Scientific Studies of Reading

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Epidemiology of reading disability: A comparison of DSM-5 and ICD-11 criteria

Cécile Di Folco, Ava Guez, Hugo Peyre & Franck Ramus

To cite this article: Cécile Di Folco, Ava Guez, Hugo Peyre & Franck Ramus (2022) Epidemiology of reading disability: A comparison of DSM-5 and ICD-11 criteria, Scientific Studies of Reading, 26:4, 337-355, DOI: 10.1080/10888438.2021.1998067

To link to this article: https://doi.org/10.1080/10888438.2021.1998067

SCIENTIFIC STUDIES OF READING
2022, VOL. 26, NO. 4, 337–355
https://doi.org/10.1080/10888438.2021.1998067

Epidemiology of reading disability: A comparison of DSM-5 and ICD-11 criteria

Cécile Di Folco[a,b++], Ava Guez[a++], Hugo Peyre[a,c,d], and Franck Ramus[a]

[a]Laboratoire De Sciences Cognitives Et Psycholinguistique, Département d'Etudes Cognitives, Ecole Normale Supérieure, EHESS, CNRS, PSL University, Paris, France; [b]AgroParisTech, Paris, France; [c]Inserm UMR 1141, Université De Paris, Paris, France; [d]Department of Child and Adolescent Psychiatry, Robert Debré Hospital, APHP, Paris, France

ABSTRACT

The present study performed a systematic comparison of DSM-5 and ICD-11 diagnostic criteria for reading disability. We quantitatively investigated the consequences of using DSM-5 or ICD-11, and of the different ways of implementing each diagnostic criterion on the prevalence of reading disability. We did so in a representative sample of the population of French sixth-graders (N = 25,000), using a reading comprehension test to assess reading ability. A compromise set of criteria and thresholds yielded a prevalence of 6.6% according to DSM-5 and 3.5% according to ICD-11. Factors that had the greatest influence on prevalence estimates were the criteria relative to IQ and to interference with academic performance. Compared with the reference population, children with reading disability were more likely to be boys (sex ratio≈1.6), to be schooled in a disadvantaged area (OR≈2.1), and to have lower SES (d≈-0.7), non-verbal IQ (d≈-0.4 – -0.9), and math scores (d≈-1.4). Our results emphasize that the choice of classification and the operationalization of diagnostic criteria have a large impact on who is diagnosed with reading disability.

Introduction

Over the last decades, reading disability has received various names and has been given different definitions. According to the World Health Organization's (WHO) most recent International Classification of Diseases (ICD-11), it is currently known as "developmental learning disorder with impairment in reading" and defined by "significant and persistent difficulties in learning academic skills related to reading, such as word reading accuracy, reading fluency, and reading comprehension." It is not due to "a disorder of intellectual development, sensory impairment (vision or hearing), neurological disorder, lack of availability of education, lack of proficiency in the language of academic instruction, or psychosocial adversity." Although it is acknowledged as a real public health issue, the exact prevalence of reading disability remains a vexed issue, with estimates varying from 2% to 20% across studies.

There are a number of reasons why estimates may vary. One is of course the use of different definitions of reading disability. Interestingly, the two most widely-used classifications have recently been updated: the APA's Diagnostic and Statistical Manual in its 5th version (DSM-5) (American Psychiatric Association, 2013), and the International Classification of Diseases in its 11th version (ICD-11) (World Health Organization, 2018). Both classifications have chosen to group disorders of reading, writing and mathematics under the umbrella terms "specific learning disorder" (DSM-5) and "developmental learning disorder" (ICD-11). Nevertheless, they have also recognized the existence of

CONTACT Franck Ramus franck.ramus@ens.psl.eu Ecole Normale Supérieure, 29 Rue d'Ulm, 75005 Paris, France
+Equal contributors.

cases with a more specific form of learning disorder, and have provided the appropriate specifiers: "with impairment in" reading, written expression, or mathematics. Thus, the former notions of reading disability readily map onto "learning disorder with impairment in reading," whether or not there is comorbid impairment in written expression or mathematics (which would be signaled with additional specifiers). In this paper, we use "reading disability" throughout as a convenient shorthand. As we will see below, although the WHO has overall followed the APA quite closely, the two classifications crucially differ in certain criteria used in the definition of reading disorder, potentially leading to the identification of different individuals, and to different prevalence estimates. It is therefore important to evaluate the potential discrepancies between the two classifications, and the consequences of using one or the other. We are not aware of any previous study doing so. This will therefore be the main purpose of the present study.

Secondly, prevalence estimates depend on the severity threshold chosen for the definition of a "significant" difficulty. Of course, given a certain distribution of reading ability across the population, any cutoff is arbitrary (Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992; Stuebing et al., 2002). However, prevalence does not just mathematically follow from the chosen threshold, for several reasons that are often under-appreciated. Indeed, the distribution of reading ability has been reported to be normal in many but not all studies (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Rodgers, 1983; Share, McGee, McKenzie, Williams, & Silva, 1987; Stevenson, 1988). Although the normal distribution is generally a reasonable approximation of observed distributions, the latter may deviate from normality for a variety of reasons, including inclusion/exclusion criteria, test properties, and educational policies and methods (notably relative to poor readers). Thus, the percentage of cases below a given threshold remains an empirical question, rather than a mathematical deduction.

Thirdly, prevalence will depend on how exactly the criteria given in official classifications are operationalized. For instance, as emphasized in the ICD-11 definition, reading ability is multifaceted, so prevalence will vary depending on 1) whether one considers reading accuracy, fluency, and/or comprehension, on 2) whether one uses one, two or three reading measures, and on 3) whether one applies the threshold to each measure separately (using an OR or an AND connector), or to a composite of the different measures. Prevalence may also vary depending on whether an absolute threshold is applied to the distribution of reading ability, or whether the threshold on reading ability is applied relative to the level of intellectual functioning. This issue of whether reading disability should be defined according to a discrepancy or not is highly controversial (Stanovich, 2005), and this is one important criterion on which DSM-5 and ICD-11 crucially differ.

Finally, definitions of reading disability are not exclusively based on reading ability, but also include a number of exclusion criteria, which contribute to reducing the prevalence compared to the raw number of individuals with a significant difficulty in reading.

Most epidemiological studies of reading disability were conducted in English-speaking countries, and provide an instructive illustration of how prevalence may vary depending on the factors mentioned above. In the United Kingdom, Yule, Rutter, Berger, and Thompson (1974) reported 3.61% of dyslexic children in the Isle of Wight (N = 1,134) and 9.26% in London (N = 1,643), using a reading comprehension test, a regression-based IQ-discrepancy criterion, and a 2-standard-error from predicted value threshold. Rodgers (1983) found a prevalence of 2.3% (N = 8,836) in a representative sample of the UK population, using an array of diverse reading tests, regression-based discrepancy, and a −2 SD threshold. Using similar methods and criteria, Lindgren, De Renzi, and Richman (1985) compared the prevalence of reading disability between American and Italian student samples, using a reading comprehension test, 3 different definitions involving either 1) low-IQ exclusion, 2) absolute reading-IQ discrepancy, or 3) regression-based discrepancy with a – 1 standard deviation threshold on reading ability and a 2-standard-error from predicted value threshold in the regression method. They reported prevalences varying between 4.5% (regression) to 12% (exclusion) in the United-States and between 3.6% (regression) to 8.5% (exclusion) in Italy, the difference being attributed to differences in orthographic transparency. In the United-States, Badian (1999a) reported a prevalence of 2.7%, using a measure of reading comprehension, a regression-based discrepancy criterion on listening comprehension, a 25th percentile absolute threshold on reading ability, and a −1.5 SD threshold from predicted reading ability. Shaywitz, Shaywitz, Fletcher, and Escobar (1990) used a composite reading ability measure, a regression-based discrepancy criterion, and a 1.5 SD threshold from predicted value, and reported about 7.5% of 3rd-graders to be reading disabled. Katusic, Colligan, Barbaresi, Schaid, and Jacobsen (2001) used four different methods on a sample of 5,718 children from a birth cohort and obtained prevalences ranging from 5.3% (with a regression-based discrepancy criterion and a 1.5 SD threshold) to 11.8% (with a low-achievement criterion: reading standard score≤90 and IQ≥80). Reading tests were not specified.

