

Margherita Pelleriti

Item Writing and Dyslexic Students

Promoting Best Practices

Abstract

In language testing, item writing is an important process which can be creative, rewarding but also challenging and sometimes frustrating. Quality items are fundamental for test validity and item writers can be formally trained in order to improve what they write (Rossi and Brunfaut 2019).

The present study analyses some items specifically written for dyslexic students as part of a high-stakes, Internet-based B1 English test administered to undergraduate students in an Italian university, namely the University of Modena and Reggio Emilia, where the number of dyslexic students enrolled has increased exponentially in recent years. The aim was to investigate whether the accommodations suggested by the item reviewer succeeded in removing unnecessary barriers which represent an unintended bias, while preserving the test construct and consequently its validity (Pelleriti 2018; Kormos and Smith 2012).

For the present investigation, a questionnaire was devised, and four certified dyslexic students were interviewed on a voluntary basis. The data collected during the semi-structured interviews confirmed that some of the suggested accommodations would be beneficial, whereas others would be detrimental to candidates with dyslexia. As a consequence, some conclusions have been drawn, with the aim of sharing best practice among the language testing community—in particular language testers, test developers, and item writers. Nevertheless, this research has also confirmed that item writing undoubtedly deserves more scholarly attention, in an attempt to shed light on this pivotal aspect, which is at times neglected despite being one of the pillars in language testing.

Keywords: accommodations, dyslexic students, item writing, language testing, validity

What you give, Momo, is yours forever.

What you keep is lost for all time!

(Eric-Emmanuel Schmitt, Monsieur Ibrahim and the flowers of the Qur'an)

Language testing implies measuring a candidate's performance (i.e., competence) in a foreign/second language. Measuring this competence can be challenging since it implies that cognitive processes are expected to be made visible, despite being abstract and intangible,

since they occur in the candidate's brain and we can only observe the result, namely their performance (Brown and Abeywickrama 2010). As a result, the test construct is key if we are to measure what we want to measure; it is equally clear that item writing is of paramount importance in order to administer a fair and just test to candidates. Indeed, this is an ethical issue (McNamara 2006) which concerns all the stakeholders involved due to the social implications it entails (McNamara and Roever 2006; Kunnan 2000) and it can become even more burning when assessing students with SpLDs—Specific Learning Difficulties or “Differences,” as Kormos and Smith (2012) prefer to define them. In particular, the present contribution investigated dyslexic students and item writing, to analyse whether the way items are written may adversely affect the performance of students with dyslexia.

Dyslexia is a neurological disorder, a neurodiversity, which concerns language processing and can cause limited working memory, difficulties in reading, reduced attention control; it may co-occur with other learning difficulties (i.e., comorbidity), such as dysorthography, dysgraphia, dyscalculia (Pelleriti 2018; D'Este and Ludbrook 2013; Kormos and Smith 2012). Furthermore, dyslexia also involves different levels of severity, namely what Turner and Pughe define as the “continuum of dyslexia” (2003, 5-6). According to the British Dyslexia Association, “ten percent of the population are believed to be dyslexic.”¹ The significant percentage of the British population diagnosed with dyslexia might be explained by the fact that English is an opaque language, as opposed to transparent languages such as Italian, for instance (Ludbrook 2018; Pelleriti 2018; D'Este and Ludbrook 2013). The opaque nature of the English language means that learning English as a Foreign Language (and being assessed in EFL) can be extremely demanding and challenging for dyslexic students.

Consequently, in the test used for the present study, some items had been specifically written with dyslexic students in mind, in an effort to remove unnecessary hurdles (Pelleriti 2018; Kormos and Smith 2012) from the start. These B1 English items were part of a high-stakes, Internet-based B1 English exam administered to undergraduate students in an Italian university, namely the University of Modena and Reggio Emilia, where the number of students with a diagnosis of dyslexia has increased exponentially in the last decade, from the first 3 dyslexic students enrolled in the 2007-2008 academic year² to the 718 dyslexic students enrolled in the 2020-2021 academic year.³ As a consequence, the need to focus more on these students

¹ www.bdadyslexia.org.uk/dyslexia. Last visited 01/10/2022.

² Even before 2010, when the Law 170/2010—which officially recognises dyslexia in Italy—was promulgated (Pelleriti 2018).

