EQUITY OR ADVANTAGE? THE EFFECT OF RECEIVING ACCESS ARRANGEMENTS IN UNIVERSITY EXAMS ON STUDENTS WITH SPECIFIC LEARNING DIFFICULTIES (SpLD).

CARDIFF UNIVERSITY

Helen Duncan
School of Healthcare Sciences
Cardiff University

This dissertation is submitted for the degree of Doctor of Philosophy

Supervisors
Dr Catherine Purcell
Mrs Sally Scott-Roberts
Dr Heather Hurst

November 2021

Acknowledgements

This thesis has been the most exacting, meaningful, challenging and exciting project that I have ever embarked on and would not have been possible without the support, encouragement and guidance of the following exceptional people:

Firstly, I am hugely grateful to the participants in this research. Without their generosity in sharing their data and deeply personal experiences with me, this project would not have been possible. I will strive to ensure that the trust they placed in me results in impactful outcomes.

Secondly, I owe the most enormous debt of gratitude to my extraordinary Director of Studies and supervisor, Dr Catherine Purcell, who has been so astonishingly generous with her time, expertise and professional guidance. Dr Purcell came into my life like a comet, and shone. I could never have completed this project without her. I am forever grateful for all that she has done for me, her kindness, her humour and her unflagging enthusiasm for my work.

I am also sincerely grateful to my supervisors, Mrs Sally Scott-Roberts and Dr Heather Hurst. Together with Dr Purcell they have formed my powerhouse of academic support. Their guidance, expertise and enthusiasm for the project has been phenomenal and I am more grateful than I have words to say.

I would like to acknowledge the endless support of my sister, Susan, my son, Alexander, and my partner, Terry, whose unceasing faith in me is startling. Without their encouragement and all the extra work they took on to allow me the time to focus on this project I would not have had the opportunity to pursue this research.

Finally, I am also deeply grateful to John Harding, Head of the Disability Resource Centre, for arranging the resources and connections that made this project viable.

To all these exceptional people I want to say the most heartfelt ‘thank you’. This thesis is the product of your collective support, encouragement, hard work and generosity with your time (as well as good measures of patience and good-humour). Without you all this project would never have come to fruition.

Publications

This research project has given rise to the following publications:

Duncan, H., & Purcell, C. (2017). Equity or Advantage? The effect of receiving access arrangements in university exams on Humanities students with Specific Learning Difficulties (SpLD). Widening Participation *and Lifelong Learning,* 19(2), 6-26.

Duncan, H., & Purcell, C. (2020). Consensus or contradiction? A review of the current research into the impact of granting extra time in exams to students with specific learning difficulties (SpLD). *Journal of Further and Higher Education*, 44 (4), 439 – 453.

Contribution to the field

This research project was triggered by the longstanding national debate about the impact that exam access arrangements have on the exam performance of students with SpLD; namely, do the adjustments promote equity, and so place students with SpLD on a level playing field with their typically developing (TD) peers, or confer an advantage by unfairly inflating their scores? The current body of research exploring this question is characterised by vastly conflicting views and no overall consensus. At the same time the extant body of research lacks ecological validity; consisting mainly of primary research that compares the performance of participants with SpLD and their TD peers in generalised mathematics tests or comprehension tests which were constructed for that particular research project – experimental conditions that are significantly different to the experience of real-life, subject specific, closed-book, timed, written, university exams. Thus, policymakers in educational institutions are required to make decisions relating to the appropriateness of exam access arrangements without sufficient empirical, ecologically valid, evidence to inform their decisions.

This research project uniquely addresses this lack of ecologically valid, empirical data by evaluating the actual exam performances of students with SpLD who were granted exam access arrangements and comparing this to the actual exam performance of their TD peers who took the same exam under standard conditions. This comparison shows whether students with SpLD who receive exam arrangements perform similarly or differently in real-life exams to their peers who sat the same exams under standard conditions. In addition, by comparing the exam marks of the participants with SpLD with their dissertation marks (cross referenced with a similar comparison of the exam marks and dissertation marks of their TD peers) this project uniquely allows for a more granular analysis of any differences between the potential of students with SpLD (as measured by their coursework performance) and their performance in exams with adjustments. This research project, by analysing real-life exam results, identified that, contrary to the belief that the exam arrangements either confer equity or advantage, an awarding gap between students with SpLD and their TD peers persists despite the implementation of exam adjustments. In addition, an exploration of the experiences of the students with SpLD themselves who receive exam adjustments has further deepened our understanding of the drivers of this awarding gap, as well as providing suggestions for ways forward. Thus, this research project has elicited unique empirical evidence that can now be used to inform decisions relating to assessment practice and policy in relation to students with SpLD in Higher Education.

Abstract

Formal, timed, written, closed book examinations are a common feature of the educational system in UK Universities and aim to assess the subject knowledge, skill and ability of all the candidates in the cohort in a uniform way. Adjustments, such as extra time and use of a word processor, may be made to the administrative conditions of exams for students with Specific Learning Difficulties (SpLD) where standard exam conditions may unfairly disadvantage this particular cohort. These adjustments are intended to remove construct-irrelevant barriers and promote parity of assessment, where time is not the construct being measured by the exam. However, granting exam access arrangements violates the principle of assessing all candidates in a uniform way (by changing the administrative conditions for some candidates) and so raises concerns that, rather than creating parity of assessment, these adjustments may inflate the results of those students to whom they are granted and thereby lead to an unfair advantage.

The purpose of this research was to identify, through an analysis of actual exam data, the impact that exam adjustments have on the exam performance of students with SpLD and thereby help reconcile this debate of 'equity or advantage?' Specifically, the length of exam script, marks, and degree classification achieved by students with SpLD were compared with those of their typically developing (TD) peers who sat the same exam under standard conditions to ascertain if any differences existed. The empirical data revealed that, although the exam adjustments improved the performance of the participants with SpLD, this cohort still achieved lower results than their TD peers who took the same exam under standard conditions. Interviews with students with SpLD illuminated this outcome by identifying the multiple and various disadvantaging barriers that pervade teaching and learning, culminating in underperformance in exams and a resultant awarding gap.

The quantitative data for this project was collected between 2016 – 18, when assessment comprised primarily timed, formal, closed book, written exams. The COVID-19 pandemic in 2020 prompted a rapid shift in university assessment processes from this traditional exam format towards open book, digital, assessments with a time span of 24 hours or more. While, post-pandemic, Universities appear committed to retaining the benefits of on-line assessment formats, most have resumed the inclusion of some timed, formal, closed book, written exams as part of their assessment portfolio, in response to concerns about academic integrity in online assessment environments (Reedy et al, 2021). Thus, the outcomes of this project continue to

be relevant to University assessment practices post-pandemic and have a unique contribution to make to post-pandemic decisions about assessment policy and practices.

APA 7th Ed. referencing style has been used in this thesis.

Table of Contents

ABBREVIATIONS

ADHD	Attention Deficit Hyperactivity Disorder
AERA	American Educational Research Association
BDA	British Dyslexia Association
DCD	Developmental Coordination Disorder
DfES	Department for Education and Skills
DSM V	The Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
fMRI	Functional Magnetic Resonance Imaging
GCSE	General Certificate of Secondary Education
HE	Higher Education
HEI	Higher Education Institution
MPT	Maximum Potential Thesis
MCQ	Multiple Choice Questions
OfS	Office for Students
SASC	SpLD Assessment Standards Committee
SPLD	Specific Learning Difficulty
TD	Typically Developing
UCAS	The Universities and Colleges Admission Service
WAIS	Wechsler Adult Intelligence Scales
WISC	Wechsler Intelligence Scales for Children

1: INTRODUCTION

This introduction chapter provides the background and context to this research project. The chapter presents the rationale for the project, including the debate that underpins the research objectives and an overview of the legal framework, along with data on the scale of the issue (in terms of numbers of individuals affected). The chapter concludes by presenting the research objectives and the hypothesis that the research project aims to investigate.

1.1 Rationale for the proposed topic

Formal, timed, written, closed book examinations remain a common method of educational assessment that many Universities employ to gauge a student's level of subject based knowledge, ability and skills following a course of study. This is based on the assumption that a student's performance in an exam is generally regarded as an objective and reliable measure of that student's learning and proficiency in the particular subject area studied (Gosh et al, 2017; López-Pastor & Sicilia-Camacho, 2017; Nieminen, 2020; Villaroel, et al, 2020). Given the high-stakes nature of university exams, it is imperative that these exams offer a fair evaluation of each candidate's learning in a way that is equitable to all candidates. This applies as much to the conditions under which the exam is conducted as the marking rubric used to gauge exam performance. Ensuring that exams are equitable to all students offers particular challenges for assessment designers. Diverse groups, including candidates with specific learning difficulties (SpLD[1]), are likely to be proportionately represented in any exam cohort[2]

[1] The term Specific Learning Difficulties is commonly used as an umbrella term to cover diagnoses such as dyslexia, dyspraxia, dysgraphia and dyscalculia, for example, but can also be used as a descriptor term in its own right (Lockwood et al, 2020; Maki & Adams, 2018; Snowling et al, 2020)

[2] Around 5% of those taking exams in Higher education are diagnosed with SpLD (Advance HE, 2020)

and it is commonly posited that standard exam conditions place students with SpLD at risk of being unfairly disadvantaged (Asghar et al, 2018; Chong, 2018; MacFarlane, 2019; Payne-Tsoupros, 2020). In order to redress any potential inequality inherent in the assessment practices that may discriminate against students with SpLD, a policy of implementing adjustments to the conditions under which the exam is conducted is generally adopted. This approach is in line with the UK Universities statutory duty under the Equality Act (2010) to make assessments more accessible to students with SpLD.

