Clinical Psychology Review 67 (2019) 45–60

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

The association between poor reading and internalising problems: A systematic review and meta-analysis

Deanna A. Francis[a,*], Nathan Caruana[a], Jennifer L. Hudson[b], Genevieve M. McArthur[a]

[a] ARC Centre of Excellence in Cognition and its Disorders, Department of Cognitive Science, Macquarie University, New South Wales 2109, Australia
[b] Centre for Emotional Health, Department of Psychology, Macquarie University, Australia

HIGHLIGHTS

- Poor readers are at moderate risk for experiencing overall internalising problems.
- Poor readers are at greater risk for problems with anxiety than depression.
- Future research is needed to improve the quality of life for these poor readers.

ARTICLE INFO

Keywords:
Poor reading
Internalising
Anxiety
Depression

ABSTRACT

Numerous studies have demonstrated an association between learning disabilities and internalising problems such as anxiety and depression. However, our understanding of this association for people with specific types of learning disability – such as poor reading – is poorly understood. Here, we present the first systematic review and meta-analysis of studies that have examined associations between poor reading and internalising problems – including anxiety and depression – in children, adolescents, and adults. Our systematic search identified 34 studies comprising 16,275 participants (N = 2491 poor readers). Our meta-analysis revealed statistically significant differences between poor readers and typical readers on general measures of internalising problems (d = 0.41), as well as specific measures of anxiety (d = 0.41) and depression (d = 0.23). These outcomes suggest that poor readers are at moderate risk for experiencing internalising problems compared to typical readers, which appears to stem from a greater risk for anxiety than depression.

1. Introduction

1.1. Internalising problems, anxiety, and depression

Traditionally, "internalising" has been an umbrella term used to refer to inwardly focused emotional problems that contrast with outwardly focused "externalising" behavioural problems (Achenbach, 1966; Achenbach & Edelbrock, 1978). Internalising problems include numerous heterogeneous disorders such as anxiety, depression, trauma, and dissociative disorders. Defining these disorders is complex. The Diagnostic and Statistical Manual (DSM: now in its 5th edition; *DSM-5;*American Psychiatric Association [APA], 2013) defines such disorders based on clusters of symptoms that co-occur within a clinical population, and a diagnosis is based on the number and duration of symptoms that meet certain criteria. The DSM further categorises numerous subtypes of internalising disorders (i.e., generalised anxiety, separation anxiety, major depression), which are defined by constellations of unique and overlapping internalising symptoms (i.e., worry, low mood, withdrawal). In the current review, we focus on two of the most common internalising disorders in modern western society - anxiety and depression (Baxter, Scott, Vos, & Whiteford, 2012; Kessler, Chiu, Demler, & Walters, 2005).

There are numerous subtypes of anxiety such as generalised anxiety, separation anxiety, social anxiety, specific phobias, and panic disorder (APA, 2013), as well as other types of anxiety such as trait anxiety, state anxiety (Spielberger, Gorusch, Lushene, Vagg, & Jacobs, 1983), and test anxiety (Beidel, 1988). These anxiety problems comprise general symptoms of anxiety (i.e., avoidance, worry, physical symptoms) as well as more specific and defining anxiety symptoms (i.e., fear of a specific object). In terms of anxiety disorder subtypes, generalised anxiety is defined by pervasive worry in many areas for more days than not, and is typically accompanied by physical symptoms and distress. Separation anxiety is characterised by worry about separation from parents or caregivers, distress at the time (or ahead of time) of

* Corresponding author.
E-mail address: deanna.francis@mq.edu.au (D.A. Francis).

https://doi.org/10.1016/j.cpr.2018.09.002
Received 19 July 2017; Received in revised form 13 September 2018; Accepted 14 September 2018
Available online 15 September 2018

separation, reassurance seeking, and sleep disturbance. For social anxiety, the symptoms are predominantly fear of negative evaluation in social or performance situations, distress before or during the social situation, and avoidance of situations where negative evaluation is possible. Specific phobias are characterised by intense fear to a specific object or situation that is avoided or endured with severe distress. Panic disorder is defined by rapid and intense fear in the absence of an immediate feared object or situation (APA, 2013). In terms of other types of anxiety, state anxiety is defined as anxiety experienced at any given moment, while trait anxiety is defined as a person's tendency to interpret situations as threatening (Spielberger et al., 1983). Finally, test anxiety is defined as excessive fear or worry surrounding test situations, with particular worry about the consequences of performing poorly on a test (Beidel, 1988). Research also suggests that fear of negative evaluation is a core feature of test anxiety, and this type of anxiety has also been associated with generalised and social anxiety (Beidel & Turner, 1988; Bogels et al., 2010).

There are also numerous subtypes of depression such as major depressive disorder, persistent depression disorder (previously termed "dysthymia), and disruptive mood dysregulation disorder (APA, 2013). Major depressive disorder is defined by feeling very down or sad, lack of interest in previously enjoyed activities, change in appetite and weight, poor sleep, slowed motor movements, lack of energy, and poor concentration for more days than not - with these symptoms occurring over a two-week period. Persistent depression disorder is defined by similar symptoms but the symptoms occur for one year. Disruptive mood dysregulation disorder is defined by anger or temper outbursts that involve aggression towards another person for no good reason, the outbursts occur three or more times per week for at least a full year, and occur in different contexts such as home, school, or public places (APA, 2013).

Recent studies have shown that people with general learning disabilities are at higher risk for these internalising problems compared to the typical population (for a review, see Nelson & Harwood, 2011a, 2011b). However, it is not yet clear whether children with specific learning disabilities – such as poor reading, poor spoken language, poor attention – are at higher risk for certain types of internalising problems (e.g., anxiety, depression) or certain subtypes of such problems (e.g., separation anxiety, major depression disorder). To start clarifying this issue, this systematic review and meta-analysis examines the associations between one specific learning disability – poor reading – and two types of internalising problems – anxiety and depression.

1.2. Poor readers

The reading abilities of 16% of children fall below the average range for their age or grade, and 5% of children have significantly impaired reading skills for their age (Shaywitz et al., 1995). These "poor readers" have different kinds of reading problems, such as learning to read new words using the grapheme-phoneme correspondence (GPC) rules (i.e., poor "phonological recoding") or by recognising whole written words from memory (i.e., poor "visual word recognition"). Poor readers may also struggle to learn to read words fluently (i.e., poor "reading fluency") or with understanding the meaning of what they read (i.e., poor "reading comprehension"; Stuart & Stainthorp, 2016). While a proportion of poor readers present with just one of these reading problems (e.g., "phonological dyslexia", "surface dyslexia", "poor comprehenders"; McArthur et al., 2013; Nation, Cocksey, Taylor, & Bishop, 2010), the majority have a number of these reading problems (Castles & Coltheart, 1993; Goulandris & Snowling, 1991).

1.3. Internalising problems in poor readers

As mentioned above, it is unclear if poor reading is associated with an increased risk for internalising problems overall, or certain types or subtypes of internalising problems more specifically. To date, some studies have found that poor readers have more general internalising problems than typical readers (Boetsch, Green, & Pennington; 1996; Snowling, Muter, & Carroll, 2007), but others studies have not (Arnold et al., 2005; Miller, Hynd, & Miller, 2005). Similarly, some studies have found that poor readers have more problems with anxiety (Arnold et al., 2005; Bonifacci, Montuschi, Lami, & Snowling, 2014; Goldston et al., 2007) and depression than typical readers (Arnold et al., 2005; Daniel, Walsh, Goldston, Arnold, Reboussin, & Wood, 2006; Maughan, Rowe, Loeber, & Stouthamer-Loeber, 2003), while other studies have not (Grills, Fletcher, Vaughn, Barth, Denton, & Stuebing, 2014; Martinez & Semrud-Clikeman, 2004; Nelson & Gregg, 2012). Thus, there appears to be considerable inconsistencies between studies investigating internalising problems for poor readers.

Within the scientific field of reading research, inconsistent findings between studies about the characteristics of poor readers is typically a red flag for heterogeneity. More specifically, these mixed results suggest that only a proportion or "subgroup" of poor readers may have internalising problems. Thus, we propose eight potential moderators – six theoretically important moderators and two methodologically important moderators – that may explain these inconsistent findings (see Table S1 in Supplementary Materials for a more detailed rationale for examining these moderators). One is anxiety disorder subtype because there are numerous anxiety and depression disorders, and poor reading might be associated with some disorder subtypes (e.g., social anxiety) but not others (e.g., generalised anxiety). Another is poor reading subtype since there are numerous reading problems (i.e., poor visual word recognition, poor reading fluency) that may differentiate poor readers with and without internalising problems. A third is attention because poor attention is independently associated with both poor reading (Willcutt & Pennington, 2013) and internalising problems (Levy, Hay, Bennet, & McStephen, 2005), and research has shown that some children with internalising problems and poor reading also have problems with poor attention (Barbosa, Tannock, & Manassis, 2002; Barriga et al., 2002; Grills-Taquechel, Fletcher, Vaughn, Denton, & Taylor, 2013). A fourth is sex as females tend to experience more problems with anxiety and depression than males (Bruce et al., 2005; Mclean, Asnaani, Litz, & Hofmann, 2011). A fifth potential moderator is age, since the prevalence of internalising disorders varies markedly across childhood, adolescence, and adulthood (Ford, Goodman, & Meltzer, 2003; Kessler et al., 2005). A sixth is ethnicity because some ethnic minorities (i.e., Native American, Latino American, Asian American, African American) experience higher rates of internalising problems than European Americans (for a review see, Anderson & Mayes, 2010), and within these groups research suggests that Latino American youth in particular experience problems with anxiety (Ginsburg & Silverman, 1996), and depression (Umana-Taylor & Updegraff, 2007). A seventh is type of informant as information on internalising problems can be gathered from young people, parents, and teachers. Each informant offers a different perspective on the internalising difficulties experienced, and these reports can also be inconsistent particularly between children and parents (Grills & Ollendick, 2002; Safford, Kendall, Flannery-Schroeder, Webb, & Sommer, 2005). The final potential moderator is type of internalising measure, because clinical interviews (i.e., categorical measures) assess for the presence or absence of an anxiety or depression disorder, while questionnaires (i.e., dimensional measures) assess for constellations of general internalising symptoms (Krueger & Eaton, 2015).

1.4. The current study

To date, there have been two literature reviews of studies that have tested poor readers for internalising problems (Maughan & Carroll, 2006; Mugnaini, Lassi, La Malfa, & Albertini, 2009). Mugnaini et al. reviewed studies on general learning disabilities (including poor reading) and general internalising problems (including anxiety and depression) and examined whether attention and age influenced these

associations. They found that poor readers of all ages are at risk for internalising problems, that poor reading is specifically associated with both anxiety and depression, and that poor readers with attention problems are at higher risk for internalising problems than poor readers without attention problems.

Maughan and Carroll (2006) reviewed three studies of poor reading and anxiety (Arnold et al., 2005; Carroll et al., 2005; Carroll & Iles, 2006), and two studies of poor reading and depression (Arnold et al., 2005; Carroll et al., 2005). Regarding the former, Carroll et al. found that poor readers were at greater risk for generalised anxiety and separation anxiety but not specific phobias, and that this increased risk was not attributable to a shared association with inattention. Similarly, Arnold et al. reported that adolescent poor readers experienced higher rates of depression, trait anxiety, and somatic problems than controls, and that this association was also not attributable to problems with inattention. Finally, Carroll and Iles discovered that poor readers had high state, trait, and social anxiety. In terms of depression, Carroll et al. found no association between depression and poor reading, while Arnold et al. found that poor reading was associated with self-reported depression, which again was not attributable to poor attention. From these findings, Maughan and Carroll suggested that poor readers are at higher risk for anxiety than typical readers, and that this risk is not explained by poor attention. However, the association between poor reading and depression is less clear since one study found higher depression in poor readers while another did not, and it is also possible that this relationship may be moderated by problems with attention.

