

Testing EFL students with dyslexia: classroom approaches to inclusion with special reference to the Polish educational context

Testowanie uczniów z dysleksją uczących się języka angielskiego jako obcego: podejścia do inkluzji w klasie szkolnej ze szczególnym uwzględnieniem polskiego kontekstu edukacyjnego

Monika Łodej

UNIwersytet PEDAGOGICZNY IM. KOMISJI EDUKACJI NARODOWEJ W KRAKOWIE

Keywords

test, dyslexia, EFL, usability, inclusion

Słowa kluczowe

test, dysleksja, angielski jako język obcy, użyteczność, inkluzja

Abstract

To date the research on principles and procedures for testing EFL students has tended to focus on regular achievers rather than special educational needs students. This paper aims to fill this gap by analyzing procedures and practices applied to testing dyslexic learners of English. This is also the first study reporting on the usability of dyslexia friendly tests for Polish teachers of English. The results show that teachers extend the use of dyslexia friendly tests to a broader SEN population as well as non-SEN students. However, the usability of these tests is limited by incongruent testing policies for low- and high-stakes examinations in Poland.

Abstrakt

Dotychczasowe badania nad sposobami i procedurami testowania przyrostu wiedzy uczniów uczących się języka angielskiego jako obcego koncentrowały się na uczniach typowych raczej niż na uczniach o specjalnych potrzebach eduka-

cyjnych. Niniejsze badanie, poprzez analizę procedur i praktyk stosowanych przy testowaniu ucznia z dysleksją, stara się tę lukę zapłacić. Jest to pierwsze badanie podejmujące próbę analizy użyteczności testów przeznaczonych dla uczniów z dysleksją i ich wykorzystania przez polskich nauczycieli języka angielskiego. Analiza danych pokazuje, iż nauczyciele rozszerzają wykorzystanie takich testów o grupę uczniów ze specjalnymi potrzebami edukacyjnymi jak i o uczniów bez specjalnych potrzeb. Równocześnie, użyteczność tych testów jest ograniczana z powodu braku spójnych regulacji dotyczących testowania języka angielskiego uczniów z dysleksją podczas sprawdzianów śródrocznych oraz na egzaminach końcowych.

Testing EFL students with dyslexia: classroom approaches to inclusion with special reference to the Polish educational context

1. Introduction

The percent share of testing in the teaching and learning process has been increasing over the past several decades. A growing number of tests carried out in schools results from the presupposition that tests offer an instant and clear insight into students' learning. Both criterion- and norm-referenced tests are seen as reliable indicators of students' progress in learning as well as effectiveness of teachers' work. However, the type of tests and the purpose they serve differ not only between educational systems but also between local authorities and individual schools (The Glossary of Education Reform, 2015; U.S. Congress, 1992). An overview of testing practices (Salaky, 2018) shows that in the UK students by the time they turn 16 take about fifteen substantial, both low-stakes and high-stakes examinations. Similarly, in the US a student is required to take at least one standardized test per year. Whereas in Finland or in China students take one standardized test at the end of a high school. This elevates the value of the high-stakes test as the results of this assessment can have far-reaching future consequences for the test taker. Polish testing policies fit in this general trend as a student takes a considerable number of low- and high-stakes tests by the time they graduate.

There has been an ongoing educational debate on the value of standardized testing in schools. The proponents of standardized tests assert that such tests are valid and objective measurements of students' achievement. Tests of this type focus on the important skills and knowledge students have to master. They also claim that standardized testing helps schools to set appropriate educational targets (Fullan, 1991; Popham, 1999; see the entry of "Is the Use of Standardized Tests Improving Education in America?" in *Pros & Cons of Controversial Issues*, 2019). The opponents, on the other hand, raise concerns that relate to the fact that tests are highly discriminatory for students with special educational needs and that teaching to the test causes a decline in teaching higher order thinking skills as well as the problem that exam preparation consumes instruction time (Sizer, 1995; Hillocks, 2002; Dixon, 2018). Standardized tests are seen as inclusive and non-discriminatory measures for regular achievers, however, they tend to lose their inclusive power when used with special educational needs students (Gandal &

McGiffert, 2003; Koretz, 1995). Consequently, test formats usually fail to recognize the needs of SEN students which are often subject or academic specific. Fletcher et al. (2006: 138) observe that “many studies provide accommodations without considering the relation of the accommodation to the area of academic difficulty”. Test accommodations are generic in nature i.e. they rely on extended time, bigger font or an altered grading system (Wasburn-Moses, 2003; ADE, 2005). When SEN students are tested for their learning progress in the EFL classroom testing, validity and reliability are overlooked which, in turn, hampers application of inclusive testing (Pohl et al., 2016). In light of these findings, an insight into testing practices used by English teachers of special educational needs students may offer an insight into the types of approaches to inclusive testing and to the degree of their usability.

2. Testing in education

The history of testing traces its roots to ancient China. During the Han Dynasty (206 BC – 220 AD), first exams were introduced to select people who were intellectually and ethically superior to others and were deemed most suitable to work as civil servants. The tests relied mainly on memorization of Confucian texts and evaluated knowledge of literary styles (see the entry of “Chinese examination system” in Encyclopaedia Britannica). Interestingly, this testing approach emphasized rote learning rather than validated practical skills (Biggs, 2001). It was believed that “those who excel in their study should become officials” (Analects Book XIX in Zeng, 1999: 21). It was not until the early 19th century when standardized tests were introduced in Europe by the British.