This approach, however, is not without its critics. Those who challenge the principle of granting exam access arrangements argue that any adjustments that alter the standard testing conditions for a particular group potentially change what the exam measures for those candidates. Thus, they posit, the granting of exam arrangements may undermine the integrity and validity of the exam and compromise the comparability of exam results across the cohort (Bachan, 2017; Sokal & Wilson, 2017; Spenceley & Wheeler, 2016).

Additional time and the use of a word processor are the most frequent exam access arrangements that students with SpLD receive in a time constrained exam situation (Ofqual, 2020). In the UK, 25% is the usual amount of extra time granted, with over 97% of all candidates who were awarded extra time in GCSE and A levels[3] receiving this amount (Ofqual, 2020), and this is similarly reflected in University exams[4]. It should be borne in mind, that the principle of granting a standard amount of extra time, rather than determining a 'tailor made' time extension calculated in an individualised manner (based on the needs of the individual and the demands of the assessment), presupposes a degree of homogeneity across all students with SpLD (Weis et al, 2016; Spenceley & Wheeler, 2016). That is to say, this accepted practice of 'one size fits all' infers that most students with SpLD lack any material variation in their needs in the context of exams. However, rather than representing a homogenous group, the impact of having an SpLD (including the impact on exam performance) can vary across a spectrum (Callens & Brysbaert, 2019; Casini et al, 2017; Hatcher, Liebel and Nelson 2017; Goswami et al, 2019). Stein, (2019), for example, argues that the severity of the impact on study of the student's SpLD broadly correlates with the underlying continuum of cognitive

---

[3] Both GCSEs (General Certificate of Secondary Education) and 'A' levels (General Certificate of Education Advanced Level) are UK, subject based, pre-HE examinations. GCSEs are offered as part of compulsory Secondary education and are often a prerequisite for taking 'A' levels. 'A' levels are taken at the post compulsory Secondary education stage and are usually required for University entrance.

[4] Data obtained from the HESA database HEIDI Plus (2021).

impairment (with each profile and manifestation being individual). However, these variations between students with SpLD, which may warrant differing amounts of additional time in exams, are not reflected in the amount of extra time granted to each individual student.

At the same time, a lack of empirical evidence exists that justifies the calculation of 25% extra time specifically, as opposed to an alternative percentage of extra time, with 'decisions about extended time involving apparently arbitrary amounts of additional time' (Lewandowski, Cohen & Lovett, 2013, p327). This lack of any standardized method of calculating the amount of extra time that is appropriate for any one given individual is clear from the fact that the amount of extra time granted as standard to students with SpLD varies considerably from country to country, with the USA, for example, most commonly granting 50% or 100% extra time as standard options in comparison to 25% in the UK (Gosh et al, 2017; Periera et al, 2017; Villaroel, et al, 2020). It appears that the determination for 25% extra time as the standard 'norm' in the UK is based on practical considerations[5] (such as the administrative convenience of conducting the exam within available resources) rather than any distinct SpLD evidence based calculation (Spenceley & Wheeler, 2016; Weis et al, 2016 Niemimen, 2020). Thus, the convention of granting 25% extra time to the majority of students with SpLD, regardless of their individual variations in cognitive profiles or study based needs, appears to be an administrative imperative that has become accepted practice, rather than one informed by empirical evidence or theory.

Nonetheless, the granting of a uniform amount of extra time (usually 25%) remains common practice in the UK for students with SpLD (JCQ, 2020). The practice of granting extra time (rather than other types of exam adjustments) to this cohort is based on the assumption that students with a diagnosis of SpLD share common deficits in the way in which information is stored, retrieved and processed, and are likely to demonstrate a cognitive profile that includes slow speed of information processing and deficits in working memory (Biotteau et al., 2017; Blank et al, 2019; MacCullagh, Bosanquet & Badcock, 2017; O'Byrne, Jagoe & Lawler, 2019; Snowling et al, 2020). These cognitive deficits play a significant role in inhibiting students with SpLD from demonstrating their full potential when sitting exams under standard time as

[5] The amount of extra time allowed seems to be influenced by the length of the exam (i.e. the longer the standard exam time, the shorter the amount of extra time allowed) rather than established by any evidence based rubric. In HEIs in the USA, for example, exams are commonly 1 hour in duration, allowing for 50% or 100% extra time, whereas in the UK standard time for exams in HE is commonly 3 hours and so, to avoid excessive exam time, only 25% extra time is commonly granted.

the impaired processing speed[6] and working memory[7] difficulties that are characteristic of SpLD may render these students slower in articulating their subject knowledge in writing and thus unable to complete the task in the designated time (Asghar et al, 2018; Broggi et al, 2019; Gooch et al, 2016). As a consequence, they are disadvantaged in examination situations in comparison to their typically developing (TD) peers. Access arrangements are therefore granted with the purpose of compensating for this inequality in the context of a formal, timed, closed book, written examinations and aim to place students with SpLD on a level playing field with their TD peers so that they can demonstrate their knowledge in the exam, rather than their disability (Camilleri et al, 2019; Chetcuti et al, 2019; Falzon, 2020).

1.2 The debate

Advocates of the granting of additional time in examinations argue that if students with SpLD are tested under standard time conditions they are potentially subject to construct-irrelevant variance in their exam results (Falzon, 2020; Kalfe, 2015; Shaw & Anderson, 2017). That is to say, results across the student cohort being tested will vary, to some extent, because of the variability in students' processing speeds and working memory, rather than exclusively due to the variability in their skills, subject knowledge and understanding (i.e. the construct that the exam is actually aiming to measure). Given that the purpose of exam arrangements is to 'minimize the impact of test-taker attributes that are not relevant to the construct that is the primary focus of the assessment' (Berger & Lewandowski, 2013, p301), the granting of exam access arrangements, such as additional time, aims to rectify the construct-irrelevant variance by ensuring that unimpaired processing speed and working memory are not essential requirements for accessing the exam (Shaw & Anderson, 2017; Shaw & Anderson, 2018). Thus the objective of exam adjustments is to allow candidates with SpLD access to the exam 'without changing the demands of the assessment' (JCQ, 2020, p3) and so reduce the barriers that prevent students with SpLD from demonstrating their potential in exam situations.

---

[6] Kaufman & Lichtenberger (2008) define processing speed as 'the speed with which one or more basic processes can be coordinated and applied, often in conjunction with the accessing of lexicons under the direction of executive functions' (Kaufman & Lichtenberger, 2008, p 239).

[7] Kaufman & Lichtenberger (2008) define working memory as 'the capacities involved in the initial registration and holding of information and the mental manipulation of information that is being held in mind' (Kaufman & Lichtenberger, 2008, p225)

It is also argued that, as each candidate has a ceiling to their knowledge, skills and understanding, then extra time, if it is not needed to mitigate the effects of a disability, does not confer an advantage to any student (McKimm, 2012; Shaw et al, 2017). This is based on the assumption that a candidate is presumed unable to exceed the ceiling of their individual subject skill level even if they are given more time[8]. In other words, a student whose skill level is ranked at a 2:1 will not achieve a 1st class degree simply because they are given extra time. This view is termed the 'maximum potential thesis' (MPT) and supports the argument that only those candidates with a disability will benefit from additional time as those students who do not have a disability will be able to achieve their 'maximum potential' under standard time conditions.

Notwithstanding these arguments, the granting of exam access adjustments, particularly additional time, is contentious (Fullarton & Duquette, 2016). Although the purpose of exam adjustments is 'to meet the candidate's particular needs without affecting the integrity of the assessment' (JCQ, 2020, p4), critics of the practice argue that, rather than acting as a mechanism for fair treatment of students with SpLD, the granting of additional time, especially in conjunction with the use of a word processor, confers an unfair advantage and compromises the comparability of exam results across the student cohort (Deckoff-Jones, 2018; Lewandowski et al, 2013; Zuriff, 2000). This pivots on the assumption that all students achieve better grades if they are permitted additional time in exams. That is to say, additional time enables all students to produce better quality responses by more fully addressing the question, including a greater wealth of detail and, depending upon the type of exam, complete more of the questions on the paper. It is also posited that, for those students who are able to word process at a fast rate (i.e. faster than normal handwriting speed), permission to use a word processor in addition to extra time in exams confers a further 'time gain' (Deckoff-Jones, 2018; Krebs, 2019).