In their day, the reviews by Maughan and Carroll (2006) and Mugnaini et al. (2009) provided important and formative synopses of the existing evidence for the association between poor reading and internalising problems. However, these reviews are now 12- and 8-years-old respectively, and each synthesised the data using a narrative analysis instead of a meta-analysis due to lack of studies. Further, these reviews considered just one potential moderator of the association bertween poor reading and anxiety and depression (i.e., inattention), and neither reported the selection criteria for studies and hence cannot be replicated. Given the limitations of these foundational reviews, the primary goal of the current systematic review and meta-analysis was to determine if there is a reliable association between poor reading and internalising problems with acceptable heterogeneity between studies (Aim 1), and if so, evaluate if this reliable association is moderated by theoretical (i.e., anxiety disorder subtype, poor reading subtype, attention subtype, sex, age, ethnicity) or methodological important moderators (i.e., type of internalising measure, type of informant; Aim 2).

2. Method

This review was designed and reported in line with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Shamseer et al., 2015). The protocol for this review was published on the international prospective register of systematic reviews (PROSPERO; Reference: CRD42016049219, available from http://www.crd.york.ac.uk/PROSPERO).

2.1. Eligibility criteria

2.1.1. Participants

This review included studies with “poor readers” who were children aged 6–12 years, adolescents aged 13–18 years, and adults aged 18 + years. A person or group was considered to have poor reading if their score or mean score on a reading test was (1) at least one standard deviation below the average level for a person's age; (2) at least one year below the average level for a person's school grade; (3) statistically significantly lower than a score on an intelligence test; (4) statistically significantly poorer than a typical reading control group. A person or group was also considered to have poor reading if they met the diagnostic criteria for reading problems specified in any edition of the DSM, such as showing: (1) a specific difficulty in learning or using academic skills specific to reading; (2) performance on a standardised reading test that is significantly lower than the level of most students the same age; (3) a reading difficulty that is apparent during the early years of schooling; and (4) a reading difficulty that has no plausible explanation such as intellectual disability or sensory impairment.

In keeping with most studies of poor reading, and the DSM criteria, we only included studies where poor readers (1) had no known medical, psychological, or neurological problem that might explain their reading difficulty – with the exception of poor attention (see below); (2) had no known general developmental or intellectual or learning delay – with the exception of additional specific learning difficulties (i.e., mathematical problems such as dyscalculia) or language problems (i.e., specific language impairment) since many poor readers experience problems in these areas, and many studies do not screen for these additional learning difficulties; (3) spoke English as their primary language - since some languages (i.e., transparent languages such as Dutch) are easier to learn to read accurately than others (i.e., opaque languages such as English; Patel, Snowling, & De Jong, 2004); and (4) had been recruited based on reading test performance rather than self report - since research suggests that some poor readers may under-report reading difficulties (Snowling, Dawes, Nash, Hulme, 2012). As mentioned above, we included studies of poor readers with attention problems - since we wished to ascertain if this was a moderating factor. We recorded if scores on tests of inattention, hyperactivity, or the two combined were (1) if were significantly higher than a control group, (2) at least one standard deviation above the average range compared to age or grade norms, or (3) met criteria according to the DSM of any edition.

2.1.2. Reading data

We included studies that reported reading test data for poor phonological recoding (i.e., letter-sound identification, nonword reading accuracy), visual word recognition (i.e., sight word or irregular word reading accuracy), reading fluency (i.e., nonword or sight word fluency), and reading comprehension (i.e., understanding the meaning of text). Studies that *only* included data for poor reading comprehension were excluded since poor reading comprehension can stem from a problem with spoken language rather than a reading difficulty per se (Oakhill, Cain, & Bryant, 2003). However, we did include studies of poor readers with poor reading comprehension who also showed evidence of poor phonological recoding, visual word recognition, or reading fluency.

2.1.3. Internalising, anxiety, and depression data

We included studies that measured these disorders using clinical interviews or questionnaires that provided a mean raw or standardised score for general internalising symptoms (i.e., emotional distress, composite measure of anxiety and depression), anxiety symptoms (i.e., worry, state or trait anxiety) or anxiety disorder subtypes (i.e., generalised anxiety disorder), or depression symptoms (i.e., sadness) or depression disorder subtypes (i.e., major depression disorder). We considered a group of poor readers to have higher internalising problems if their mean raw or standardised score on the clinical interview or questionnaire was significantly higher than a (1) control group, (2) clinical cut-off point, or (3) the average level for a person's age according to normative data. We also considered poor readers to have higher internalising problems if they met criteria for a diagnosis of an anxiety or depression disorder (DSM any edition).

2.1.4. Exclusionary criteria

We excluded studies that (1) reported single case studies because effect sizes from case studies can have a disproportionate effect on mean effect sizes compared to group studies; (2) reported data in dissertations if the same data was presented in a peer reviewed paper; (3)

did not report reading test scores, which precluded us from determining whether or not poor readers' met our inclusion criteria; (4) reported composite measures of internalising and externalising problems; and (5) reported data from the same population in separate publications. In the last instance, we included the study with the largest sample size to maximise power for our analysis.

2.1.5. Information sources

We ran the searches for this review in July 2016 and August 2018. In July 2016, we used PsycINFO (Ovid, 1860 to July 2016), MEDLINE (Ovid, 1902 July 2016), EMBASE (Ovid, 1902 July 2016), WILEY, PubMed, Google Scholar, and PsycEXTRA for all available years. In August 2018, we used PsycINFO (Ovid, July 2016 to August 2018), MEDLINE (Ovid, July 2016 to August 2018), EMBASE (Ovid, July 2016 to August 2018), WILEY (July 2016 to August 2018), PubMed (July 2016 to August 2018), PsycEXTRA (July 2016 to August 2018), and Google Scholar for all available years. The searches were limited to human participants and studies published in English. The search strategy was adapted to meet the truncation and Boolean operations of each database. The following search strategy, which was conducted in PsycINFO, is provided as an example: "dyslexi$.tw or reading adj1 (disord$ or impair$ or defict$ or delay$ or dysfunction$ or achievement$ or difficult$ or problem$) or (poor adj1 (read$ or literacy)) or (specific adj1 learning dis$) AND (anxi$.tw or depress$.tw or psychosocial or internali?ing or psych$ comorbid$ or suicide$"). The list of full search terms is provided in Appendix A.

2.1.6. Study selection

We screened the studies in six steps, using the SysRev program for data management to organise and screen the citations. In Step 1, we developed an initial set of search terms. DF (first author) developed the search terms in consultation with reading and emotional health experts, and refined the search terms through the iterative search strategy shown in Fig. 1. In Step 2, we entered the search terms into the databases. DF conducted the searches and retrieved all articles to be reviewed. In Step 3, we retrieved the studies and screened the title and abstract of all studies. Two reviewers (DF and GM [fourth author]) read the title and abstract of all studies. The reviewers accepted any study that appeared to measure the association between poor reading and internalising, anxiety, or depression based on the information provided in the title and abstract, and rejected any study that clearly did not. The reviewers were blind to the journal titles, study authors, institutions, and publication status.

In Step 4, DF downloaded the full-text portable-document-format (PDF) files from the world wide web (www) for studies that were accepted by both reviewers, as well as studies that were accepted by one reviewer but rejected by the other. If a PDF of a study could not be found on the www, a request was made through Macquarie University Library resources, or by contacting the corresponding author of the paper in question. If there was no response from the corresponding author then we contacted the co-author. The study was excluded if the full-text could not be retrieved. DF and GM examined the full text PDFs of studies retrieved in Step 4. A study was accepted if it met the aforementioned inclusion and exclusion criteria, and rejected if it did not. All decisions were recorded in an excel spreadsheet. DF compared the decisions made by the two reviewers for the accepted and rejected studies. Studies accepted by one review but rejected by the other were discussed between the two reviewers to reach a joint decision regarding inclusion. The studies that were accepted by both reviewers were included in the review.

In Step 5, we identified the number of studies that measured internalising problems, anxiety, or depression in poor readers (for precision; see Fig. 1), and the number of studies that did not measure internalising problems, anxiety, or depression in poor readers (i.e., for specificity; see Fig. 1). DF compared the accepted and rejected studies from the two reviewers.

In Step 6, we screened the reference lists of the identified studies to determine if we had missed any critical studies with the search terms (i.e., sensitivity). We identified 13 missed studies (Aman, 1979; Boetsch, 1997; Chapman, Tunmer, & Prochnow, 2004; Daniel et al., 2006; Hughes et al., 2013; MacPhail, 2013; Martinez & Semrud-Clikeman, 2004; McGee, Williams, Share, Anderson, & Silva, 1986; Mercer, 2005; Murray, 1978; Sanson, Prior, & Smart, 1996; Scarborough & Parker, 2003; Snowling et al., 2007). We therefore instigated a reiterative process whereby we revised the search terms and repeated Steps 1 to 6 to see if we could capture these studies. In the first reiterative phase, we captured 9 of the 13 missing studies. In the second reiterative phase, we captured all 13 of the missing studies (see Fig. 1).

2.1.7. Data extraction

The data items extracted from each included study are displayed in Table 1. DF independently recorded the data in a customised excel data extraction form, and a second reviewer (NC; the second author) independently read the selected studies and checked the data for accuracy. There were no major discrepancies between the two reviewers. In all cases, DF and NC discussed the discrepancies with careful reference to the original article and a final decision was made. In terms of missing data, we contacted the corresponding author to request the data. If there was no response from the author in question, the co-author was contacted. If there was no response from either author, or the data was unavailable, then the study was excluded from the systematic review and meta-analysis.

2.1.8. Appraisal of methodological quality

Two reviewers (DF and NC) independently assessed methodological quality and risk of bias for all studies included in this review using a modified cross-sectional rating scale of the Newcastle-Ottawa Scale (NOS; Wells et al., 1996). The NOS is recommended by the Cochrane Collaboration to evaluate risk of bias for non-randomised trials (Higgins & Green, 2008). The scale comprises eight multiple-choice questions that were tailored to the characteristics of the studies included in this review. The two reviewers rated the risk of bias for each outcome (i.e., internalising problems, anxiety, depression) according to the following criteria: (1) selection of groups, which assessed the representativeness and characteristics of the sample; (2) exposure, which assessed the comparability of participants and level of control; and (3) outcome, which assessed the quality of the outcomes and the statistical analysis of the outcomes. A maximum of nine points were assigned per study per outcome. These ratings were: low risk of bias (7 to 9 points); moderate risk of bias (4 to 6 points); and high risk of bias (0 to 3 points). In other words, low scores indicated high risk of bias. A priori, we decided that studies with low and moderate risk of bias would be included in the meta-analysis, and studies with high risk of bias would be excluded from the meta-analysis. There was moderate agreement between the two reviewers on the risk of bias ratings (kappa = 0.44), and any discrepancies were satisfied upon discussion between the two reviewers and reference to the original study.