³ Data provided by the Office operating at the University of Modena and Reggio Emilia and devoted to providing support to learners with SpLDs.

has increased as well and become urgent; this explains why some B1 items had been specifically written for these candidates. However, while revising them, the item reviewer suggested some accommodations, namely changes which were not meant to affect the test construct and therefore its validity (Pelleriti 2018; Kormos and Smith 2012). Indeed, accommodations should only be a tool to provide equal opportunities to dyslexic test takers and must not give an unfair advantage to this target population under testing conditions (Li and Suen 2012). To shed light on the impact these item accommodations would have and ascertain their validity, qualitative investigation was carried out thanks to a multiple case study (Dörnyei 2007, 152). The aim of the present study was to investigate how students with dyslexia—whose L1 is Italian—tackle some B1 English items, so that greater insight can be gained, in particular about the delicate issue of testing their competence in English as a Foreign Language. Indeed, it is worth remembering the above-mentioned concept, that is, EFL poses more challenges to dyslexic students, since it is an opaque language (Ludbrook 2018; Pelleriti 2018; D'Este and Ludbrook 2013) while Italian is not.

1. Methodology

The high-stakes, Internet-based B1 English exam taken into consideration in this study is meant to assess the following skills and competencies: listening, reading, grammar, and vocabulary. Some B1 English items were written specifically for dyslexic students; then, they were revised by an item reviewer who combined expertise, knowledge of dyslexia and her perceptions of what might have been an unintended obstacle to the fair assessment of the candidates' competence in EFL. Nevertheless, the need to involve dyslexic students soon became apparent in order to gain insight from them and shed light on what may be perceived as fair or unfair. The aim was therefore to reveal strengths and weaknesses of dyslexic students with regard to the English language and give insight on the assessment of these candidates.

Four officially diagnosed dyslexic students took part in this multiple case study on a voluntary basis. In Italy, being officially diagnosed means that students have been evaluated by an official public body, which is part of the Italian National Health Service. These four students were enrolled at the aforementioned university where the Internet-based test under investigation is administered. As a requisite, they were asked to have already passed a B1 English exam (i.e., at university or as an international certification), in order to avoid any unintentional bias due to a possible inadequate CEFR (Common European Framework of Reference for Languages) level for the scope of the present study. These informants, who can be regarded as a purposive sampling (Dörnyei 2007, 153), were administered semi-structured interviews based on a

questionnaire specifically written for them, along with slides showing two versions of the same B1 English items or sentences⁴ under scrutiny; in particular, option a) was the original version, whilst option b) was the one providing the accommodations suggested by the item reviewer. To avoid any potential, unwanted bias, the respondents were shown the two versions of each item/sentence always in the same order: the original option first, followed by the accommodated version. Nevertheless, it is worth noting that minor revisions had been applied to the original versions before administering them to the informants. In this way, on the one hand, testing security was preserved, whilst on the other hand, the foci of the study were still the same as in the items/tasks originally written by the item writer.

Although the sample size used in this study is small (i.e., four university students), it should be considered that dyslexic students are a sub-group of test takers (Pelleriti 2018), namely a limited (by its very nature) number of candidates. In addition, for the present study, prospective participants had been asked to satisfy the following two requisites: having already passed a B1 English exam at university or as an international certification, and being willing to be interviewed. This might have caused a further reduction in the number of potential interviewees, consequently the small sample size of four respondents. Nonetheless, there are studies on dyslexic candidates conducted with only one informant (e.g., D'Este and Ludbrook 2013).

2. The interview

The first part of the questionnaire was meant as an ice-breaker and as a way to get some information about the interviewees (e.g., age, studies, levels of severity, etc.); the second part included the 11 items/sentences under investigation which were shown to the informants as slides. In particular, for each single item/sentence they were first administered option a) only, then option b) only, finally the two options together on the same slide. In this way, the respondents were given the chance to focus on a single option at a time and then, in order to help them recall and decide between the two options, both versions were provided on the same slide. They were expected to decide which option they could read (i.e., decode the written text) more easily and why, and which option they could understand more easily and why.