In addition, critics of the concept of exam accommodations argue that the extent to which exam arrangements threaten the validity of the exam is directly related to the impact the exam adjustment has on the construct being measured (Bachan, 2017; Sokal & Wilson, 2017; Lovett & Lewandowski, 2015). This raises the question; 'does extra time itself introduce construct

[8] This also assumes there are a finite number of marks allocated to each exam question and the exam time limit is determined such that the exam is fully achievable for TD candidates in the standard time allocated.

irrelevant variance if the accommodation changes the construct being measured?' (Gregg & Nelson, 2012, p134). Thus, academic staff in university settings are concerned that, rather than achieving parity with their peers, students with SpLD who are granted additional time in exams, as well as those using a word processor, are gaining an advantage over their TD peers thereby undermining the robustness and validity of the exam process (Deckoff-Jones, 2018; Krebs, 2019; Niemimen, 2020).

1.3 The legal framework in the HEI setting

Universities have a statutory duty under the Equality Act (2010) to make reasonable adjustments for disabled students in order to ensure that they are not placed at a substantial disadvantage compared to non-disabled peers (Equality Act 2010). This includes adjustments to the methods by which students' competences in their subject are assessed[9]. Part 6 of the Equality Act (2010) covers the duty imposed on Higher Education Institutions (HEIs) in relation to assessment methods and requires HEIs to make 'reasonable adjustments' to exams in order to make them more accessible to students with impairments or specific learning difficulties. In addressing this statutory duty, universities develop local policies in relation to granting exam access arrangements with the following purposes:

- 'to relieve a student of any substantial disadvantage that might arise as a result of their disability if the examination were to be conducted under standard conditions
- to ensure fairness to other candidates; individual arrangements should not give one student an advantage over others
- to preserve the integrity of the examination for the full cohort'[10]

These policies aim to ensure that universities meet their statutory duty to redress any disadvantage experienced by an individual student on account of a disability, while simultaneously establishing a framework that ensures fairness to the whole cohort.

[9] It is important to differentiate between the competence standard itself (for which there is no duty to make adjustments) and the assessment of that standard (to which the duty to provide reasonable adjustments applies)
[10] 'Staff Guidance Notes on Examination Access Arrangements' (Student Registry, University of Cambridge, 2019)

1.4 Models of disability

At this juncture its worth being aware of the different ways in which disability is conceptualised and how these different views impact on underpinning approaches when Universities determine policies for meeting their statutory duty under the Equality Act to avoid disadvantaging students with SpLD in exams.

The World Health Organisation (WHO) defines disability as the outcome of the interaction between the features of an individuals' impairment (in this case, specific learning difficulties) and features of the environment that the individual is participating in (in this case, closed book, timed, written exams) (WHO, 2021, p11). This interaction causes restrictions of participation (in this case, difficulties being able to access the exams under standard conditions), which are thus experienced as disabling:

> 'Disability is an evolving concept and results from the interaction between persons with impairments and attitudinal and environmental barriers that hinders their full and effective participation in society on an equal basis with others' (UN Convention of the Rights of persons with Disabilities, 2006).

Models of disability have been developed over many years as a means to conceptualise societal views and approaches towards disability (Lawson & Beckett, 2020; Shakespeare 2015). While there are a number of models of disability that have been posited, each with varying characteristics (some overlapping, some complementary, and some oppositional), most are located along a continuum that spans the medical model and social models of disability. At one end of this continuum, the medical model perceives the disability arising from the individual's inherent limitations (independent of their environment), while towards the other end of the continuum the social model perceives the disability arising from social or physical barriers in the environment that prevent full participation, rather than disability being an attribute of the individual. The medical model focuses on curing the individual of their impairment so that they can function more closely to the 'norm' in society, while the social model focuses on changing the environment to remove societal barriers and so enable full and equal participation (Barnes, 2012; Haegele & Hodge, 2016; Oliver & Barnes, 2012). It is worth noting that these models form part of a continuum that may overlap or converge in models that suggest that impairments can have a significant disabling impact, which changes to the environment may not be fully able to ameliorate, and so the disability arises from a combination of biological, psychological and social factors; albeit social and environmental barriers may play the most significant role

(Levitt, 2017). A comparison of the key characteristics of the medical and social models of disability is shown in Table 1.1

Table 1.1: Comparison of the medical and social models of disability

Model	Medical	Social
Philosophy	In person deficit model (the individual is 'faulty' and disabled by their bodies). Disability is the result of the impairments that are intrinsic to the individual. The in-person deficits are not 'normal', cause disadvantage, and so exclude them from full participation Disability is negative Disabled individuals are dependent on society	Social construct model (the individual is disabled by society) Disability is a form of social oppression that is imposed on individuals with an impairment. Barriers in the environment prevent individuals with an impairment from fully participating in that environment Having an impairment is neither positive nor negative Disabled individuals are contributors to society
Ontology	The in-person deficits need to be 'cured' or 'normalised' Change is the responsibility of the individual	Societal practices disable the individual, not their impairment. The individual is valued. Change is the responsibility of the environment
Location of disability	In the person	In the interaction between the person and their environment
Remedy	Adjustments to the person, to 'cure' or 'normalise the person	Adjustments to the environment to remove the barriers
Agent of Remedy	The professional	The individual, or advocate, who effects change to the barrier in the environment

1.4.1. Social and medical model perspectives of exam access arrangements

Relating these models to the context of timed, written closed book exams, the medical model would view the difficulty that individuals with SpLD experience in timed, written exams as the result of their impairment which imposes limitations, irrespective of the environment (the focus being on in-person deficits that cause the problem). The social model, by contrast, would view the difficulty that individuals with SpLD experience in timed, written exams as a result of an exam environment that is inaccessible for individuals with this cognitive profile and places barriers in their path (the focus being on barriers in the environment that cause the problem). A comparison of the medical model and social model's perspectives of exam adjustments is shown in Table 1.2.

Table 1.2. A comparison of the medical model and social model perspectives on exam arrangements.

	Medical model	Social model
Philosophy/ ontology	The individual's impairment disables them. Their deficits cause them to underperform in timed, written, closed book exams. Timed, closed book, written, exam environment is predicated on a non-disabled 'norm'. Students with SpLD need to function in this existing environment. Add on adjustments to the standard exam conditions to allow access to individual candidates with SpLD. Different exam arrangements for specific 'disabled' groups. Adjustments seen as potentially undermining robustness and reliability of exams.	The barriers that timed, written, closed book exams present to candidates with SpLD are disabling, and so they are under-rewarded in this environment. A move towards universal design for assessment and diversified assessment, predicated on an assessment environment that is fully accessible for all. Change the assessment environment to one that promotes equity in assessment and reduces barriers. Assessment environment needs to be designed to be fully inclusive for all; overcoming need to implement 'add-on' adjustments for any given individual. Same assessment environment for all. Timed, closed book, written exams seen as structurally inaccessible. Exam adjustments aim to level the playing field but are unable to overcome the inherently inequitable mode of assessment.
Location of agency	Adjustments determined by numerical data (standard scores below mean in diagnostic assessment). Diagnosis driven. Decisions about adjustment privileges the voice and authority of professionals in the field (exclusionary, 'curing' practices). Locating the professional as the expert in what the student needs.	Decisions about assessment based on students' experience of impact. Experience driven. Assessment environment and diverse models seek the voice of the student (participatory, co-design, student centred practices). Locating the student as the expert in their own experience.
Effect on individual	Othering; segregation; stigmatising. High administrative time burden (bureaucracy involved in acquiring diagnostic evidence and administration for applying for exam arrangements).	Inclusion; participation; empowering. No additional time cost as there is no requirement to apply for different assessment conditions to that of the main cohort.
Language	Attainment gap	Awarding gap

	Under-performing Individual adjustments to standard exam practices	Under-rewarded Inclusive modes of assessment for all

Many UK Universities (including the case-study institution used in this project) assess students through timed, closed book exams and make separate adjustments (commonly extra time and the use of a word processor) for disabled students (Nieminen, 2020; Villaroel, et al, 2020). The philosophy underpinning the implementation of exam adjustments for candidates with SpLD is, arguably, closer to the medical model than the social model on the continuum. That is to say, this approach assumes a non-disabled 'norm' of standard exam practices with the requirement for students with SpLD to function in this existing environment. To obtain adjustments in order to access the exam, the student is required to obtain a medical model diagnosis showing a sufficient set of deficits to warrant the adjustments, and they then receive individual treatment based on those diagnosed deficits. By contrast, disability practitioners within universities (including this researcher) frame their practice through the theoretical lens of social model thinking (Marshall, 2021). Since 2000, National bodies setting codes of practice for disability practitioners working in Disability Services in universities have explicitly encouraged practitioners working in disability services to move away from practices that privilege a medical model approach, such as ad hoc adjustments, which may be viewed as being insufficient to achieve social justice. Instead, disability practitioners are encouraged to focus on social model approaches that aim to reduce or remove barriers to full participation, centralise the students' voice, and alter inherently disabling environments (NADP[11], ADSHE[12]).