2.1.9. Data analysis

The results of the individual studies measuring the relationship between poor reading and internalising problems, anxiety, and depression were meta-analysed using the Comprehensive Meta-Analysis program (Borenstein, Hedges, Higgins, & Rothstein, 2006). We calculated Cohen's *d* effect sizes using the means and standard deviations for each outcome for the poor reading and control groups. If the means and standard deviations were not reported, we used the procedures of the Comprehensive Meta-Analysis program to calculate the effect size using the available data (i.e., odds ratios or proportions). If the study did not report an appropriate control group, and a standardised test with normative data was administered, we imputed the normative data (i.e., St Sc., $M = 100$, $SD = 15$) and sample size of the normative population in the analysis. Hedges *g* was also calculated in an attempt to correct for

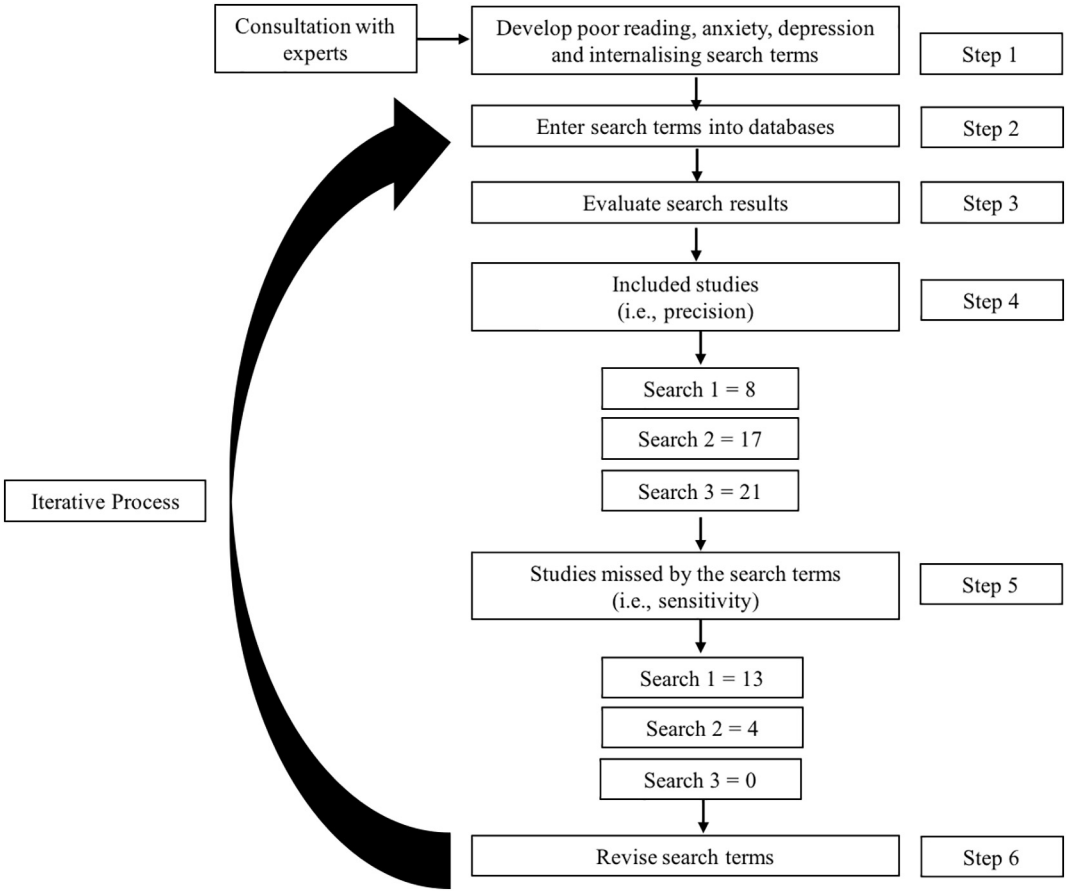

Fig. 1. The reiterative search strategy process.

Table 1
Category, name, and description of the data extracted from each study.

Category	Name	Description
Source	Citation	Author, publication year, title, journal
	Publication type	Journal article, dissertation, "other" unpublished work
Eligibility	Inclusion criteria	Recruitment criteria and any neurological/medical/psychological comorbidities
	Exclusion criteria	Record why the study was excluded
Participants	Study design	Longitudinal, cross-sectional, case-control, correlation
	Sample size	Total and group sample sizes
	Poor reading	Criteria used to define reading ability of the groups
	Age and grade	Range, mean (*M*), standard deviation (*SD*) of the groups
	Sex	Number of females and males in each group
	Type of school	Public, private, learning specialist, university
	Ethnicity	Division of ethnicity for the groups
	Sociodemographic status	Highest level of education (adult participants), or mother and father education, or household income (child participants)
	Intelligence	Test scores from intelligence assessment
Reading	Tests	Tests used to measure reading ability
	Impaired	*M, SD,* and effect sizes on tests showing impairment
	Unimpaired	*M, SD,* and effect sizes on reading tests not showing impairment
Internalising	Tests	Tests used to measure internalising problem
	Internalising problem	Type of internalising, anxiety, or depression
	Subtype	The subtype measured (e.g., generalised anxiety, major depression)
	Informant	Self-, parent-, or teacher report
	Criteria	Criteria used to define internalising problems (i.e., clinical cut off score or normative data)
	Impaired	*M, SD,* and effect sizes on tests showing impairment
	Spared	*M, SD,* and effect sizes on reading tests not showing impairment
	Type of measure	Questionnaire or clinical interview
Theory	Quotes	Quotes and theoretical motivation for the study

small sample sizes (Hedges & Olkin, 1985). However, because the values of *g* and *d* were almost identical, we only reported Cohen's *d* as it is regularly reported in meta-analyses (Higgins & Green, 2008). Positive effect sizes were interpreted as showing higher internalising problems, anxiety, or depression for poor readers compared to controls, and Cohen's *d* was interpreted as small (0.30), moderate (0.50), and large (0.80; Cohen, 1988).

We expected considerable heterogeneity between studies because of methodological differences between the studies included in this review (i.e., the studies adopted different criteria and administered different

tests to measure reading and internalising problems in poor readers). Therefore, we planned to use a random effects model that accommodates the variation in effect sizes between studies (Borenstein et al., 2006). As a test of heterogeneity, we also calculated the I^2 statistic. We interpreted I^2 as showing no heterogeneity (0%), low heterogeneity (25%), moderate heterogeneity (50%), and high heterogeneity (75%; Higgins, Thompson, Deeks, & Altman, 2003).

In some studies, more than one test was administered to measure internalising problems, anxiety, or depression. For studies that reported data for multiple tests per outcome, we (1) calculated Cohen's *d* for each test; and (2) calculated an average Cohen's *d* for each outcome. Thus, there was one overall effect size for each sample in each study that measured internalising problems, anxiety, or depression. For longitudinal studies that reported outcome data across multiple time points, we calculated Cohen's *d* using the earliest time point at which poor reading and internalising problems, anxiety, or depression were measured concurrently. This procedure was followed to counteract the effect of any intervention on the outcomes. Finally, some studies included data for multiple groups of poor readers. In some of these studies, these groups were compared to one control group. In these cases, we combined the average data of all the poor reading groups and calculated the average Cohen's *d* for each outcome. In other studies, each group of poor readers had its own control group. In these cases, we considered the poor reading group and the respective control group as a single study. These procedures were followed to prevent violating the assumption of independent data points, where greater weight is assigned to studies with multiple effect sizes (Lipsey & Wilson, 2001).

Publication bias was assessed for each outcome by visually inspecting the funnel plots, examining Egger's test (Egger, Davey Smith, Schneider, & Minder, 1997), and following the trim and fill procedure (Duval & Tweedie, 2000). The funnel plot displays the standard error on the *y*-axis, and the standardised mean difference on the *x*-axis. Egger's test determines the asymmetry of the funnel plot. And the trim and fill procedure estimates the number of missing studies ("trim"), computes artificial studies to add to the funnel plot ("fill"), and produces an unbiased estimate of the effect.

We also planned to use a series of subgroup analyses to determine if any statistically reliable association with acceptable heterogeneity was moderated by any of the eight variables outlined in the Introduction. This analysis was only planned if there were 10 or more studies per moderator. If this was not possible due to insufficient data, a narrative synthesis of the results was planned instead.

2.1.10. Coding of internalising problems

We extracted data for each study that measured poor reading and general internalising problems (i.e., combined anxiety and depression, general internalising problems), anxiety symptoms and disorders (i.e., overall anxiety; generalised anxiety), and depression symptoms and disorders (i.e., sadness; major depression disorder).

2.1.11. Coding of moderators

2.1.11.1. Anxiety disorder subtype. Studies were separated into five groups based on the anxiety disorder subtype, including (1) generalised anxiety; (2) separation anxiety; (3) social anxiety; (4) specific phobias; and (5) panic disorder. It is noteworthy, that while this review includes studies of general anxiety, trait anxiety, state anxiety, and test anxiety, these types of anxiety are not considered to be disorders according to the DSM-5, and hence were not included as subgroups of anxiety disorder subtype.

2.1.11.2. Depression disorder subtype. Studies were separated into three groups based on the depression disorder subtype examined, including (1) major depression disorder, (2) persistent depression disorder; and (3) disruptive mood dysregulation disorder.

2.1.11.3. Poor reading subtype. Studies were separated into four groups based on the type of reading problem: (1) phonological recoding; (2) visual word recognition; (3) reading fluency; and (4) mixed poor reading (i.e., any combination of 1 to 3).

2.1.11.4. Attention subtype. Studies were separated into three groups based on the type of attention problems that poor readers experienced: (1) inattention, defined by difficulty maintaining attention, short attention span, and distractibility; (2) hyperactivity, defined by impulsivity and difficulty regulating attention; and (3) combined attention problems, defined by difficulties with inattention and hyperactivity.

2.1.11.5. Sex. Studies were separated into three groups based on differences between the numbers of males (M) and females (F) in the sample: (1) > 10 males than females (M > F); > 10 females than males (F > M); and (3) equal numbers of males and females (M = F). We used 10 as a minimum difference between sexes since it ensured a substantial imbalance in the number of males and females in the samples included in this review.

2.1.11.6. Age. Studies were separated into three groups based on the age of participants: (1) children aged 6 to 12 years; (2) adolescents aged 13 to 18 years; (3) adults aged 18 years and above; and (4) mixed age, comprising any combination of 1 to 3.

2.1.11.7. Ethnicity. Studies were separated into four groups based on the number of participants identified as (1) > 60% of the sample were Caucasian; (2) > 60% of the sample were African-American; (3) > 60% of the sample were Hispanic; and (4) there were mixed ethnicities reported in the sample (i.e., any combination of 1 to 3).

2.1.11.8. Type of informant. Studies were separated into four groups based on the informant, including (1) self-reports, where the individual reports on their own internalising problems; (2) parent-reports, where the parent reports on their child or adolescent's internalising problems; (3) teacher-reports, where the teacher reports on a child or adolescent's internalising problems; and (4) multi-informant reports, where a combination of 1 to 3 was administered.

2.1.11.9. Type of internalising measure. Studies were separated into two groups based on the type of measure administered, including (1) questionnaires that assess for general internalising symptoms, and (2) clinical interviews that assess for the presence or absence of anxiety or depression disorders.

3. Results

3.1. Study selection

In the initial search in July 2016, our search identified a total of 5058 articles. Having removed 1448 duplicates, we screened the titles and abstracts of the remaining 3610 articles against the inclusion criteria. We identified 283 potentially relevant articles, and excluded 3327 articles as irrelevant. We excluded a further 250 articles. One of these articles included four separate studies (Boetsch et al., 1996), which were included as four separate studies (Boetsch et al., 1996a; 1996b; 1996c; 1996d). This left us with 30 studies from 27 articles.

In the search in August 2018, our search identified a total of 1272 articles. We removed 101 duplicates, leaving us with 1171 titles and abstracts to compare against the inclusion criteria. We excluded 1143 articles, leaving 28 potentially relevant articles. We excluded a further 25 articles. This left us with three articles. One of these articles (Wu, 2018) included two separate studies, which were included as two separate studies (Wu, 2018a; 2018b). Considered together, the July 2016 and August 2018 searches left us with 34 studies in total from 30 articles (see Fig. 2 for the articles identified from the searches).