Similarly to the Chinese format, first tests run in England were aimed at selecting candidates for governmental jobs. Eventually, in the 1850s standardized tests made its way to British universities and schools. The first examinations were introduced in 1858 by the University of Cambridge Local Examinations Syndicate – now known as Cambridge Assessment (Cambridge Assessment, 2019). Their creation responded to educational needs expressed by schools to help them mark their pupils’ attainment. In the 1850s educational tests evaluated solely foundational knowledge therefore students were expected to learn specific information by heart. Rote learning was favoured as the test format expected students to draw from memory the names of entities such as kings or queens, mountains, lakes or rivers (cf. Łodej, 2017a). Andrew Watts, acting Director of the Cambridge Assessment Network observed that

In the Syndicate's early History, Geography, Science and Scripture papers, question after question asks for the recall of facts, often from set books or periods of history, or sections of the Bible... This reflects a view of the educated person as being a collector of knowledge (in Weir et al., 2013: 198).

Soon, feedback obtained from the test reports revealed that students neither had any opportunity to demonstrate their understanding of the facts nor apply knowledge to practice. The concept of standardized testing spread quickly as it recognized a person's competence and did not favour social standing. From Britain standardized tests travelled to mainland Europe and America. It is to be noticed that the foundations for modern day testing were mainly laid by experimental works in psychology (Gallagher, 2003; Gregory, 2004). Wilhelm Wundt, James McKeen Cattell, Alfred Binet and Theodore Simon and Frank Parsons were among the pioneers in experimental testing.

Wilhelm Wundt, a German psychologist, contributed to the development of testing procedures with experimental studies that measured mental processes. He used the so called 'thought meter' to trace attention shifts. He observed that it took 1/10 of a second to shift a person's attention from the sound of a bell to the position of the pendulum. On the basis of this experiment he concluded that this relates to the shift time from one thought to another (Hergenhahn, 2005). The works of Wundt inspired James McKeen Cattell, an American student, who decided to travel overseas to Leipzig to complete his doctorate under Wundt. On the completion of his studies, McKeen Cattell returned to the University of Pennsylvania and established his laboratory where he worked on various mental tests and measurements. A 10 test battery designed by Cattell measured, among others, strength of a hand grip, naming time for colours, reaction time for sound, or recall time for number and letters. To account for the utilitarian use of the tests Cattell wrote in 1893 that

In conjunction with the ordinary school examination such tests would show whether the course of study is improving or blunting the fundamental processes of perception and mental life. Tests made at the beginning and end of the day, week, and session would show whether the student is exhausted by the required curriculum. They could be used in comparing different systems of education, different schools... They would show whether girls are able, without injury to health, to follow the same course of study as boys. (Benjamin, 2019: 63)

Also Alfred Binet and Theodore Simon, French scholars, contributed to mental testing in 1905. They designed a 30-item scale with an increasing level of difficulty that aimed to evaluate school children in Paris schools. Children were asked to answer questions to the point where they knew the answers.

The point at which they were unable to provide an answer determined their mental stage (Klee & Moore, 2014). The intelligence test was individually administered and used to identify learning deficits in “slow children who would not profit significantly from schooling” (Walsh & Betz, 1995 in Gallagher, 2003: 86). Around the same time in Boston, Massachusetts, Frank Parsons devised a first scientific counseling instrument to support people in their vocational choices. This method involved seven steps: personal data, self-analysis, the person’s own choice and decision, counselor’s analysis, outlook on the vocational field, induction and advice, general helpfulness in fitting in the chosen work (Gysbers et al., 2014). The purpose of the scientifically based counseling was “to aid young people in choosing an occupation, preparing themselves for it, finding an opening in it, and building up a career of efficiency and success” (Parsons, 1908 in Jones 1994: 288).

3. Testing and assessment in foreign languages

Carroll (1954) notes that from the beginning of examinations in foreign languages both ancient and modern test formats included exercises in translation and composition. These comprised of literal translation of texts and simple questions about grammar. As early as in 1916 Hanus who worked on the measures that would test learning progress in Latin, based his approach on three language elements that is vocabulary, translation and grammar. The selection of these three areas allowed “to measure the growth of power in these three elements of Latin [which were] assumed to be fundamental” (Hanus, 1916: 342). Along with the development of experimental testing in psychology, outlined above, and the search for more reliable and objective measures language tests themselves started to take on a more standardized approach. In the 1920’s Handschin (1920: 217) observed that “we have reached a point in our modern language teaching where we feel the need of standard objective scales by which we measure the product of our teaching”. Consequently, he identified four language skills which should be tested i.e. hearing, speaking, seeing (reading) and writing where hearing tested listening and speaking oral skills. The next step was to propose eight principles for construction of a foreign language test. The test he created was the first standardized modern language test verified with results obtained from French and Spanish.