In France, however, no such research has been done. A few prevalence estimates can be extracted from studies that did not primarily aim at calculating it. These studies defined reading disability as under-achievement, without any discrepancy criterion. Estimates include 12% (N = 199 7-year-old children) according to Plaza et al. (2002), and 7.5% (N = 485 8-year-old children) according to Zorman, Lequette, and Pouget (2004), who added a phonological deficit criterion. The limited size and non-representativeness of the samples make any generalization to the French population impossible.

Thus, the aims of the present study are to 1) estimate the prevalence of reading disability in France, using both DSM-5 and ICD-11 criteria, comparing them for the first time and evaluating their concordance; 2) evaluate the impact of each diagnostic criterion on prevalence estimates; 3) characterize the population of reading disabled individuals, and how this varies according to diagnostic criteria. We do so on a large representative sample of French 6th grade pupils.

Methods

Participants

The *Direction de l'Evaluation, de la Prospective et de la Performance* (DEPP), French Ministry of Education, conducted a large prospective survey following students from the beginning (grade 6) to the end of French middle school (grade 9). The National Council for Statistical Information (CNIS) approved this study, ensuring public interest and conformity with ethical, statistical and confidentiality standards. This gave this survey the same status as a national census. Participation was compulsory as part of the evaluation policy of the French Ministry of National Education, although there was no penalty for not responding. Secondary research on this data was made possible as part of the public release of national statistical data. The sample consisted of about 35,000 randomly selected students among the 760,000 entering middle school in 2007, with a higher sampling rate for schools in disadvantaged areas. Only a small fraction (<1%) of children who were deemed too severely disabled to enter middle school but were placed in medical institutions were not represented in this sampling.

For the current purpose, we restricted the analysis to Grade 6, a stage at which the prevalence of reading disability stabilizes according to Katusic et al. (2001)'s longitudinal follow-up. We excluded students who had missing values for intelligence, reading or academic performance scores in grade 6. Given an unexpectedly high number of students scoring 0 at the intelligence test, suggesting either a refusal to take the test, low engagement in the task, or problems with administration and scoring, we also excluded those students (Guez, Panaïotis, Peyre, & Ramus, 2018). Children who repeated a year were not excluded. This process is summed up in the flowchart (Supplementary Figure S1). 25,041 students were thus included in the present study. Included and excluded participants differed in observed characteristics (see Supplementary Table S1).

Measures

The DEPP conducted standardized evaluations in French, mathematics, non-verbal intelligence and perceived self-efficacy, in paper/pencil format. These tests were administered collectively, supervised by teachers. Children's socio-economic status was estimated based on a questionnaire filled by parents (or legal guardians). Consistently with French law, ethnic information was not collected. The following variables were used for the present analysis. Examples of test items can be found in supplementary methods.

Measures used for diagnosis

Reading comprehension. Students were asked to read a short text, then answer 5 open-ended comprehension questions. They did so for 3 different texts, in a maximum of 12 minutes (total: 15 items).

Unfortunately, other important aspects of reading ability such as word reading accuracy and reading fluency were not assessed in the DEPP 2007 Panel. Our assessment of reading disability is therefore entirely based on reading comprehension performance. Although this was also the case in many of the previous studies reviewed above, this is a limitation whose consequences will be further discussed.

Nonverbal intelligence. The *Raisonnement sur Cartes de Chartier* test (Chartier's Reasoning Test on Playing Cards) assesses logical reasoning skills using playing cards (Chartier, 2012; Terriot, 2014). Inspired from Raven's progressive matrices, it consists of 30 items in which children must find the missing card (from a deck of 4 suits of 10 cards) in an array composed of 4 to 12 cards, within a time limit of 20 minutes.

Academic performance. At the beginning of grade 6 (entrance in middle school), all French students have to take national assessments in French and Mathematics. These tests are administered in schools by the school teachers and have no stakes for students or teachers: they are only meant to give teachers and parents an idea of the students' initial levels. We took the mean of results in French and Mathematics as an indicator of academic performance in grade 6.

Serious disease. Parents were asked to indicate if the child had any serious illness, and specify the age range of onset of the disease (before age 1, between 1 and 5, between 6 and 9, between 10 and 13). No details regarding the duration or type of the disease were reported.

Year of arrival in France. In case the child was born abroad, the year of arrival in France was reported amongst: between 1993 and 1995, between 1996 and 1998, between 1999 and 2001, between 2002 and 2004, between 2005 and 2007.

Whom the child lives with. Parents or legal guardian reported whether the child lived with his family (several options) or in child care.

Measures used for the characterization of reading impaired pupils

Phonological awareness. Participants were given a list of 5 written words and had to tick the one that did not share a common sound with the others. There were 10 trials.

Grammar. Students had to fill in the blanks in 3 short texts with logical connectors, determiners or pronouns (20 items, open-ended questions).

Mathematics. Students had to answer 48 questions (26 open and 22 multiple choice) involving problem solving, logic, mental arithmetic, notions of time and units.

Perceived self-efficacy. In grade 6, students answered questions from the Children's Perceived Self-Efficacy scales (Bandura, 1990), closely translated into French. It is a 37-item questionnaire from which factors representing perceived academic self-efficacy, social self-efficacy and self-regulatory efficacy were extracted. For each item, students had to evaluate their ability to perform a given activity using a 5-point Likert scale. The perceived academic self-efficacy score measures students' perceived ability to manage their learning, to master different academic subjects (mathematics, science, etc . . .), and to fulfill parents' and teachers' expectations. The perceived social self-efficacy score measures efficacy regarding leisure group activities, the ability to form and maintain social relationships and manage interpersonal conflicts, and self-assertiveness. Lastly, the perceived self-regulatory efficacy score measures students' perceived ability to resist peer pressure to engage in high-risk activities (alcohol, drugs, transgressive behaviors).

Motivation. Students' academic motivation was assessed in grade 6 with questions derived from the Academic Self-Regulation Questionnaire (SRQ-A) (Ryan & Connell, 1989), adapted and translated into French (Leroy, Bressoux, Sarrazin, & Trouilloud, 2013). This is a self-report measurement which assesses individual differences in motivational styles. The items asked students the reasons why they do their home and class work, and try to do well in school. Each item provides a possible reason that represents a certain motivational style (for example: "I do my classwork because I want to learn new things," or "I do my classwork because I'd be ashamed of myself if it didn't get done"). Each item was answered using a 5-point Likert scale. Three factors were extracted: intrinsic motivation, extrinsic motivation, and amotivation.

Handedness. Parents reported whether the child was right-handed (N = 20,466), left-handed (N = 2,977) or ambidextrous (N = 300). We grouped together left-handed and ambidextrous; our handedness variable is thus binary (right-handed versus non-right-handed).

Household monthly income. Parents reported the household monthly income. Given the skewed distribution, we used the natural logarithm of income.

Parental education. Parents' highest diploma was converted into years of education for each parent (from 0 to 18.5 years – 18.5 years corresponding to a graduate degree). Parental education was then estimated as the mean of the two parents' education (when one was missing, only the other one was used).

Socio-economic status (SES) was computed as the mean of z-scored household income and parental education (when one was missing, only the other one was used).

Disadvantaged school. Obtained from the ministry's central database, this variable indicated whether the child's school belonged to either *Réseau de Réussite Scolaire* or *Réseau Ambition Réussite*, two networks of disadvantaged areas targeted for priority education policy.

Analysis of diagnostic criteria

Table 1 lists each of the diagnostic criteria included in ICD-11's definition of "developmental learning disorder with impairment in reading" and in DSM-5's definition of "specific learning disorder, with impairment in reading." While the wording differs, many of the criteria are similar, and are therefore placed on the same line in the table for comparison. In the third column, we indicate whether and how we took each criterion into account in the present analysis.

In summary, criteria shared by the two classifications include 1) significance/severity of the disorder; 2) school-age manifestation and persistence; 3) specification of relevant reading skills; 4) reference to a norm for chronological age; 5) interference with academic or occupational performance; 6) exclusion of intellectual disability, sensory and neurological disorders, inadequate education, non-proficiency in instruction language, and psychosocial adversity.

Criteria where the classifications differ are 1) only the DSM-5 includes an "insufficient response to intervention" criterion; 2) only the ICD-11 includes an IQ-discrepancy criterion.

Some criteria deserve further comment.

Table 1. Systematic comparison of ICD-11 and DSM-5 criteria for reading disability, and their operationalization in the present study.