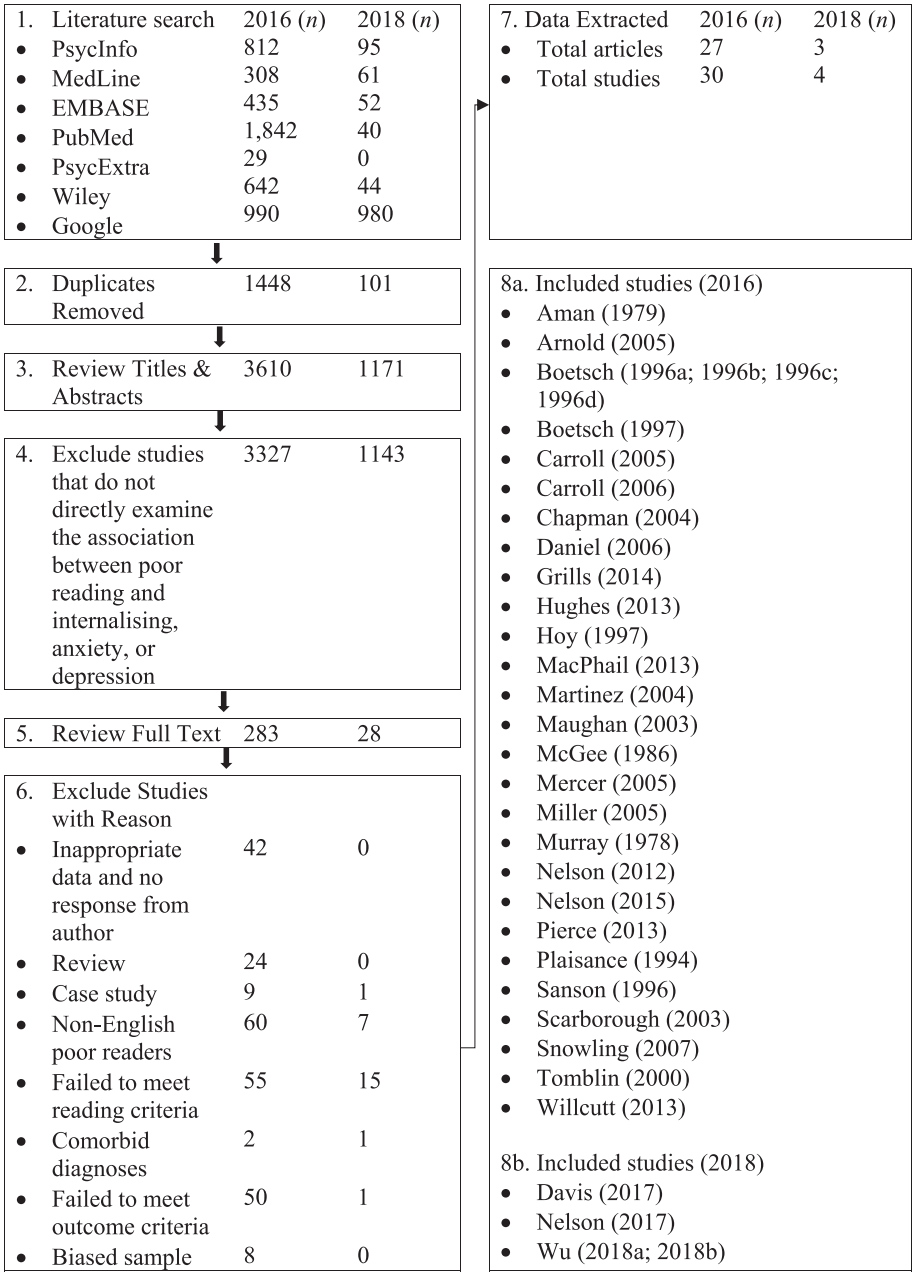

Fig. 2. The study selection process showing the number of studies retrieved, excluded, and included in the meta-analysis.

3.2. Risk of bias within studies

Two authors, DF and NC, independently evaluated the 34 studies included in this review for risk of bias. Sixty-seven risk of bias ratings were assigned, and the ratings ranged from moderate to low risk of bias (see Table 2 for the risk of bias ratings for each study). The 34 studies included in this review ranged from low to moderate risk of bias and no study was excluded for high risk of bias.

3.3. Study characteristics

The characteristics of the 34 studies included in this review are shown in Appendix C. Of these, 25 studies were conducted in the US (82%), 5 in Europe (14%), and 3 in New Zealand and 1 in Australia (11%). There were 8 longitudinal studies, 24 cross-sectional studies, and 2 treatment studies. There were 25 journal articles and 5 dissertations. Some studies tested poor readers for more than one type of internalising problem. For example, there were 14 studies comprising 12,092 participants (n = 1147 poor readers) that measured general internalising problems in poor readers. There were 22 studies comprising 11,372 participants (n = 1732 poor readers) that measured anxiety in poor readers. There were 23 studies comprising 10,714 participants (n = 1950 poor readers) that measured depression in poor readers. Table 3 summarises the size of the effects between poor reading and control groups for each of these general internalising problems, anxiety, and depression. In terms of the measures administered, most studies assessed internalising problems, anxiety, and depression using questionnaires rather than clinical interviews (with the exception of three studies: Carroll et al., 2005; Willcutt et al., 2013; Wu, 2018). There were numerous questionnaires administered. The most common measure of general internalising problems was the Child Behaviour Checklist (CBCL; Achenbach, 1991), which was administered in 7 of 14 internalising studies (Arnold et al., 2005; Boetsch et al., 1996; Chapman et al., 2004; Scarborough & Parker, 2003; Tomblin et al.,

Table 2
The risk of bias ratings assigned to each included study (higher stars indicate lower risk of bias) using the Newcastle Ottowa Scale.

	Sample characteristics					Outcome assessment			Statistical tests			Total risk of bias		
	Representativeness	Sample size	Response rate	Measurement tool	Controlled factors	Anxiety	Depression	Internalising	Anxiety	Depression	Internalising	Anxiety	Depression	Internalising
Aman, 1979	*	b	*	**	**	*			*			8		
Arnold, 2005	*	b	*	**	**	c	c	c	*	*	*	7	7	7
Boetsch, 1997	c	b	c	**	*		d			*			4	
Boetsch, 1996a	c	b	c	**	**	c	c	c	*	*	*	5	5	5
Boetsch, 1996b	c	b	c	**	**		c			*			5	
Boetsch, 1996c	*	b	c	**	**	*	c	*	*	*	*	7	6	7
Boetsch, 1996d	*	b	c	**	**	*	*	c	*	*	*	7	7	6
Carroll, 2005	*	b	*	**	**	*	*	*	*	*	*	8	8	8
Carroll, 2006	*	b	c	**	**	c			*			6		
Chapman, 2004	*	b	c	**	*	d	d	d	*	*	*	5	5	5
Daniel, 2006	*	b	*	**	**		*			*			8	
Davis, 2017	c	b	c	**	**	c			b			4		
Grills, 2014	*	b	*	**	**	c			*			7		
Hoy, 1997	c	b	c	**	**	c	c		*	*		5	5	
Hughes, 2013	c	b	c	**	*		c			*			4	
MacPhail, 2013	c	*	c	**	**		c			*			6	
Martinez, 2004	*	b	c	**	*	c	c	c	*	*	*	5	5	5
Maughan, 2003	c	*	*	**	**		c			*			7	
McGee, 1986	*	*	*	**	**	d	d		*	*		8	8	
Mercer, 2005	*	b	*	**	*	c	c		*	*		6	6	
Miller, 2005	c	b	c	**	**	*	*	d	*	*	*	6	6	5
Murray, 1978	c	b	c	**	**	c	d		*	*		5	5	
Nelson, 2017	c	b	c	**	**	c	c		*	*		5	5	
Nelson, 2015	c	b	c	**	**	c			*			5		
Nelson, 2012	*	b	c	*	**	c	c		*	*		5	5	
Pierce, 2013	*	b	c	**	**	d		d	*		*	6		6
Plaisance, 1994	*	b	b	**	**	d	*		*	*		6	7	
Sanson, 1996	c	b	*	**	**	d			*			6		
Scarborough, 2003	c	b	*	**	**			d			*			6
Snowling, 2007	c	b	*	**	**			d			*			6
Tomblin, 2000	*	b	*	**	**			d			*			7
Willcutt, 2013	*	*	c	**	**	*	*	*	*	*	*	8	8	8
Wu, 2018a	c	b	c	**	**	*	*	d	*	*	*	6	6	5
Wu, 2018b	c	*	c	**	**	d	d	d	*	*	*	6	6	6

Note. Representativeness (* = truly or somewhat representative of the average in the target population; c = selected group of users; d = no description of the sampling strategy); sample size (* = justified and satisfactory; b = not justified); non-respondents (* = satisfactory response rate > 60%; b = unsatisfactory response rate < 60%; c = no description of the response rate); poor reading (** = validated measurement tool; b = non-validated measurement tool, but the tool is available or described; c = no description of the measurement tool); control (** = English poor readers and additional factors controlled; * = English poor readers only with no additional factors controlled); outcome assessment for anxiety, depression, internalising (* = interview, or self and parent or teacher report; c = self report; d = parent or teacher report); statistical tests for anxiety, depression, internalising (* = statistical test is clearly described and appropriate; b = statistical test is inappropriate, incomplete or not described); total risk of bias rating /9 for anxiety, depression, internalising.

Table 3
The standardised mean difference (Cohen's d) and 95% confidence intervals (95% CI) for poor reading (PR) and control (C) groups for included studies that measured the association between poor reading and internalising problems, anxiety, or depression.

Study	*N*		Anxiety			Depression			Internalising		
			Cohen's *d*	95% CI		Cohen's *d*	95% CI		Cohen's *d*	95% CI	
	PR	C		Lower	Upper		Lower	Upper		Lower	Upper
Aman, 1979	28	28	0.59*	0.06	1.11						
Arnold, 2005	94	94	0.45*	0.15	0.73	0.51**	0.21	0.80	0.19	−0.09	0.48
Boetsch, 1997	28	38				0.38	−0.11	0.87			
Boetsch, 1996a	18	18				0.46	−0.20	1.122	0.88*	0.19	1.56
Boetsch, 1996b	70	67				−0.51*	−0.84	−0.16			
Boetsch, 1996c	26	26				−0.25	−0.80	0.31	0.52	−0.03	1.07
Boetsch, 1996d	98	118				−0.01	−0.27	0.26			
Carroll, 2005	289	5463	0.59*	0.14	1.02	0.08	−0.32	0.48	0.45**	0.33	0.57
Carroll, 2006	16	16	1.35*	0.56	2.15						
Chapman, 2004	38	55	0.21	−0.20	0.62	0.70**	0.27	1.13	0.40	−0.01	0.82
Daniel, 2006	94	94				0.52*	0.09	0.95			
Davis, 2017	22	21	0.86*	0.24	1.49						
Grills, 2014	73	31	−0.07	−0.49	0.34						
Hoy, 1997	184	140	0.16	0.61	0.50	0.28**	0.72				
Hughes, 2013	19	21				0.43	−0.20	1.05			
MacPhail, 2013	58	67				0.20	−0.15	0.55			
Martinez, 2004	30	30	0.14	−0.37	0.64	0.39	−0.11	0.90	0.33	−0.17	0.84
Maughan, 2003	134	1282				0.61**	0.27	0.94			
McGee, 1986	40	436	0.46	−0.01	0.93						
Mercer, 2005	25	56	−0.15	−0.62	0.32	−0.20	−0.67	0.26			
Miller, 2005	20	59	0.41	−0.06	0.90	0.18	−0.30	0.66	−0.06	−0.55	0.41
Murray, 1978	104	104	0.21	−0.06	0.48	−0.30*	−0.57	−0.02			
Nelson, 2017	110	110	−0.02	−0.28	0.25	0.23	−0.04	0.49			
Nelson, 2015	50	50	0.41*	0.02	0.81						
Nelson, 2012	90	60	0.11	−0.21	0.44	−0.09	−0.41	0.23			
Pierce, 2013	47	1983	0.91**	0.62	0.20				0.61**	0.32	0.90
Plaisance, 1994	50	58	0.88**	0.48	1.27	0.56*	0.18	0.95			
Sanson, 1996	232	42	0.97**	0.57	1.37						
Scarborough, 2003	44	28							0.35	−0.24	0.93
Snowling, 2007	21	17							0.86*	0.18	1.53
Tomblin, 2000	174	2368							0.35**	0.19	0.50
Willcutt, 2013	429	419	0.52**	0.21	0.81	0.93**	0.49	1.37	0.56**	0.40	0.72
Wu, 2018a	9	141	−0.25	−0.93	0.42	−0.01	−0.68	0.67	0.59	−0.09	1.27
Wu, 2018b	121	244	0.44*	0.22	0.66	0.32*	0.09	0.54	0.11	−0.01	0.42

* < 0.05.
** < 0.001.