The first two major large-scale studies which aimed at measuring objectivity of foreign language tests were conducted by Wood and Henmon in the late 1920s. Wood’s study was requested by the Board of Regents of the State of New York and aimed at validation of paper and pencil tests that measured

vocabulary, grammar and reading comprehension in Spanish and French. The tests were administered to high school students in New York. Each of the test questions was individually analyzed with regard to its validity and difficulty. The tests in French and Spanish were identical to allow for easy comparison of the results. The tests consisted of three parts. In the first part which was vocabulary oriented students were presented with 100 target language words. Each of the words was followed by five English words. A student's task was to choose the correct English equivalent for the tested word. In part two of the test, which dealt with reading comprehension, students were presented with sixty incomplete statements in the foreign language, and asked to complete each question with one out of five alternative endings. Whereas part three of the test aimed at checking knowledge of grammar. Students were offered sixty English sentences, each followed by an incomplete translation in the target language. A student's task was to complete the translation. No language options were given here to choose from. This format made use of a five item multiple choice testing technique and was considered time-saving and proficiency-based. This format evaluated what students learned in the classroom and outside the school context (Barnwell, 1993).

Unlike Wood's, Henmon's study was based on testing language skills acquired in Latin and French i.e. a classical and a modern language. The study relied on tests of vocabulary and sentence translation (Henmon, 1921 in Morton-Finney, 1941). These tests consisted of sixty French words and twelve sentences arranged in the order of increasing difficulty. Students were expected to translate contextless French words and phrases to English. As observed by Barnwell (1993: 17) "the parallel between Henmon's Latin test and his French test is significant, since it shows that no specific methodology for testing the vernacular aspects of the modern languages had yet been devised". Despite the fact that Henmon's work did not contribute largely to the structure of modern language testing his study was quite innovative and represented the then-current approach to testing.

In contemporary times, examination and testing practices in modern foreign languages in Europe are closely associated with the language policies of the EU and the Council of Europe. These regulations are inscribed in the Common European Framework and the European Language Portfolio (EALTA, 2019). There is a number of international institutions which provide standards, guidelines and code of practices for language testing. ILTA guidelines for good testing practice are set by the International Language Testing Association (2007), ALTE code of practice is issued by the Association of Language Testers in Europe (1994), ALTE principles of good practice are established by the Association of Language Testers in Europe (2001), ETS

standards for quality and fairness are developed by the Educational Testing Service (2014) and EALTA guidelines for good practice were formulated by the European Association for Language Testing and Assessment (2006). Interestingly, one of the ILTA guidelines, which refers to the responsibilities of test designers and test writers, advises that “Care must be taken to ensure that all test takers are treated in the same way in the administration of the test” (ILTA, 2007). This requirement contradicts the statement issued by ALTE (2007) which reads that “The examination system provides support for candidates with special needs”. The latter statement clearly suggests implementing an inclusive approach to testing. It recognizes the role of administrative support required by special educational need students. The support that, in the case of dyslexic learners, is ensured by relevant test accommodations and modifications. Such changes to language tests usually involve presentation, response, timing, and scheduling accommodations (Steinfeld, 2010: 2; NDA, 2012; ADE, 2005: 3-4).

4. Purpose of the study

Dyslexia is one of the most common language dysfunctions which results in difficulties in language learning. Such difficulties have a strong tendency to surface in foreign language learning. According to statistics, dyslexia affects one in five children, with prevalence of 5% to 10% in clinical research and 17,5% in school-identified population (Shaywitz & Shaywitz, 2003). It is estimated that in Poland symptoms of dyslexia are observed in 10-16% of school age students (Bogdanowicz & Adryjanek, 2005). Therefore, the problem of testing dyslexic students in the EFL context has been recently gaining attention from national examination boards, EFL course book publishers and regular classroom teachers.

The Polish educational scene seems not to be free of the urgent need for establishing explicit standard norms for testing language skills in dyslexic students. Voices of concern are heard from classroom teachers who request clear directives on how to test students with dyslexia. At present, there are three inconsistent regulations that govern practices of testing dyslexic students. They are included in the Ordinance of the Minister of Education for Poland of 9 August 2017 on organizing educational support for SEN students (Rozporządzenie Ministra Edukacji Narodowej, 2017), the directives of the Polish National Examination Board (CKE, 2019) and test modifications designed by various EFL book authors. It is to be observed that these sets of regulations and guidelines follow different approaches to classroom inclusion. The results of this legal environment can be seen in a situation

when a class teacher following the Ordinance of the Minister of Education for Poland provides their dyslexic students with altered low-stakes English language tests with a full awareness that the format of high-stakes English test is the same for dyslexic and non-dyslexic students. On a deeper level, the altered tests are either provided by the EFL book publisher if the book teachers use includes dyslexia friendly versioning and if not teachers may alter the format of the tests intuitively. At the same time they bear in mind that the format of high-stakes English tests is the same for all learners and not accommodated. The current study looks at the testing practices of EFL teachers in Poland. In specific, the study aims to outline teachers' choices and decisions regarding the scope and purpose of using accommodated EFL classroom tests with dyslexic students.

5. Method

5.1 Study design and procedure

The study sought to validate the usability of dyslexia friendly tests which are available for use to teachers who teach with the *English Class* EFL course book. Two aspects of the usability were put under scrutiny i.e. frequency and degree of use. In other words, the study aimed to find out a) if the EFL primary school teachers use dyslexia friendly tests and b) how they use them. In order to meet this goal, the study design employed an online survey method. Both qualitative and quantitative methods were used in this investigation. The survey was targeted at EFL primary school teachers in Poland who use accommodated tests for dyslexic students designed by the author of this paper and published by Pearson. An online disseminated questionnaire which included two open-ended questions required respondents to provide information on a) whether they use the *English Class* dyslexia versioning tests (Łodej, 2017b, 2018) in their teaching practice exclusively with dyslexic students or with whole classes and b) what is their opinion of the accommodated tests.