ICD-11 criteria	DSM-5 criteria	Operationalization	Comment
Persistent difficulties	Presence of at least one of the following symptoms that have persisted for at least 6 months. The learning difficulties begin during school-age years	Application in 6th grade meets both persistence and school-age criteria.	
	despite the provision of interventions that target those difficulties	Not taken into account	No data available.
in learning academic skills related to reading, such as word reading accuracy, reading fluency, and reading comprehension.	With impairment in reading: Word reading accuracy Reading rate or fluency Reading comprehension	Score below threshold on reading comprehension measure	No other reading measure available.
The individual's performance in reading is markedly below what would be expected for chronological age	The affected academic skills are substantially and quantifiably below those expected for the individual's chronological age,	Reading score *x* SD below the mean of students born in 1996 and in 6th grade in 2007	x = 1, 1.25, 1.5 (default), 1.75, 2
and level of intellectual functioning		Reading score *x* SD below that predicted by linear regression of reading on non-verbal intelligence.	x = 1, 1.25, 1.5 (default), 1.75, 2
and results in significant impairment in the individual's academic or occupational functioning.	and cause significant interference with academic or occupational performance, or with activities of daily living	Score at 6th grade national evaluations *y* SD below mean	y = 0, 0.5 (default), 1
Developmental learning disorder with impairment in reading is not due to a disorder of intellectual development,	The learning difficulties are not better accounted for by intellectual disabilities,	Score above −2SD on non-verbal IQ	
sensory impairment (vision or hearing),	uncorrected visual or auditory acuity,	Not taken into account	No data available.
neurological disorder,	other mental or neurological disorders,	Not taken into account	No data available.
lack of availability of education,	or inadequate educational instruction.	Exclusion of pupils with "serious illness" before age 9.	No information available on "adequacy" of instruction, or on other causes of school drop-out
lack of proficiency in the language of academic instruction,	lack of proficiency in the language of academic instruction,	Exclusion of pupils arrived in France after age 6.	No direct information available on French language proficiency
or psychosocial adversity.	psychosocial adversity,	Exclusion of pupils in care.	No information available on other indicators of psychosocial adversity

- Regarding the IQ-discrepancy threshold (ICD-11), we applied regression-based discrepancy. We investigated 1, 1.25, 1.5, 1.75, 2 SD below the reading score linearly predicted by non-verbal IQ. Because the variance of reading scores decreases as non-verbal intelligence increases (a Breusch-Pagan test confirmed heteroscedasticity: $p < .0001$), we used weighted least squares to estimate the variance of the residuals of reading comprehension on non-verbal intelligence score, as a function of non-verbal intelligence scores (assuming that the variance varies linearly with IQ scores). We first regressed the squared residuals of the OLS regression on intelligence score. The fitted values from this regression constitute our estimate of residual variance. We then ran weighted least squares using the inverse of this estimate as weights in order to produce efficient regression parameters.
- The "interference with academic or occupational performance" is a ubiquitous criterion in medical classifications, but is never applied in research and only partly so in clinical settings (beyond the fact that consultation implies some degree of perceived interference). Here, we applied this criterion, using a threshold on the mean score of 6th grade national evaluations

(including both French and Mathematics), to represent interference with academic performance. We investigated thresholds 0, 0.5 and 1 SD below the mean, with 0.5 SD as the default threshold, reasoning that "interference" does not imply as stringent thresholds as "disorder".

Of course, as specified in the DSM-5, in clinical settings, these criteria would be "based on a clinical synthesis of the individual's history (developmental, medical, family, educational), school reports, and psychoeducational assessment." In a population study where clinical diagnoses are unavailable, each diagnostic criterion is evaluated or approximated using the data that was collected, and some criteria have to be ignored by lack of data. Nevertheless, clinical diagnoses also have to relate measures of performance to norms and to apply thresholds depending on various methodological choices. It is on these specific aspects that the present study may be instructive.

Statistical analyses

Prevalence of reading disability was calculated using the full set of available criteria and each threshold for both ICD-11 and DSM-5. To recall, the full set of criteria is:

- Reading comprehension score x standard deviations below the mean of French pupils born in 1996 and schooled in 6th grade in 2007.
- Reading comprehension score x SD below that predicted by linear regression of reading on non-verbal intelligence. (ICD-11 only).
- Score at 6th grade national evaluations y SD below mean.
- Score above −2 SD on non-verbal IQ.
- Exclusion of pupils with "serious illness" before age 9.
- Exclusion of pupils arrived in France after age 6.
- Exclusion of pupils in care.

The impact of the various criteria on prevalence was then investigated by:

- Varying the severity threshold x (from 1 to 2 SD below mean). For ICD-11, the regression-based IQ-discrepancy threshold was kept equal to the severity threshold, for consistency.
- Varying the "interference with academic performance" threshold y (from 0 to 1 SD below mean).
- Relaxing each criterion one at a time (using the −0.5 SD default threshold for interference).

Agreement between ICD-11 and DSM-5 diagnoses was compared using a contingency table (using default thresholds).

The reading-disabled and control populations were then compared on all the available variables: reading comprehension, grammar, phonology, mathematics, non-verbal intelligence, mean score at the national evaluations, self-efficacy, motivation, SES, schooling in priority education area, sex, handedness, and grade repetition, using weighted t-tests and chi-square tests.

Finally, given the interest in "twice exceptional children" who show both high IQ (>130) and a learning disability, and given uncertainties on their prevalence (Brody & Mills, 1997; Lovett & Lewandowski, 2006; Toffalini, Pezzuti, & Cornoldi, 2017; Van Viersen, Kroesbergen, Slot, & De Bree, 2016), we also estimated the prevalence of reading disability separately in high-IQ and in other pupils.

In order to account for differences in the likelihood to respond to the survey, we computed non-response propensity weights, which we combined with the initial weights from the exhaustive baseline survey, in order to adjust the results to a representative sample of the population. Non-response propensity scores were first computed with a logistic regression, and inversed in order to obtain the non-response propensity weights. Both the baseline weights and the non-response propensity weights

Table 2. Prevalence of dyslexia according to ICD-11 and DSM-5, at different thresholds, using all criteria and relaxing one criterion at a time. Default thresholds are in bold. Source: MENESR DEPP, Panel 2007.

ICD-11								
	All criteria							
	Achievement threshold			Relaxing 1 criterion				
Severity threshold	0 SD	**−0.5 SD**	−1 SD	IQ (absolute threshold + discrepancy)	Interference with academic performance	Serious illness	Language proficiency	Psychosocial adversity
−1 SD	9.6	**7.3**	4.7	15.0	12.3	7.6	7.7	7.4
−1.25 SD	6.4	**5.1**	3.5	11.2	8.0	5.3	5.4	5.2
−1.5 SD	**4.3**	**3.5**	**2.5**	**7.9**	**5.1**	**3.7**	**3.7**	**3.6**
−1.75 SD	2.6	**2.2**	1.7	5.1	3.0	2.3	2.4	2.2
−2 SD	2.0	**1.7**	1.3	5.1	2.3	1.8	1.8	1.7

DSM-5								
	All criteria							
	Achievement threshold			Relaxing 1 criterion[b]				
Severity threshold	0 SD	**−0.5 SD**	−1 SD	IQ (absolute threshold)	Interference with academic performance	Serious illness	Language proficiency	Psychosocial adversity
−1 SD	17.2	**13.1**	8.4	15.0	21.4	13.6	13.7	13.3
−1.25 SD	12.0	**9.7**	6.6	11.2	14.3	10.0	10.1	9.8
−1.5 SD	**7.9**	**6.6**	**4.8**	**7.9**	**9.3**	**6.9**	**7.0**	**6.7**
−1.75 SD	4.8	**4.2**	3.3	5.1	5.4	4.3	4.4	4.2
−2 SD[a]	4.8	**4.2**	3.3	5.1	5.4	4.3	4.4	4.2

[a]The last 2 lines' showing identical numbers is due to the discreteness of reading comprehension scores, and the absence of any cases between −1.75 SD and −2 SD.
[b]When one criterion other than interference with academic performance is relaxed, the threshold for academic performance is −0.5 SD below mean.

were then scaled; the product of these two scaled weights constitutes our final weights. All analyses were conducted in R (survey, questionr and weights packages). Analysis scripts are available on https://osf.io/kejcp/?view_only=c97c136728fa4a10ba805b2c35e6f07c

Results

Descriptive statistics for the included participants and comparison with the excluded participants are provided in Table S1. Weighted descriptive statistics in the working sample and measures' reliability are reported in Table S2. The distribution of reading comprehension scores is shown in Supplementary Figure S2. Kurtosis was close to normality (2.92) and skew was negative (−0.52). Examination of the histogram suggests that the asymmetry is due to the test not having enough difficult items.

