Check for updates

Screening for learning disabilities in Oman: confirmatory factor analysis of the Arabic version of the learning disabilities diagnostic inventory

Mahmoud Mohamed Emam, Rashid Almehrizi, Ehab Omara* and Ali Mahdi Kazem

Department of Psychology, College of Education, Sultan Qaboos University, Muscat, Oman

Students at risk for learning disabilities (LD) are overidentified in elementary schools in Oman due to the absence of adequate instruments which teachers can use in validating their observations. Teachers need valid instruments so that their judgment of students' behaviours can help in making academic and non-academic decision. The Learning Disabilities Diagnostic Inventory (LDDI) is widely used to examine manifestations of LD as an intrinsic processing disorder. This study aimed to verify the six-factor structure with confirmatory factor analysis, and test its measurement invariance across child gender. Teachers completed the Arabic version of the LDDI for an Omani sample of 1564 children aged 7 to 11. Overall, the six-factor model showed an acceptable fit after performing some post-hoc modifications that were justified on a theoretical ground. Results indicated metric invariance across gender. Zero-order correlations, however, were highly significant, which reflected that the Arabic version of the LDDI showed poor discriminant validity compared to the original LDDI. The Arabic version of the LDDI, however, can be a handy tool to screen for the LD manifestations and help in responding to teachers' academic concerns about students in Key stages 1 and 2 in Oman.

Keywords: screening, learning disabilities diagnostic inventory, Arab context, Oman

The diagnosis of children with learning disabilities (LD) has been a common point of interest among professionals from education, medicine, psychology, and mental health. The terms specific learning disorders and specific learning disabilities are used interchangeably, with the former being used mostly by professionals working in psychiatry and clinical professions, and for whom the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) serves as the canonical text and the latter being used by professionals from education, psychology and mental health (Al-Yagon et al. 2013). Epidemiological research shows that about 5–7% of school aged children struggle with reading, writing, and mathematics. Recently, there has been sufficient research evidence resulting from the use of neuroimaging techniques, which unraveled the genetic background of LD (Lagae 2008). The definition by the National Joint Committee on Learning Disabilities (NJCLD) is perhaps the most cited definition in literature. Generally, an SLD is defined as an unexpected, specific, and persistent failure to acquire competent academic skills despite conventional instruction, adequate intelligence, and sociocultural opportunity (Colker 2011).

Within the medical model of disability SLD diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders-IV, and IV-TR were subject to a number of changes which appeared in DSM-5 (Cavendish 2013; DuPaul *et al*. 2013; Scanlon 2013). Generally, however, the recent three releases of DSM classify learning problems as developmental problems. Recent research evidence shows that learning problems can already be identified at an earlier preschool or early school age (Hulme and Snowling 2013; Pennington *et al*. 2012; Butterworth *et al*. 2011). According to DSM-5, SLD is a type of neurodevelopmental disorders, and included three sub-categories SLD with impairment in reading (dyslexia), SLD with impairment in written expression, SLD with impairment in mathematics (Dyscalculia) (Katchergin 2016)

Teachers are often regarded as a rich source for obtaining information which assists in the assessment of various aspects of students' behaviour (Liljequist and

Correspondence to: Mahmoud Mohamed Emam, Department of Psychology, College of Education, Sultan Qaboos University, P.O. Box: 32 Al-Khoud, PC123, Muscat, Oman. Email: memam@squ.edu.om
*Department of Educational Psychology - College of Education, Ain Shams University, Cairo, Egypt.

© The British Society of Developmental Disabilities 2019
DOI 10.1080/20473869.2019.1683367

Renk 2007). Therefore, teachers' ratings are often used for the identification of students suspected of having various disabilities; and their evaluations are regarded as being core predictors of students' placement (Tenenbaum and Ruck 2007). Recently, teachers' involvement in the identification process of students with LD have received support within the use of the responsiveness-to-intervention (RTI) model of identification in North America (Haager *et al*. 2007; Vaughn and Fuchs 2003).

The availability of several teacher rating scales that can be used in screening for LD raised questions with regard to which instrument is the most efficient and valid for use by practitioners in schools (Kettler and Albers 2013). While there is a proliferation of teachers' survey instruments of LD in the Western context, the lack of such instruments in Arab countries emerges as an issue that affects adequate screening, identification and referral procedures in schools. In Oman, teachers' nomination of students at risk for LD does not follow standard widely known procedures and protocols and do not rely on the use of valid and reliable instruments through which teachers' observations can be validated.

The Learning Disabilities Diagnostic Inventory (LDDI, Hammill and Bryant 1998) is a teacher rating scale that received empirical evidence for use in the identification of LD in Western cultures. Grounded in the information processing approach to the assessment of LD, the LDDI attempts to systematically identify SLD based on a student's intrinsic processing difficulties by observing the student's day-to-day academic endeavors. The current study aimed to examine the adequacy of the LDDI structure for validating its use in another context, namely Oman. The validation of the LDDI was part of a larger research project that aimed at the development of an optimal framework for the identification of students with LD in key stages 1 (grades 1–4) and 2 (grades 5–6) in Oman.

Previous research on teacher ratings for the identification of SLD

Several studies indicated that teacher ratings commonly assist in identifying SLD (Kettler and Albers 2013; Kettler and Elliott 2010; Kettler *et al*. 2008; Kettler 2007). These studies found that teacher judgments can be accurate predictors of students' needs for academic assistance. Two early studies (Gresham *et al*. 1987; Gerber and Semmel 1984) suggested valuing teacher judgments because teachers have daily contacts with students in a meaningful context in which to evaluate academic performance. Gresham *et al*. (1987), in an early and classical study, indicated that teachers correctly identified 96% of students diagnosed with a learning disability. Additionally, Gresham *et al*. (1997) reported that teachers were competent in identifying students with LD, students with low achievements, and students with low IQs.

A number of studies tested the criterion and predictive validity of teacher ratings for identifying LD. These studies compared between teachers' rating scales and other performance measures including curriculum based measures (Martin and Shapiro 2011; Begeny *et al*. 2011; Eckert *et al*. 2006) and state achievement proficiency tests (Kettler and Albers 2013; Kettler and Elliott 2010; Kettler *et al*. 2008). These studies showed that teacher rating correlated with other measures for identifying LD. The studies, however, differed in terms of the accuracy of teachers' judgments ranging from low to large. Alternatively, however, Vaughn and Fuchs (2003) found that teachers in general education settings were less reliable when providing ratings related to LD than special education teachers in special education settings. Overall, the studies that tested the feasibility, reliability and validity of teacher ratings supported the use of teacher rating scales in screening for LD in school settings.