2000; Wu, 2018). The three most common measures of anxiety were the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1983; Arnold et al., 2005; Carroll & Iles, 2006; Hoy, 1997), the Multidimensional Anxiety Scale for Children (MASC; March 1997; Grills et al., 2014; Mercer, 2005; Murray, 1978), and the CBCL (Arnold et al., 2005; Chapman et al., 2004; Wu, 2018) – each administered in three anxiety studies. The most common measure of depression was the Child Depression Inventory (CDI; Kovacs, 1992), which was administered in 5 of 23 depression studies (Boetsch et al., 1996c; Boetsch et al., 1996d; Boetsch, 1997; Hughes, 2013; Miller et al., 2005). It is important to note that different measures were used to measure these internalising problems, and these differences can make comparisons between studies difficult, which we consider further in the Discussion.

3.4. Participant characteristics

In studies that reported age, the mean age of poor readers was 15.07 years. Most studies included samples of both male and female participants, but the total proportion of poor readers included more male (68.37%) than female participants (31.62%). Few studies reported the ethnicity of participants included in their sample. For studies that did report this information, studies recruited poor readers who were mostly Caucasian (35.30% of included studies), African-American (5.88% of included studies), Hispanic (2.94% of included studies), or mixed ethnicity (8.82% of included studies). It is difficult to determine whether any ethnic groups were underrepresented in the included studies, as 16 of 34 studies (48.05%) failed to report ethnicity of the sample. As defined by the Eligibility Criteria, all studies included in this review recruited participants with English as their primary language.

3.5. Meta-analysis of the association between poor reading and internalising problems

3.5.1. Main analysis

Fourteen included studies provided data on the association between poor reading and internalising problems. Fig. 3 shows the 14 effect sizes from 14 studies comparing the standardised mean difference and 95% CIs between poor reading and control groups on measures of internalising problems (n poor reading groups = 14, sample size = 1147, mean sample size = 81.93; n control groups = 14, total sample size = 10,945, mean sample size = 781.79). The overall effect was positive, moderate, and statistically significant (d = 0.41), 95% CI [0.31, 0.51], $p < 0.001$, and low heterogeneity was also present I^2 = 35.07%, Q (13) = 20.02, $p > 0.05$.

3.5.2. Subgroup analysis

We could not conduct any of the planned subgroup analyses because there were fewer than 10 studies per subgroup, precluding any reliable comparison of effect sizes between subgroups for each potential moderator. The studies that examined the association between poor reading and internalising problems, and the number of studies examining each moderator are shown in Table S2 in the Supplementary Materials. We

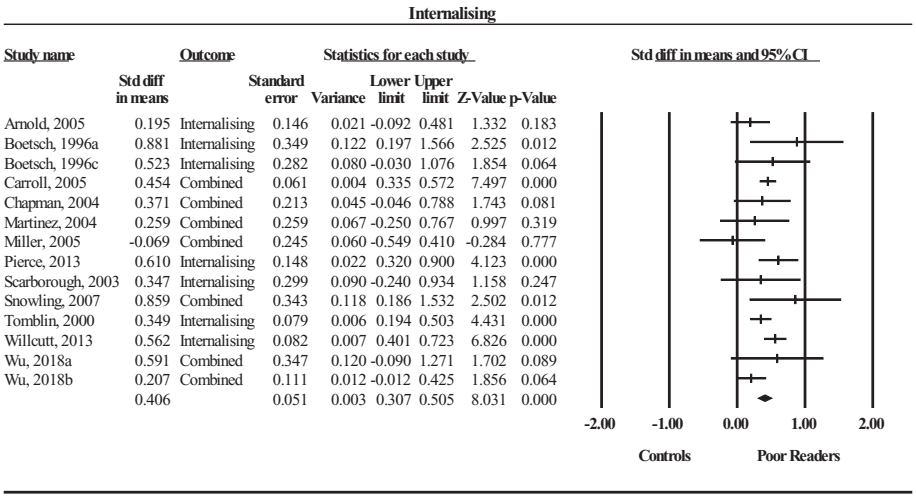

Fig. 3. Random effects forest plot showing the standardised mean difference and 95% CIs for each study on the association between poor reading and general internalising problems.

provide a narrative analysis of the evidence for the studies representing these moderators in the Discussion.

3.5.3. Meta-analysis of the association between poor reading and anxiety

3.5.3.1. Main analysis. Twenty-two included studies provided data regarding the association between poor reading and anxiety. Fig. 4 shows the 22 effect sizes comparing the standardised mean difference and 95% CIs between poor reading and control groups on measures of anxiety (*n* poor reading groups = 22, sample size = 1732, mean sample size = 78.73; *n* control groups = 22, sample size = 9640, mean sample size = 438.18). The overall effect was moderate, positive, and statistically significant ($d = 0.41$), 95% CI [0.26, 0.55], $p \leq 0.001$, and there was also evidence for moderate heterogeneity, $I^2 = 68.39\%$, Q (21) = 66.43, $p < 0.001$.

3.5.3.2. Subgroup analyses. It was not possible to conduct any of the planned subgroup analyses because there were fewer than 10 studies per subgroup for each moderator. The studies that examined the association between poor reading and anxiety, and the number of studies examining each moderator variable are shown in Table S3 in the Supplementary Materials. Again, we provide a narrative synthesis of the evidence for these moderators on the relationship between poor reading and anxiety in the Discussion.

3.5.3.3. Meta-analysis of the association between poor reading and depression

3.5.3.3.1. Main analysis. Twenty-three included studies provided data on the association between poor reading and depression that were included in the meta-analysis. Fig. 5 shows the 23 effect sizes comparing the standardised mean difference and 95% CIs between poor reading and control groups on measures of depression (*n* poor reading groups = 23, sample size = 1950, mean sample size = 84.78; *n* control groups = 23, sample size = 8764, mean sample size = 381.04). The overall effect was very small but significant ($d = 0.23$), 95% CI [0.07, 0.37], $p < 0.05$, and there was also evidence for high and significant heterogeneity, $I^2 = 74.51$, Q (22) = 86.33, $p < 0.001$.

3.5.3.3.2. Subgroup analyses. We could not carry out the planned subgroup analyses on the potential moderators of the association between poor reading and depression because the small mean effect size, though statistically significant, was associated with high and

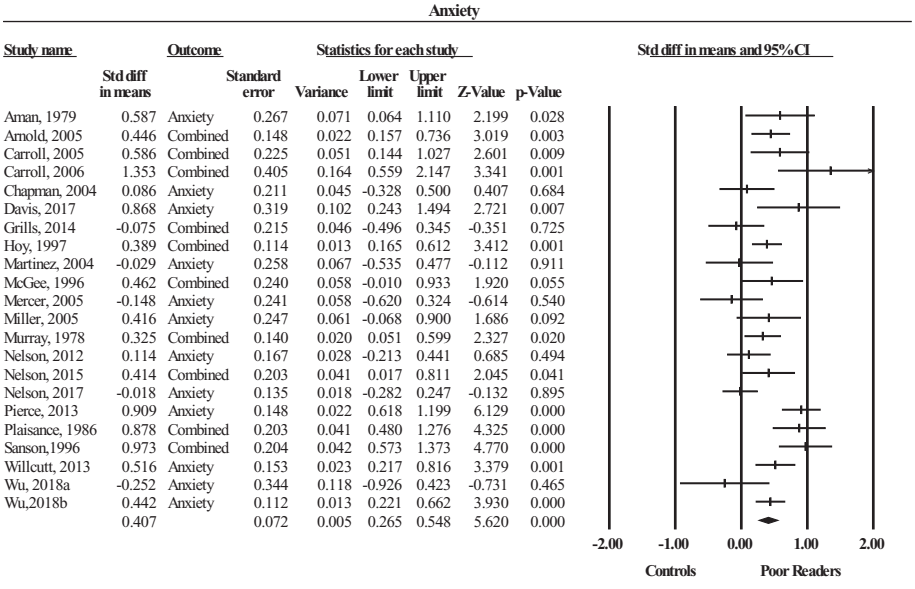

Fig. 4. Random effects forest plot showing the standardised mean difference and 95% Cis for each study on the association between poor reading and anxiety.

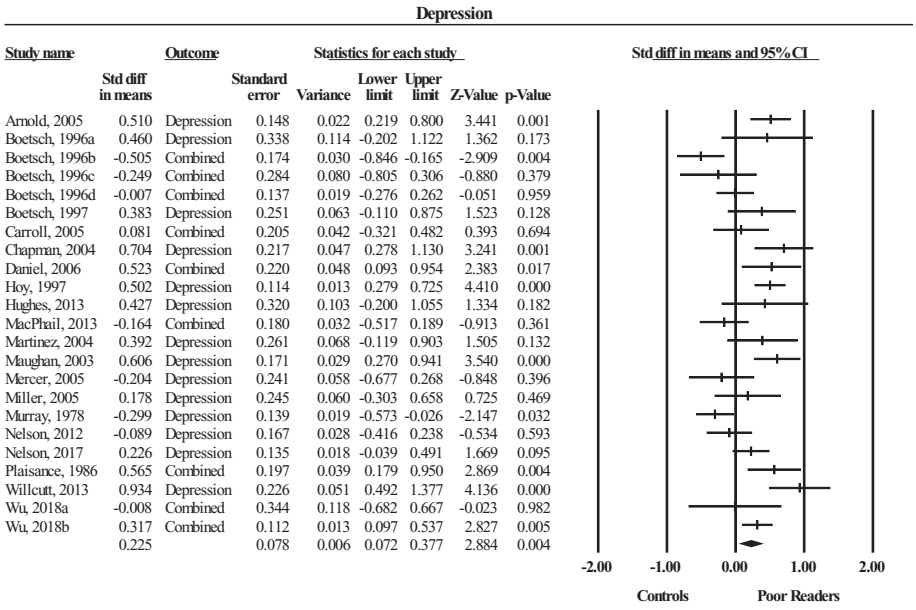

Fig. 5. Random effects forest plot showing the standardised mean difference and 95% CIs for each study on the association between poor reading and depression.

significant heterogeneity found between studies.

3.5.4. Risk of bias across studies: publication bias

We evaluated the likelihood of publication bias by visually inspecting the random effects funnel plots for studies that measured the association between poor reading and general internalising problems, anxiety, and depression, with the standard error plotted on the *y*-axis and the standardised mean difference plotted on the *x*-axis (see Fig. 6). The funnel plots showed symmetry, and evaluation of Egger's test showed that publication bias was not present for internalising problems, t (12) = 0.05, p > 0.05, 95% CI [−1.42, 1.35], anxiety, t (20) = 0.25, p > 0.05, 95% CI [−2.22, 2.85], or depression, t (21) = 0.10, p > 0.05, 95% CI [−2.69, 2.98]. This suggests that there was no systematic difference between studies that found stronger and weaker associations between poor reading and internalising problems, anxiety, or depression.

4. Discussion

The primary goal of this systematic review and meta-analysis was to determine if there were reliable associations between poor reading and internalising problems, anxiety, and depression (Aim 1), and if so, examine if any association was moderated by theoretical (i.e., anxiety disorder subtype, poor reading subtype, attention subtype, sex, age, ethnicity) or methodological moderators (i.e., type of internalising measure, type of informant; Aim 2). Below, we use the outcomes of the meta-analysis to evaluate the association between poor reading and internalising problems, poor reading and anxiety, and poor reading and depression, respectively. We also offer a narrative discussion of the moderators evaluated in this review, discuss the clinical and theoretical implications of these outcomes, as well as discuss the potential limitations of the current review. We offer suggestions for how these limitations might be addressed by future research.

4.1. Poor reading and internalising problems

Although based on a relatively small number of studies (n = 14), the results of the meta-analysis revealed a statistically significant and moderate association between poor reading and internalising problems, with moderate heterogeneity between studies. These results suggest that poor readers may, on average, experience elevated internalising problems compared to people without reading difficulties. Our results are consistent with the previous narrative reviews that examined associations between general learning disabilities (including poor reading) and internalising problems (including anxiety and depression; Maughan & Carroll, 2006; Mugnaini et al., 2009). Together, the outcomes of these reviews suggest that poor readers, on average, are at increased risk for experiencing overall general internalising problems.

4.2. Poor reading and anxiety

Again based on a small number of studies (n = 22), we found a statistically-significant and moderate association between poor reading and anxiety, with moderate heterogeneity between studies. Similar to the outcomes for general internalising problems, these results indicated that poor readers experience greater problems with anxiety than typical readers. This is again consistent with the previous narrative reviews in this field, which found that poor readers of all ages are at higher risk for anxiety than typical readers (Mugnaini et al., 2009), and that poor readers are also at risk of anxiety independent of problems with poor attention (Maughan & Carroll, 2006).