Impact of classification, thresholds and criteria on reading disability prevalence

Table 2 presents the prevalence obtained using the two classifications, depending on severity thresholds, and interference with academic performance threshold. Furthermore, we investigated the impact of each of the other criteria by reporting the prevalence when this criterion is relaxed (at the default achievement threshold −0.5 SD). For instance, according to ICD-11, when using a −1.5 SD threshold on reading ability and a −0.5 SD threshold on academic performance, the prevalence of reading disability is 3.5%. When relaxing the IQ criterion (IQ>70 and reading < −1.5 SD score predicted by IQ), the prevalence raises to 7.9%. Using the same thresholds, the prevalence of reading disability according to DSM-5 is 6.6%. When relaxing the IQ criterion (IQ>70), the prevalence raises to 7.9%, just like according to ICD-11, since the two classifications differ only in the IQ criterion in the present study (the other difference, absence of response to intervention, could not be taken into account). Using the same thresholds, when the academic performance criterion is relaxed, prevalence raises from 3.5 to 5.1% according to ICD-11 and from 6.6 to 9.3% according to DSM-5. All the other criteria (as implemented here) have a much lower impact on prevalence.

Figure 1 shows the relation between reading comprehension and non-verbal IQ scores across the entire population (r = 0.46), with lines showing both severity and IQ thresholds according to each classification. It illustrates the main difference between the two classifications: In ICD-11, reading-disabled individuals are in the bottom right trapezium (in red), while in DSM-5 they are in the bottom right rectangle (in yellow). Individuals in the yellow triangle in Figure 1C are those diagnosed as reading disabled by DSM-5 but not by ICD-11, because their reading score is not sufficiently discrepant with their IQ.

Concordance between the two classifications

Table S3 shows the contingency table of the classification into disabled and normal reader according to both classifications. Using the default thresholds (−1.5 SD for reading severity and −0.5 SD for achievement), DSM-5 diagnoses almost twice as many individuals as ICD-11 (884 for ICD-11, 1662 for DSM-5). All the individuals diagnosed by ICD-11 are also diagnosed by DSM-5 (in the absence of the response to intervention criterion), but DSM-5 additionally includes a large number of individuals with a reading score above the reading-IQ discrepancy threshold (778 individuals).

Characteristics of individuals diagnosed as reading-disabled

Table 3 reports the results on all the available variables for individuals diagnosed or not with reading disability, according to ICD-11 and DSM-5. By definition, disabled readers scored much lower than normal readers in reading comprehension (Cohen's d≈-2.8). They also scored lower in other measures of French language (grammar: d≈-1.5; phonology: d≈-0.9), in Mathematics (d≈-1.4), and in overall

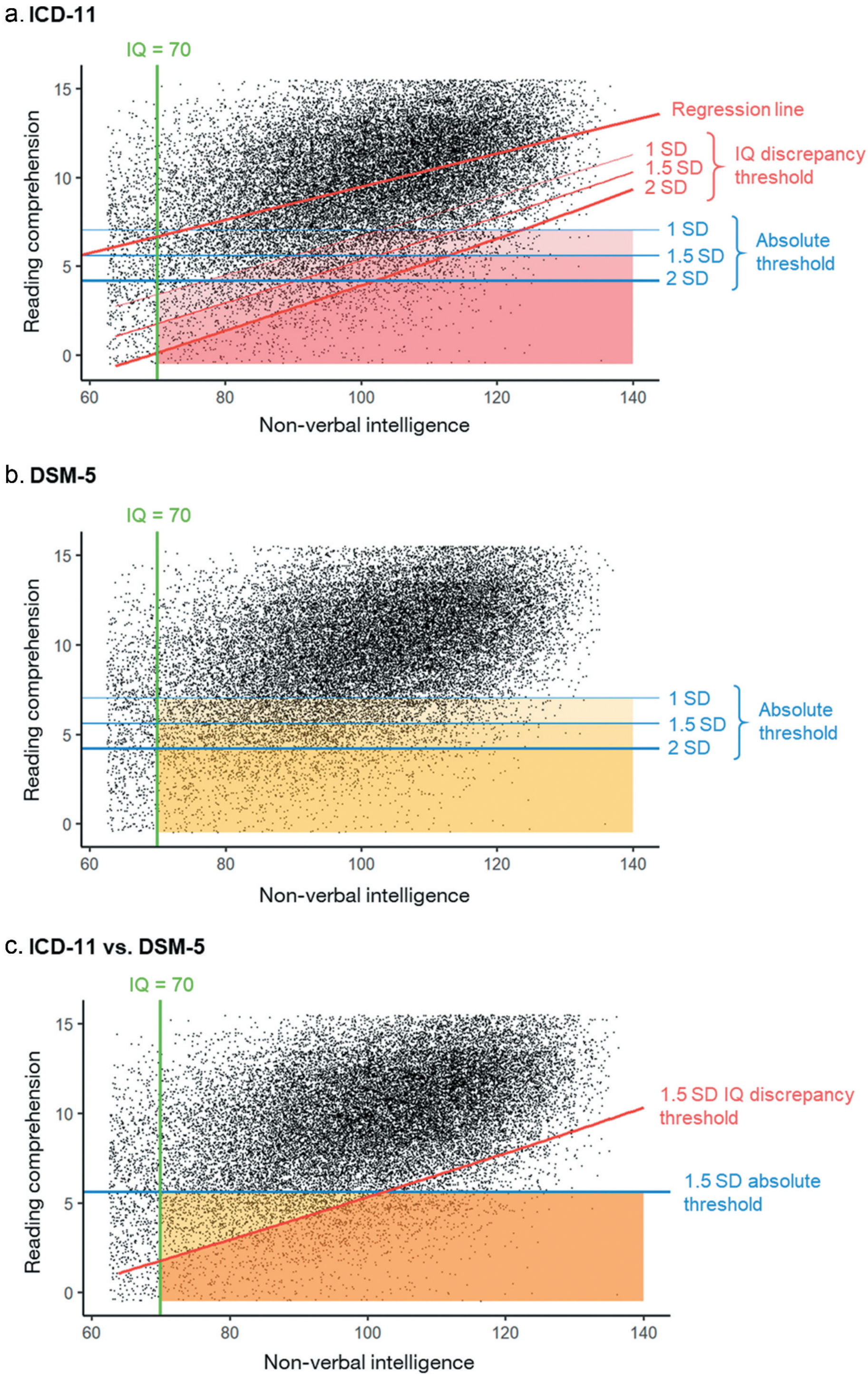

Figure 1. Reading comprehension as a function of non-verbal intelligence in the entire population. Lines represent different severity thresholds and their application by ICD-11 and DSM-5. A) ICD-11. B) DSM-5. C) Comparison of ICD-11 and DSM-5 criteria. Regions in shades of red represent individuals diagnosed by ICD-11. Regions in shades of yellow represent individuals diagnosed by DSM-5. The region in Orange (C) represents individuals diagnosed by both classifications. Scores on both dimensions are jittered for better visibility. Source: MENESR DEPP, Panel 2007.

Table 3. Descriptive statistics for individuals diagnosed by ICD-11, DSM-5 and normal readers, computed using the −1.5 SD severity criterion, the −0.5 SD achievement criterion, and weighted data. Source: MENESR DEPP, Panel 2007.