The LDDI is one good example of using evidence based teacher rating scales to assist in the identification of SLD. The LDDI was used in several studies in Western countries. Lock and Layton (2001b) modified the LDDI and used it in self-report format in order to make it a self-discovery tool to measure the student's strengths and weaknesses. As such, the authors used it to help assess college students' ability of self-advocacy. Lock and Layton (2002) examined in-service teachers' sensitivity to differences between typical second language acquisition issues and the presence of learning disabilities. Thirty general educators without specific training in second language acquisition development were selected to complete the LDDI. The general education teachers used the LDDI to evaluate a total of 121 students. Sixty students were English speaking; and sixty-one students were identified as English language learners. Independent t-tests indicated significant differences between the two groups for all subscales on the LDDI at the .0001 level of confidence, except for the writing subscale that was significant at the .002 level of confidence. A large effect size was obtained for all subgroups. Additionally, chi square analysis indicated that students who were English language learners but not labeled as learning disabled were significantly over-identified by the teacher using the LDDI. The Chi-square was significant at the .05 level. Results from these analyses could indicate that the LDDI is not a valid instrument for a population of students who are English language learners. The authors concluded, however, that the teachers' ratings might have been inadequate. It appeared that the teachers demonstrated a lack of sensitivity towards the issues of typical second language acquisition versus those of learning disabilities. This lack of sensitivity may have resulted in an inappropriate evaluation of the students overly identifying them as learning disabled. In an attempt to address the question of the LDDI's validity

versus the teachers' lack of sensitivity, Layton and Lock (2002) conducted a study to examine the effect of training teachers on increasing their sensitivity to distinguish between learning disability indicators and second language acquisition using the LDDI. The authors reported that training teachers' improved their use of the LDDI to differentiate between English language learners and those who are English language learners with learning disabilities. Padeliadu and Sideridis (2008) developed the Learning Disabilities Screening Scale (LDSS) for teachers, which was identical in principle and concept to the LDDI, and found that both scales were reliable and valid. Layton and Lock (2001) used the LDDI to examine SLD in students with low vision. The authors aimed to examine the valid procedures for documenting a dual diagnosis. The authors recommended the use of a mixed methodology assessment including both quantitative and qualitative data, documenting the usefulness of the LDDI.

SLD in the Sultanate of Oman

The Sultanate of Oman is a country in Southwest Asia, on the southeast coast of the Arabian Peninsula. It borders the United Arab Emirates in the northwest, Saudi Arabia in the west, and Yemen in the southwest. The coast is formed by the Arabian Sea in the south and east, and the Gulf of Oman in the northeast. It is a member of the United Nations, the Arab League, the Gulf Cooperation Council (GCC). Oman has made a stride in expanding access to education since 1970. by providing over 1000 tuition free government schools for their citizens. The total duration of general education in Oman is 12 years, divided into two stages: Primary education, which lasts for 10 years and is composed of two key stages and secondary education that lasts for two years. Key stage 1 of primary education covers grades 1–4 while the key stage 2 covers grades 5–10. Students who complete and pass key stage 2 are allowed to participate in the last two years of study which culminates in an exam to earn the General Educational Diploma, a prerequisite qualification to attend higher education. Available research evidence on the achievement of Omani students indicates that there is continuous gender achievement gap. The gap has been reported for students at the different phases of general education (Akiba and Liang 2014) as well as higher education (Alfarhan and Dauletova 2017; Jansen *et al*. 2016; Martin *et al*. 2012; Mullis *et al*. 2016). Omani female students generally outperform male students on international assessments such as Trends in International Mathematics and Science Study (TIMSS) (Akiba and Liang 2014).

Currently, one of the most challenging issues that exist in the educational system in Oman is the overidentification of students with LD for special education services. Sensitizing Omani teachers to the issues that indicate a learning disability versus the typical achievement problems that result from lack of quality teaching, the absence of achievement motivation or any other reason is of paramount importance. The National Joint Committee on Learning Disabilities (2001) defines SLD in part as a ''general term that refers to heterogeneous groups of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning, or mathematical abilities''. The definition, which is widely approved by policy makers, researchers, and practitioners in the Arab world, identifies these disorders as intrinsic and occurring throughout the individual's lifetime. Statistics of the Department of Special Education at the Ministry of Education (MoE) in Oman shows that the number of pupils with LDs who receive special education services in inclusive schools has increased significantly in the last decade. In addition, schools have become incapable of providing special education services to more pupils with LDs due to lack of both human and learning resources (Author 2016).

An Educational Mandate was issued in 2007 in accordance with the Omani Children with Disabilities Care and Rehabilitation Act in 1996 that was reauthorized in 2008. Through the mandate children with LDs became eligible for receiving additional support in a Learning Support Unit (LSU) in schools by LD teachers who hold a specialized one-year Higher Diploma in LDs, a program which is tailored by Sultan Qaboos University to fulfill the training needs of the Ministry of Education professionals in the Sultanate of Oman (Author 2016). In general, however, the issue of determining a learning disability remained particularly elusive for general educators in Oman who nominate students for referrals to LDUs. Authors (2015) found that teachers were unable to differentiate between students with emotional and behavioural difficulties and students with LD in Oman. Similarly, Authors (2015) were also unable to differentiate between students with attention deficit hyperactivity disorder (ADHD) and those with LD. The two studies showed that teachers are unable to ascertain the difference between symptoms of comorbid disorders such as emotional and behavioural difficulties (EBD) and ADHD from intrinsic processing disorders (learning disabilities). Additionally, both studies referred to the prevalence of inadequate practices in the evaluation process leads to inappropriate diagnosis. Based on the aforementioned review of literature and description of the research context, the purpose of the present study was to use confirmatory factor analysis (CFA) to validate the factor structure of the LDDI with a national sample of elementary school students in Oman. The current study sought to answer the following two research questions:

- Does the data on the Arabic version of the LDDI support the originally hypothesized six-factor structure?

- For studying gender differences when screening for LD, Is the measurement by the LDDI invariant across gender?

Methods

Participants

A total of 2450 students in key stages 1 and 2 schools in Oman were selected for the study as part of the larger strategic research project described above. Key stage 1 extends from grade 1 to 4 covering ages between 6 and 10 while key stage 2 includes grades 5 and six covering ages from 11 to 12. The sample selection procedures, although not epidemiological, was representative of the all the geographical areas in Oman, covering all the 11 governorates of Oman. Stratified sampling technique guided the selection of the study participants. With stratified sampling, we divided the population into separate 11 geographic main regions in Oman. We, then, used a probability simple random sampling technique for each geographic area. Teachers were requested to rate the students and return the forms. Of these, 1564 or 63.8% of the students were rated by their teachers, and 886 or 36.2% of the students were rated by the LD teachers who serve as tutors in the LSU. Participants included both genders (48.3% male, 51.7% female) with an age range of 7–12 years. The response rate was relatively good as the MoE formally supported the recruitment for the strategic research project. In addition, the estimated time for completing the LDDI was reasonable, ranging from 10 to 20 min.