4.3. Poor reading and depression

The association between poor reading and depression was also based on a small number of studies (n = 23). The results of the meta-analysis revealed a small but statistically significant association between poor reading and depression that was associated with high heterogeneity between studies. This finding suggests that the association between poor reading and depression is less reliable and more complex than the association between poor reading and internalising problems or anxiety. It also suggests that the moderate and reliable association between poor reading and internalising problems (d = 0.41, p < 0.001, I^2 = 35.07%) may more closely reflect the moderate but stable association between poor reading and anxiety (d = 0.41, p < 0.001, I^2 = 68.38%) than the weak and unstable (yet statistically reliable) association between poor reading and depression (d = 0.23, p < 0.05, I^2 = 74.51%). This evidence converges and clarifies the results from the previous narrative reviews in this field, whereby Mugnaini et al. (2009) suggested that poor readers are at higher risk for depression, while Maughan and Carroll (2006) suggested that the relationship between poor reading and depression is less clear. The

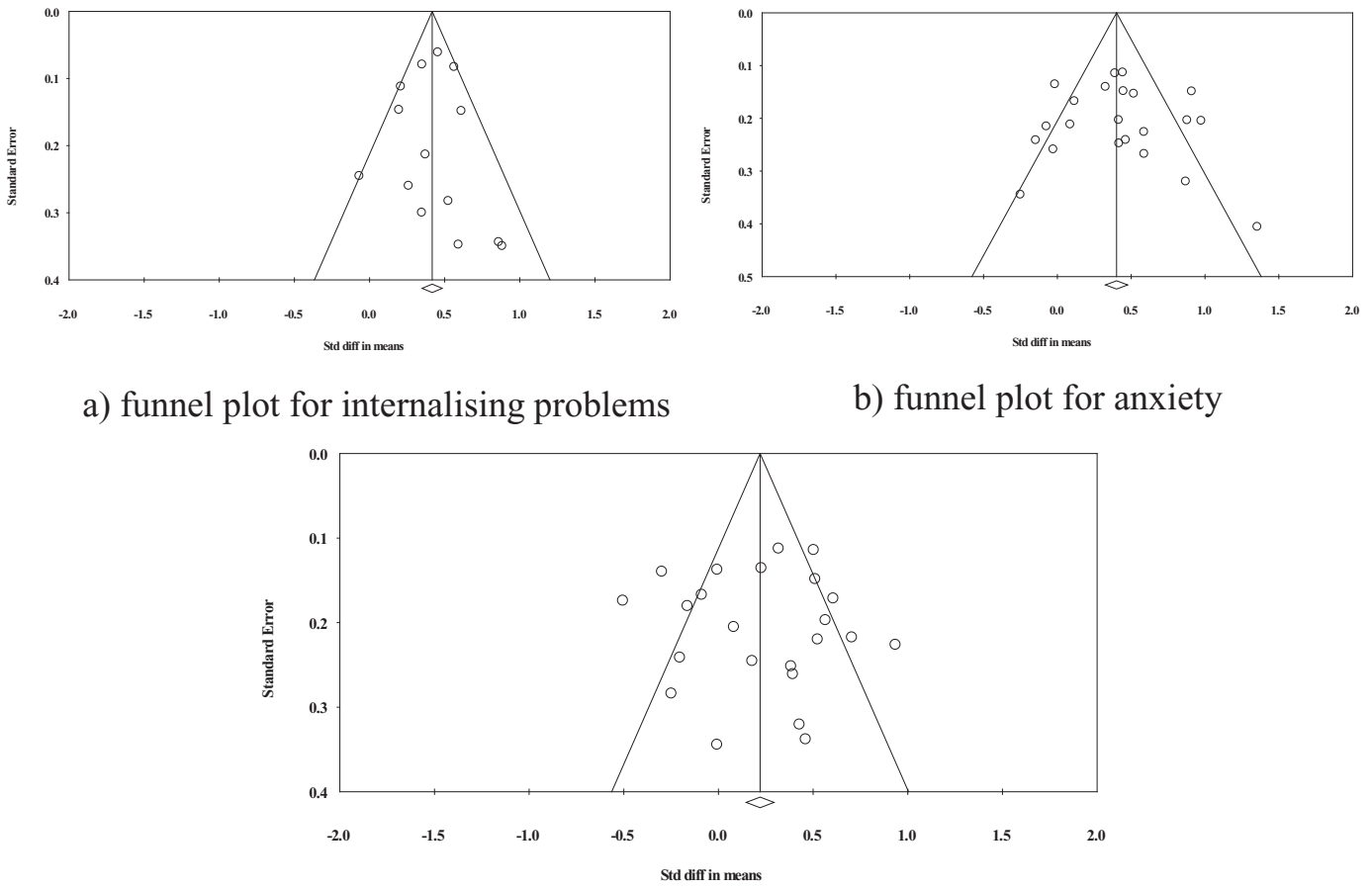

Fig. 6. Funnel plots showing symmetrical distribution of studies measuring internalising problems (a), anxiety (b), and depression (c).

outcomes of all these reviews converge to suggest that the association between poor reading and depression may be more complex than the association with anxiety.

4.4. Theoretical moderators

The secondary aim of this review was to conduct subgroup analyses to investigate potential moderators of any reliable associations between poor reading and internalising problems with acceptable heterogeneity between studies. In the current review, we found such an association between poor reading and anxiety. However, we were unable to conduct any subgroup analyses because there were fewer than 10 studies per subgroup for each moderator. We therefore offer a cautious synopsis of the existing evidence relating to potential moderators of the association between poor reading and anxiety.

4.4.1. Anxiety disorder subtype

Of the 22 anxiety studies, only three examined anxiety disorder subtypes: All three studies examined generalised anxiety (Carroll et al., 2005; Willcutt et al., 2013; Wu, 2018), and one also examined separation anxiety (Carroll et al., 2005). Two studies found higher anxiety for poor readers than controls (Carroll et al., 2005; Willcutt et al., 2013), while one did not (Wu, 2018). This evidence, albeit extremely limited, suggests that poor reading is associated with both generalised and separation anxiety disorder subtypes.

4.4.2. Poor reading subtype

All of the anxiety studies in this review reported the type of poor reading. Eighteen studies included samples with mixed poor reading (Aman, 1979; Arnold et al., 2005; Chapman et al., 2004; Davis et al., 2017; Grills et al., 2014; Hoy, 1997; Martinez & Semrud-Clikeman, 2004; McGee et al., 1986; Miller et al., 2005; Murray, 1978; Nelson et al., 2015; Nelson & Gregg, 2012; Nelson & Liebel, 2017; Plaisance, 1994; Sanson et al., 1996; Willcutt et al., 2013; Wu, 2018). Two studies tested poor readers for visual word recognition only (Carroll et al., 2005; Mercer, 2005) and two studies tested their reading fluency only (Carroll & Iles, 2006; Pierce et al., 2013). There were no studies that tested poor readers for phonological recoding only. All bar six of the 18 "mixed" reading studies found higher anxiety for poor readers than controls (Chapman et al., 2004; Grills et al., 2014; Martinez & Semrud-Clikeman, 2004; Nelson & Gregg, 2012; 2017; Wu, 2018), as did one study that tested visual word recognition only (Mercer, 2005). Both studies that tested reading fluency only found higher anxiety for poor readers than controls. Overall, the existing evidence does not suggest that poor reading subtype moderates the association between poor reading and anxiety.

4.4.3. Attention subtype

Of the 22 anxiety studies, 11 provided details on whether poor readers had problems with attention. Three of these studies included poor readers with inattention problems (Aman, 1979; Pierce et al., 2013; Wu, 2018), two with hyperactivity problems (Carroll et al., 2005; Wu, 2018), and five with combined attention problems (Arnold et al.,

2005; Chapman et al., 2004; McGee et al., 1986; Plaisance, 1994; Willcutt et al., 2013). One study excluded poor readers with poor attention (Davis et al., 2017). All bar one of the inattention and hyperactivity studies found an association between poor reading and anxiety (Wu, 2018), and all bar one of the combined inattention and hyperactivity studies (Chapman et al., 2004) reported higher anxiety for poor readers than controls. Considered together, the weight of evidence does not suggest that attention subtype has a moderating role in the association poor reading and anxiety, but does suggest that there is an association between poor reading, anxiety, and poor attention overall.

4.4.4. Sex

There were 18 anxiety studies that reported the sex of poor readers in their sample. Eleven of these studies recruited more males than females (Aman, 1979; Arnold et al., 2005; Carroll et al., 2005; Hoy, 1997; McGee, 1986; Mercer, 2005; Miller et al., 2005; Murray, 1978; Pierce, 2013; Plaisance, 1994; Willcutt et al., 2013), two recruited more females than males (Carroll & Iles, 2006; Nelson & Liebel, 2017), and five recruited similar numbers of males and females (Grills et al., 2014; Martinez & Semrud-Clikeman, 2004; Nelson et al., 2015; Nelson & Gregg, 2012; Sanson et al., 1996). All bar one study with majority of males (10 of 11 studies). One study with more females found higher anxiety for poor readers than controls (Carroll & Iles, 2006), while the other did not (Nelson & Liebel, 2017). The existing evidence therefore suggests that sex does not moderate the association between poor reading and anxiety. However, this suggestion is made with great caution given the extremely limited number of studies that recruited mostly females.

4.4.5. Age

All of the studies included in this review specified the age of their sample. There were 11 studies that recruited children aged 6 to 12 years (Aman, 1979; Chapman et al., 2004; Davis et al., 2017; Grills et al., 2014; McGee et al., 1986; Murray, 1978; Pierce et al., 2013; Plaisance, 1994; Sanson et al., 1996; Wu, 2018), three that recruited adolescents aged from 13 to 18 years (Arnold et al., 2005; Martinez & Semrud-Clikeman, 2004; Mercer, 2005), four that recruited adults aged 18 years and over (Carroll et al., 2006; Hoy, 1997; Nelson et al., 2015; Nelson & Liebel, 2017), and four that recruited mixed samples of children and adolescents (Carroll et al., 2005; Miller et al., 2005; Nelson & Gregg, 2012; Willcutt et al., 2013). All bar three of the child studies (Chapman et al., 2004; Grills et al., 2014; Wu, 2018), and two of the adolescent studies (Martinez & Semrud-Clikeman, 2004; Mercer, 2005), found higher anxiety for poor readers than controls. All bar one of the mixed age studies reported an association between poor reading and anxiety (Nelson & Gregg, 2012). All bar one of the adult studies found higher anxiety for poor readers than controls (Nelson & Liebel, 2017). En masse, this evidence supports Mugnaini et al.'s (2009) conclusion that poor readers of all ages are at risk for anxiety.

4.4.6. Ethnicity

Fourteen studies described the ethnicity of their sample. There were 10 studies that recruited predominantly Caucasian participants (Aman, 1979; McGee, 1986; Miller et al., 2005; Nelson et al., 2015; Nelson & Gregg, 2012; Nelson & Liebel, 2017; Plaisance, 1994; Willcutt et al., 2013; Wu, 2018), one study that recruited predominantly African-American participants (Grills et al., 2014), and Hispanic participants (Pierce et al., 2013), and two studies that recruited samples with mixed ethnicities (Arnold et al., 2005; Martinez & Semrud-Clikeman, 2004). The studies with mixed ethnicities and Hispanic samples found higher anxiety for poor readers than controls, as did all bar three studies with majority Caucasian participants (Nelson & Gregg, 2012; Nelson & Liebel, 2017; Wu, 2018). The study with a mostly African-American sample did not find evidence for an association between poor reading and anxiety (Grills et al., 2014). Altogether, this evidence suggests that ethnicity may not moderate the association between poor reading and anxiety.