While the purpose of this research project is to investigate whether or not the current system of granting exam access arrangements to students with SpLD achieves its purpose of providing equity of opportunity in the exam or confers an advantage to this cohort, it is through this lens of the contrasting models of disability (i.e. medical model and social model) that the literature, the data, and this researcher's position, may be viewed.

[11] The National Association of Disability Practitioners (NADP)
[12] The Association of Dyslexia Specialists in Higher Education (ADSHE)

1.5 The scale of the issue

Firstly, however, given the tension between the statutory duty to grant reasonable adjustments on the one hand and universities' obligations to safeguard the validity and robustness of the overall examination process on the other, it is worthwhile considering the scale of the potential issue. Specifically, are the incidences of students with SpLD being granted access arrangements of a magnitude that warrants an in-depth investigation?

Over the past 8 years the numbers of Higher Education (HE) students who disclosed a specific learning difficulty increased substantially from 110,195 in 2013/14 to 119,600[13] in 2019/20. This represents almost 5% of the entire HE student body in the UK and is the largest single group (36%) of all students disclosing a disability (Advance HE, 2020)[14]. These students may exercise their rights under the Equality Act 2010 to receive 'reasonable adjustments' in the form of adjustments to examination administration conditions, including permission to receive additional time and the option to use a word processor. The increase in number of HE students disclosing SpLD (and potentially eligible for exam adjustments) is shown diagrammatically in Figure 1.1.

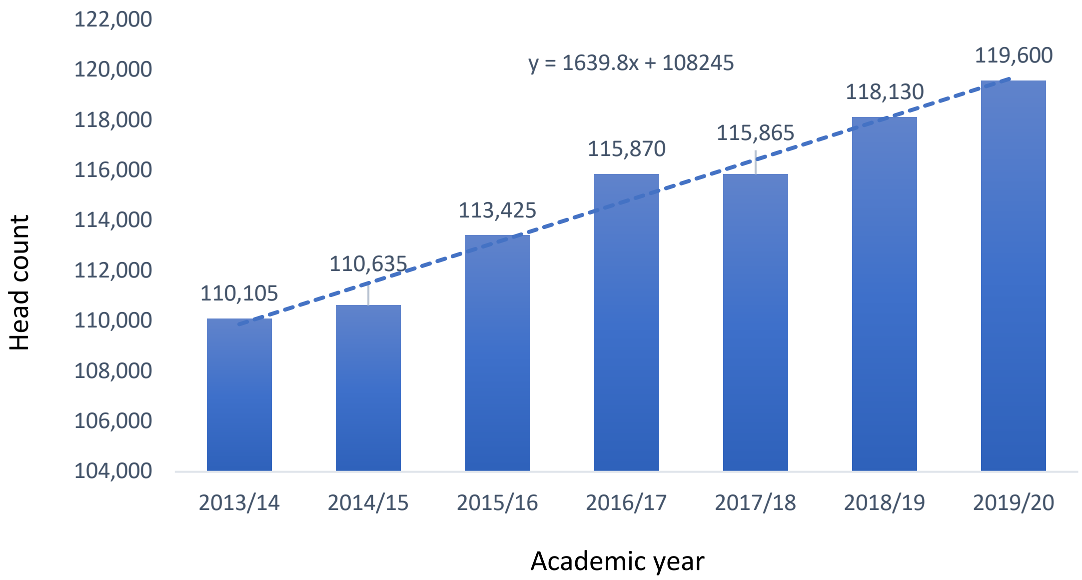


Figure 1.1: Numbers of HE students disclosing SpLD between 2013/14 & 2019/20[15]

---

[13] 90,410 of the 119,600 HE students disclosing SpLD were undergraduate students

[14] This increase in numbers of students with SpLD enrolled at HEIs is likely to be the result of a combination of factors, such as increasing rates of diagnosis in the general population, a higher number of students disclosing their condition and widening participation initiatives implemented by HEIs aimed at broadening the diversity of the student cohort.

[15] Data obtained from the HESA database HEIDI Plus (2021).

Thus, the increasing numbers of students disclosing SpLD and receiving exam access arrangements at university continues to fuel the debate about the awarding of extended time in exams and whether this promotes equity or confers an advantage to students with SpLD. This research project explores this debate by investigating whether the exam results of students diagnosed with SpLD who are permitted additional time, or use a word processor in addition to extended time, demonstrate an unfair advantage or whether their exam results suggest that the exam access arrangements have simply ensured parity by relieving the 'student of any substantial disadvantage that might arise as a result of their disability' (Equality Act, 2010).

1.6 Research objectives and Hypothesis

Having identified that uncertainties exist in relation to the effect of receiving exam access arrangements on university students with SpLD, this research project aims to explore this debate by investigating the following two-tailed hypothesis:

> The length of answers and results achieved for students with SpLD who are granted exam access arrangements (either 25% extra time or 25% extra time and the use of a word processor) will differ significantly from their TD peers who have taken the same exam under standard conditions

The independent variable is group [SpLD candidates; TD candidates] and the dependent variable is exam outcome [number of words; exam results].

The null hypothesis for this research project is:

> H_O: there will be no significant difference in the length of answer or the results achieved between the SpLD group (who have been granted 25% extra time or 25% extra time and the use of word processor) and the TD group (who have taken the exam under standard conditions)

It is also hypothesised that there will be a relationship between the independent variables (SpLD candidates; TD candidates) and the dependent variables (number of words; exam results). However, a possible outcome of this research is that the data suggests no relationship exists between the variables.

This two-tailed hypothesis is operationalised by addressing the following sub-questions:

1. Do students with SpLD who are granted 25% extra time, or 25% extra time and use of a word processor, produce significantly more words on an exam paper than their peers who have taken the same exam under standard conditions?

2. Is there a significant difference between the exam results for students with SpLD who have been granted 25% extra time (or 25% extra time and use of a word processor), and the exam results of their TD peers who have taken the same exam under standard conditions?

3. Is there a relationship between the number of words produced in an exam and the quality of exam performance, as indicated by exam grade?

4. Do students with SpLD who are diagnosed part-way through their course achieve significantly higher exam marks after they received exam adjustments than they achieved when they sat their exams under standard conditions and do any changes in their exam marks differ to the 'normal' changes in marks of their TD peers over a similar period of time?

The investigation of the hypotheses above is further augmented by the additional, qualitative phase of research that aims to further illuminate the data arising from the quantitative phases through the research question; 'what are the perceptions of students with SpLD of the impact that exam arrangements have on their exam performance?' In this way, this research project seeks to address the current lack of evidence on the impact of the use of a word processor and/or receiving 25% extra time on the exam performance of students with SpLD and consider whether the granting of these exam access arrangements places candidates with SpLD on a level playing field with their TD peers or provides them with an unfair advantage. While this project focuses on data from a single University in the UK, it is anticipated that the findings may be generalisable to all HEIs that conduct closed book, time constrained, written exams.

2: WHAT IS SPLD AND HOW DOES IT IMPACT ACADEMIC STUDY AND EXAM PERFORMANCE?

This chapter explores the definition of the term 'Specific Learning Difficulty' in order to provide an explanation of the inclusion and exclusion criteria that have been applied to this research project. The chapter starts by examining the definition of dyslexia, followed by an examination of the definition of dyspraxia, and then considers the ways in which the impact of these two diagnoses are similar (or different) within the context of study at Higher Education. This aims to provide a rationale for the inclusion of both participants with a diagnosis of dyslexia as well as those with a diagnosis of dyspraxia within this research project under a single heading of 'SpLD'.

2.1 Introduction

In any consideration of the appropriateness of adjustments granted to students with SpLD, the first question is 'what is meant by SpLD and who is included within this category'? The term 'Specific Learning Difficulty' is used to describe individuals who experience difficulties in specific areas of cognitive functioning, as opposed to general cognitive functioning difficulties, and is commonly used as an umbrella term to cover diagnoses such dyslexia, dyspraxia, dyscalculia and dysgraphia[16] (Grigorenko et al, 2019; Lockwood et al, 2021; Snowling et al, 2020; Wagner et al, 2020). For the purposes of this research project, the term 'SpLD' is being used to indicate a diagnosis of either dyslexia or dyspraxia or a dual diagnosis of both.

[16] The term Specific Learning Difficulties can also be used as a descriptor term in its own right (Lockwood et al, 2020; Maki & Adams, 2018; Snowling et al, 2020)

Dyslexia and dyspraxia are currently classified as neurodevelopmental disorders and are characterised by deficits in short-term and working memory[17] as well as in visual and verbal processing speed (Gooch et al, 2016; Fletcher et al, 2019; Kearns et al, 2019; Snowling et al, 2020). Individuals with either dyslexia and/or dyspraxia can, to varying degrees, experience difficulties with motor skills[18], and executive functioning (affecting concentration, time-management, and organisation) (Brimo et al, 2021; Fletcher & Miciak, 2019; Mountford et al, 2021). However, individuals with dyslexia experience deficits in phonological awareness (Goswami, 2019; Layes et al, 2020; Moshkani et al, 2017; Snowling et al, 2020), while those with dyspraxia present with deficits in the motor domains of motor planning, motor learning, proprioception, sensorimotor coordination, and poor postural control (Barnett & Prunty, 2020; Brown-Lum et al, 2020; Mountford et al, 2021, Tseng et al, 2019).