Instrument

The LDDI is a rating scale which was developed by Hammill and Bryant (1998) and published by PRO.ED. The scale was designed to identify intrinsic processing disorders and SLD in students aged 8–17. It was designed for the single purpose of helping professionals screen for and identify SLD in individuals. The authors surveyed the available instruments that can be used for evaluating the symptoms of SLD as an intrinsic processing disorder (Dean and Burns 2002) and found that the LDDI is adequate and evidence based as it was in several previous studies to indicate the risk status of students in elementary school. The LDDI is not an ability or achievement measure; it does not tell how well or how poorly students read, write, listen, speak, calculate mathematics, or think. Instead, it will identify the extent to which students' skill patterns in a particular area are consistent with individuals known to have a learning disability in that area (reading disorder, writing disorder, or mathematics disorder). Thus, using the LDDI shifts the diagnostic emphasis away from interpreting norm-referenced ability test scores toward studying an individual's skill patterns, especially patterns that are indicative of individuals who are known to have SLD.

The LDDI is a 90 item rating scale that is composed of six independent scales, Listening, Speaking, Reading, Writing, Mathematics, and Reasoning. Each scale included 15 items describing one of the areas listed in the definition by the U.S. Office of Education and the National Joint Committee on Learning Disabilities, a definition that is used widely in several contexts worldwide including the Arabic context. Teachers rate students' skills on a nine-point response scale (1,2,3=frequently, 4,5,6 = sometimes, and 7,8, 9=rarely). According to the LDDI user manual, the LDDI was built with the American Psychological Association's standards for technical adequacy clearly in mind. The test was normed on 2,152 students with learning disabilities residing in forty-three states and the District of Columbia. The demographic characteristics of the normative sample are representative of the population of students who have learning disabilities in the United States as a whole with regard to gender, race, ethnicity, urban/rural residence, family income, educational attainment of parents, and geographical distribution.

According to the user manual, the LDDI has excellent psychometric properties. Internal consistency reliability coefficients exceeded .90 for all scales. Evidence for stability and interscorer reliability is provided, and coefficients ranged between .80 and .90. Validity indicators showed that the LDDI scores have content description, criterion-prediction, and construct-identification validity. Validity testing involved extensive item selection and differentiation examinations, which included confirmatory factor analysis, and the examination of the LDDI's relationship to age, academic achievement, group differentiation, gender, and ethnicity, all of which support the validity of the LDDI scores. Students are identified as being at risk for having LD based on the use of stanines (1–9) which are converted from raw scores. A profile of a typical student with LD includes at least one score below stanine 6 (i.e. 5 or lower) and at least one score above 6 (i.e. 7 or higher). This pattern which contains both low and high scores indicates the presence of a significant intraindividual difference in that at least one score indicates an intrinsic processing disorder while another score shows normal performance. Alternatively, profiles composed of consistently low or high scores do not indicate the risk status for LD (Hammill and Bryant 1998).

Procedures

Data were collected began in the Spring semester and continued to the Spring semester of the following year. Ethical approval was obtained from Sultan Qaboos University Research Ethical Board (SQUREB) prior to commencing the data collection phase. Similarly, approval was obtained from the Ministry of Education to conduct the study at the selected schools prior to

data collection. Our target sample was 3000 children but we could obtain consent from schools as well as from the families for 2450. Families gave assent by signing a consent form prior to their participation in the study indicating that they were willing to have their children included in the study in order to be assessed with regard to their academic skills. Families and teachers were informed that participation was voluntary and that confidentiality of the data obtained would continue even after the study results are disseminated. Teachers were selected based on their good knowledge of the selected students. We recruited a cohort of 30 teachers for the strategic project and received training on the administration and scoring procedures of the LDDI. This central cohort of trainers delivered the same training to local cohorts of trainers in the different 11 governorates of Oman. Teachers were given enough time to fill out the LDDI so that the study did not intervene with their daily schedule.

We translated the LDDI using Modern Standard Arabic (MSA) language. MSA is the most widely used version of Classical or Standard Arabic (C/SA). It is used in newspapers, books and everyday media. Alternatively, C/SA is the oldest version of Arabic and which still exists to date because it the language in which the Quran was written. We used the consecutive and back-translation method (Hambleton *et al.* 2004) by the authors. A third translator who had background in educational psychology conducted the back-translation while having no access to the original questionnaires. Two translations were sufficient to achieve a desired adaptation. However, we modified the following items to match the structure, morphology and phonology of the Arabic language. These items originally included examples from the English language and were replaced by examples from the Arabic language. We modified the examples provided in the following items: (1) Item 3 in Listening 'Has difficulty recognizing that sentences with different syntax can mean the same thing (e.g. "The dog chased the cat" means the same as "The cat was chased by the dog"), (2) item 11 in Speaking "cannot say common blends (e.g./bl/,/cr/,/fl/)", (3) item 4 in Reading "substitutes phonetically similar words while reading aloud (e.g. "chair" for cheer, "then" for when)", (4) item 10 in Reading "confuses words that appear similar (e.g. "bread" for broad)", and (5) item 11 in Reading "Omits inflectional endings (i.e. -s, -ed, -ing)" when reading aloud.'

Data analysis

The six-factor structure of the LDDI was assessed using confirmatory factor analysis (CFA). The analysis was conducted using Amos 22. CFA is one of the most widely used techniques to test the factor structure of multidimensional psychological measures (Dana 2000; Weiner and Craighead 2010). The method depends on comparing a theoretical model with the observed data which is collected from a representative sample. Additionally, invariance of the factor structure across gender was tested using multiple group comparisons.

CFA was used to compare the extent to which the Omani sample data fit the LDDI six factor structure that was established in the norming of the instrument (Hammill and Bryant 1998). CFA was used because it enables testing specific hypotheses and determines a priori the structure of the instrument as it is theoretically designed (Harrington 2009). Hence, items that belong to a factor are allowed to load on that specific factor. Invariance of structure across gender was also tested. The original six factor model proved to be robust in the norming of the instrument and therefore the exact items from the LDDI used for norming were included in the present study. The authors followed the guidelines for conducting SEM/CFA (Jackson *et al.* 2009). The data were examined for normality using skewness and kurtosis. Both values indicate normality when they are between (−2 and +2). We manipulated the missing data by using the mean substitution method after deleting cases with incomplete data. All parameter estimates were performed using covariance matrices and maximum-likelihood estimation to test the LDDI structure goodness of fit. In addition, the latent variable covariances were free to be estimated. Following the CFA, the correlations were calculated to determine the relations between the factors. In testing this model, three indexes of model fit were computed in addition to the χ^2 statistic, χ^2– degrees of freedom ratio. The three indexes include: Bentler's (1990) comparative fit index (CFI), Tucker and Lewis's (1973) index of fit (TLI), and Browne and Cudek's (1993) root mean square error of approximation (RMSEA). These indices assess different aspects of model fit and have varying criteria for a model demonstrating good fit. Recent methodological research indicated that the CFI and TLI, values should be at or above .90 to indicate a good fitting model (Bryne 2010), with values close to 1 indicating a very good fit on any of these indexes. An RMSEA of less than .10 indicates a reasonable fit, and an RMSEA of about .08 or less indicates a close fit of the model in relationship to the degrees of freedom. If the ratio of χ^2 to degrees of freedom is less than 4 the model can be considered a fitting model (Bowen and Guo 2011).