4.5. Methodological moderators

4.5.1. Type of informant

Each of the anxiety studies specified the type of informant that reported anxiety. Eleven of these studies relied on self-report (Arnold et al., 2005; Carroll & Iles, 2006; Davis et al., 2017; Grills et al., 2014; Hoy, 1997; Martinez & Semrud-Clikeman, 2004; Mercer, 2005; Murray, 1978; Nelson et al., 2015; Nelson & Gregg, 2012; Nelson & Liebel, 2017), two on parent-report (Willcutt et al., 2013; Wu, 2018), three on teacher-report (Aman, 1979; Chapman et al., 2004; Pierce et al., 2013), and six on multi-informant reports (Carroll et al., 2005; McGee et al., 1986; Miller et al., 2005; Plaisance, 1994; Sanson et al., 1996; Wu, 2018). All bar five of the self-report studies found higher anxiety for poor readers than controls (Grills et al., 2014; Martinez & Semrud-Clikeman, 2004; Mercer, 2005; Nelson & Gregg, 2012; Nelson & Liebel, 2017), as did one of the teacher-report studies (Chapman et al., 2004). One of the parent report studies found an association between poor reading and anxiety (Willcutt et al., 2013) while the other did not (Wu, 2018). All of the studies with mutli-informant reports found an association between poor reading and anxiety. Thus, the current evidence suggests that the type of informant may not moderate the association between poor reading and anxiety.

4.5.2. Type of anxiety measure

Finally, all of the anxiety studies described the type of anxiety measure used to assess levels of anxiety. There were 19 studies that measured anxiety using questionnaires (Aman, 1979; Arnold et al., 2005; Carroll & Iles, 2006; Chapman et al., 2004; Davis et al., 2017; Grills et al., 2014; Hoy et al., 1997; Martinez & Semrud-Clikeman, 2004; McGee et al., 1986; Mercer, 2005; Miller et al., 2005; Murray, 1978; Nelson et al., 2015; Nelson & Gregg, 2012; Nelson & Liebel, 2017; Plaisance, 1994; Pierce et al., 2013; Sanson et al., 1996; Wu, 2018), and three studies that measured anxiety using clinical interviews (Carroll et al., 2005; Willcutt et al., 2013; Wu, 2018). All bar six of the 19 questionnaire studies found higher anxiety for poor readers than controls (Chapman et al., 2004; Grills et al., 2014; Martinez & Semrud-Clikeman, 2004; Mercer, 2005; Nelson & Liebel, 2017; Nelson & Gregg, 2012), and all bar one of the clinical interview studies reported higher anxiety for poor readers than controls (Wu, 2018). These findings suggest that type of anxiety measure is not a moderator of the association between poor reading and anxiety.

4.6. Clinical implications

This review has clinical implications regarding accurate assessment of anxious poor readers. It suggests that some poor readers reliably experience anxiety, and hence clinicians should assess poor readers for problems with anxiety. This could be done by assessing for general symptoms of anxiety (i.e., questionnaires that correspond with DSM anxiety disorders; MASC; March et al., 1997; Spence Children's Anxiety Scale; Spence, 1998), as well as assessing for clinical anxiety disorders (i.e., clinical interviews such as the ADIS: IV-C/P; Silverman & Albano, 1996). However, it is unfortunately unclear from our synopsis whether any of the moderators could be used to predict which poor readers might also have problems with anxiety.

The outcomes of this review also have clinical implications in terms of intervention for anxious poor readers. Specifically, the results clearly demonstrate that poor readers are at increased risk for anxiety. While only one treatment study was identified by this review (Grills et al., 2014), the overall results suggest that anxious poor readers should be referred for treatment that targets their poor reading and anxiety. For instance, if a poor reader has a specific fear of reading, then clinicians should consider teaching children strategies to lower their anxiety before engaging in reading training. This approach will equip poor

readers with strategies to face their fear of reading in a gradual way, and lower their arousal before learning new reading skills. We are currently using this rationale to develop a combined reading and anxiety intervention for anxious poor readers.

4.7. Limitations of the current literature

The studies included in this review have a number of limitations in common. As previously mentioned, poor reading is a heterogeneous disorder that comprises various reading problems that differ in nature and severity, and there are no gold standard criteria to define different types of poor reading. In line with this, the studies included in this review varied in their sample selection and definition of poor reading, and most studies selected poor readers based on performance on general measures of reading that assessed both regular and irregular words. Given this variability between studies, it would be helpful if future research conducted more detailed reading assessments with poor readers. This would allow us to examine the possible moderating role of poor reading subtypes on the association between poor reading and anxiety.

There were also very few studies included in this review that reported the full characteristics of their sample. For instance, only 14 of the 22 anxiety studies reported their participants' ethnicity. In future research, recruiting samples of different ethnicities would allow for more direct examination of the possible moderating role of ethnicity. The included studies also underrepresented female poor readers, with only two studies including more female than male participants. Again, including samples with predominantly female participants would allow for more direct comparison of the possible moderating role of sex on this association.

A third limitation of the studies in this review is that most used dimensional questionnaires to measure anxiety symptoms. Such measures do not provide information about anxiety disorders, which can only be diagnosed using categorical clinical interviews. Two studies that used clinical interviews discovered that poor readers have separation anxiety (Carroll et al., 2005) or generalised anxiety (Carroll et al., 2005; Willcutt et al., 2013), while the third study found no difference in generalised anxiety between poor readers and controls (Wu, 2018). We need more studies using clinical interviews to ascertain if poor readers are at higher risk for these anxiety disorders.

4.8. Limitations of the current review

The main limitation of this review is the modest number of included studies (i.e., 34). Given the importance of understanding the emotional health of poor readers for their quality of life, this highlights the need for more studies on internalising problems in poor readers.

A second limitation – again stemming from a limited literature – is that we were unable to conduct the planned subgroup analyses. As Tables S2 and S3 show, even if a more conservative criterion was applied (i.e., 5 studies per subgroup), there would have been insufficient numbers for the subgroup analyses. This highlights the need for future studies to examine potential moderating variables on the association between poor reading and internalising problems – and particularly anxiety.

This review also focussed on the association, or “correlation”, between poor reading and internalising problems. While correlational studies are a sensible starting point for understanding the relationship between two variables, they cannot inform us about the direction of causation between those two variables. For example, in the context of this review, such studies cannot reveal if poor reading causes internalising problems, if the reverse is true, if there is a bidirectional relationship between poor reading and internalising problems, or if another factor has a causal effect on both reading and internalising that creates a “faux” association between the two. Given that the evidence to date supports a statistically reliable association between poor reading and internalising problems – particularly for anxiety – intervention studies are needed to test the causal mechanisms that might be responsible for this association. Such studies are extremely rare at this point in time (Grills et al., 2014).

5. Conclusions

This is the first systematic meta-analytic review of the association between poor reading and internalising problems, poor reading and anxiety, and poor reading and depression. The meta-analysis suggests that there is a statistically reliable association between poor reading and internalising problems, and between poor reading and anxiety. The association between poor reading and depression is also statistically reliable, yet smaller in size and less stable. When considered together, the effect sizes suggest that the association between poor reading and internalising problems is driven by, or most closely reflects, an association between poor reading and anxiety – rather than an association between poor reading and depression. The outcomes from this review guide clinicians to make informed decisions about how to assess poor readers for problems with anxiety, and suggests that we now need to investigate *why* this association exists.

Role of funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interest

Conflicts of interest: none.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cpr.2018.09.002.

References

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs: General and Applied, 80*, 1–37. https://doi.org/10.1037/h0093906.

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile.* Burlington, VT: University of Vermont Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin, 85*, 1275–1301. https://doi.org/10.1037/0033-2909.85.6.1275.

Aman, M. G. (1979). Cognitive, social, and other correlates of specific reading retardation. *Journal of Abnormal Child Psychology, 7*, 153–168. https://doi.org/10.1007/BF00918896.

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

Anderson, E. R., & Mayes, L. C. (2010). Race/ethnicity and internalizing disorders in youth: A review. *Clinical Psychology Review, 30*, 338–348. https://doi.org/10.1016/j.cpr.2009.12.008.

Arnold, E. M., Goldston, D. B., Walsh, A. K., Reboussin, B. A., Daniel, S. S., Hickman, E., & Wood, F. B. (2005). Severity of emotional and behavioural problems among poor and typical readers. *Journal of Abnormal Child Psychology, 33*, 205–217. https://doi.org/10.1007/s10802-005-1828-9.

Barbosa, J., Tannock, R., & Manassis, K. (2002). Measuring anxiety: Parent-child reporting differences in clinical samples. *Depression and Anxiety, 15*, 61–65. https://doi.org/10.1002/da.10022.

Barriga, A. Q., Doran, J. W., Newell, S. B., Morrison, E. M., Barbetti, V., & Robbins, B. D. (2002). Relationships between problem behaviours and academic achievement in adolescents: The unique role of attention problems. *Journal of Emotional and Behavioural Disorders, 10*, 233–240. https://doi.org/10.1177/10634266020100040501.

Baxter, A. J., Scott, K. M., Vos, T., & Whiteford, H. A. (2012). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine, 43*, 897–912. https://doi.org/10.1017/S003329171200147X.

Beidel, D. C. (1988). Psychophysiological assessment of anxious emotional states in children. *Journal of Abnormal Psychology, 97*, 80–82. https://doi.org/10.1037/0021-843X.97.1.80.

Beidel, D. C., & Turner, S. M. (1988). Comorbidity of test anxiety and other anxiety disorders in children. *Journal of Abnormal Child Psychology, 16*, 275–287. https://doi.