	ICD-11						DSM-5						Difference	
	Reading disabled		Normal readers		Difference		Reading disabled		Normal readers		Difference			
	(N = 911)		(N = 24.130)		(Disabled – normal readers)		(N = 1.726)		(N = 23.315)		(Disabled – normal readers)		Reading disabled ICD-11 – DSM-5	
	M(or %)	*SD*	*M(or %)*	*SD*	*d or OR*	*p*	*M(or %)*	*SD*	*M(or %)*	*SD*	*d or OR*	*p*	*d or OR*	*p*
Reading comprehension	2.80	1.55	9.67	2.91	−2.95	<.0001	3.57	1.48	9.85	2.79	−2.81	<.0001	−0.51	<.0001
Grammar	3.33	2.55	8.74	4.26	−1.54	<.0001	3.56	2.53	8.91	4.20	−1.54	<.0001	−0.09	0.0282
Phonology	4.75	2.26	6.79	2.16	−0.92	<.0001	4.86	2.25	6.85	2.13	−0.91	<.0001	−0.05	0.1991
Mathematics	16.00	5.93	26.16	8.75	−1.36	<.0001	15.62	5.51	26.54	8.60	−1.51	<.0001	0.06	0.1398
National assessment score	33.18	10.82	59.98	16.91	−1.89	<.0001	33.83	10.20	60.85	16.42	−1.98	<.0001	−0.07	0.1149
Intelligence score (non-verbal)	94.29	11.72	100.22	15.07	−0.44	<.0001	88.77	11.54	100.82	14.89	−0.91	<.0001	0.47	<.0001
Socio-economic index	−0.61	0.87	0.02	1.00	−0.67	<.0001	−0.60	0.88	0.04	1.00	−0.68	<.0001	−0.02	0.7200
Priority Education (%. OR)	26.93		15.18		2.06	<.0001	28.11		14.69		2.27	<.0001	1.06	0.5694
Repeaters (%. OR)	55.20		16.15		6.40	<.0001	53.37		14.94		6.51	<.0001	0.93	0.3733
Boys (%. OR)	64.65		50.58		1.79	<.0001	61.33		50.34		1.56	<.0001	0.86	0.0971
Non-right-handed (%. OR)	14.85		13.81		1.09	0.4354	16.44		13.66		1.24	0.0029	1.13	0.3768
Self-regulation	−0.86	1.09	0.02	0.98	−0.86	<.0001	−0.82	1.11	0.05	0.96	−0.83	<.0001	−0.04	0.3360
Social self-efficacy	−0.31	1.18	0.00	0.98	−0.29	<.0001	−0.28	1.16	0.01	0.98	−0.27	<.0001	−0.03	0.6139
Academic self-efficacy	−0.68	1.08	0.00	0.98	−0.66	<.0001	−0.63	1.09	0.02	0.98	−0.62	<.0001	−0.05	0.3744
Intrinsic motivation	−0.10	0.98	−0.01	0.92	−0.10	0.0149	−0.08	0.97	−0.01	0.92	−0.08	0.0083	−0.02	0.6350
Extrinsic motivation	0.02	0.90	−0.01	0.82	0.02	0.5206	0.05	0.88	−0.01	0.82	0.07	0.0145	−0.04	0.4061
Amotivation	0.47	1.15	0.00	0.88	0.46	<.0001	0.43	1.16	−0.02	0.87	0.44	<.0001	0.03	0.5277

academic performance (d≈-1.9). They also scored lower on nonverbal IQ (by d = −0.44 according to ICD-11 and d = −0.91 according to DSM-5). On average, they had a lower SES (d≈-0.7), they were twice as likely to be in a priority education area, 3.5 times as likely to have repeated a grade, 70% more likely to be a boy, and 4 to 25% more likely to be non-right-handed. They also showed lower ratings in various measures of self-efficacy and motivation. In all these measures, differences between the two groups were generally similar between the two classifications. However, due to the difference on the discrepancy criterion, reading-disabled participants were significantly more impaired in reading comprehension (d = −0.51) yet had higher non-verbal IQ (d = 0.47) according to the ICD-11 than to the DSM-5 definition.

Table S4 reports a more detailed analysis of how the sex-ratio varied as a function of each diagnostic criterion. Whereas the sex-ratio (M/F) in the entire population was 1.04, it increased to 1.6 for poor readers (reading score below −1.5 SD). Furthermore, the more severe the threshold, the higher the sex-ratio, from 1.47 at −1 SD to 1.76 at −2 SD. Most exclusion criteria did not affect sex-ratio, however the exclusion of individuals with intellectual disability decreased the sex ratio slightly to 1.54 (at −1.5 SD), suggesting that boys were slightly overrepresented among the very low IQ poor readers. Overall, the sex-ratio for DSM-5 criteria remained close to that for poor readers. However, the addition of the IQ-discrepancy (ICD-11 criteria) significantly increased the sex-ratio, to 1.53 at −1 SD and to 2.22 at −2 SD, suggesting that boys were particularly overrepresented among the discrepant poor readers.

Table S5 provides a more detailed analysis of handedness as a function of diagnostic criteria. Overall, non-right-handedness was slightly more prevalent in poor readers (15.7% at −1.5 SD) and in DSM-5 disabled readers (16.4%) than in the entire population (13.8%). Prevalence of non-right-handedness increased with the severity threshold (from 15% at −1 SD to 16.5% at −2 SD). However, the application of the IQ discrepancy criterion lowered the prevalence of non-right-handedness (to about 15%), making it non-significantly different from the base rate.

We then considered whether this increased prevalence of non-right-handedness in reading disability might be due to the sex-ratio in favor of boys, or to the lower IQ. Indeed, we found non-right-handedness in 15.6% of boys and 12.0% of girls. Furthermore, we found that non-right-handers scored on average 2 IQ points lower than right-handers. In order to eliminate biases due to sex and IQ, we ran a logistic regression with diagnostic category as dependent variable, and handedness, sex, IQ and their interactions as independent variables (Supplementary Results). We found significant effects of IQ, of sex (in DSM-5 only), and of the interaction between sex and IQ on diagnostic category. However, there was no significant effect of handedness, nor any interaction involving handedness, suggesting that the prevalence of handedness does not differ between disabled and normal readers, once sex and IQ are controlled.

Finally, we found a prevalence of reading disability of 0.4% [0.05; 3.0] in high IQ students (nonverbal intelligence > 130) with both DSM-5 and ICD-11 criteria, compared to 3.7% [3.4; 4.0] (ICD-11) and 6.9% [6.6; 7.0] (DSM-5) in the non-high IQ population.

Discussion

We evaluated the prevalence of reading disability in France in a large representative population of 6th grade pupils, according to the two widely-used international classifications, and at five different severity thresholds on a reading comprehension test. Prevalence estimates range from 1.3% (ICD-11, −2 SD severity threshold, – 1 SD achievement threshold) to 17.2% (DSM-5, −1 SD severity threshold, 0 SD achievement threshold). Using reasonable compromise thresholds (−1.5 SD for severity, −0.5 SD for achievement), prevalence estimates are 3.5% with ICD-11 and 6.6% with DSM-5.

These numbers are consistent with expectations based on applying a −1.5 SD threshold onto a normal distribution (corresponding to 6.68%), and the fact that the IQ discrepancy criterion is expected to exclude a substantial number of cases, particularly given the correlation between IQ and reading comprehension. These numbers are also within the range described in the literature, although no previous study to our knowledge used exactly the same set of criteria approximating most closely the definitions in DSM-5 and ICD-11. Beyond the numbers that are to some extent arbitrary, the present study's most important contribution lies in the investigation of the consequences of each diagnostic criterion on prevalence estimates and on the characteristics of the population diagnosed with reading disability.

Impact of each criterion on prevalence

Regression-based IQ-reading discrepancy

We found that DSM-5 yielded systematically higher prevalence rates than ICD-11, and this is directly attributable to the one criterion that differed between the two: the IQ-reading discrepancy threshold, required by ICD-11 but not by DSM-5. Compared with the DSM-5, applying this threshold additionally excluded from the ICD-11 diagnosis all individuals who had an IQ above 70, and who had a reading comprehension score below the severity threshold, but above the discrepancy threshold (illustrated by the region in yellow in Figure 1C). Depending on the specific thresholds, this had the effect of excluding up to half of all individuals diagnosed with reading disability by DSM-5. Whether or not one implements such a discrepancy criterion has therefore important consequences for many individuals whose reading ability is not substantially below that predicted by their non-verbal intelligence. However, it should also be noted that our reading score being based on reading comprehension must have inflated the correlation between reading and IQ (compared to a reading measure based on reading accuracy and/or fluency), thus may overestimate the impact of the IQ-reading discrepancy criterion.

As a side remark, it may seem paradoxical that the DSM-5 chose to emphasize specificity in the name "specific learning disorder" while simultaneously shunning the IQ-discrepancy criterion that operationalized cognitive specificity, when the ICD-11 preserved this specificity criterion, without emphasizing it in the name of the disorder ("developmental learning disorder").