Results

Descriptive statistics

Table 1 shows the means, standard deviations, skewness and kurtosis of each of the items of the LDDI. It can be noted that none of the items has a severe skewness or kurtosis. All values were between (−2 and +2). These values are considered acceptable to claim the normality of item's responses (Blunch 2012).

Table 1. Means, standard deviation, skewness and kurtosis for the six scales of the LDDI.

		Listening			Speaking			Reading			Writing			Math			Reasoning		
Items	N	M (SD)	Sk	Ku	M (SD)	Sk	Ku	M (SD)	Sk	Ku	M (SD)	Sk	Ku	M (SD)	Sk	Ku	M (SD)	Sk	Ku
S1	1507	6.11 (2.34)	−.39	−.86	5.51 (2.49)	−.16	−1.17	5.98 (2.58)	−.37	−1.14	5.75 (2.52)	−.25	−1.15	5.65 (2.48)	−.19	−1.13	5.02 (2.38)	.14	−1.01
S2	1507	7.21 (2.11)	−1.16	.45	5.46 (2.46)	−.09	−1.15	4.81 (2.71)	.20	−1.30	5.19 (2.54)	.00	−1.19	5.66 (2.5)	−.20	−1.15	5.27 (2.37)	−.03	−1.04
S3	1507	6.29 (2.24)	−.50	−.71	5.71 (2.52)	−.27	−1.14	4.71 (2.72)	.25	−1.29	4.75 (2.62)	.29	−1.19	5.52 (2.49)	−.19	−1.14	5.02 (2.42)	.10	−1.06
S4	1507	6.12 (2.29)	−.35	−.93	5.37 (2.45)	−.01	−1.16	5.8 (2.45)	−.23	−1.11	5.93 (2.57)	−.30	−1.17	5.59 (2.55)	−.20	−1.16	5 (2.44)	.13	−1.11
S5	1507	6.27 (2.2)	−.47	−.74	5.4 (2.51)	−.09	−1.20	5.64 (2.42)	−.16	−1.09	4.78 (2.67)	.20	−1.28	5.17 (2.59)	−.04	−1.22	5.48 (2.46)	−.13	−1.16
S6	1507	5.94 (2.24)	−.33	−.81	5.21 (2.46)	.02	−1.14	5.86 (2.47)	−.27	−1.07	4.88 (2.51)	.21	−1.17	5.21 (2.7)	−.08	−1.29	4.84 (2.54)	.16	−1.16
S7	1507	6.42 (2.11)	−.53	−.60	5.67 (2.3)	−.16	−.98	4.91 (2.58)	.14	−1.21	5.46 (2.51)	−.06	−1.22	5.44 (2.45)	−.16	−1.09	4.81 (2.49)	.18	−1.12
S8	1507	5.77 (2.34)	−.25	−.99	5.72 (2.3)	−.18	−.98	5.63 (2.55)	−.17	−1.21	5.87 (2.47)	−.29	−1.10	5.8 (2.56)	−.31	−1.12	4.89 (2.52)	.17	−1.15
S9	1507	6.3 (2.35)	−.59	−.72	6.01 (2.33)	−.36	−.91	6.35 (2.45)	−.60	−.81	5.27 (2.71)	−.02	−1.33	5.06 (2.68)	−.01	−1.28	5.15 (2.47)	.02	−1.16
S10	1507	5.4 (2.35)	−.05	−1.03	5.55 (2.43)	−.15	−1.08	6.1 (2.4)	−.40	−.96	5.3 (2.68)	−.02	−1.32	5.26 (2.54)	−.07	−1.20	5.01 (2.43)	.09	−1.08
S11	1507	6.19 (2.33)	−.43	−.87	5.86 (2.34)	−.29	−.93	5.88 (2.47)	−.29	−1.06	5.57 (2.6)	−.11	−1.26	5.35 (2.5)	−.10	−1.16	5.3 (2.59)	−.06	−1.24
S12	1507	6.14 (2.35)	−.43	−.93	5.95 (2.34)	−.33	−.92	5.92 (2.58)	−.34	−1.13	5.73 (2.56)	−.21	−1.20	5.01 (2.6)	.00	−1.23	5.47 (2.52)	−.15	−1.18
S13	1507	6.01 (2.32)	−.34	−.97	5.52 (2.45)	−.11	−1.14	5.73 (2.59)	−.22	−1.19	5.89 (2.46)	−.30	−1.07	4.8 (2.62)	.11	−1.22	4.91 (2.49)	.15	−1.16
S14	1507	5.68 (2.44)	−.19	−1.12	5.7 (2.41)	−.25	−1.04	4.21 (2.7)	.46	−1.12	5.67 (2.65)	−.24	−1.24	5.34 (2.56)	−.09	−1.21	4.85 (2.53)	.16	−1.16
S15	1507	5.54 (2.41)	−.13	−1.07	6.94 (2.18)	−.98	.04	5.09 (2.56)	.12	−1.17	4.72 (2.71)	.23	−1.28	4.58 (2.7)	.23	−1.26	5.58 (2.5)	−.18	−1.16

Confirmatory factor analysis

At first, we handled the missing or incomplete data using IBM SPSS 23. Of the total sample, the response forms of twenty students were found incomplete. We used case deletion practice because we had a small percentage of the number of cases with missing data (Osborne 2013). Then, we conducted CFA on the remaining 1507 cases. The aforementioned fit indices were examined to evaluate the overall fit of the six factor model. First, the results showed that the model had poor fit to the data ($\chi^2 = 22815.15$, df $= 3900$, $\chi^2/df = 5.85$, CFI $= 0.89$, TLI $= 0.89$, RMSEA $= 0.06$).

Investigating the sources of poor fit

According to Harrington (2009) there is a number of sources that could result in poor fit models. Big or too small models, and ill-defined theoretical models are expected to have poor fit indices. The CFA model in this study includes 6 latent variables with 15 items allowed to load on each factor. Also, the modification indices of CFA indicated that there were some indicators with correlated errors within latent variables rather than different variables. Measurement errors could have emerged as a result of a similarity in meaning in both positive and negative statements (Harrington 2009), or due to the respondent's desire to agree with factors that would affect all item scores (Polit and Beck 2008). Additionally, Munro (2005) argues that correlated measurement errors could occur when the variables are not defined or measured directly, which can affect the participants' responses to these items.