There is agreement that both dyslexia and dyspraxia are lifelong conditions, from childhood through adulthood, but the characteristics can, to varying degrees, be masked and modulated by the development of coping strategies and other environmental factors, delaying in some cases the diagnosis to adolescence or adulthood (Camilleri et al, 2019; Camilleri et al, 2020; Snowling et al, 2020; Walker et al, 2021). Current studies suggest that between 3 - 20% of the population may be affected by dyslexia and between 1.4% - 19% by dyspraxia, although around 5 – 8% appears to be the most commonly accepted prevalence rate for both conditions[19] (Adlof & Hogan, 2018; American Psychiatric Association, 2013; Biotteau et at, 2017; Gentle et al, 2021; Meachon, 2017; Visser et al, 2020; Wagner et al, 2020). At the same time, high rates of co-morbidity and overlapping traits between the two conditions are commonly reported[20] (Brimo et al, 2021; Meachon, 2017). Both dyslexia and dyspraxia appear to be multi-factorial, complex, disorders with family, twin and adoption studies strongly suggesting a genetic factor related to the conditions (Centanni, 2020; Guidi et al, 2018;

---

[17] Working memory refers to an individual's ability to process information and store the products of that processing over a brief period of time while carrying out a task (Ghani and Gathercole, 2013).
[18] It is uncertain whether or not the motor difficulties identified in individuals with dyslexia arise from comorbidity or are a feature of dyslexia itself (Alamargot, 2020; Brimo et al, 2021).
[19] These wide differences in prevalence estimations arise from the lack of a universally agreed definition of SpLD, as well as variances in diagnostic methods & diagnostic criteria (Lockwood, 2021; Fletcher & Miciak, 2019).
[20] Some studies suggest that the 'susceptibility genes' identified in dyslexia are associated with a generalised genetic liability to other SpLDs, which accounts for high co-morbidity between dyslexia and dyspraxia (Brimo et al, 2021; Goulardins et al, 2015; Meachon, 2017).

Mountford et al, 2021), with recent genome studies identifying a range of candidate genes and variants that might help explain the pathology (particularly in the case of dyslexia) (Beider et al, 2020; Brown-Lum et al, 2020; Centanni, 2020).

2.2 Defining Dyslexia

The history of establishing an agreed, categorical, definition of dyslexia has been one beset with difficulties, due to the heterogeneous, multifaceted, continuum nature of the condition that is complex in its manifestation. This complexity has resulted in disagreement about the precise nature and extent of the presenting traits of dyslexia (including the severity as well as inclusion or exclusion criteria), albeit there is agreement that dyslexia is characterised by deficits in the processes underpinning efficient literacy skills (Gooch et al, 2016; Moll et al, 2020; Snowling et al, 2020) Dyslexia, however, encompasses broader difficulties than simply those related to reading and spelling words, including weaknesses in executive functioning, motor coordination, and some aspects of mathematical skills, for example, and for some it is experienced as difficulties with the automaticity and fluency of literacy skills, rather than an actual inability to decode or encode at the single word level (Broggi et al, 2019; Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020). As such, there is disagreement about which features are core to the diagnosis and which are secondary. There is however broad consensus regarding the constellation of difficulties that contribute to a diagnosis of dyslexia, particularly in the domains of phonological awareness, short-term and working memory, visual and auditory processing speed and temporal processing[21] - manifesting as difficulties in the acquisition, accuracy, automaticity and fluency of literacy skills (Casini et al, 2017; Goswami, 2019; Snowling & Hulme, 2016; Snowling et al, 2020; Stein, 2019; Wagner et al, 2020). However, there is disagreement about the respective weighting of these traits; which are included and which are excluded, where the 'cut off' points on the continuum are located[22], and whether the deficits identified need to constitute absolute

[21] Temporal processing is defined as 'the accurate timing of auditory and visual sensory input' (Casini et al, 2017, p3) and is needed for the auditory sequencing of phonemes and the visual sequencing of graphemes (Casini et al, 2017; Goswami et al, 2019; Stein, 2019)

[22] The traits associated with dyslexia, including literacy difficulties, are reported to be continuously distributed in the population (Fletcher et al., 2019; Wagner et al, 2020).

weaknesses or relative weaknesses[23] (Fletcher et al., 2019; Ryder & Norwich, 2018; Snowling et al, 2020; Wagner et al, 2020).

This lack of consensus gives rise to a range of different definitions that focus on different aspects of the condition, ranging from a narrow focus on difficulties with literacy acquisition at the word level (World Health Organization, 2010), through the broader category of 'Specific Learning Disorders' with designated specifiers describing sub-skills used by the DSM-5[24], to the focus on an uneven cognitive profile and difficulty with the automaticity of literacy skills adopted by the British Dyslexia Association (BDA), which is currently the most commonly used definition in the UK (British Dyslexia Association, 2010)[25]:

> Dyslexia is a learning difficulty that primarily affects the skills involved in accurate and fluent word reading and spelling. Characteristic features of dyslexia are difficulties in phonological awareness, verbal memory and verbal processing speed. Dyslexia occurs across the range of intellectual abilities. It is best thought of as a continuum, not a distinct category, and there are no clear cut-off points. Co-occurring difficulties may be seen in aspects of language, motor co-ordination, mental calculation, concentration and personal organisation, but these are not, by themselves, markers of dyslexia. A good indication of the severity and persistence of dyslexic difficulties can be gained by examining how the individual responds or has responded to well-founded interventions.
>
> In addition to these characteristics the British Dyslexia Association (BDA) acknowledges the visual and auditory processing difficulties that some individuals with dyslexia can experience and points out that dyslexic readers can show a combination of abilities and difficulties that affect the learning process. Some also have strengths in other areas, such as design, problem solving, creative skills, interactive skills and oral skills (BDA, 2010).

The BDA definition of dyslexia acknowledges a wide variation in literacy skills, allowing for the inclusion not only of those who have poor functional literacy, but also those whose literacy skills are developed but lack automaticity, as is commonly the case

---

[23] i.e. whether the performance in reading accuracy or spelling, for example, needs to be weaker than the norm of the population (so a deficit in absolute terms), or weak relative to the individual's own performance in other, unaffected, areas (so a deficit in relative terms).

[24] The Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) is an internationally recognised guide to mental disorders and contains descriptions, symptoms and criteria for diagnosis of mental health disorders, including neurodevelopmental disorders (American Psychiatric Association, 2013)

[25] The SpLD diagnostic assessment reports conducted in the UK commonly use the BDA definition of dyslexia, rather than the DSM-5 or ICD-10 criteria (SpLD Assessment Standards Committee, 2020)

for students in HE (Adlof & Hogan, 2018; Livingston, Siegal & Ribary, 2018; Wagner et al, 2020). At the same time the definition recognises a range of co-occurring difficulties, including difficulties associated with executive functioning as well as some aspects of motor-coordination[26], which speaks to the view that dyslexia and other SpLDs, such as dyspraxia, present with many overlapping traits (Cheng et al, 2011; Visser et al, 2020). However, the broader, more holistic, definitions tend to be critiqued for a lack of precision and accuracy (and so the possibility of over diagnosis due to wider inclusion criteria), while the narrower definitions are critiqued for an over simplification of the complex nature of dyslexia and the omissions of many of its associated features. This over-simplification then carries the risk of underdiagnosis and the possibility that those who had a childhood diagnosis of dyslexia may no longer identify with the diagnosis in adulthood if they have made sufficient improvements in single word reading and spelling accuracy over time, despite still being affected by other associated difficulties underpinning academic skills. These wide differences in exclusionary and inclusionary criteria engender a significant variation in the estimated prevalence of dyslexia from 3% to 20% (Adlof & Hogan, 2018; Biotteau et at, 2017; Peterson & Penning, 2012; Spencer et al, 2014; Wagner et al, 2020) and also have consequences when determining which individuals are, and are not, eligible to receive support (such as exam adjustments) during their education.

Nonetheless, despite the contentious nature of defining dyslexia and the lack of consensus, there is a general agreement that dyslexia is a genetically based[27] neurodevelopmental condition, that impedes the ability to acquire, or fluently use, specific academic skills that underpin many areas of academic learning (particularly language and literacy skills, albeit other academic skills may also be affected) (Byrne, 2018; Falzon, 2020; Snowling & Hulme, 2016; Snowling et al, 2020). Core to the

---

[26] Motor coordination difficulties appear to be present in a sub-group of individuals with dyslexia However, it is unclear whether these motor difficulties are inherently linked with dyslexia itself or arise from a comorbid, albeit undiagnosed, profile of dyspraxia (Alamargot et al, 2020; Brimo et al, 2021; Visser et al, 2020).