The interviews were conducted in Italian, since the interviewees' L1 is Italian; in this way, the four informants were free to express themselves, without the potential bias of the language

⁴ In language testing terminology, only questions 5 and 9 can be considered items; the remaining questions are sentences which deserved investigation, for instance because they are part of an item (e.g., the stem, an option) or a sentence taken from a reading passage, and so forth.

barrier. The interviews were carried out with a single informant at a time; as a result, independent answers were provided which were not influenced by those of the other informants; each interview was expected to last approximately 30 minutes. Furthermore, each interviewee had been previously informed about (1) the aim of the research, (2) who would have access to their data and (3) how anonymised data would be treated. Table 1 below provides an outline of the interviewees, whose age is that reported at the time of the interview. Each informant was asked to choose the name of a famous dyslexic person among the ones provided by the interviewer; consequently, their anonymity was guaranteed thanks to a pseudonym which is shown in Table 1 below:

Name	Age	Gender	Course of Studies
Jennifer Aniston	19	F	Bachelor's degree
Cher	21	F	Single-cycle Master's degree
Agatha Christie	23	F	Master's degree
Erin Brockovich	23	F	Master's degree

Tab. 1: Interviewees: an outline

Despite the fact that the interview was expected to take about 30 minutes, it took longer with each of the four informants; it seems plausible that the reason might lie in the informants' willingness to provide more information, namely their desire to communicate and contribute to this research in order to help other fellow students with dyslexia. Table 2 below illustrates the exact duration of each interview, along with supplementary information about the courses of study of the four informants.

Name	Interview Duration	Name of Course
Jennifer Aniston	50 minutes	European Languages and Cultures
Cher	40 minutes	Primary Education Sciences
Agatha Christie	50 minutes	Teaching and Communicating Sciences
Erin Brockovich	60 minutes	Civil Engineering

Tab. 2: Interview duration and name of course

During the interview, the four informants were asked to define their dyslexia according to the

“continuum of dyslexia” (Turner and Pughe 2003, 5-6). Three of them defined it as mild, whereas only Agatha Christie (AC) defined hers as of medium severity and despite this, she had attained an international certification in English at C1 level (proficient user) of the CEFR. Concerning severity, it is worth noting that Erin Brockovich (EB) was diagnosed and certified as medium but she defined her dyslexia as mild “based on how I’ve managed it myself” (quotation from the interview with EB). Furthermore, all four informants were affected by comorbidity, in particular dyscalculia and dysorthography and none of them used a synthesiser to study at university (this was referred to any subject, not just English). In addition, all of them started to study English at an early age; in particular, Jennifer Aniston (JA) and AC at 3 years old, whereas EB and Cher (C) had been studying English since they were 6. The interviewees were also posed a question about their perception of test fairness; in particular, they were asked whether they considered as fair being administered the same English test (i.e., the B1 exam made up of the following components: listening, reading, grammar, and vocabulary) as their fellow students who are not affected by dyslexia. Despite their major difficulties, all the informants agreed by stating it was fair, “even for inclusion” (interview with JA), “maybe it is more difficult, but like the others” (interview with C); this last remark was also confirmed by EB (“I can do anything that the others do”). Based on their own personal perception, the four interviewees were also asked to rate the level of difficulty regarding the receptive skills and competencies evaluated in the B1 English test. Three informants declared that listening comprehension was the easiest skill, whereas EB mentioned reading comprehension as the easiest for her; EB was also the informant who rated vocabulary as the hardest part, especially if she were expected to retrieve it by herself; a similar idea was expressed by C, whose first answer was that grammar is the hardest, but then she felt the need to clarify that writing vocabulary would be more difficult for her. Finally, even AC rated grammar as the most difficult. All of the above-mentioned ratings (Listening = L; Reading = R; Grammar = G; Vocabulary = V) are illustrated in Table 3 below, where it is worthy of notice that JA rated grammar and vocabulary as equally difficult, choosing the same figure, namely 2, for both language competencies:

Name	L	R	G	V
	(1 = the easiest; 4 = the most difficult)			
Jennifer Aniston	1	3	2	2
Cher	1	3	4	2
Agatha Christie	1	2	4	3
Erin Brockovich	3	1	2	4

Tab. 3: Rating language skills and competencies

In order to gain greater insight into how dyslexic students tackle an English test, the four informants were shown slides with two versions of the same B1 English items/sentences. The items and tasks had been written specifically for dyslexic students; then, the item reviewer had specifically selected some items and sentences from tasks which needed revision, namely accommodations for dyslexic students. The items/sentences shown were meant to investigate different aspects, such as potential difficulty (i.e., bias) in the way a listening introduction is presented, in the way items are written (e.g., interrogative versus declarative sentence in the stem), in decoding names, nouns, figures, and so forth. The aim of the investigation was therefore to validate the accommodations suggested by the item reviewer—as illustrated in the next section—and to promote best practice in item writing to be shared with the stakeholders involved—in particular, language testers, test developers, and item writers.