The respondents posted their comments in response to the Facebook disseminated post. The survey post was published in eighteen closed FB groups created for English teachers in Poland. These groups included *Nauczyciele angielskiego*, *Nauczyciele języków obcych ze Śląska*, *Dysleksja a języki obce*, *Light my fire-nauczanie języków obcych dla młodzieży*, *Lektorzy i nauczyciele języków obcych*, *Anglomaniacs-angliści z pasją*, *Metodyka Nauczania języków obcych u Teresy Perkołub*, *Nauczyciele języka angielskiego u Teresy Perkołub*, *Nauczanie języka angielskiego-pomysły na kreatywne i dynamiczne lekcje*, *Szkolenia, kursy, konferencje dla nauczycieli języka angielskiego*, *Nauczyciele/*

Lektorzy z Małopolski, Nowoczesne lekcje, Lektorzy i nauczyciele języków obcych, Nauczyciele, Nauczyciele angielskiego w Lublinie, Teachers help Teachers, Nauczyciele angielskiego Kielce, and Teaching English as a FL to students with Special Educational Needs. The post was published between 8-16 May 2019 and all the comments were collected in the ensuing days. To ensure objectivity, the teachers were not informed of the fact that they commented on the series of dyslexia friendly test books authored by the researcher. The survey was distributed among teachers who were native speakers of Polish. As a consequence, the questions of the survey were posted in Polish as well. The responses provided by the respondents were also in Polish (see Appendix for a complete list of the teachers' commentaries). For the purposes of the present discussion the responses were translated into English.

5.2 Results and discussion

By the end of the survey period, data had been collected from 19 individuals. The teachers' comments on the usability of dyslexia friendly versions of EFL tests were scrutinized qualitatively and quantitatively. As presented in Table 1, the information obtained from the qualitative analysis of the teachers' commentaries on the use of the dyslexia friendly versions of tests allowed for a categorization of these statements into four strands a) use of the tests is restricted to dyslexic students b) use of the tests is extended to other SEN students, c) use of the tests is extended to non-SEN groups, d) a teacher restrains from using these tests.

Type of usability	Frequency of usability	Degree of usability
a) use of the tests is restricted to dyslexic students	5	
b) use of the tests is extended to other SEN students	5	Asperger syndrome, mild mental retardation, SEN students, students with other learning needs, students on the lower end of intellectual capacity
c) use of the tests is extended to non-SEN groups	6	less proficient students, the whole class, students who work slower
d) a teacher restrains from using these tests	3	not congruent with the high-stakes examination format

Table 1. Type, frequency and degree of usability of dyslexia friendly tests by EFL teachers

The teachers who constitute the first group revealed that their use of dyslexia friendly tests was restricted to dyslexic students only (see comments 3, 4, 11, 12 and 17).

(3) **I also use them!** [dyslexia friendly]. They're fantastic! Incredibly motivate students to study. Because of their perfect layout **my students do these tests without much trouble.**

(4) I like these new Pearson tests which accompany their new course book. **I teach a Year 4 student. I gave him two tests. A non-accommodated one and dyslexia friendly. He cried working on the first one. He did the other test with two mistakes.**

(11) **I use these tests only for dyslexic students.** In my opinion these tests are too easy for students without a dyslexia statement. I use them with students who re-sit exams. For the majority of students, especially the gifted ones, they are definitely not demanding enough.

(12) **Only for dyslexics. The students are happy so I am satisfied too.** I follow the recommendations from the Dyslexia Diagnostic Center. **The students get better grades, fewer exercises means more time for thinking...**

(17) **I teach Year 4, sometimes I use dyslexia friendly tests,** otherwise they [dyslexic students] are unable to understand test instructions.

The teachers cited above use dyslexia friendly textbooks exclusively with dyslexic students. In addition, the respondents provide information on the impact these decisions have on students, as well as what motivates teachers' decisions of how and when to use these tests. An observation is made that dyslexic students perform significantly better on the dyslexia versioning. The students are reported to receive better grades, show gains in satisfaction with testing experience, perform on these tests with more ease, and display signs of decreased tension and stress. Most importantly, the test formats enable dyslexic students to access the content of the test. However, the respondents point to contradicting SEN support requirements at low- and high-stakes examinations as the major reason for the limited use of the accommodated tests.

The collected data have shown that the extension of the use of dyslexia friendly tests to other SEN students constitutes a separate type of test usability which can be seen in comments 2, 15, 16, 18 and 19.

(2) With little alteration they work well with **a student with Asperger syndrome and mild mental retardation.**

(15) I use them for dyslexics and for **students with other learning needs**.

(16) I use these tests **for students with SEN statements and SEN certificates**. I know which of my students are unable to do non-accommodated tests. I don't want to discourage them.

(18) I use them as an accommodation for students with dyslexia but for other **students with learning issues** as well.

(19) I use them also for less advanced students. I mean, **students who are on the lower end of intellectual capacity**. They do well on such a test and it does not discourage them from learning.

Interestingly, the special educational needs group comprises of students with both SEN statements and certificates. This category of test takers includes students with Asperger syndrome, mild mental retardation, lower intellectual ability, and other learning needs not further specified. It is to be emphasized that the dyslexia friendly tests were targeted at dyslexic students who hold a SEN statement. Therefore, the tests were accommodated not modified. That means that the changes were made to how students were tested but not to what was tested. Whereas, SEN certificates require test modifications.