[27] A number of studies suggest that the proteins expressed by 'susceptibility' genes (particularly DYX1C1, KIAA0319, and DCDC2) affect neuronal migration, which then determines brain connectivity (particularly between the temporal and parietal regions), resulting in the cognitive difficulties that characterise dyslexia. This neuronal migration deficit hypothesis appears to be the most commonly cited theory in the literature on the genetic basis of dyslexia (Beider et al, 2020; Carron-Castillo et al, 2017; Centanni, 2020; Gostic et al, 2019; Guidi et al, 2018; Martinez-Garay et al, 2017).

definition of dyslexia is that these learning difficulties are not accounted for by difficulties with general development, lack of learning opportunities or sensory impairments, so the difficulties are unexpected and are present from childhood, persisting into adulthood (Fletcher et al., 2019; Maki & Adams, 2018; Snowling & Hulme, 2016; Snowling et al, 2020). Dyslexia is considered to be a cross-cultural condition, although its presentation can be affected by cultural differences (such as differences in language and orthography) and developmental changes (Maki & Adams, 2018; Meachon, 2017; Snowling & Hulme, 2020). Although dyslexia can present differently in different individuals (and the nature, extent and the impact of dyslexia on the individual can vary according to the nature of the task, the educational context, and the range of compensatory strategies acquired) what is agreed is that the deficits generally seen as associated with dyslexia can have a material negative impact on the individual's educational achievements and examination performance (Byrne, 2018; Kendall, 2018; MacCullagh, Bosanquet & Badcock, 2017; Majer, 2018; O'Byrne, Jagoe & Lawler, 2019; Shaw & Anderson, 2018).

The lack of a universal definition of dyslexia, arising from the heterogeneous nature of the condition and individual variations across the spectrum, is further problematised by an insufficient understanding of the precise aetiology of dyslexia. The neurobiological mechanisms underlying dyslexia are not fully understood, despite much research in this area, and remain the matter of continued debate, with each theoretical position failing to account for all the cognitive and behavioural differences identified (Callens & Brysbaert, 2019; Fletcher & Miciak, 2019; Georgiou et al, 2018; Gostic et al, 2019; Goswami 2019; Kimel & Ahissar, 2020; Paracchini et al. 2016). Nonetheless, a range of studies exploring the cognitive profile of Higher Education students with dyslexia identify deficits in spelling accuracy (especially in continuous writing), writing speed, pseudo-word reading, text reading speed and accuracy, visual and verbal processing speed, phonological awareness, visual short-term memory and auditory working memory as common characteristics of the profile (Callens & Brysbaert, 2019; Casini et al, 2017; Hatcher, Liebel and Nelson 2017; Goswami et al, 2019; Stein, 2019). These studies concluded that these range of deficits affect the underpinning skills that students in HE rely on to access the various elements of their academic course of study (such as taking notes in lectures, reading texts & making notes, planning, organising & structuring ideas, writing essays, revision and exam performance) and thus negatively impact the educational performance of students with dyslexia, both during their course (in teaching and learning) as well as during timed, closed book, written examinations.

2.3 Defining Dyspraxia[28]

Like dyslexia, dyspraxia is a heterogenous condition (Cacola & Lage, 2019; Sumner et al, 2016; Wade & Kazeck, 2018; Wilson et al, 2017), with some studies suggesting that, rather than representing a single category, dyspraxia may be more accurately described as a series of more homogenous sub-groups, sharing common traits and causation (Visser et al, 2020; Wade & Kazeck, 2018). The Dyspraxia Foundation defines dyspraxia as a visual perceptual and motor co-ordination difficulty, without a defined neurological or physiological cause, that manifests as 'an impairment of the organisation of movement, which has associated problems of language, perception and thought' (dyspraxiafoundation, 2019). Dyspraxia UK and Movement Matters developed the following descriptor of dyspraxia in 2018:

> Developmental coordination disorder, also known as dyspraxia in the UK, is a disorder affecting movement and coordination. The difficulties are present in childhood and persist into adulthood, although non-motor difficulties may become more prominent as expectations and demands change over time. A range of co-occurring difficulties can have a substantial impact on life…including difficulties with time management, planning, organisation. (Movementmatters.uk.org, 2018).

This definition acknowledges that, whilst the core features of dyspraxia relate to motor impairment (including the precursory aspects of planning of the movement), dyspraxia also affects 'planning and organisation at the internal cognitive level, involving the organisation and structuring of language, thought and ideas' (Movementmatters.uk.org, 2018; Gentle et al, 2020; Meachon, 2017; Visser et al, 2020; Walker et al, 2021; Wilson

---

[28] The term 'dyspraxia' rather than 'Developmental Coordination Disorder' is used throughout this thesis. The diagnostic assessments that constitute evidence for the inclusion of the participants in this study were conducted by educationally trained diagnostic assessors (such as educational psychologists or specialist teachers) who focus on the cognitive profile of 'dyspraxia' (such as short term and working memory, processing speed, etc) and the impact this profile has on education, rather than by medically trained assessors (such as Occupational Therapists) who assess the impairment in movement, physical coordination and motor skills necessary for a diagnosis of Developmental Coordination Disorder. Educational assessments use the term 'dyspraxia' as they are not medically qualified to formally assess physical motor and coordination skills. As the focus of this study is on the impact of the cognitive profile on education (primarily in exams), the educational assessment of 'dyspraxia' is more relevant in this context.

et al, 2017). Executive functioning, including time management, planning and organisation are common challenges (Blank et al, 2019; Cacola & Lage, 2019; Jascenoka et al., 2015; Wade & Kazeck, 2018; Walker et al, 2021) as are problems with visual perception, speed of processing, visual short-term & working memory, retrieval of factual information and sequencing. These deficits cause problems with academic activities including difficulty sorting incoming information quickly and imposing structure on tasks (Blank et al, 2019; Cacola & Lage, 2019; Dewey & Bernier, 2016; Meachon, 2017; Sumner et al, 2021; Visser et al, 2020; Walker et al, 2021; Wilson et al, 2017). The Dyspraxia Foundation also identifies that individuals with dyspraxia can experience problems with literacy acquisition as automatic and efficient literacy skills (such as accurate spelling when writing continuous prose under time pressure) can be affected by deficits in working memory, while the impaired speech processing system of those with dyspraxia who experience persistent speech difficulties can affect phonological tasks, such as segmenting and blending sounds, which underpins accurate literacy skills (dyspraxiafoundation.org.uk, 2021). Individuals with dyspraxia can have difficulty organising the individual parts of an activity into a whole as assembling, manipulating, sequencing and organising both verbal and visual information can be challenging (Cacola & Lage, 2019; Gentle et al, 2020; Sumner et al, 2021; Visser et al, 2020; Walker et al, 2021;Wilson et al, 2017). While individuals with dyspraxia may have unimpaired ability to decode text (i.e. to read in isolation), they experience difficulties with the comprehension of the texts (i.e. reading in context) unless additional time is provided as, despite sufficient levels of accuracy at the single word level, bringing the parts of the text together to create meaning and enable understanding takes unexpected time and effort (Cacola & Lage, 2019; Leonard & Hill, 2015; Meachon, 2017; Sumner et al, 2021; Walker et al, 2021). Individuals with dyspraxia may also spell with accuracy at the single word level, but have a tendency to make unnoticed 'clerical' errors when writing in continuous prose. At the same time, organising thoughts, sequencing ideas, making connections in written composition, developing a logical argument, and creating structure when writing can be challenging and take significant time and effort, thereby negatively affecting essay composition skills (Blank et al, 2019; Sumner et al, 2021; Visser et al 2020; Walker et al, 2021).

As the Dyspraxia UK & Movement Matters (2018) definition highlights, the non-motor difficulties associated with dyspraxia 'may become more prominent as expectations and demands change over time' (Movementmatters.uk.org, 2018). Studies suggest that although the impact of gross motor difficulties associated with dyspraxia may appear to diminish with brain maturation, and as a consequence of remedial intervention during childhood, difficulties with organisation and time management, academic activities (including reading speed and essay writing), attention, planning and the speed and legibility of handwriting persist, and are particularly acute when the student is faced with a new task. Thus the difficulties of students with dyspraxia in HE in the context of their course of study tend to manifest as problems with taking notes and simultaneously listening in lectures, difficulties absorbing meaning from texts at first reading (resulting in the need to re-read copiously), a slow rate of reading, problems summarising from texts, 'clerical' errors in writing, difficulties organising and sequencing ideas in writing, editing and proofreading difficulties and generally working at a slower pace. This presentation is similar to that seen in students with dyslexia and may be attributed to the slow processing, working memory deficits, sequencing difficulties, slow handwriting and deficits in the automatisation of skills shared by both diagnoses (Biotteau et al., 2017; Blank et al, 2019; Cacola & Lage, 2019; Leonard & Hill, 2015; Walker et al, 2021). The underlying deficits of memory common to both dyspraxia and dyslexia (short term and working memory) mean that information cannot be held for long in the mind, despite effort, and so rapidly decays (Kirby, 2012; Sumner et al, 2021). This results in further disorganisation as the individual is unable to recall what has just been read, needing to frequently re-read in order to prompt the memory into retaining information (Blank et al, 2019; Meachon, 2017; Sumner et al, 2016; Sumner et al, 2021; Walker et al, 2021). Slow processing speeds further undermine the execution of reading and writing tasks as information must be held in the mind for longer in order to complete a task, placing further demands on weak working memory (Blank et al, 2019; Cacola & Lage, 2019; Visser et al 2020; Wade & Kazeck, 2018; Walker et al, 2021).