Insufficient response to intervention

Our finding of a systematically higher prevalence in DSM-5 should be taken with caution, given that we were unable to evaluate the impact of the "insufficient response to intervention" criterion, which is required in DSM-5 only, and which should logically decrease the reported prevalence for DSM-5. We are not aware of any previous study implementing this diagnostic criterion. Therefore, how much DSM-5 prevalence would decrease if the diagnostic criteria were fully implemented is a matter of speculation.

Interference with academic performance

Together with IQ-reading discrepancy, this was the other criterion that had a major impact on prevalence estimates. In ICD-11, prevalence ranged from 2.5% with a −1 SD threshold on achievement to 5.1% without such a threshold. In DSM-5, prevalence similarly ranged from 4.8% to 9.3%. Furthermore, the use of reading comprehension as the only reading measure likely inflated the correlation between reading and academic performance, thereby attenuating the impact of the cutoff on academic performance. Thus, again, whether one applies such a criterion has a major impact on who will or will not receive a diagnosis. The individuals concerned are those who manage, in one way or in another, to reach at least normal academic performance despite their reading difficulties. Some people consider that learning disability may sometimes be masked by high intellectual ability, such that even normal and over-achievers may sometimes deserve a diagnosis of reading disability and require help (Brody & Mills, 1997; Van Viersen et al., 2016). In contrast, medical classifications consider that there can be no diagnosis without sufficient evidence of interference with one's functioning. The present study quantifies the impact of adopting one view or the other.

Exclusion criteria

The exclusion criteria that we were able to partly apply were intellectual disability, lack of education, lack of proficiency in the language of instruction and psychosocial adversity. The intellectual disability criterion excluded between 1 and 2% of children from a diagnosis of reading disability. All other exclusion criteria had a small impact (at most 0.6%). The exclusion criteria that we were not able to apply were sensory impairment (auditory or visual), neurological disorder, and inadequate instruction (DSM-5), although the population sampling relying on entrance into middle school did exclude a number of severely disabled children. Applying the first two criteria might decrease prevalence by at most another 1%, at least in countries like France where such disorders are infrequent. Arguably, excluding cases with inadequate instruction might have a larger impact, depending on how "adequate instruction" is defined (e.g., more or less strict phonics methods) and implemented in the country of interest.

Characteristics of reading-disabled pupils according to ICD-11 and DSM-5

Non-verbal intelligence

Pupils diagnosed with reading disability had on average slightly lower non-verbal IQ than the rest of the population, and this was more pronounced according to DSM-5 (d = −0.91) than according to ICD-11 (d = −0.44). This difference follows from the reading-IQ regression-based discrepancy criterion, whereby ICD-11 excluded a number of pupils with relatively low-average IQ that were included by DSM-5 (see also Stevenson, 1992; Stuebing, et al., 2002 for similar results). This lower IQ contrasts with the widespread stereotype of the "normal or superiorly intelligent" dyslexic child, but is a logical consequence of the correlation between general intelligence and reading ability. It is consistent with some previous population-based studies (Katusic et al., 2001; Stevenson, 1992; Stuebing, et al., 2002) but not others (Rutter & Yule, 1975). In the present study, the IQ difference may have been amplified by the application of the "interference with academic performance" criterion, which was not used by most previous studies, and by the use of a reading comprehension measure for reading ability.

Academic performance

By definition, children with reading disability scored much lower than normal readers in reading comprehension (d≈-2.8). They also showed lower performance on grammar (d≈-1.5), phonology (d≈-0.9), mathematics (d≈-1.4), and on overall academic performance (d≈-1.9). They were also 3.5 times as likely to have repeated a grade (Zorman et al., 2004). These results are partly due to our application of the "interference with academic performance" criterion. They may also partly be due to the fact that all those academic skills were tested in writing exclusively. And they may be amplified by relying solely on a reading comprehension measure. Nevertheless, they are consistent with a large literature showing that measures of IQ, oral language, and all academic skills are positively correlated in the population, and that poor readers typically score lower in all these domains (Badian, 1999b; Dirks, Spyer, & Sonneville, 2008; Gross-Tsur, Manor, & Shalev, 1996; Lewis, Hitch, & Walker, 1994; Moll et al., 2014). Although the two classifications largely agreed in this respect, it may be noted that mean scores in mathematics and overall academic performance were slightly lower for pupils diagnosed with reading disability by DSM-5 than by ICD-11, and their incidence of grade repetition was higher. This may be interpreted as a consequence of including more low-IQ children in the DSM-5 diagnosis.

Sociological variables

On average, pupils diagnosed with reading disability had a lower SES (d≈-0.7), and they were twice as likely to be in a priority education (i.e., disadvantaged) area. This result may have been amplified by relying solely on a reading comprehension measure. Nevertheless, this is consistent with a large literature on the impact of social factors on reading ability (Fluss, et al., 2009; Plaza et al., 2002; Shaywitz et al., 1999; Watier, Dellatolas, & Chevrie-Muller, 2006). Interestingly, although it has been hypothesized that the IQ discrepancy criterion allowed one to better identify cases of reading disability with a predominantly biological (rather than social) origin (e.g., Fluss et al., 2009), in the present study the mean SES of the diagnosed population did not differ between the two classifications.

Sex ratio

The Programme for International Students Assessment (PISA) has now well documented sex differences in reading achievement: in France in 2015, boys' reading performance was 29 points lower than girls', with other OCDE countries showing the same trend (27 point mean difference) (OECD, 2016). Consistently, many studies have reported that reading disability affected boys more than girls. Shaywitz et al. (1990) suggested that estimations of sex ratio were inflated when they were based on schools or clinics, due to referral bias: indeed, they found that while the sex ratio could reach 4 when based on school identification, it was around 1.26–1.5 when applying uniform criteria to an entire population. Yet, most population-based studies have reported uneven sex-ratios, although of varying magnitude.

For instance, the Isle of Wight study (Rutter & Yule, 1975) reported a sex ratio of about 3.3 boys to 1 girl for specific reading retardation. Badian (1999a) similarly reported 3.2 boys to 1 girl for discrepant reading disability. Neither of these studies used a lower IQ threshold, potentially leading them to overestimate the sex-ratio, but they used a discrepancy criterion, which should have limited the problem. Since then, sex ratios have tended to decrease with successive studies: 1.3–1.4 in (Flynn & Rahbar, 1994), around 2 in (Flannery, Liederman, Daly, & Schultz, 2000; Katusic et al., 2001); 1.4 to 3 in (Rutter et al., 2004); 1.6–2.4 in (Quinn & Wagner, 2015); close to 1 in (Landerl & Moll, 2010; Moll et al., 2014) for poor reading (and 1.2–1.4 for poor spelling). Interestingly, it has been noted since Rutter and Yule (1975) that the sex-ratio was higher for specific reading disorder (involving a discrepancy criterion) than for poor reading (they reported 1.3) (see also Badian, 1999a; Quinn & Wagner, 2015). Our data are consistent with this observation: we find a sex ratio of 1.8 according to the ICD-11 definition, vs. 1.6 according to the DSM-5 or to a poor reading threshold. Furthermore, the more stringent the severity threshold, the larger the sex ratio (up to 2.2 at −2 SD, consistent with Quinn & Wagner, 2015). Given that these results were obtained in a fully representative population, this confirms that boys are at higher risk of reading disability, particularly so of severe and specific reading disability, and that their higher prevalence in clinical practice cannot entirely be due to referral bias.

Handedness

Historically, Geschwind and Behan (1982) reported a significant association between handedness, reading disability and other disorders. They proposed an explanatory model involving fetal testosterone that is still controversial, and replications of the findings have been mixed (Pennington, Smith, Kimberling, Green, & Haith, 1987; Tonnessen, Løkken, Høien, & Lundberg, 1993). Another explanation could be pathological left-handedness, a condition linking some cases of left-handedness to intellectual retardation, due to birth trauma (Soper & Satz, 1984), but this is also controversial (Hardyck, Petrinovich, & Goldman, 1976). However, meta-analyses suggested that the association between handedness and reading disability was very small, if real at all (Bishop, 1990; Eglinton & Annett, 1994). In the present study, we found a slight excess of non-right-handedness in reading disability, which was significant using DSM-5 criteria but not using ICD-11. However, analyses adjusting on sex and IQ suggested that there was no specific association between reading disability and non-right-handedness, in agreement with Bishop (1990).