The next group of respondents reported on the administration of the dyslexia friendly tests to non-SEN students (see comments 1, 5, 6, 7, 8 and 14).

(1) I use the accommodated tests for students with dyslexia but **they also work well with less proficient students**. They are graphically clear and I think they will be suitable for all students.

(5) Because I teach a **less proficient group, I usually download two versions of the test and decide which exercises to use**. In result, I create two versions of the test: one for students with special educational needs which consists of the exercises from the dyslexia friendly test book and exercises from the original version of the test which I accommodate myself **and one version of the test for the remaining students** which consists of several exercises from the accommodated test and several from the original, basic version.

(6) I **give the same test to the whole class** because otherwise non dyslexic students ask why dyslexic students have a different or easier test. **All the students get the same version of dyslexia friendly test and everybody is content with such a solution**.

(7) I have been using these tests since the last school year and **they work really well with dyslexic and dysgraphic students** (less writing, more matching and copying) **as well as with the students who are less proficient**

or work slower because they test the same content in a form that would allow them to do the test within the assigned time. In addition, these tests give them a sense of achievement in learning the language.

(8) **In Year 4 I used dyslexia friendly tests for the whole class just for a good start** so as to help them get used to being tested. But later I gave them non-accommodated tests. This worked for me.

(14) My sister's teacher sometimes **uses them with the whole class even though there are no students with dyslexia statements.**

In the comments above, the respondents reported on the use of the test with less advanced EFL learners and those students who work at a slower pace. Also, the teachers used the dyslexia friendly tests for the whole class on a number of specified occasions. They motivated their decisions with the fact that the dyslexia versioning has the same content as the generic version of the test. They made an observation that it is only the format that is the distinguishing factor between the two versions of the tests. Additionally, a need for testing fairness was stressed as one of the reasons which could lead to an extended use of the dyslexia friendly tests in the classroom.

The final group of respondents restrained themselves from using dyslexia friendly tests with dyslexic and non dyslexic students (see comments 9, 10, and 13).

(9) **I use a non-accommodated version.** There is no mercy on external exams.

(10) **I give non-accommodated tests as well.** I allow more time to write a test. **The same during external exams.**

(13) **I used them in Year 4 for dyslexic students. I don't use them in Year 5** in the first semester and I did not lower their grades if they made mistakes – but they had to correct the mistakes in their notebooks. Instead, I always print a Word version of the test, which I accommodate visually. I print a maximum of two exercises on one page. But, in the second semester, I already lowered their grades if they made mistakes. **On external exams dyslexic students are just offered extended testing time.** I talked with the parents about the need for additional study time at home, [students] are expected to prepare crosswords or word cards on their own. They have to memorize how the word is written.

Likewise, these responses stress the issue of inconsistent testing procedures at low- and high-stakes tests. Consequently, teachers are made to prioritize and decide whether they are working towards the low or high-stakes exams. This, in turn, determines their testing choices.

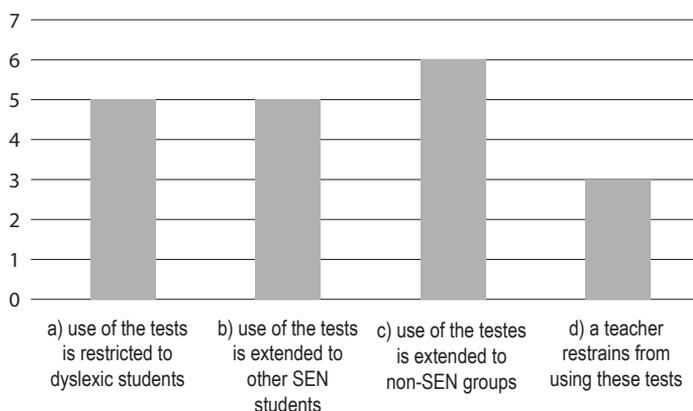


Figure 1. Type and frequency of usability of dyslexia friendly tests by EFL teachers

A quick quantitative calculation of the data obtained from nineteen Polish EFL teachers (see Figure 1) shows that the number of teachers ($N=16$) who use the dyslexia friendly tests and those who refrain from using them ($N=3$) represent respectively 84% and 16% of the surveyed sample. Within the group of sixteen test users five teachers use the tests solely with dyslexic students and eleven teachers extend their use to other students. In the same vein, out of eleven teachers who extend the use of dyslexia friendly tests five extend their use to the SEN population while six teachers use these tests with non-SEN students.

6. Conclusions

The present study reveals a pressing need for the development of a consistent set of testing principles, guidelines and practices for dyslexic EFL students. Currently, it is not clear which regulations should guide teachers' assessment practices. In the Polish context, it may be assumed that the regulations that are outlined in the Ordinance of the Minister of Education for Poland should determine the actions taken by the Polish National Examination Board and consequently set standards for book publishers. Some reassuring cohesion between testing practices can be observed between the regulations set by the Minister of Education for Poland and testing practices used by book publishers. However, this is not seen in the testing policy developed by the Polish National Examination Board.