As can be seen from the descriptions above, many of the characteristics associated with dyspraxia similarly occur in students with dyslexia, especially those related to executive functioning, processing speed and memory deficits (Broggi et al, 2019; Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020). Even difficulties with handwriting skills are traits often shared by both diagnoses, albeit potentially for differing reasons. For example, the literature posits that the handwriting difficulties prevalent in students with dyslexia may relate to visual perceptual deficits, graphomotor difficulties, and the impact of inefficient linguistic processes on motor execution, whereas the handwriting difficulties identified in the students with dyspraxia may relate primarily to deficits in motor planning & sequencing, coordination and motor control (Gosse & Reybroek, 2020; Suarex-Coalla et al, 2020). Nonetheless, the behavioural manifestation of handwriting difficulties results in similar challenges and increased cognitive burden in exams (due to the lack of automaticity and greater cognitive effort involved) for both students with dyslexia and those with dyspraxia, regardless of the underpinning causation, due to the need to handwrite responses (Barnett & Brunty, 2020; Connelly et al, 2005; Ukwueze, 2015). In addition, students with dyspraxia, as well as those with dyslexia, report problems with working memory which affects recall and the ability to take notes while listening as well as difficulties with absorbing text-based information at speed (resulting in a need to re-read copiously in order to take in meaning), as well as problems with attention, organisation, time management and with integrating information (Blank et al, 2019; Biotteau et al., 2017; Cacola & Lage, 2019; Kirby et al, 2010). Thus the research suggests that the difficulties that HE students with dyslexia and those with dyspraxia experience in academic activities as a result of their diagnosis have more in common than is different.

2.4 Dyslexia and Dyspraxia in Higher Education: More commonalities than differences

As outlined above, both dyslexia and dyspraxia are a constellation of difficulties and, in the context of students enrolled in HE, share more commonalities than differences. Homotypic comorbidity within SpLDs is also high with 50% to 80% of individuals with dyslexia also presenting with significant motor coordination difficulties[29] (Dewey & Bernier, 2016; Margari et al, 2013; Moll et al, 2020), while around 70% of individuals diagnosed with dyspraxia present with significant reading problems (Biotteau, 2017; Meachon, 2017; Wade & Kazeck 2018). This is greater than would be expected by chance. However, while the high frequency of co-occurring traits between dyslexia and dyspraxia is well established, the reason for this co-occurrence and

[29] Margari et al., (2013) suggest that the cerebellum, that is responsible for motor control, is also involved in automating overlearned tasks (such as literacy skills), creating an overlap in the deficits observed between dyslexia and dyspraxia in motor control.

association is contested. Some studies suggest that the two conditions share a number of aetiological bases that affect cognitive processes such as executive functioning, memory, automaticity of learned tasks, attention, visuo-spatial regulation, language and motor skills (Brimo et al, 2021; Dewey & Bernier, 2016; Moll et al, 2020; Nicholson & Fawcett, 2011). Others, however, suggest that, while there is potentially some sharing of genes involved in the development of the neural network between dyspraxia and dyslexia, there are other genes that are unique to each diagnosis with the result that the surface similarities in traits shared by the two conditions may, at times, arise from different underlying causes (Cacola & Lage, 2019; Kirby et al., 2010; Wilson et al, 2017).

The overlapping traits shared between dyslexia and dyspraxia, regardless of aetiology, are particularly evident in the skills used in academic study. Key areas that are associated with both dyslexia and dyspraxia in the context of academic learning and assessment relate to poorer automaticity[30] of literacy skills (affecting spelling and reading under time pressure), reading speed (and the need to frequently re-read in order to absorb information), handwriting fluency, retention and recall of information (short term & working memory and retrieving verbal information from long-term memory), processing speed, executive functioning (including organisation, planning and sequencing of ideas, time management, and attention), as well as with the academically related skills of proofreading, skimming & scanning and summarising texts (Broggi et al, 2019; Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020; Wade & Kazeck 2018; Wilson et al, 2017). Deficits in these areas are shared between students with dyslexia and those with dyspraxia and impact the abilities of both groups of students to access their course efficiently, revise their material and perform to their potential in exams. Thus, the evidence suggests, students with dyslexia experience similar challenges with the access skills needed to demonstrate their potential in their academic course and exams to those experienced by students with dyspraxia, even if there is a lack of consensus about whether these difficulties arise from similar or different underpinning aetiological bases. These shared traits and similarities are shown in Table 2.1.

[30] Automaticity is beneficial since it allows tasks to be executed more efficiently and quickly.

Table 2.1: Dyslexia and dyspraxia: similarities and differences[31]

Diagnostic Category	Underpinning deficit	Impact on academic activity in HE
Dyslexia only	Phonological awareness	Decoding and encoding accuracy
Dyspraxia only	Motor planning; motor learning; proprioception, sensorimotor coordination; sensory sensitivities; poor postural control	Handwriting speed and legibility (may also be experienced by students with dyslexia)
Both dyslexia & dyspraxia	Short term and working memory; retrieval of information from long-term memory; visual and auditory processing speed; rapid automatised naming; executive functioning; visual perception; automatisation of skills; motor coordination	Prose reading accuracy and speed; spelling accuracy in continuous writing and 'clerical' errors; comprehension under time pressure (need for frequent re-reading); handwriting fluency; planning, organising and sequencing ideas in speech and writing; memorisation and recall; time management; proofreading; skimming and scanning; summarising; note-taking & note-making; written expression; sustained concentration & focus; distractibility; general organisation; writing speed; imposing structure on tasks; sorting incoming information at speed; assembling, manipulating, sequencing and organising verbal and visual information.

As the definitions of dyslexia and dyspraxia cited by the BDA and Dyspraxia UK earlier suggest, the constellation of difficulties associated with dyslexia and dyspraxia become arguably more subtle as the individual moves into adulthood, albeit improvements that occur with maturation are not sufficient to adequately level the playing field with their TD peers (Biotteau, 2017; Snowling et al, 2020; Wade & Kazeck 2018). This may be even more acutely apparent in students in HE whose presenting characteristics may be less associated, for example, with significant impairments in gross motor control (in the case of dyspraxia[32]) or with functional literacy (in the case of dyslexia), and more associated with the automaticity of these skills (Blank et al, 2019; Cacola & Lage, 2019; Callens & Brysbaert, 2019; Stein, 2019). Thus, for students in Higher Education, the presenting difficulties of the two

[31] Table contents represents a summary of the research by Biotteau, 2017; Brimo et al, 2021; Broggi et al, 2019; Callens & Brysbaert, 2019; Cacola & Lage, 2019; Dewey & Bernier, 2016; Falzon, 2020; Meachon, 2017; Moll et al, 2020; Sayeski, 2019; Snowling et al, 2020; Wade & Kazeck 2018; Wilson et al, 2017.

[32] It should be noted that the participants with dyspraxia in this study generally have an educational diagnosis, rather than a medical one. As a result, the focus of the assessment is on the cognitive profile rather than any direct assessment of motor and coordination skills. The motor coordination difficulties are primarily identified through a background history in the educational assessment, rather than a direct motor assessment.

conditions in the context of academic study are potentially more closely aligned with each other than may be seen during childhood or is evident in the wider population. Given that comorbidity is the norm and that the characteristics of each of the two conditions, in the context of HE, have more in common than is different, for the purposes of this research project students with a diagnosis of either dyslexia and or dyspraxia have been included as a single, heterogeneous, group under the umbrella term of 'SpLD'. This aims to reflect the 'real world' situation of students in HE[33] in order to make the research exploring the performance of students with dyslexia and/or dyspraxia in exams generalisable to the cohort as a whole. To create a sample group of students with a sole diagnosis of dyslexia, for example, would fail to reflect the true nature of students in HE with SpLD (where co-morbidity is the norm) and thereby reduce the application of the outcomes of this research project to the real-life situation in HEIs.