Giftedness

Finally, our results are consistent with Toffalini et al. (2017), in suggesting that high-IQ children are at much lower, not higher, risk of reading disability. This result may have been amplified by relying solely on a reading comprehension measure, which is more highly correlated with IQ than other reading measures.

Overall, although the disabled readers identified by ICD-11 and DSM-5 differ in reading comprehension (lower in ICD-11) and in non-verbal IQ (lower in DSM-5) by construction, it is interesting to note that they do not significantly differ in any other aspect. In other words, the phenotype of reading disability seems to be the same, whether one identifies it using a discrepancy criterion or not, as noticed long ago (Stanovich, 1991).

Limitations

The main limitations of this study have already been evoked and may be summarized as follows:

- Some diagnostic criteria could not be applied because the relevant information was not available in the database (see Table 1). These include: 1) "provision of interventions that target those difficulties" (DSM-5); 2) sensory impairment; 3) mental or neurological disorders. The first criterion would be particularly likely to modify prevalence rates and some of the comparisons between the two classifications if applied. However, we are not aware of any previous study applying it.
- Some diagnostic criteria were only approximated using available data. These include 1) lack or inadequacy of instruction, for which only serious illnesses were taken into account; 2) lack of proficiency in the language of academic instruction, for which only age of arrival in France was taken into account; 3) psychosocial adversity, for which only pupils in care were excluded.
- The only reading measure was reading comprehension, which has a number of consequences.
  - First, this provides only a partial assessment of reading ability, ignoring reading accuracy and reading fluency. Of course, the three aspects of reading ability are substantially correlated (Cirino et al., 2013), but not identical. Basing the diagnosis only on reading comprehension is likely to have underrepresented purely decoding-based and fluency-based poor readers who would nevertheless show reading comprehension within the normal range. According to Cirino et al. (2013), 28.4% of struggling readers did not have comprehension problems. How exactly this limitation affects prevalence estimates depends on how thresholds would be implemented, should reading accuracy and fluency measures be available. If the severity threshold was applied separately to each reading component, then requiring a discrepancy in all three components would likely decrease the prevalence, while requiring a discrepancy in only one (or two) of the three components would likely increase it, compared to the present study based on only one component. If the severity threshold was applied to a composite measure of the three components, the specific children diagnosed would marginally change: more decoding- and fluency-based, and fewer comprehension-based poor readers would be diagnosed. It is difficult to predict in what direction overall prevalence estimates would change, if at all.
  - Another consequence of the comprehension-based reading assessment is that it may have inflated the reading-IQ correlation and the impact of IQ-based criteria, thereby increasing the discrepancy between ICD-11 and DSM-5 classifications, decreasing the mean IQ of pupils diagnosed with reading disability, and decreasing the prevalence of reading disability among high-IQ pupils.
  - Similarly, this may have inflated the correlation between reading and academic performance, thereby decreasing the mean achievement of reading-impaired pupils, and attenuating the impact of the criterion on interference with academic performance.
  - Finally, relying on only one reading measure made it less reliable than would be desirable, which would be a problem for individual diagnosis, but is less so for the present purposes.
- Reading and IQ tests were administered collectively in class, making them less reliable than individual evaluation by a trained professional, and perhaps inflating their correlation with school achievement, thus perhaps diminishing the impact of the achievement criterion.
- The results obtained are of course limited to the population of French 6th-graders in 2007, and are expected to vary depending on grade, language, and educational system.

It should be noted that, while most of these limitations can be overcome in smaller-scale studies in clinical and experimental settings, where it is theoretically feasible to apply every single diagnostic criterion (although none ever did), this is practically impossible in large-scale population studies that necessarily have to compromise between several different objectives.

Nevertheless, this study is to our knowledge 1) the one using the largest subset of all DSM and ICD criteria; 2) the only one to systematically compare DSM-5 and ICD-11, including the differential effects of various criteria and thresholds; 3) one of those relying on the largest population sample, using weights to adjust the results to a representative sample of the population.

Conclusion

We calculated the prevalence of reading disability in France using a large database representative of French sixth-graders, and according to two different official definitions: the ICD-11's and the DSM-5's. Applying a −1.5 SD threshold, 3.5% and 6.6% of grade 6 students were identified as reading-disabled, according to ICD-11 and DSM-5, respectively. However, the two definitions disagreed on the classification of a large number of students. Differences directly followed from the use by ICD-11 (but not DSM-5) of a discrepancy criterion between IQ and reading performance. Boys, students with lower SES, and those schooled in disadvantaged areas were overrepresented amongst those diagnosed with reading disability, regardless of the definition. Furthermore, pupils with reading disability also had lower scores in mathematics, non-verbal IQ, self-efficacy and motivation on average.

Authors' contributions

FR designed the study, CDF and AG performed the analysis under supervision of HP and FR, CDF and FR wrote the paper, all authors revised the paper and approved the final version.

Disclosure statement

We have no known conflict of interest to disclose.

Funding

This work has received support under the program "Investissements d'Avenir" launched by the French Government and implemented by ANR with the references ANR-17-EURE-0017 and ANR-10-IDEX-000102 PSL. The data used was obtained from « Panel d'élèves du second degré, recrutement 2007 2007-2013, DEPP - Ministère de l'Éducation [producteur], ADISP-CMH [diffuseur] ».

ORCID

Ava Guez http://orcid.org/0000-0002-6509-1893
Hugo Peyre http://orcid.org/0000-0001-8757-0783
Franck Ramus http://orcid.org/0000-0002-1122-5913

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). Washington, D.C.: American Psychiatric Association.

Badian, N. A. (1999a). Reading disability defined as a discrepancy betwveen listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities*, 32(2), 138–148. doi:10.1177/002221949903200204

Badian, N. A. (1999b). Persistent arithmetic, reading, or arithmetic and reading disability. *Annals of Dyslexia*, 49(1), 43. doi:10.1007/s11881-999-0019-8

Bandura, A. (1990). *Multidimensional scales of perceived self-efficacy*. Stanford, CA: Stanford University.

Bishop, D. V. (1990). *Handedness and developmental disorder*. Cambridge: Cambridge University Press.

Brody, L. E., & Mills, C. J. (1997). Gifted children with learning disabilities: A review of the issues. *Journal of Learning Disabilities*, 30(3), 282–296. doi:10.1177/002221949703000304

Chartier, P. (2012). *Evaluer les capacités de raisonnement avec les tests RCC. Paris: Eurotests Éditions.*

Cirino, P. T., Romain, M. A., Barth, A. E., Tolar, T. D., Fletcher, J. M., & Vaughn, S. (2013). Reading skill components and impairments in middle school struggling readers. *Reading and Writing*, 26(7), 1059–1086. doi:10.1007/s11145-012-9406-3

Dirks, E., Spyer, G., & Sonneville, L. D. (2008). Prevalence of combined reading and arithmetic disabilities. *Journal of Learning Disabilities*, 41(5), 460–473. doi:10.1177/0022219408321128

Eglinton, E., & Annett, M. (1994). Handedness and dyslexia: A meta-analysis. *Perceptual and Motor Skills*, 79(3_suppl), 1611–1616. doi:10.2466/pms.1994.79.3f.1611

Flannery, K. A., Liederman, J., Daly, L., & Schultz, J. (2000). Male prevalence for reading disability is found in a large sample of Black and White children free from ascertainment bias. *Journal of the International Neuropsychological Society*, 6(4), 433–442. doi:10.1017/S1355617700644016

Fluss, J., Ziegler, J. C., Warszawski, J., Ducot, B., Richard, G., & Billard, C. (2009). Poor reading in French elementary school: The interplay of cognitive, behavioral, and socioeconomic factors. *Journal of Developmental & Behavioral Pediatrics*, 30(3), 206–216. doi:10.1097/DBP.0b013e3181a7ed6c

Flynn, J. M., & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *Psychology in the Schools*, 31(1), 66–71. doi:10.1002/1520-6807(199401)31:1<66::AID-PITS2310310109>3.0.CO;2-J

Francis, D. J., Shaywitz, S. E., Stuebing, K. K., Shaywitz, B. A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88 (1), 3–17. doi:10.1037/0022-0663.88.1.3