3. Accommodating items: data analysis and discussion

The qualitative research conducted in this study in some cases confirmed the item reviewer's perceptions, while in others it revealed new aspects to be taken into consideration. For instance, regarding *decoding names and nouns*, the informants were shown the following three pairs of sentences, where option a) is the one originally written for dyslexic students (although slightly modified in order to preserve testing security), whereas option b) is the accommodated version suggested by the item reviewer:

- (1) a) Jim and Nimisha are eating an ice-cream.
b) Jim and Mary are eating an ice-cream.
- (2) a) Steve and Hitomi live in Bristol.
b) Steve and Robert live in Bristol.
- (3) a) Paul has ordered chicken tikka.
b) Paul has ordered fish and chips.

The aim was to understand whether an unusual name (i.e., “Nimisha,” “Hitomi”) or noun (i.e., “chicken tikka”)—which appeared in the original items specifically written for dyslexic students—would represent an unnecessary hurdle for them in order to decode the written text. As a consequence, options b) with more common (i.e., high-frequency) words were written and the interviewees were asked to state, for each pair of sentences, which one they could read (i.e., decode) more easily and why, and which one they could understand more easily and why. EB stated that she could read and understand both versions (a and b) easily; in particular, she

relied on capital letters in order to realise she was reading about names (i.e., “Nimisha,” “Hitomi”), whereas her strategy to deal with the third pair of sentences was focusing on the first part of the chunk only, namely “chicken” in “chicken tikka” and “fish” in “fish and chips.” This was also confirmed by AC who, in decoding the first two pairs of sentences, not only relied on upper case in order to infer names, but also on the conjunction “and.” Concerning the third example, AC tried to decode and understand “tikka,” the only word unknown to her, and came to the conclusion that it might have been an adverb, new to her, possibly American, maybe meaning “soon.” The informant C could read the alternatives provided by the item reviewer (i.e., “Mary,” “Robert”) more easily, since for example when decoding “Nimisha,” she read it as “Natasha”; in addition, she found “chicken tikka” harder to read. Finally, JA stated that she could read and understand both versions of the three pairs of sentences under scrutiny easily. Another aspect investigated in this study was *word order*, in particular whether the proximity of two adverbs (i.e., “twice,” “every morning”) in a sentence might cause unnecessary strain to dyslexic students while decoding the written text. In order to shed light on this, the four informants were administered two versions of the same sentence with a different word order, as shown below:

- (4) a) I get on the bus twice every morning.
 b) Every morning I get on the bus twice.

JA confirmed that she could read and understand the two of them easily; EB could easily read both sentences but she understood b) more easily, since it sounded more spontaneous; C and AC also considered b) easier.

Multiple-choice items were also analysed, as illustrated by example 5 below where assessing *tenses* (related to the past) by use of a discrete-point item was the original aim of the item writer who opted for the distractor “did start” in version a), whilst version b) is the one suggested by the item reviewer:

- (5) a) The film _____ by the time Lucy arrived. (did start / had started / started)
 b) The film _____ by the time Lucy arrived. (has started / had started / started)

Three informants (i.e., EB, C, AC) agreed on the fact that they would never choose “did start,” consequently this implies that the distractor seems to be ineffective; nonetheless, the new distractor suggested by the item reviewer did not seem appropriate since it was similar to the

key (i.e., “had started”). In item b) a single letter (i.e., “s” in “has” and “d” in “had”) was the only difference between the distractor “has started” and the item key “had started,” and such a slight difference can be quite demanding for dyslexic students, as reported by the informant EB, who in the end preferred version a) since the three options (i.e., “did start/had started/started”) are different, whereas—she added—options which are similar (e.g., “has started/had started”) can lead to potential mistakes. Nevertheless, it is worth remarking that JA was the only interviewee stating she could read and understand both versions easily; this could be explained by the fact that her dyslexia is mild and that she is studying “European Languages and Cultures.” Finally, AC could read and understand b) easily, although discarding “did start” in option a). It is worth noting that when candidates with SpLDs are administered grammar tests, the ideal number of options in discrete-point multiple-choice items is three (Ludbrook 2018).