We conducted the exploratory factor analysis (EFA) because we assumed that its results may suggest additional factors that can interpret the shared variance between the indicators that the latent variable does not explain (Brown 2015). We conducted the EFA with maximum likelihood extraction method for the items in each factor separately in order to explore whether there are other factors underlying the participant's responses. The results indicated that only one component with eigenvalue greater than one was extracted in each group of items. The percentages of explained variance were 63.48%, 65.14%, 74.10%, 75.51%, 76.41%, and 79.88% for the extracted components (Listening, Speaking, Reading, Writing, Mathematics, and Reasoning), respectively. Additionally, the factor loadings ranged between (0.52–0.91, 0.61–0.88, 0.81–0.88, 0.77–0.91, 0.79–0.92, and 0.82–0.92), respectively. These results support the theoretical structure of LDDI, and the unidimensionality of each group of items.

Through the textual investigation of the meaning of items in each factor, the authors found that the indicators that have correlated errors have similar logical relationships. For example, 'misunderstanding simple

Table 2. Factor loadings for the six scales of the LDDI.

	Listening		Speaking		Reading		Writing		Mathematics		Reasoning	
Items	UFL (SE)	SFL	UFL (SE)	SFL	UFL (SE)	SFL	UFL (SE)	SFL	UFL (SE)	SFL	UFL (SE)	SFL
S1	1.25 (0.04)	0.65	1.32 (0.05)	0.64	1.42 (0.04)	0.72	1.00 (0.00)	0.51	1.00 (0.00)	0.57	1.00 (0.00)	0.63
S2	0.96 (0.05)	0.52	1.47 (0.04)	0.76	1.54 (0.04)	0.77	1.23 (0.04)	0.63	1.25 (0.03)	0.69	1.27 (0.04)	0.68
S3	1.26 (0.04)	0.7	1.35 (0.05)	0.64	1.53 (0.04)	0.76	1.54 (0.04)	0.80	1.25 (0.05)	0.59	1.4 (0.04)	0.78
S4	1.35 (0.04)	0.76	1.49 (0.04)	0.79	1.35 (0.04)	0.72	1.42 (0.04)	0.71	1.54 (0.04)	0.79	1.46 (0.03)	0.82
S5	1.25 (0.04)	0.71	1.48 (0.04)	0.74	1.31 (0.04)	0.70	1.51 (0.04)	0.75	1.61 (0.04)	0.83	1.4 (0.04)	0.75
S6	1.05 (0.05)	0.54	1.48 (0.04)	0.77	1.37 (0.04)	0.73	1.38 (0.04)	0.71	1.47 (0.05)	0.67	1.47 (0.04)	0.78
S7	1.13 (0.04)	0.65	1.1 (0.05)	0.55	1.44 (0.04)	0.75	1.47 (0.04)	0.79	1.5 (0.04)	0.81	1.49 (0.04)	0.83
S8	1.36 (0.04)	0.74	1.21 (0.04)	0.63	1.44 (0.04)	0.75	1.39 (0.04)	0.74	1.55 (0.04)	0.79	1.51 (0.04)	0.84
S9	0.85 (0.06)	0.39	1.27 (0.04)	0.66	1.28 (0.04)	0.67	1.62 (0.04)	0.83	1.55 (0.05)	0.74	1.49 (0.03)	0.84
S10	1.34 (0.04)	0.72	1.36 (0.04)	0.69	1.35 (0.04)	0.75	1.6 (0.04)	0.82	1.56 (0.04)	0.81	1.46 (0.03)	0.84
S11	1.35 (0.04)	0.74	1.29 (0.04)	0.67	1.41 (0.04)	0.78	1.55 (0.04)	0.82	1.59 (0.04)	0.86	1.54 (0.04)	0.82
S12	1.37 (0.04)	0.74	1.41 (0.04)	0.77	1.41 (0.04)	0.72	1.49 (0.04)	0.78	1.63 (0.04)	0.83	1.42 (0.04)	0.74
S13	1.36 (0.04)	0.75	1.5 (0.04)	0.79	1.45 (0.04)	0.74	1.4 (0.04)	0.75	1.6 (0.04)	0.81	1.48 (0.04)	0.82
S14	1.22 (0.04)	0.64	1.43 (0.04)	0.75	1.24 (0.04)	0.61	1.5 (0.04)	0.75	1.59 (0.04)	0.82	1.52 (0.04)	0.83
S15	1.00 (0.00)	0.56	1.00 (0.00)	0.50	1.00 (0.00)	0.58	1.53 (0.04)	0.74	1.65 (0.04)	0.80	1.39 (0.04)	0.73

Note: UFL: Unstandardized Factor Loading and SFL: Standardized Factor Loading.

Table 3. Correlations among the latent factors of LDDI, omega model-based reliability, and cronbach's alpha coefficients.

	1 List	2 Speak	3 Read	4 Writ	5 Math	6 Reas
Listening	(.71/.96)					
Speaking	.85**	(.76/.97)				
Reading	.78**	.80**	(.76/.98)			
Writing	.72**	.76**	.93**	(.80/.98)		
Mathematics	.54**	.53**	.62**	.63**	(.82/.98)	
Reasoning	.66**	.70**	.73**	.75**	.68**	(.85/.98)

spoken sentences and questions (item 4) may cause "misunderstanding spoken directions (item 5)" in Listening.' A student who "has difficulty speaking spontaneously (item 5)" may have "trouble paraphrasing information presented orally (item 6)". In Mathematics, a student who 'takes time to complete calculations (item 12) may "have difficulty with multi-step problems" (item 13) and "word problems" (item 15) which depend also on calculations (item 15).' Hence, the correlation between these item errors could be a result of the relationships between the meanings of the items. The authors, therefore, used the covariances between errors only for items that have a logical relationship.

Fitting of the modified model

We conducted CFA of the modified model. The CFA for the modified model provided an adequate fit to the data ($\chi^2 =$ 17749.57, df = 3877, $\chi^2/df =$ 4.58, CFI = 0.922, TLI = 0.920, RMSEA = 0.049). As noted, although the value of χ^2 was high due to the large sample size, the χ^2/df was less than 5. This indicates that the model had an adequate fit with the collected data and observations.