org/10.1007/BF00913800.
Boetsch, E. A. (1997). *A longitudinal study of the relationship between dyslexia and socio-emotional functioning in young children.* Unpublished doctoral dissertationUnited States: University of Denver.
Boetsch, E. A., Green, P. A., & Pennington, B. F. (1996). Psychosocial correlates of dyslexia across the lifespan. *Development and Psychopathology, 8*, 539–562. https://doi.org/10.1017/S0954579400007264.
Bogels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: Questions and answers for the DSM-V. *Depression and Anxiety, 27*, 168–189. https://doi.org/10.1002/da.20670.
Bonifacci, P., Montuschi, M., Lami, L., & Snowling, M. J. (2014). Parents of children with dyslexia: Cognitive, emotional and behavioural profile. *Dyslexia, 20*, 175–190. https://doi.org/10.1002/dys.1469.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2006). *Comprehensive meta-analysis.* Englewood, NJ: Biostat.
Bruce, S. E., et al. (2005). Influence of psychiatric comorbidity on recovery and recurrence in generalized anxiety disorder, social phobia, and panic disorder: A 12-year prospective study. *American Journal of Psychiatry, 162*, 1179–1187. https://doi.org/10.1176/appi.ajp.162.6.1179.
Carroll, J. M., & Iles, J. E. (2006). An assessment of anxiety levels in dyslexic students in higher education. *British Journal of Educational Psychology, 76*, 651–662. https://doi.org/10.1348/000709905X66233.
Carroll, J. M., Maughan, B., Goodman, R., & Meltzer, H. (2005). Literacy difficulties and psychiatric disorders: Evidence for comorbidity. *Journal of Child Psychology and Psychiatry, 46*, 524–532. https://doi.org/10.1111/j.1469-7610.2004.00366.x.
Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149–180. https://doi.org/10.1016/0010-0277(93)90003-E.
Chapman, J. W., Tunmer, W. E., & Prochnow, J. E. (2004). Repressed resilience? A longitudinal study of reading, self-perceptions and teacher behaviour ratings of poor and average readers in New Zealand. *Journal of the International Academy for Research in Learning Disabilities, 22*, 9–15. http://www.iarld.com/the-journal-thalamus.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.
Daniel, S. S., Walsh, A. K., Goldston, D. B., Arnold, E. M., Reboussin, B. A., & Wood, F. B. (2006). Suicidality, school dropout and reading problems among adolescents. *Journal of Learning Disabilities, 39*, 507–514. https://doi.org/10.1177/00222194060390060301.
Davis, K., Margolis, A. E., Thomas, L., Huo, A., & Marsh, R. (2017). Amygdala sub-regional functional connectivity predicts anxiety in children with reading disorder. *Developmental Science*e12631. https://doi.org/10.1111/desc.12631.
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*, 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x.
Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*, 629–634. https://doi.org/10.1136/bmj.315.7109.629.
Ford, T., Goodman, R., & Meltzer, H. (2003). The British child and adolescent mental health survey 1999: The prevalence of DSM-IV disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*, 1203–1211. https://doi.org/10.1097/00004583-200310000-00011.
Ginsburg, G. S., & Silverman, W. K. (1996). Phobic and anxiety disorders in Hispanic and Caucasian youth. *Journal of Anxiety Disorders, 10*, 517–528.
Goldston, D. B., et al. (2007). Reading problems, psychiatric disorders, and functional impairment from mid- to late adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*, 25–33. https://doi.org/10.1097/01.chi.0000242241.77302.f4.
Goulandris, N. K., & Snowling, M. (1991). Visual memory deficits: A plausible cause of developmental dyslexia? Evidence from a single case study. *Cognitive Neuropsychology, 8*, 127–154. https://doi.org/10.1080/02643299108253369.
Grills, A. E., & Ollendick, T. H. (2002). Issues in parent-child agreement: The case of structured diagnostic interviews. *Clinical Child and Family Psychology Review, 5*, 57–83. https://doi.org/10.1023/A:10145737.
Grills, A. E., Fletcher, J. M., Vaughn, S. R., Denton, C. A., & Taylor, P. (2013). Anxiety and inattention as predictors of achievement in early elementary school children. *Anxiety, Stress, & Coping, 26*, 391–410. https://doi.org/10.1080/10615806.2012.691969.
Grills, A. E., Fletcher, J. M., Vaughn, S., Barth, A., Denton, C. A., & Stuebing, K. K. (2014). Anxiety and response to reading intervention among first grade students. *Child & Youth Care Forum, 43*, 417–431. https://doi.org/10.1007/s10566-014-9244-3.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.
Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions.* Chichester, UK: John Wiley & Sons Ltd.
Higgins, J., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*, 557–560. https://doi.org/10.1136/bmj.327.7414.557.
Hoy, C., Gregg, N., Wisenbaker, J., Manglitz, E., King, M., & Moreland, C. (1997). Depression and anxiety in two groups of adults with learning disabilities. *Learning Disability Quarterly, 20*, 280–291. https://doi.org/10.2307/1511226.
Hughes, J. A., Phillips, G., & Reed, P. (2013). Brief exposure to a self-paced computer based reading programme and how it impacts reading ability and behaviour problems. *PLoS One, 8*, e77867. https://doi.org/10.1371/journal.pone.0077867.
Kessler, R. C., Chiu, W. T., Demler, O., & Walters, E. E. (2005). Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the national comorbidity survey replication. *Archives of General Psychiatry, 62*, 617–629. https://doi.org/10.1001/archpsyc.62.6.617.
Kovacs, M. (1992). *Children's Depression Inventory.* North Tonawanda, NY: Multi-Health System.
Krueger, R. F., & Eaton, N. R. (2015). Transdiagnostic factors of mental disorders. *World Psychiatry, 14*, 27–29. https://doi.org/10.1002/wps.20175.
Levy, F., Hay, D. A., Bennett, K. S., & McStephen, M. (2005). Gender differences in ADHD subtype comorbidity. *Journal of the American Academy of Child & Adolescent Psychiatry, 44*, 368–376. https://doi.org/10.1097/01.chi.0000153232.64968.c1.
Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage Publications.
MacPhail, S. S. (2013). *Perceived competence and depressive symptoms in children with reading problems.* Unpublished doctoral thesisNew York: Fordham University.
March, J. (1997). *Multidimensional anxiety Scale for Children.* North Tonawanda, NY: Multi-Health Systems.
Martinez, R. S., & Semrud-Clikeman, M. (2004). Emotional adjustment and school functioning of young adolescents with multiple versus single learning disabilities. *Journal of Learning Disabilities, 37*, 411–420. https://doi.org/10.1177/00222194040370050401.
Maughan, B., & Carroll, J. (2006). Literacy and mental disorders. *Current Opinion in Psychiatry, 19*, 350–354. https://doi.org/10.1097/01.yco.0000228752.79990.41.
Maughan, B., Rowe, R., Loeber, R., & Stouthamer-Loeber, M. (2003). Reading problems and depressed mood. *Journal of Abnormal Child Psychology, 31*, 219–229. https://doi.org/10.1023/A:1022534527021.
McArthur, G., Kohnen, S., Larsen, L., Jones, K., Anandakumar, T., Banales, E., & Castles, A. (2013). Getting to grips with the heterogeneity of developmental dyslexia. *Cognitive Neuropsychology, 30*, 1–24. https://doi.org/10.1080/02643294.2013.784192.
McGee, R., Williams, S., Share, D. L., Anderson, J., & Silva, P. A. (1986). The relationship between specific reading retardation, general reading backwardness and behavioural 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. https://doi.org/10.1111/j.1469-7610.1986.tb00185.x.
McLean, C. P., Asnaani, A., Litz, B. T., & Hofmann, S. G. (2011). Gender differences in anxiety disorders: Prevalence, course of illness, comorbidity and burden of illness. *Journal of Psychiatric Research, 45*, 1027–1035. https://doi.org/10.1016/j.jpsychires.2011.03.006.
Mercer, K. L. (2005). *Relations of self-efficacy to symptoms of depression and anxiety in adolescents with learning disabilities* (Unpublished doctoral thesis).United States: The University of British Columbia.
Miller, C. J., Hynd, G. W., & Miller, S. R. (2005). Children with dyslexia: Not necessarily at risk for elevated internalising symptoms. *Reading & Writing, 18*, 425–436. https://doi.org/10.1007/s11145-005-4314-4.
Mugnaini, D., Lassi, S., La Malfa, G., & Albertini, G. (2009). Internalising correlates of dyslexia. *World Journal of Pediatrics, 5*, 255–264. https://doi.org/10.1007/s12519-009-0049-7.
Murray, M. E. (1978). The relationship between personality adjustment and success in remedial programs in dyslexic children. *Contemporary Educational Psychology, 3*, 330–339. https://doi.org/10.1016/0361-476X(78)90035-8.
Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *The Journal of Child Psychology and Psychiatry, 51*, 1031–1039. https://doi.org/10.1111/j.1469-7610.2010.02254.x.
Nelson, J. M., & Gregg, N. (2012). Depression and anxiety among transitioning adolescents and college students with ADHD, dyslexia or comorbid ADHD/dyslexia. *Journal of Attention Disorders, 16*, 244–254. https://doi.org/10.1177/1087054710385783.
Nelson, J. M., & Harwood, H. R. (2011a). A meta-analysis of parent and teacher reports of depression among students with learning disabilities: Evidence for the importance of multi-informant assessment. *Psychology in the Schools, 48*, 371–385. https://doi.org/10.1002/pits.20560.
Nelson, J. M., & Harwood, H. (2011b). Learning disabilities and anxiety: A meta-analysis. *Journal of Learning Disabilities, 44*, 3–17. https://doi.org/10.1177/0022219409359939.
Nelson, J. M., & Liebel, S. W. (2017). Socially desirable responding and college students with dyslexia: Implications for the assessment of anxiety and depression. *Dyslexia, 24*, 44–58. https://doi.org/10.1002/dys.1563.
Nelson, J. M., Lindstrom, W., & Foels, P. A. (2015). Test anxiety among college students with specific reading disability (Dyslexia): Nonverbal ability and working memory as predictors. *Journal of Learning Disabilities, 48*, 422–432. https://doi.org/10.1177/0022219413507604.
Oakhill, J. V., Cain, K., & Bryant, P. E. (2003). The dissociation of word reading and text comprehension: Evidence from component skills. *Language and Cognitive Processes, 18*, 443–468. https://doi.org/10.1080/01690960344000008.
Patel, T. K., Snowling, M. J., & de Jong, P. F. (2004). A cross-linguistic comparison of children learning to read in English and Dutch. *Journal of Educational Psychology, 96*, 785–797. https://doi.org/10.1037/0022-0663.96.4.785.
Pierce, M. E., Wechsler-Zimring, A., Noam, G., Wolf, M., & Katzir, T. (2013). Behavioural problems and reading difficulties among language minority and monolingual urban elementary school students. *Reading Psychology, 34*, 182–205. https://doi.org/10.1080/02702711.2011.626108.
Plaisance, R. L. (1994). *Psychiatric comorbidity in children with disabilities.* Unpublished doctoral thesisUnited States: Athens Georgia.
Safford, S. M., Kendall, P. C., Flannery-Schroeder, E., Webb, A., & Sommer, H. (2005). A longitudinal look at parent-child diagnostic agreement in youth treated for anxiety disorders. *Journal of Clinical Child & Adolescent Psychology, 34*, 747–757. https://doi.org/10.1207/s15374424jccp3404_16.
Sanson, A., Prior, M., & Smart, D. (1996). Reading disabilities with and without behaviour problems at 7-8 years: Prediction from longitudinal data from infancy to 6 years. *Journal of Child Psychology and Psychiatry, 37*, 529–541. https://doi.org/10.1111/j.

1469-7610.1996.tb01439.x.
Scarborough, H. S., & Parker, J. D. (2003). Matthew effects in children with learning disabilities: Development of reading, IQ, and psychosocial problems from grade 2 to grade 8. *Annals of Dyslexia, 53*, 47–71. https://doi.org/10.1007/s11881-003-0004-6.
Shamseer, L., et al. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *British Medical Journal, 349*, g7647. https://doi.org/10.1136/bmj.g7647.
Shaywitz, B. A., et al. (1995). A Matthew effect for IQ but not for reading: Results from a longitudinal study. *Reading Research Quarterly, 30*, 894–906. https://doi.org/10.2307/748203.
Silverman, W. K., & Albano, A. M. (1996). *Manual for the ADIS-IV-C/P.* New York, NY: Psychological Corporation.
Snowling, M. J., Muter, V., & Carroll, J. (2007). Children at family risk of dyslexia: A follow-up in early adolescence. *Journal of Child Psychology and Psychiatry, 48*, 609–618. https://doi.org/10.1111/j.1469-7610.2006.01725.x.
Snowling, M., Dawes, P., Nash, H., & Hulme, C. (2012). Validity of a protocol for adult self-report of dyslexia and related difficulties. *Dyslexia, 18*, 1–15. https://doi.org/10.1002/dys.1432.
Spence, S. H. (1998). A measure of anxiety symptoms among children. *Behaviour Research and Therapy, 36*, 545–566. https://doi.org/10.1016/S0005-7967(98)00034-5.
Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the state-trait anxiety inventory.* Palo Alto, CA: Consulting Psychologists Press.
Stuart, M., & Stainthorp, R. (2016). *Reading development and teaching.* London: SAGE.
Tomblin, J. B., Zhang, X., Buckwalter, P., & Catts, H. (2000). The association of reading disability, behavioural disorders, and language impairment among second-grade children. *Journal of Child Psychology and Psychiatry, 41*, 473–482. https://doi.org/10.1111/1469-7610.00632.
Umana-Taylor, A. J., & Updegraff, K. A. (2007). Latino adolescents' mental health: Exploring the interrelations among discrimination, ethnic identity, cultural orientation, self-esteem, and depressive symptoms. *Journal of Adolescence, 30*, 549–567. https://doi.org/10.1016/j.adolescence.2006.08.002.
Wells, G. A., et al. (1996). The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available at http://www ohri ca/programs/clinical_epidemiology/oxfordasp. Accessed September 2, 2016.
Willcutt, E. G., et al. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46*, 500–516. https://doi.org/10.1177/0022219413477476.
Wu, S. S. (2018). *Measurement and classification issues in mathematics and reading disorders.* Unpublished doctoral thesisBoulder: The University of Colorado.

Authors Francis and McArthur developed the methods of this study, including the study selection and data extraction criteria. Francis developed the search strategy in collaboration with McArthur and Hudson, who provided expertise on poor reading and internalising problems, respectively. Francis conducted the search. Francis and McArthur screened the studies identified by the search. Francis and Caruana extracted the data. Francis conducted the meta-analysis. Francis composed the first draft of the manuscript and all authors provided comment and feedback and have approved the final manuscript.