The lack of consistency at the national level directly affects schools and classroom practice. Due to the absence of clear and definite testing standards for low- and high-stakes English language tests for SEN students independent

policies and decisions are made by school teachers. The results of the study show that these decisions tend to be intuitive rather than research-driven. The single most striking observation to emerge from the teachers' commentaries was that the number of teachers who extend the use of dyslexia friendly tests to the non-SEN population is larger than the number of teachers who use them only with SEN students. This, in turn, leads to an observation that teachers intuitively apply full testing inclusion to their classrooms.

The evidence from this discussion suggests that more investigation is required in the area of testing practices of students with dyslexia who learn English as a foreign language. In particular, an examination of testing approaches and formats that would meet the needs of both special educational needs students and regular achievers. Appropriate test formats would be required to follow the universal design principles.

7. Limitations of the study

The aim of the study was to assess frequency and use of dyslexia friendly tests in the EFL classroom. Although the study used a focused survey, the findings in this report are subject to two limitations that relate to the form and time of the collection of data. These factors had an impact on the number of the responses to the post questions which was significantly lower than expected.

First, the present methodology applied an on-line survey disseminated via Facebook. Despite the fact that there were between 1.000 to 9.000 members in each group the post was shared to the actual reach might have been smaller. Members of Facebook groups often belong to several groups which attract practitioners of a particular profession.

Secondly, the post was distributed in May which is the time when similar requests to complete on-line questionnaires are frequently published. The members of the groups, who are mainly in-service teachers, work towards their additional teaching qualifications and collect their research data. The number of requests, in turn, makes Facebook users less sensitive to such posts. In spite of the fact that, the questionnaire request was published not in the form of a separate link to an on-line questionnaire but as a post with two questions which displayed in the groups automatically, this did not contribute to an increase in the number of responses. It is expected that these points will help future researchers to avoid facing the same shortcomings.

References

- ADE. Arkansas Department of Education. (2005). Guidelines for Assessment Accommodations For Students with Disabilities. Special Education Unit Arkansas Department of Education & Academic Standards and Assessment Unit. Retrieved from [https://arksped.k12.ar.us/documents/specialprojects/Guidelines%20for%20Assessm ent%20Accommodations.pdf](https://arksped.k12.ar.us/documents/specialprojects/Guidelines%20for%20Assessm%20ent%20Accommodations.pdf)
- ALTE. Association of Language Testers in Europe. (1994). *ALTE code of practice*. Retrieved from: https://www.alte.org/resources/Documents/code_practice_en.pdf
- ALTE. Association of Language Testers in Europe. (2001). ALTE principles of good practice for ALTE examinations. Retrieved from https://www.testdaf.de/fileadmin/Redakteur/PDF/TestDaF/ALTE/ALTE_good_practice.pdf
- ALTE. Association of Language Testers in Europe. (2007). Minimum standards for establishing quality profiles in ALTE examinations. Retrieved from https://www.alte.org/resources/Documents/minimum_standards_en.pdf
- Barnwell, D. P. (1993). *Problems of Articulation and Testing: Lessons from the 1920s*. Paper presented at a Conference on Research Problems in Adult Language Learning (Ohio State University, Columbus, OH. Washington, D.C.). Distributed by ERIC Clearinghouse pp. 1-44. Retrieved from [https://eric.ed.gov/?q=%22%22&ff1=sub Standardized+Tests&ff2=pub+Historical+Materials&id=ED365108](https://eric.ed.gov/?q=%22%22&ff1=sub+Standardized+Tests&ff2=pub+Historical+Materials&id=ED365108) [Accessed 15th August 2019]
- Benjamin, L. T. (2019). *A brief history of modern psychology*. Hoboken, NJ : John Wiley & Sons, Inc.
- Biggs, J. (2001). Assessment of student learning: Where did we go wrong? *Assessment Update*, 13(6), 6-9.
- Bogdanowicz, M., & Adryjanek, A. (2005). *Uczeń z dysleksją w szkole: poradnik nie tylko dla polonistów*. [Dyslexic students at school. A guideline not only for teachers]. Gdynia: Operon.
- Cambridge Assessment. (2019). How have schools exams changed over the past 150 years? Retrived from <https://www.cambridgeassessment.org.uk/news/how-have-school-exams-changed-over-the-past-150-years/>
- Carroll J. B. (1954). *Notes on the measurement of achievement in foreign languages*. Retrieved from <http://www.ealta.eu.org/documents/resources/J.B.%20Carroll%20Notes%201954.pdf> [Accessed 3rd August 2019].
- Chinese examination system. (2019). In Encyclopaedia Britannica. Retrieved July 20, 2019, from <https://www.britannica.com/topic/Chinese-examination-system>
- CKE. Centralna Komisja Egzaminacyjna. (2018). *Arkusze*. [National Examination Board. Tests]. Retrieved from <https://cke.gov.pl/egzamin-osmoklasisty/arkusze/>
- Dixon, B. (2018). The Testing Emperor Finally Has No Clothes. *Education Digest*, 83(9), 45-50.
- EALTA. European Association for Language Testing and Assessment. (2006). EALTA guidelines for good practice by European Association for Language Testing