The unstandardized and standardized factor loadings for the six factors of the LDDI are presented in Table 2. All loadings were statistically significant ($p < 0.01$). According to Harrington (2009), loadings greater than .50 are considered fairly high. The Listening item loadings ranged between 0.39 and 0.76, the Speaking item loadings ranged between 0.50 and 0.79, and the Reading item loadings were between 0.58 and 0.78. In addition, the Writing and Mathematics item loadings ranged between 0.51 and 0.83 and between 0.57 and 0.86 respectively. Finally, eight items of the Reasoning scale had factor loadings greater the 0.8 and the remaining items loaded over 0.6.

Correlations among latent factors

Table 3 provides the correlations among the six latent factors of LDDI. All correlation coefficients were significant ($p < .000$), ranging from .53 (for Speaking and Mathematics) to .93 (for Reading and Writing). Also, the 95% confidence intervals for the correlation indicated that none of the six factors overlapped with the others (all correlations were less than one). Table 3 also reports the omega model-based reliability and Cronbach's alpha Coefficients. The first reliability indicator relies on latent factors' scores while the second one depends on raw scores. The omega coefficients for the six factors ranged from .71 for the Listening scale to .85 for the Reasoning scale, indicating that the six factors were internally consistent. Cronbach's alpha coefficients indicate high reliability of the LDDI scales.

Invariance across gender

Gender measurement invariance (M) test was examined across gender. We conducted a series of invariance tests beginning with M1 as the baseline model which allowed all parameters to be estimated freely. In M2, factor loadings were constrained to be equal across gender. All three models which were tested had a good fit

Table 4. Goodness of fit indices of invariance across gender.

Model	χ^2	DF	P	χ^2/DF	CFI	RMSEA	$\Delta\chi^2$	ΔDF	P*
M1: Unconstrained	23575.62	7754	0.00	3.04	0.91	.04	–	–	–
M2: Measurement weights	23679.12	7838	0.00	3.02	0.91	.04	103.50	84	0.07
M3: Measurement intercepts	24022.61	7928	0.00	3.03	0.91	.04	446.98	174	0.00
Independence model	187758.51	8010	0.00	23.44	0.00	.12	–	–	–

Note. $\Delta\chi^2$: Chi-square difference, ΔDF : degrees of freedom difference, p*: Significant of the difference. All the previous indicators computed by assuming unconstrained model to be correct.

to the data. Table 4 presents a summary of the goodness of fit indices for unconstrained and constrained models. The unconstrained model showed a good fit, with $\chi^2 = 23575.62$, df = 7754, χ^2/df = 3.04, CFI = 0.91, RMSEA = 0.04. In the measurement weights model, the measurement weights (i.e. regression coefficients or factor loadings) are constrained to be equal. The results showed that there was no significant difference between the baseline model (M1) and the measurement model (M2) ($\Delta\chi^2 = 103.50$, ΔDF = 84, $p = 0.07$). These results indicated that the factor loadings were invariant across gender. Such finding confirms that the items included in the model measured the respective factor with equal validity for both boys and girls.

In M3, item intercepts were additionally constrained to be equal. As shown in Table 4, the goodness of fit indices for M3 had a good fit to the data. However, the χ^2- differences between M3 model and the baseline model indicate that the item intercepts were not invariant across gender. The overall results of the study show that the measurement aspect of the structural equation model was invariant across gender, providing the minimum requirement for mean comparison across gender. Thus, mean comparison can be made between boys and girls particularly at the item level.

Discussion

The LDDI is a teacher rating scale that has been shown to be suitable for use in Western societies (e.g. Padeliadu and Sideridi 2008; Layton and Lock 2002; Layton and Lock 2001; Hammill and Bryant 1998). The present study, to our knowledge and based on searching the scientific databases-is the first study to test its use for cross-national comparisons. The study tested whether the LDDI was also suitable for the Sultanate of Oman, and potentially, other countries of the Gulf region of the Arab World as well as countries of the Middle East and North Africa region, by testing the factor structure of the instrument and the gender invariance on national Omani sample of 1546 students in Oman. The results indicated that the underlying six factor model of the LDDI fit the study's data fairly well. All six factors were also found to be internally consistent. The LDDI items loaded significantly on the factors that were originally reported by the authors of the instrument (Hammill and Bryant 1998). It is noteworthy to mention that the Arabic adapted translation of the LDDI could have an impact on the participants' responses as several items conveyed similar or associated meanings, which resulted in having an initial poor model fit with the data. The structure of the LDDI was replicated in Oman in the same way as in previous research studies which used the LDDI (Padeliadu and Sideridi 2008; Layton and Lock 2002; Layton and Lock 2001; Lock and Layton 2001a, 2001b; Hammill and Bryant 1998)

After examining the sources of poor model fit and using the modification fit indices, the results of the CFA indicated that the formation of the six-factor structure of the LDDI is supported by the data for the Omani sample. Three indices (CFI, TLI, and RMSEA) supported the fit of the model to the data within the established optimum range of fit. Additionally, all of the subscale factor loadings were moderate to very large and significantly different from zero. All the zero-order correlations were significant and also fell within a moderate to very large range (.53–.93). These data compare quite favorably to the correlation coefficients reported for the normative sample that ranged between .66 and .77, with a mean of .71 (Hammill and Bryant 1998).

These results indicate that the six factors of the LDDI (Listening, Speaking, Reading, Writing, Mathematics, and Reasoning) can form an effective model for the construct of SLD in students with LD or at risk for LD in key stages 1 and 2 in Oman, which is embraced by professionals working in educational and clinical settings. This six-factor model provides broader information about manifestation and symptoms of LD than other informal assessment used with children who have SLD, and is consistent with the description of the definition of SLD by the Joint Committee on Learning Disabilities (2001). Other research has reported effective assessment of SLD using the LDDI (Layton and Lock 2003; Layton and Lock 2001). These studies together with the current study make the LDDI a valuable and broad-based representation of SLD from which intervention and educational planning for children with SLD can be developed.

Limitations and future directions

The purpose of this study was to provide evidence on the construct validity of the Arabic version of the LDDI for use in screening for young children with LD in an Arab context. LDDI structure was validated on a

national, broad, geographically representative Omani sample. A number of limitations, however, suggest avenues for future research on the LDDI in Oman and the Arab region. As with all new assessment instruments, more research needs to be conducted on the psychometric properties of the LDDI in other Arabic contexts and cross cultural examination of its factor structure and invariance across cultures should be investigated. Additionally, future studies in the Arab region should target testing the predictive and criterion validity of the LDDI as well as investigating its sensitivity and specificity indexes. The LDDI should be used with other outcome measures including formative assessments such as curriculum based measurements, norm-referenced tests although these are not available in Oman, and criterion-referenced tests such as teacher made tests or tests developed and used by the Ministry of Education in Oman to assess the academic skills of students in grades 1–4. Such studies would provide strong evidence for its clinical and educational uses in Oman as well as in the Arab region. Furthermore, other LD rating scales should be validated and compared with the LDDI on similar Omani samples as well as on samples from other Arab countries. Examples include Performance Screening Guides (PSG; Elliott and Gresham 2007), and Learning Disability Evaluation Scales (LDES-R2; McCarney and Arthaud 2007). Also, a possible limitation of the current study is that we focused on a non-clinical sample. Future studies could include both clinical and non-clinical samples.