Since the Internet-based B1 English exam includes reading and listening as well, the informants were also administered a few items/sentences in order to shed light on how they tackle these receptive skills. As far as *reading comprehension* is concerned, the interviewees were administered the two versions of a sentence taken from a reading passage (example 6) and the two versions of an item option (example 7):

- (6) a) Banks offer small businesses, often run by young people, loans of about € 5,000.
 b) Banks offer loans of about € 5,000 to small businesses, often run by young people.
- (7) a) Not having to pay to stay in the hotel in Spain was a great opportunity.
 b) The hotel for free in Spain was a great opportunity.

Regarding example 6, JA preferred version a) since it was easier to read, whereas she considered version b) as being structured more chaotically; a preference for a) was also shared by AC, who relied a lot on punctuation in order to read and understand the sentence. On the contrary, EB’s preference was for version b) since the main clause was in the first part of the sentence and in version a) she would not have read the aside written between the two commas (i.e., “often run by young people”), which was intended to provide extra information to the reader. Finally, a preference for version b) was also expressed by informant C. This means that no rule of thumb can be applied to example 6, since these interviewees relied on different strategies to tackle the written text. Concerning example 7—which is an item option of a reading comprehension task—JA opted for b) since this version was easier to read and understand; this same preference was also expressed by both EB and C, the latter stating that b) was easier because it was shorter, although she could also understand version a). Conversely, AC could read b) more easily but it

was easier for her to understand a), due to the fact she was uncertain about the chunk “for free,” since it was unclear to her whether it referred to “Spain,” to “opportunity,” or to “hotel.” Finally, in L1 it is important to remark that reading (i.e., decoding) involves mental representation of the text and automaticity (i.e., the ability to predict the text), but reading comprehension processes in a foreign language are not so automatized (Brunfaut et al. 2021); this can be even more problematic with dyslexic students, who struggle more in decoding a written text and whose working memory is limited, even in their first language.

As far as *listening comprehension* is concerned, a good rule of thumb (also in order to avoid construct-irrelevant variance) is that the items should not be more difficult than the aural input, namely “written language should be as simple and clear as possible” (Buck 2001, 123). This is of paramount importance since listening comprehension implies immediate processing and relying on working memory, which should not be overloaded when dealing with a listening comprehension task (Brunfaut 2016). Consequently, this can be attained, for instance, by avoiding vocabulary in the items (i.e., both in the stem and in the options) which might be inappropriate given the CEFR level under measurement (e.g., B1 for the scope of the present study). In order to validate this aspect, the four interviewees were asked to deal with the following, which is an item option:

- (8) a) He writes about the proceedings of the American Senate.
- b) He writes about what happens in the American Senate.

EB preferred option b) since it was easier to read and understand; C could easily understand both versions, also because she already knew the word “proceedings”; nevertheless, she stated that b) was easier to read due to the more common words used in this version. JA shared the same opinion, namely, she could read and understand both versions easily; nonetheless, she deemed b) easier to understand—it is worth noting that when stating this, very likely JA had other dyslexic students in mind. Finally, AC—although able to understand “proceedings” since her proficiency level in English was C1—agreed on the fact the b) was easier to read and to understand, also for the reason that words such as “what,” “whom” and the like can help dyslexic students to read a given sentence, since they are useful in order to decode the text. However, it should be underlined that “proceedings” was not in the aural input, but just in the item option originally written by the item writer.

In an effort to avoid any construct-irrelevant variance, another area of investigation was represented by numbers while tackling a listening comprehension task. In particular, the aim was to shed light on whether numbers should be written in letters or in figures in the language

testing items meant to elicit listening comprehension, so that any unintended bias would be avoided. Consequently, the four informants were asked to express their preference—pretending they were tackling a listening task—between the two versions of the item shown below:

- (9) a) Nick has written _____ [*three/four/seven*] articles this week.
 b) Nick has written _____ [*3/4/7*] articles this week.

The interviewee EB preferred b) because “it is more visual” and she added that distinguishing “three” from “tree” would be harder for dyslexic students, so she felt that figures should be highly recommended. A preference for option b) was also expressed by JA, despite the fact that she could easily read and understand both. Option b) was also preferred by C who could easily read figures despite being affected by dyscalculia. Furthermore, C added the following noteworthy piece of information: in a reading comprehension task—so focusing on a different skill—a parallel version would be highly recommended both in the text and in the item(s), namely if numbers are presented in letters in the reading passage, numbers should be presented in letters in the item(s) too, and vice versa (i.e., figures in the text and figures in the item). A similar remark about reading tasks was also expressed by AC, who also added that figures are the best option when dealing with listening comprehension.