- and Assessment. Retrieved from <http://www.ealta.eu.org/documents/archive/guidelines/English.pdf>
- EALTA. European Association for Language Testing and Assessment. (2019). *Rationale*. Retrieved from <http://www.ealta.eu.org/index.htm>
- ETS. Educational Testing Services. (2014). ETS standards for quality and fairness by Educational Testing Services. Retrieved from <https://www.ets.org/s/about/pdf/standards.pdf>
- Fletcher, J. M., Francis, D. J., Boudousquie, A., Copeland, K., Young, V., Kalinowski, S., & Vaughn, S. (2006). Effects of Accommodations on High-Stakes Testing for Students With Reading Disabilities. *Exceptional Children*, 72(2), 136-150.
- Fullan, M. G. (1991). *The new meaning of educational change*. New York: Teachers College Press.
- Gandal, M., & McGiffert, L. (2003). The power of testing. *Educational Leadership*, 60(5), 39-42.
- Gallagher, C. J. (2003). Reconciling a Tradition of Testing with a New Learning Paradigm. *Educational Psychology Review*, 15(1), 83-99.
- Gregory, R. J. (2004). *Psychological testing: History, principles, and applications*. Needham Heights, MA, US: Allyn & Bacon.
- Gysbers, N. C., Heppner, M. J., & Johnston, J. A. (2014). *Career counseling: Holism, diversity, and strengths*. Alexandria, VA: American Counseling Association.
- Handschin, C. H. (1920). Tests and Measurements in Modern Language Work. *The Modern Language Journal*, 4(5), 217-225.
- Hanus, P. H. (1916). Measuring Progress in Learning Latin. *The School Review*, 24(5), 342-351. https://www.jstor.org/stable/1077913?seq=1#metadata_info_tab_contents
- Hergenhahn, B. R. (2005). *An introduction to the history of psychology*. Australia: Thomson/Wadsworth.
- Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College, Columbia University.
- ILTA. International Language Testing Association. (2007). *Guidelines for practice*. Retrieved from https://cdn.ymaws.com/www.iltaonline.com/resource/resmgr/docs/ilta_guidelines.pdf
- Is the Use of Standardized Tests Improving Education in America? (2019). Pros & Cons of Controversial Issues. Retrieved from <https://standardizedtests.procon.org/> [Accessed 1st July 2019].
- Klee, A., & Moore, B. A. (2014). *EPPP fundamentals: Review for the examination for professional practice in psychology*. New York, NY : Springer Pub. Company.
- Koretz, D. M. (1995). Sometimes a cigar is only a cigar, and often a test is only a test. In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessment?* (pp. 154-166). Washington, DC: Brookings Institute.

- Jones, K. J. (1994). Frank Parsons' contribution to career counseling. *Journal of Career Development*, 20(4), 287-294.
- Łodej, M. (2017a). Overlearning i algorytmy w nauczaniu leksyki języka angielskiego [Overlearning and algorithms in teaching English lexis]. *Języki Obce w Szkole*, 3, 20-26.
- Łodej, M. (2017b). *English Class. Level 1 Test Book. Dyslexia versioning*. Warszawa: Pearson Central Europe.
- Łodej, M. (2018). *English Class. Level 2 Test Book. Dyslexia versioning*. Warszawa: Pearson Central Europe.
- Morton-Finney, P. R., (1941). The Transfer of Training from Latin to French and Spanish. Graduate Thesis Collection. 377. <https://digitalcommons.butler.edu/grtheses/377>
- NDA. National Disability Authority. (2014). *The seven principles*. Dublin: Center for Excellence in Universal Design. Retrieved from <http://universaldesign.ie/What-is-Universal-Design/The-7-Principles/>
- Pohl, S., Südkamp, A., Hardt, K., Carstensen, C. H., & Weinert, S. (2016). Testing Students with Special Educational Needs in Large-Scale Assessments – Psychometric Properties of Test Scores and Associations with Test Taking Behavior. *Frontiers in Psychology*, 7(154), 1-14.
- Popham W. J. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, 58, 8-16.
- Rozporządzenie Ministra Edukacji Narodowej z dnia 9 sierpnia 2017 r. w sprawie zasad organizacji i udzielania pomocy psychologiczno-pedagogicznej w publicznych przedszkolach, szkołach i placówkach. Dziennik Ustaw poz. 1591 tom 1 [Journal of Laws of the Republic of Poland No. 1592, Vol. 1]. Retrieved from <http://dziennikustaw.gov.pl/du/2017/1591/1>
- Salaky, K. (2018). What standardized tests look like in 10 places around the world. *Insider*. Retrieved from <https://www.insider.com/standardized-tests-around-the-world-2018-9#canada-has-standardized-tests-that-can-vary-by-province-9>
- Shaywitz, S. E., & Shaywitz, B. A. (2003). The science of reading and dyslexia. *Journal of Aapos*, 7(3), 158-166.
- Sizer, T. R. (1995). Will national standards and assessments make a difference? In D. Ravitch (Ed.), *Debating the future of American education: Do we need national standards and assessment?* (pp.33-39). Washington, DC: Brookings Institute.
- Steinfeld, E. (2010). Advancing Universal Design. In J. L. Maisel, (Ed.), *The state of the science in universal design: Emerging research and developments* (pp. 1-19). Sharjah, U.A.E.: Bentham Science Publishers.
- The Glossary of Education Reform. (2015). High stakes test. Retrieved from <https://www.edglossary.org/high-stakes-testing/>
- U.S. Congress, Office of Technology Assessment. (2012). Testing in American Schools: Asking the Right Questions, OTA-SET-519 Washington, DC: U.S. Government Printing Office, February 1992. Retrieved from http://govinfo.library.unt.edu/ota/Ota_1/DATA/1992/9236.PDF

- Wasburn-Moses, L. (2003). What Every Special Educator Should Know About High-Stakes Testing. *Teaching Exceptional Children*, 35(4), 12-15.
- Weir, C. J., Vidakovic, I., & Dimitrova-Galaczi, E. (2013). *Measured constructs: a history of Cambridge English language examinations*. Cambridge: Cambridge University Press.
- Zeng, K. (1999). *Dragon Gate: Competitive Examinations and Their Consequences*. London: Cassell.