Some of the results do raise questions about the overall construct being measured by the LDDI. If the LDDI scales all measure some aspect of LD, they should be significantly intercorrelated, but not too highly. Correlations that are too high indicate that those scales do not contribute significant unique variance to the overall construct being measured, and thus are redundant with each other. Our results showed that nine of the intercorrelations among the six scales were greater than .70, indicating possible poor discriminant validity among the six scales.

That is, the LDDI provides valid measures of each construct; however, the measures of different constructs overlap substantially. Various problems and difficulties tend to co-occur (Willcutt and Pennington 2000); thus, it is not surprising that the LDDI six scales overlap. The implication of the relatively low level of factor distinctiveness is that the LDDI is not appropriate for individual diagnosis but can be used for referral and screening purposes. We believe that the use of the LDDI in future studies in the Arab region may provide further evidence on this finding. Additionally, the manifestations and symptoms assessed by the six scales may be associated frequently and therefore teachers were not able not able to distinguish among the items adequately. This finding replicates a general remark by Layton and Lock (2002) who argued that continuous training of teachers on the use of LDDI is necessary in order to increase its sensitivity to identifying children with academic challenges and/or SLD.

Clinical and educational implications

We have shown elsewhere in this study the need for an instrument like the LDDI in Oman, and the Arab world. It has been noted that, without accurate and reliable prevalence data on conditions such as LD, it is difficult to provide programs and services to meet individuals' needs (Al-Thani 2006). There have been calls to use survey instruments to screen for LD and to provide information about prevalence rates of students at risk for LD in schools in Oman (Al-Mahrezi *et al*. 2016).

The study has a number of implications for educational and clinical practitioners who could use the LDDI. First The LDDI can be a handy tool to screen for the LD manifestations and help in responding to teachers' academic concerns about students in Key stages 1 and 2 in Oman. One important use of the LDDI for Omani teachers is that can help them validate their daily observations regarding the academic performance of various students particularly struggling learners. Second, there is research evidence that the LDDI was used to facilitate advocacy and success in postsecondary settings (Lock and Layton 2002; Lock and Layton 2001a; Lock and Layton 2001b), reflecting that the LDDI can lead to improved outcomes in transition planning outcomes as well as in intervention outcomes. Therefore, future researchers and practitioners in Oman and the Arab region can similarly test and use the Arabic version for similar purposes.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by a grant from His Majesty's Research Trust Fund (HMRTF), The Sultanate of Oman, under grant number (SR/EDU/PSYC/12/01). The authors, therefore, acknowledge with thanks HMRTF technical and financial support.

ORCID

Mahmoud Mohamed Emam http://orcid.org/0000-0003-4451-9237

References

Akiba, M. and Liang, G. 2014. Teacher qualification and the achievement gap: A cross-national analysis of 50 countries. In: J.V Clark, ed. *Closing the achievement gap from an international perspective: Transforming STEM for effective education*. Dordrecht: Springer, pp. 21–40.

Alfarhan, U. F. and Dauletova, V. 2017. Revisiting the gender academic achievement gap: Evidence from a unique environment. *Gender and Education*, 1–22.

Al-Mahrezi, A., Al-Futaisi, A. and Al-Mamari, W. 2016. Learning disabilities: Opportunities and challenges in Oman. *Sultan Qaboos University Medical Journal*, 16, e129–e131.

Al-Thani, H. 2006. Disability in the Arab region: Current situation and prospects. *Journal of Adult Education and Development*, 3, 4–9.

Al-Yagon, M., Cavendish, W., Cornoldi, C., Fawcett, A. J., Grünke, M., Hung, L.-Y., Jiménez, J. E., Karande, S., van Kraayenoord, C. E., Lucangeli, D., Margalit, M., Montague, M., Sholapurwala, R., Sideridis, G., Tressoldi, P. E. and Vio, C. 2013. The proposed changes for DSM-5 for SLD and ADHD: International perspectives—Australia, Germany, Greece, India, Israel, Italy, Spain, Taiwan, United Kingdom, and United States. *Journal of Learning Disabilities*, 46, 58–72.

Author. 2016. Management of inclusive education in Oman: A framework for Action. *Support for Learning*, 31, 296–312.

Authors. 2015. Comorbidity of learning disorders and attention deficit hyperactivity disorder in a sample of Omani schoolchildren. *Sultan Qaboos University Medical Journal*, 15, 528–533.

Authors. 2015. Teachers' perceptions of the concomitance of emotional behavioural difficulties and learning disabilities in children referred for learning disabilities in Oman. *Emotional and Behavioural Difficulties*, 20, 302–316.

Begeny, J. C., Krouse, H. E., Brown, K. G. and Mann, C. M. 2011. Teacher judgments of students' reading abilities across a continuum of rating methods and achievement measures. *School Psychology Review*, 40, 23–38.

Blunch, N. 2012. *Introduction to structural equation modeling using IBM SPSS statistics and AMOS*. London: Sage Publications.

Bowen, N. K. and Guo, S. 2011. *Structural equation modeling*. Oxford: Oxford University Press.

Butterworth, B., Varma, S. and Laurillard, D. 2011. Dyscalculia: From brain to education. Science (New York, N.Y.), 332, 1049–1053.

Cavendish, W. 2013. Identification of learning disabilities: Implications of proposed DSM-5 criteria for school-based assessment. *Journal of Learning Disabilities*, 46, 52–57.

Colker, R. 2011. The learning disability mess. *Journal of Gender Society, Policy, and the Law*, 20, 81–106.

Dana, R. H. (Ed.). 2000. *Handbook of cross-cultural and multicultural personality assessment*. Hillsdale, NJ: Erlbaum.

Dean, V. J. and Burns, M. K. 2002. Inclusion of intrinsic processing difficulties in LD diagnostic models: A critical review. *Learning Disability Quarterly*, 25, 170–176.

DuPaul, G. J., Gormley, M. J. and Laracy, S. D. 2013. Comorbidity of LD and ADHD: Implications of DSM-5 for assessment and treatment. *Journal of Learning Disabilities*, 46, 43–51.

Eckert, T. L., Dunn, E. K., Codding, R. S., Begeny, J. C. and Kleinmann, A. E. 2006. Assessment of mathematics and reading performance: An examination of the correspondence between direct assessment of student performance and teacher report. *Psychology in the Schools*, 43, 247–265.

Elliott, S. N. and Gresham, F. M. 2007. *Social skills improvement system: Performance screening guides*. Bloomington, MN: Pearson Assessments.