Finally, the *stem in a listening task* was also under scrutiny in order to realise which form—interrogative versus declarative—should be preferred when writing items where the one or the other form can be used in the item stem. As a consequence, the two options below were shown to the four respondents:

- (10) a) What does Michael say about actors?
 b) Michael says actors...

On the one hand, EB could read and understand a) more easily, especially because *wh-* questions (e.g., “what,” “where,” “who”) can help a lot since they provide contextualisation. C agreed on this aspect (i.e., she understood a) more easily) since she was made aware of what to focus her attention on. Nevertheless, in a listening comprehension task, she would prefer option b) since it is shorter, so it would take her less time to read it. On the other hand, JA could read and understand both item stems easily, whereas AC preferred option b).

To conclude, a final area of investigation was represented by the few lines meant as an *introduction* in a *listening task*, which provide useful contextualisation before tackling the aural input. An example is shown below:

- (11) a) You will hear a journalist called Jennifer Long telling a group of students about her work as a newspaper journalist.
- b) You will hear a journalist called Jennifer Long. She tells a group of students about her work.

All four informants preferred option b); in particular, C could read option b) more easily thanks to punctuation, namely the fact that the information provided had been divided into two sentences. This point of view was shared by JA, who found option b) clearer and simpler thanks to punctuation, which helps while decoding the written text. AC also opted for b), explaining that while she felt it would be easier especially for B1 test takers, it might sound artificial to more proficient candidates (an opinion based on the fact that she had already obtained a C1 certificate from an international examination board).

Finally, at the end of the interview, all the informants were asked whether they wanted to provide further remarks or additional information. On the one hand, EB reported an interesting aspect regarding the fact that she experiences days/periods in which her dyslexia gets worse, and her explanation was that very likely her high level of fatigue in those days/periods negatively affects her. On the other hand, C concluded by saying she appreciated the fact that someone (i.e., the interviewer) was interested in them (i.e., dyslexic students) and that she was happy to help other students with dyslexia.

4. Conclusions and implications for future research

This study aimed to investigate whether some accommodations to the items or sentences from a task might prove fruitful for dyslexic students, and whether some general conclusions might be drawn and then applied in the future as best practice. Consequently, the present qualitative research aimed to shed light on weaknesses and strengths, along with the strategies pursued by dyslexic students when dealing with a B1 English test. Indeed, these strategies can be very effective in overcoming potential unnecessary difficulties, which might be a bias undermining the test construct and therefore its validity.

The data analysed revealed that only some of the item reviewer's accommodations—suggested in order to avoid any potential unintended bias in the items—were confirmed, as proved by the answers to questions 4 and 7 (3 informants), 8, 9, and 11 (all the respondents); whereas other questions provided unexpected answers: a 50% agreement for question 6; a 75% agreement that the accommodation for question 5 was unsuccessful due to the slight difference between “had” and “has.” Moreover, the data analysed also revealed some of the strategies put into action by

dyslexic students, for instance when relying on upper case (e.g., questions 1 and 2) or when focusing just on the first part of the chunk (e.g., question 3). Consequently, it seems plausible to state that sometimes a common ground was found among the interviewees, namely they all agreed on some questions (e.g., 8, 9, 11), whilst for other questions (e.g., 6, 10) no independent, commonly shared consensus was reached. This means that more investigation, with a larger sample of interviewees—as advocated in the guidelines promoted by Cardinaletti (2018)—should be conducted in order to see whether some general rules of thumb may be confirmed, in an attempt to promote best practice when writing English items which might be administered to dyslexic candidates in high-stakes tests. Nevertheless, the preliminary data reported herein offer insight for stakeholders such as language testers, test developers, and item writers and a starting point for future work aimed at generalising these findings.