Geschwind, N., & Behan, P. (1982). Left-handedness: Association with immune disease, migraine, and developmental learning disorder. *Proceedings of the National Academy of Science U.S.A*, 79(16), 5097–5100. doi:10.1073/pnas.79.16.5097

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology*, 38(1), 25–33. doi:10.1111/j.1469-8749.1996.tb15029.x

Guez, A., Panaïotis, T., Peyre, H., & Ramus, F. (2018). Predictors of the IQ-achievement gap in France: A longitudinal analysis. *Intelligence*, 69, 104–116. doi:10.1016/j.intell.2018.05.008

Hardyck, C., Petrinovich, L. F., & Goldman, R. D. (1976). Left-handedness and cognitive deficit. *Cortex*, 12(3), 266–279. doi:10.1016/S0010-9452(76)80008-1

Katusic, S. K., Colligan, R. C., Barbaresi, W. J., Schaid, D. J., & Jacobsen, S. J. (2001). Incidence of reading disability in a population-based birth cohort, 1976–1982, Rochester, Minn. *Mayo Clinic Proceedings*, 76(11), 1081–1092. doi:10.4065/76.11.1081

Landerl, K., & Moll, K. (2010). Comorbidity of learning disorders: Prevalence and familial transmission. *Journal of Child Psychology and Psychiatry*, 51(3), 287–294. doi:10.1111/j.1469-7610.2009.02164.x

Leroy, N., Bressoux, P., Sarrazin, P., & Trouilloud, D. (2013). Un modèle sociocognitif des apprentissages scolaires: Style motivationnel de l'enseignant, soutien perçu des élèves et processus motivationnels. *Revue Francaise De Pedagogie*, n° 182(1), 71–92. doi:10.4000/rfp.4008

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry*, 35(2), 283–292. doi:10.1111/j.1469-7610.1994.tb01162.x

Lindgren, S. D., De Renzi, E., & Richman, L. C. (1985). Cross-national comparisons of developmental dyslexia in Italy and the United States. *Child Development*, 56(6), 1404–1417. doi:10.2307/1130460

Lovett, B. J., & Lewandowski, L. J. (2006). Gifted students with learning disabilities: Who are they? *Journal of Learning Disabilities*, 39(6), 515–527. doi:10.1177/00222194060390060401

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., Schulte-Körne, G., & Kroesbergen, E. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS ONE*, 9, 7. doi:10.1371/journal.pone.0103537

OECD. (2016). *PISA 2015 results (Volume I). Excellence and Equity in Education* 11/12/2020. https://ictlogy.net/bibliography/reports/projects.php?idp=3204

Pennington, B. F., Smith, S. D., Kimberling, W. J., Green, P. A., & Haith, M. M. (1987). Left-handedness and immune disorders in familial dyslexics. *Archives of Neurology*, 44(6), 634–639. doi:10.1001/archneur.1987.00520180054016

Plaza, M., Chauvin, D., Lanthier, O., Rigoard, M. T., Roustit, J., Thibault, M., & Touzin, M. (2002). *Validation longitudinale d'un outil de dépistage des troubles du langage écrit. Etude d'une cohorte d'enfants dépistés en fin de CP et réévalués en fin de CE1* 81, (Glossa (Paris),) 22-33. .

Quinn, J. M., & Wagner, R. K. (2015). Gender differences in reading impairment and in the identification of impaired readers results from a large-scale study of at-risk readers. *Journal of Learning Disabilities*, 48(4), 433–445. doi:10.1177/0022219413508323

Rodgers, B. (1983). The Identification and Prevalence of Specific Reading Retardation. British Journal of Educational Psychology, 53(3), 369-373. https://doi.org/10.1111/j.2044-8279.1983.tb02570.x

Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology*, 53(3), 369–373. doi:10.1111/j.2044-8279.1983.tb02570.x

Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., . . . Carroll, J. (2004). Sex differences in developmental reading disability—New findings from 4 epidemiological studies. *Jama-Journal of the American Medical Association*, 291(16), 2007–2012. doi:10.1001/jama.291.16.2007

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal of Child Psychology and Psychiatry*, 16 (3), 181–197. doi:10.1111/j.1469-7610.1975.tb01269.x

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, 57(5), 749–761. doi:10.1037//0022-3514.57.5.749

Share, D. L., McGee, R., McKenzie, D., Williams, S., & Silva, P. A. (1987). Further evidence relating to the distinction between specific reading retardation and general reading backwardness. *British Journal of Developmental Psychology*, 5(1), 35–44. doi:10.1111/j.2044-835X.1987.tb01039.x

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading ability. *New England Journal of Medicine*, 326(3), 145–150. doi:10.1056/NEJM199201163260301

Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., . . . Shaywitz, B. A. (1999). Persistence of dyslexia: The connecticut longitudinal study at adolescence. *PEDIATRICS*, 104(6), 1351–1359. doi:10.1542/peds.104.6.1351

Shaywitz, S. E., Shaywitz, B. A., Fletcher, J. M., & Escobar, M. D. (1990). Prevalence of reading disability in boys and girls. Results of the connecticut longitudinal study. *Journal of the American Medical Association*, 264(8), 998–1002. doi:10.1001/jama.1990.03450080084036

Soper, H. V., & Satz, P. (1984). Pathological left-handedness and ambiguous handedness: A new explanatory model. *Neuropsychologia*, 22(4), 511–515. doi:10.1016/0028-3932(84)90046-0

Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly*, 26(1), 7–29. doi:10.2307/747729

Stanovich, K. E. (2005). The future of a mistake: Will discrepancy measurement continue to make the learning disabilities field a pseudoscience? *Learning Disability Quarterly*, 28(2), 103–106. doi:10.2307/1593604

Stevenson, J. (1988). Which aspects of reading ability show a 'hump' in their distribution? *Applied Cognitive Psychology*, 2(1), 77–85. doi:10.1002/acp.2350020107

Stevenson, J. (1992). Identifying sex differences in reading disability: Lessons from a twin study. *Reading and Writing*, 4 (4), 307–326. doi:10.1007/BF01027711

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal*, 39(2), 469–518. doi:10.3102/00028312039002469

Terriot, K. (2014). Testons les tests: Le RCC (Raisonnement sur Cartes de Chartier). *ANAE-Approche Neuropsychologique Des Apprentissages Chez l'Enfant*, 26(129), 179–183.

Toffalini, E., Pezzuti, L., & Cornoldi, C. (2017). Einstein and dyslexia: Is giftedness more frequent in children with a specific learning disorder than in typically developing children? *Intelligence*, 62, 175–179. doi:10.1016/j.intell.2017.04.006

Tonnessen, F. E., Løkken, A., Høien, T., & Lundberg, I. (1993). Dyslexia, left-handedness, and immune disorders. *Archives of Neurology*, 50(4), 411–416. doi:10.1001/archneur.1993.00540040063016

Van Viersen, S., Kroesbergen, E. H., Slot, E. M., & De Bree, E. H. (2016). High reading skills mask dyslexia in gifted children. *Journal of Learning Disabilities*, 49(2), 189–199. doi:10.1177/0022219414538517

Watier, L., Dellatolas, G., & Chevrie-Muller, C. (2006). Difficultés de langage et de comportement à 3 ans et demi et retard en lecture au cours élémentaire. *Revue d'Épidémiologie et de Santé Publique*, 54(4), 327–339. doi:10.1016/S0398-7620(06)76728-1

World Health Organization. (2018). *International statistical classification of diseases and related health problems* (*11th Revision* (Geneva: World Health Organization)Accessed 18 12 2019 ed.). https://icd.who.int/browse11/l-m/en

Yule, W., Rutter, M., Berger, M., & Thompson, J. (1974). Over- and under-achievement in reading: Distribution in the general population. *British Journal of Educational Psychology*, 44(1), 1–12. doi:10.1111/j.2044-8279.1974.tb00760.x

Zorman, M., Lequette, C., & Pouget, G. (2004).Dyslexies: intérêt d'un dépistage et d'une prise en charge précoce à l'école. Évaluation du BSEDS 5-6 [Dyslexia: Significance of early screening and remediation at school. The BSEDS 56 test]. In Metz-Lutz, M. N. ed. .*Développement Cognitif Et Troubles Des Apprentissages [Cognitive Development and Learning Disabilities]* (Marseille: Solal), 245–270.