and Psychiatry, 43,* 1004-1017. doi.:10.1111/1469-7610.00228

McGee, R., Williams, S., Share, D., Andersen, J., & Silva, P. (1986). The relationship between specific reading retardation, general reading backwardness, behaviour problems in a large sample of Dunedin boys: A longitudinal study from five to eleven years. *Journal of Child Psychology and Psychiatry, 27,* 597-610. doi.:10.1111/1469-7610.00228

Medford, E., & McGeown, S. P. (2016). Social, emotional and behavioural influences on young children's pre-reading and word reading development. *Journal of Applied Developmental Psychology, 43,* 54-61. doi.:10.1016/j.appdev.2015.12.008

Meltzer, H., Gatward, R., Goodman, R., & Ford, T. (2000) *Mental health of children and adolescents in Great Britain.* London: The Stationary Office.

Mundy, L., Canterford, L.l Tucker, D., Bayer, J., Romaniuk, H., Sawyer, S., Lietz, P., Redmond, G., Proimos, J., Allen, N., & Patton, G.. (2017). Academic Performance in Primary School Children with Common Emotional and Behavioral Problems. *Journal of School Health, 87,* 593-601. doi.:10.1111/josh.12531

Neale, M.D., Christophers, U., & Whetton, C. (1989) *Neale Analysis of Reading Ability – Revised British Edition.* Windsor: NFER-Nelson.

NFER-Nelson (1985). Group Reading Test 6-12. Windsor: NFER-Nelson.

Plewis, I., & Hurry, J. (1998). A multilevel perspective on the design and analysis of intervention studies. *Educational Research and Evaluation, 4,* 13-26. doi.:10.1076/edre.4.1.13.13014

Puura, K., Almqvist, F., Piha, J., Moilanen, I., Tamminen, T., Kumpulainen, K., Räsänen, E., & Koivisto, A-M. (1998). Children with symptoms of depression—What do the adults see? *Journal of Child Psychology and Psychiatry, 39,* 577-585. doi.:10.1017/S0021963098002418

Rabiner, D., Malone, P., & the Conduct Problems Prevention Research Group (2004). The impact of tutoring on early reading achievement for children with and without attention

problems. *Journal of Abnormal Child Psychology, 32*, 273-284. doi.:10.1023/B:JACP.0000026141.20174.17

Riccio, C., & Jemison, S. (1998). ADHD and emergent literacy: Influence of language factors. *Reading and Writing Quarterly, 14*, 43-58. doi.:10.1080/1057356980140103

Roberts, G., Rane, S., Fall, A-M., Denton, C., Fletcher, J., & Vaughn, S. (2015). The Impact of Intensive Reading Intervention on Level of Attention in Middle School Students. *Journal of Clinical Child Adolescent Psychology, 44*, 942–953. doi.:10.1080/15374416.2014.913251

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology, 46*, 995-1007. doi.:10.1037/a0018880

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *Journal of Child Psychology and Psychiatry 8*, 1-11. doi.:10.1111/j.1469-7610.1967.tb02175.x

Rutter, M., Tizard, J., & Whitmore, K. (Eds.) (1970). *Education. health, and behaviour.* London: Longman.

Schmitt, M.C., Askew, B.J., Fountas, I.C., Lyons, C.A. & Pinnell, G.S. (2005). *Changing Futures: The influence of Reading Recovery in the United States.* Worthington, OH: Reading recovery Council for North America.

Seethaler, P., & Fuchs, L. (2005). A drop in the bucket: Randomized contact trials testing reading and math interventions. *Learning Disabilities Research and Practice, 20,* 98-102. doi.:10.1111/j.1540-5826.2005.00125.x

Shaywitz, B., Fletcher, J., & Shaywitz, S. (1995). Defining and classifying learning disabilities and attention deficit/hyperactivity disorder. *Journal of Neurology, 10* (Suppl), S50-S57. doi.:10.1177/08830738950100S111

Siperstein, G.N., Wiley, A.L., & Forness, S.R. (2011). School context and the academic and behavioral progress of students with emotional disturbance. *Behavioral Disorders*, *36,* 172-184.

Spira, E., & Fischel, J. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: A review. *Journal of Child Psychology and Psychiatry, 46,* 755-773. doi.:10.1111/j.1469-7610.2005.01466.x

Stevenson, J., Asheron, P., Hay, D., Levy, F., Swanson, J., Thapar, A., & Willcutt, E. (2005). Characterizing the ADHD phenotype for genetic studies. *Developmental Science, 8,* 115-121. doi.:10.1111/j.1467-7687.2005.00398.x

Stevenson, J., & Graham, P. (1993). Antisocial behaviour and spelling disability in a population sample of 13-year-old twins. *European Child and Adolescent Psychiatry, 2,* 179-191. doi.:10.1007/BF02098577

Suggate, S. P. (2016). A Meta-Analysis of the Long-Term Effects of Phonemic Awareness, Phonics, Fluency, and Reading Comprehension Interventions. *Journal of Learning Disabilities, 49,* 77-96. doi.:10.1177/0022219414528540

Sylva, K., Scott, S., Totsika, V., Ereky-Stevens, K., & Crook, C. (2008). Training parents to help their children read: A randomized control trial. *British Journal of Educational Psychology, 78,* 435-455. doi.:10.1348/000709907X255718

Tannock, R., Frijters, J. C., Martinussen, R., White, E. J., Ickowicz, A., Benson, N. J., & Lovett, M. W. (2018). Combined modality intervention for ADHD with comorbid reading disorders: A proof of concept study. *Journal of Learning Disabilities, 51*, 55 – 72. doi.:10.1177/0022219416678409

Terras, M.M., Thompson, L.C., & Minnis, H. (2009). Dyslexia and psycho-social functioning: An exploratory study of the role of self-esteem and understanding. *Dyslexia, 15,* 304-27. doi.:10.1002/dys.386

Trzesniewski, K., Moffitt, T., Caspi, A., Taylor, A., & Maughan, B. (2006). Revisiting the association between reading achievement and antisocial behavior: New evidence of an environmental explanation from a twin study. *Child Development, 77,* 72-88. doi:10.1111/j.1467-8624.2006.00857.x

Tsang, Vicky, Sam, K-L Sam, Wong, W-K, W-K Libby Cheng, Sin, K-F, Ho, F-C. (2016). A paradigm shift in intervention approaches for children with attention-deficits / hyperactivity disorders: a systematic review of psycho-behavioral interventions. *Acta Psychopathologica, 2,* 44-76. doi.:10.4172/2469-6676.100044

Vellutino, F., Scanlon, D., Zhang, H., & Schatschneider, C. (2008) Using response to kindergarten and first grade intervention to identify children at-risk for long-term

reading difficulties. *Reading and Writing, 21,* 437-480. doi: 10.1007/s11145-007-9098-2

Watson, B., & Askew, B. (2009). *Boundless horizons: Marie Clay's search for the possible in children's literacy.* Rosedale, New Zealand: Pearson Education.

Willcutt, E., & Pennington, B. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry, 41,* 1039-1048. doi.:10.1111/1469-7610.00691

Williams, S., & McGee, R. (1994). Reading attainment and juvenile delinquency. *Journal of Child Psychology and Psychiatry, 35,* 442-461. doi.:10.1111/j.1469-7610.1994.tb01733.x

Young, D. (1983). *Parallel Spelling Tests.* London: Hodder and Stoughton.