Appendix

The list of respondents' comments

The original survey post:

English Class Users! Mam prośbę do osób, które wykorzystują testy dla uczniów z dysleksją do podręcznika English Class. Czy używacie tych testów tylko do testowania uczniów z dysleksją czy też całej klasy? Jak Wam się z nimi pracuje?

(1) Używam głównie dla uczniów z dysleksją, ale świetnie sprawdzają się też w słabszych grupach. Testy są przejrzyste i myślę, że będą dobre dla wszystkich uczniów.

(2) Po drobnej przeróbce sprawdza się dla dziecka z aspergerem i z upośledzeniem lekkim.

(3) Też z nich korzystam! Są rewelacyjne. Ogromnie motywują uczniów do działania, ponieważ dzięki doskonałemu rozplanowaniu testu, radzą sobie bez problemu z rozwiązywaniem zadań!

(4) Podoba mi się testy Pearsona do nowych podręczników. Pracuję z dzieckiem z 4 klasy. Dałam mu dwa testy zwykły i dla dyslektyków. Przy pierwszym się popłakał. Drugi rozwiązał robiąc dwa błędy.

(5) Jako że uczę słabszą grupę, zazwyczaj pobieram obie wersje testu i sama decyduję, które zadania umieścić na sprawdzianie. W efekcie tworzę dwie wersje: jedną dla uczniów ze specjalnymi potrzebami (gdzie są zadania z dostosowanego testu oraz odpowiednio dostosowane przeze mnie zadania z wersji podstawowej) oraz jedną wersję dla pozostałych dzieci (zawierającą trochę zadań z dostosowanego testu i trochę z wersji podstawowej).

(6) Ja daję wszystkim ten sam bo potem reszta pyta czemu tamci piszą inny albo łatwiejszy. Wszyscy mają wersję dla dyslektyków i jest spokój.

(7) Korzystam od zeszłego roku, bardzo dobrze sprawdzają się u dyslektyków, dysgrafików (mniej pisania własnego, więcej łączenia i przepisywania) i uczniów słabszych i wolniej pracujących, bo nadal sprawdzamy ten sam zakres materiału, ale w formie, która pozwoli im wykonać zadanie w wyznaczonym czasie, co dodatkowo daje im poczucie pewnego sukcesu w nauce języka.

(8) W klasie 4 zrobiłam testy dla dys. wszystkim na start żeby się zaklimatyzowali. Natomiast później już robiłam normalne. Ta taktyka dobrze się u mnie sprawdziła.

(9) Normalnie zazwyczaj. Na egzaminie końcowym nie ma zmiłuj.

- (10) Ja też daję normalne. I pozwalam pisać dłużej. Tak samo jak później na egzaminie.
- (11) Ja testy dyslektyczne używam tylko dla dyslektyków. Moim zdaniem dla uczniów bez opinii te testy są zazwyczaj zbyt łatwe. Stosuję je również dla uczniów poprawiających sprawdzian. Dla większości, a zwłaszcza dla tych zdolniejszych są zdecydowanie za mało wymagające.
- (12) Tylko dla dyslektyków. Uczniowie są zadowoleni to ja też jestem, pracuję zgodnie z zaleceniami poradni, uczniowie uzyskują lepsze noty, mniej ćwiczeń to także więcej czasu na myślenie... .
- (13) Ja stosowałam w kl 4 – dla dyslektyków. W klasie 5 już nie używam – w semestrze 1 nie obniżałam oceny za błędy- ale musieli poprawić błędy w zeszytach. Natomiast zawsze drukuję dla nich wersję Worda, którą przystosowuję graficznie- czytelnie, czyli np 2 zadania na stronie. W drugim semestrze już obniżałam oceny- na egzaminie dyslektyk ma jedynie wydłużony czas. Z rodzicami rozmawiam o konieczności dodatkowej pracy w domu- przygotowują sami krzyżówki, fiszki. Muszą się „opatrzyć” z zapisem.
- (14) Nauczycielka mojej siostry używała czasami tej grupy dla całej klasy mimo że nikt nie ma oświadczenia o dysleksji.
- (15) Ja używam dla dyslektyków oraz dla dzieci ze specjalnymi trudnościami w uczeniu.
- (16) Ja używam tych testów dla uczniów z orzeczeniem lub opinią. Wiem, które osoby nie są w stanie napisać zwykłej wersji. Nie chce uczniów zniechęcać.
- (17) Ja mam 4 klasę sześciolatków, czasem biorę zadania dla dys., bo inaczej nie są w stanie zrozumieć poleceń.
- (18) Ja używam jako dostosowanie nie tylko dla dyslektyków ale dla innych dzieci z trudnościami też.
- (19) Ja korzystam również dla uczniów słabszych. Mam na myśli takich, którzy są na granicy normy intelektualnej. Z takim testem sobie radzą i nie zniechęcają się do nauki.