Gerber, M. M. and Semmel, M. I. 1984. Teacher as imperfect test: Reconceptualizing the referral process. *Educational Psychologist*, 19, 137–148.

Gresham, F. M., MacMillan, D. L. and Bocian, K. M. 1997. Teachers as "tests": Differential validity of teacher judgments in identifying students at-risk for learning difficulties. *School Psychology Review*, 26, 47–60.

Gresham, F. M., Reschly, D. J. and Carey, M. 1987. Teachers as "tests": Classification accuracy and concurrent validation in the identification of learning disabled students. *School Psychology Review*, 16, 543–563.

Haager, D., Calhoon, M. B. and Linan-Thompson, S. 2007. English language learners and response to intervention: Introduction to special issue. *Learning Disability Quarterly*, 30, 151–153.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (Eds.). (2004). Adapting educational and psychological tests for cross-cultural assessment. Psychology Press.

Hammill, D. D. and Bryant, B. R. 1998. *Learning disabilities diagnostic inventory*. Austin, TX: Pro-Ed.

Harrington, D. 2009. *Confirmatory factor analysis*. Oxford: Oxford University Press.

Hulme, C. and Snowling, M. J. 2013. *Developmental disorders of language learning and cognition*. John Wiley & Sons.

Jackson, D. L., Gillaspy, J. A. and Purc-Stephenson, R. 2009. Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14, 6–23.

Jansen, P., Zayed, K. and Osmann, R. 2016. Gender differences in mental rotation in Oman and Germany. *Learning and Individual Differences*, 51, 284–290.

Jordan, N. C., Kaplan, D., Nabors Oláh, L. and Locuniak, M. N. 2006. Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, 77, 153–175.

Katchergin, O. 2016. The DSM and learning difficulties: Formulating a genealogy of the learning-disabled subject. *History of Psychiatry*, 27, 190–207.

Kaur, N. and Narkeesh, A. 2011. Effect of combined integrated learning programme (CILP) in children with learning disabilities. *Journal of Exercise Science and Physiotherapy*, 7, 50.

Kettler, R. J. 2007. Screening for early learning problems within an urban population: The brief academic competence evaluation screening system. *Journal of Urban Learning, Teaching, and Research*, 3, 47–58.

Kettler, R. J. and Albers, C. A. 2013. Predictive validity of curriculum-based measurement and teacher ratings of academic achievement. *Journal of School Psychology*, 51, 499–515.

Kettler, R. J. and Elliott, S. N. 2010. A brief broadband system for screening children at risk for academic difficulties and poor achievement test performance: Validity evidence and applications to practice. *Journal of Applied School Psychology*, 26, 282–307.

Kettler, R. J., Elliott, S. N. and Albers, C. A. 2008. Structured teacher ratings to identify students who need help: Validation of the brief academic competence evaluation screening system. *Journal of Psychoeducational Assessment*, 26, 260–273.

Lagae, L. 2008. Learning disabilities: Definitions, epidemiology, diagnosis, and intervention strategies. *Pediatric Clinics of North America*, 55, 1259–1268.

Layton, C. A. and Lock, R. H. 2002. Sensitizing teachers to English language learner evaluation procedures for students with learning disabilities. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 25, 362–367.

Layton, C. A. and Lock, R. H. 2003. The impact of reasoning weaknesses on the ability of postsecondary students with learning disabilities to select a college major. *NACADA Journal*, 23, 21–29.

Layton, C. and Lock, R. 2001. Determining learning disabilities in students with low vision. *Journal of Visual Impairment & Blindness (Blindness)*, 95, 288–299.

Liljequist, L. and Renk, K. 2007. The relationships among teachers' perceptions of student behaviour, teachers' characteristics, and ratings of students' emotional and behavioural problems. *Educational Psychology*, 27, 557–571.

Lock, R. H. and Layton, C. A. 2001. Succeeding in postsecondary education through self-advocacy. *Teaching Exceptional Children*, 34, 66–71.

Lock, R. H. and Layton, C. A. 2001b. The efficacy of the learning disabilities diagnostic inventory in postsecondary settings. *College and University*, 77, 3.

Lock, R. H. and Layton, C. A. 2002. Isolating intrinsic processing disorders from second language acquisition. *Bilingual Research Journal*, 26, 383–394.

Lock, R. H. and Layton, C. A. 2002. The efficacy of the learning disabilities diagnostic inventory in postsecondary settings. *College and University*, 77, 3.

Martin, M. O., Mullis, I. V., Foy, P. and Stanco, G. M. 2012. *TIMSS 2011 International results in science.* Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

Martin, S. D. and Shapiro, E. S. 2011. Examining the accuracy of teachers' judgments of DIBELS performance. *Psychology in the Schools*, 48, 343–356.

McCarney, S. B. and Arthaud . 2007. Learning disability evaluation scale-– Second Edition (LDES-R2). Available at: http://ebukumu.bugs3.com/pdf/What_is_a_Learning_Disability/LEARNING_DISABILITY_EVALUATION_SCALE__RENORMED_Second__/22_pdf

Mullis, I. V., Martin, M. O., Foy, P. and Hooper, M. 2016. *TIMSS 2015 International Results in Science*. Amsterdam, The Netherlands: International Association for the Evaluation of Educational Achievement.

National Joint Committee on Learning Disabilities. 2001b. Operationalizing the NJCLD definition of learning disabilities for ongoing assessment in schools. In: *Collective perspectives on issues affecting learning disabilities.* 2nd ed. Austin, TX: Pro-Ed, pp. 39–53. (Original work published 1997).

Padeliadu, S., and Sideridis, G. D. (2008). Learning disabilities screening for teachers. EPEAEK II Action, 1.

Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D. R., Samuelsson, S., Byrne, B., Willcutt, E. G. and Olson, R. K. 2012. Individual prediction of dyslexia by single versus multiple deficit models. *Journal of Abnormal Psychology*, 121, 212.

Polit, D. F., and Beck, C. T. (2008). Nursing research: Generating and assessing evidence for nursing practice. Philadelphia: Lippincott Williams & Wilkins.

Scanlon, D. 2013. Specific learning disability and its newest definition: Which is comprehensive? And which is insufficient?. *Journal of Learning Disabilities*, 46, 26–33.

Tenenbaum, H. R. and Ruck, M. D. 2007. Are teachers' expectations different for racial minority than for European American students? A meta-analysis. *Journal of Educational Psychology*, 99, 253–273.

Vaughn, S. and Fuchs, L. S. 2003. Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research and Practice*, 18, 137–146.

Weiner, I. B. and Craighead, W. E. 2010. *The Corsini encyclopedia of psychology*. New York: John Wiley & Sons, Inc.

Willcutt, E. G. and Pennington, B. F. 2000. Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry*, 41, 1039–1048.