As Kormos reminded during her plenary lecture at the 17th EALTA Conference in 2021, the underlying principle inspiring future studies should be “*Nothing about us without us.*” Keeping in mind this mantra of the movement for the empowerment of those with SpLDs can be of great value in improving language testing (even more so than in language learning), especially when administering high-stakes tests. Indeed, item writing—which is challenging by its very nature—can become even harder when writing items which need to be fair and just to dyslexic candidates, for instance by providing accommodations meant to remove unnecessary barriers (Pelleriti 2018; Kormos and Smith 2012), while preserving the test construct and therefore its validity, in an attempt to avoid any construct-irrelevant variance. Consequently, in order to gain insight in an area which deserves more scholarly attention, it is of paramount importance to involve dyslexic students in language testing research, since their voices can unveil some interesting features of the cognitive processes involved in language acquisition and testing—processes which are invisible since they happen in the human brain—and can shed light on the strategies put into action in order to deal with a language test. This information would be of great help in writing fair and just items for dyslexic candidates who are expected to take an English test.

Acknowledgements: the author would like to express her sincere gratitude to the four informants of this study for their willingness to be interviewed, for understanding the importance of providing honest answers, and for contributing to valuable insight in an area still in need of scholarly attention.

Margherita Pelleriti *after a nine-year experience as a certified EFL teacher in University*

Language Centres and state schools, now works at the Language Centre of the University of Modena and Reggio Emilia, Italy. Her research interests include language testing, dyslexia in relation to language learning and language testing. She was awarded an MA Language Testing degree from Lancaster University; she is a UKALTA member and was in charge of a CLIL course.

Works cited

- Brown, H. Douglas and Priyanvada Abeywickrama. *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education, 2010.
- Brunfaut, Tineke. "Assessing Listening." *Handbook of Second Language Assessment*. Edited by Dina Tsagari and Jayanti Banerjee. Berlin: De Gruyter Mouton, 2016. 97-112.
- Brunfaut, Tineke, et al. "Testing Young Foreign Language Learners' Reading Comprehension: Exploring the Effects of Working Memory, Grade Level, and Reading Task." *Language Testing* 38.3 (2021): 356-377.
- Buck, Gary. *Assessing Listening*. Cambridge: Cambridge University Press, 2001.
- Cardinaletti, Anna. "Linee guida per test linguistici accessibili e prospettive future." *Test linguistici accessibili per studenti sordi e con DSA: pari opportunità per l'accesso all'università*. Edited by Anna Cardinaletti. Milan: FrancoAngeli, 2018. 317-324.
- D'Este, Claudia and Geraldine Ludbrook. "Fairness and Validity in Testing Students with SpLDs: A Case Study from Italy." *Assessing L2 Students with Learning and Other Disabilities*. Edited by Dina Tsagari and George Spanoudis. Newcastle upon Tyne: Cambridge Scholars Publishing, 2013. 169-188.
- Dörnyei, Zoltán. *Research Methods in Applied Linguistics*. Oxford: Oxford University Press, 2007.
- Kormos, Judit and Anne Margaret Smith. *Teaching Languages to Students with Specific Learning Differences*. Clevedon: Multilingual Matters, 2012.
- Kunnan, Antony John, edited by. *Fairness and Validation in Language Assessment*. Cambridge: Cambridge University Press, 2000.
- Li, Hongli and Hoi K. Suen. "Are Test Accommodations for English Language Learners Fair?" *Language Assessment Quarterly* 9.3 (2012): 293-309.
- Ludbrook, Geraldine. "La prova di grammatica." *Test linguistici accessibili per studenti sordi e con DSA: pari opportunità per l'accesso all'università*. Edited by Anna Cardinaletti. Milan: FrancoAngeli, 2018. 253-266.
- . "La valutazione della lingua straniera: aspetti critici per gli studenti sordi e con DSA." *Test linguistici accessibili per studenti sordi e con DSA: pari opportunità per l'accesso*

- all'università*. Edited by Anna Cardinaletti. Milan: FrancoAngeli, 2018. 56-66.
- McNamara, Tim. "Validity in Language Testing: The Challenge of Sam Messick's Legacy." *Language Assessment Quarterly* 3.1 (2006): 31-51.
- McNamara, Tim and Carsten Roever. *Language Testing: The Social Dimension*. Malden: Blackwell Publishing, 2006.
- Pelleriti, Margherita. *Dyslexic Students: From Language Learning to Language Testing*. Valencia: Universitat Politècnica de València, 2018.
- Rossi, Olena and Tineke Brunfaut. "Test Item Writers." *The TESOL Encyclopedia of English Language Teaching*. Edited by John I. Liantas. Hoboken: John Wiley & Sons, 2019. 1-7.
- Turner, Elizabeth and Jayne Pughe. *Dyslexia and English*. London: David Fulton Publishers, 2003.