2.5 Conclusion

Dyslexia and dyspraxia are complex, heterogenous, conditions that share many traits and commonly co-occur (Brimo et al, 2021; Meachon, 2017). Family, twin and adoption studies suggest that the cognitive profiles associated with both dyslexia and dyspraxia are likely to be genetically influenced and certain 'susceptibility' genes have been identified that influence neural migration, which then shape the neuro-anatomy and brain connectivity in key regions (Darki, Peyard-Janid, Matsson et al, 2012; Carron-Castillo et al, 2017). Although a comprehensive picture of the pathogenesis and the contribution of gene-environmental interplay in dyslexia and dyspraxia is still not fully known, the evidence paints a compelling picture of a complex interaction of genes and environment that results in structural and connectivity brain differences in individuals with dyslexia and dyspraxia, which cause cognitive differences that then result in the presenting characteristics of dyslexia and dyspraxia (Brimo et al, 2021; Goulardins et al, 2015; Meachon, 2017). These cognitive differences persist into adulthood and can negatively affect educational experiences by impacting the efficiency of the construct-irrelevant skills that Higher Education students rely on in

[33] Many students with SpLD in HE have a dual diagnosis, and those who have a sole diagnosis of, say, dyslexia commonly present with characteristics of dyspraxia in addition and vice versa, even if these additional characteristics did not meet the threshold level for formal diagnosis (Sumner et al, 2021)

order to demonstrate their subject knowledge and ability in timed exams (including the speed with which they read and digest large quantities of complex information, the speed and efficiency of recall of information, as well as with the planning, organisation and expression of ideas in writing at speed in a way that reflects their ability). This connection between genetically influenced structural brain differences that manifest themselves in differences in cognitive performance and negatively impact the construct irrelevant skills needed to access exams, leading to the granting of exam adjustments, is shown diagrammatically in Figure 2.1.

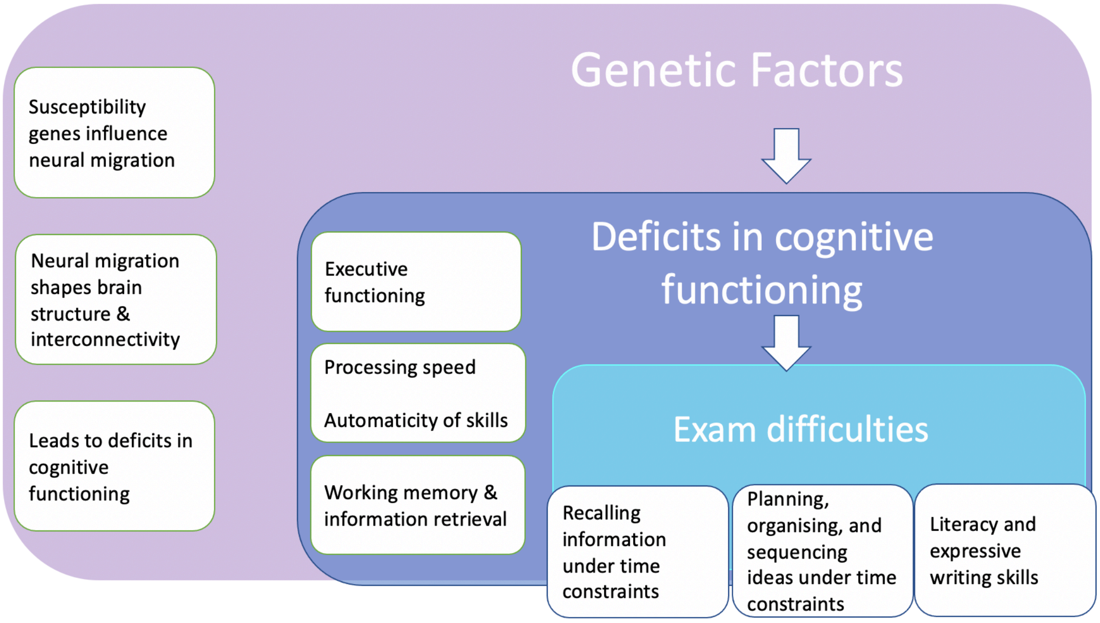


Figure 2.1: From genes to exam difficulties: The underpinning rationale for awarding exam adjustments to students with dyslexia and/or dyspraxia

As a result of their cognitive profile, students with dyslexia and/or dyspraxia need extra time in exams to read and process the question, plan, organise and sequence ideas and execute the essay and also need to use a word processor to compensate for poor working memory and slow or illegible handwriting. Without these adjustments, candidates with dyslexia and/or dyspraxia are subject to construct-irrelevant variance in their exam results, with their marks reflecting their deficits in the automaticity of their literacy skills, speed of processing, executive functioning and working memory, rather than their subject knowledge, critical thinking skills and understanding (i.e. the construct that the exam is actually aiming to measure) (Callens & Braybaert, 2019; Harrowell et al, 2018; Shaw & Anderson, 2018; Shaw et al, 2017).

In summary, the current body of literature suggests that the underlying intellectual capacity of students with SpLD entering higher education is no different to that of their TD peers, however the characteristics associated with dyslexia and/or dyspraxia negatively impact on their ability to access their learning when they are required to rapidly process and recall new information from reading (i.e. text books) and through listening (i.e. lectures and tutorials) as well as when asked to demonstrate their understanding of their subject in written form at speed (i.e. timed exams) (Callens & Braybaert, 2019; Harrowell et al, 2018; Shaw & Anderson, 2018; Shaw et al, 2017). These difficulties need to be accommodated if students with SpLD are to demonstrate their potential in their course of study, including being granted compensatory measures in the context of timed examinations, such as extra time plus the option to use a word-processor. However, in order to evaluate whether these exam arrangements compensate for the disadvantage that candidates with SpLD would otherwise experience in exams in comparison to their TD peers, due to their cognitive profile, or whether the changes to standard exam conditions confer an advantage to students with SpLD, it is necessary to analyse the empirical research that has already been conducted into the effects of granting extra time to individuals with SpLD, so that exam adjustment policy decisions in HEIs are robustly underpinned by a rigorous evidence base.

3: LITERATURE REVIEW:
THE IMPACT OF EXTRA TIME ON EXAMS

This chapter reviews the current body of research within which the quantitative phase of this research project is conceptually located, and identifies the gap that this research project aims to address: namely, the lack of consensus about the impact that extra time (or extra time and the use of a word processor) has on the exam performance of students with SpLD and whether these adjustments promote equity or confer an advantage.

3.1 Introduction

The accommodation-disability interaction paradigm establishes a theoretical framework for reasonable adjustments such as exam access arrangements (Elliot & Marquart 2004; Phillips, 1994). This paradigm defines a valid exam adjustment as one that removes the barriers to accessing the exam that a disabled candidate experiences solely due to their disability. In the case of candidates with SpLD, the accommodation-disability interaction paradigm would suggest that extra time represents an appropriate adjustment in timed exams where the standard time limits in exams act as a barrier preventing candidates with SpLD performing to their potential due to the information processing deficits associated with SpLD (provided that speed is not a competency standard being assessed by the exam). In these circumstances, additional time in exams would be an appropriate adjustment as this should compensate for the deficits in the underpinning skills needed to access the exam and so enable the candidate with SpLD to demonstrate his/her potential. The literature cited in the previous chapter suggests that individuals with SpLD present with genetically influenced neuro-anatomical differences to their TD peers which lead to cognitive deficits that negatively impact on the fluency with which a student with SpLD in Higher Education (HE) is able to express ideas and show his/her subject knowledge and critical thinking skills in timed, closed book, unseen written exams. Provided that these skills do not represent the competence standards being assessed, these skills are not the constructs that the exam is intending to measure. Thus, the accommodation-disability interaction paradigm suggests, to compensate for these construct-irrelevant deficits, a candidate with SpLD requires (and should be granted) more time in exams than TD peers to perform the range of cognitive activities required for producing a written response to an exam question.

At this junction, it is worthwhile being mindful of the medical and social models of disability referred to in the introduction, as this framework can act as a useful lens through which the literature may be viewed. As previously mentioned, the practice of applying separate adjustments for disabled students to the standard exam practice that is predicated on a non-disabled 'norm', speaks to the medical model of disability (i.e. individual treatment based on an inherent set of deficits) rather than the social model (which would advocate for a change to the format of assessment based on principles of inclusion and universal design). This medical model approach, whereby students with SpLD receive different treatment in exams, potentially leads to students with SpLD being 'othered' by institutional exam policies, and 'outed' as they are visibly set apart from their peers. This can provoke feelings of being different and lesser and negatively impact academic self-esteem (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al., 2019; Lightfoot, 2018; Majer, 2018). The literature in the field that evaluates the impact of exam arrangements in terms of achieving parity or advantage focuses solely on the outcomes for students in terms of the quantitative data (i.e. whether or not the student who receives extra time is over or under-rewarded in the exams in terms of the marks they receive) but conspicuously ignores the impact of the experience of this different treatment on the student and the resultant negative ramifications of being 'othered' and outed.

Bridgeman, Cline & Hessinger (2003) and Verleger (2016) suggest that there are two main reasons why an exam imposes time constraints. The most common reason for an exam to be time constrained is simply due to administrative expediency and cost. By imposing a time limit on exams, Higher Education Institutions (HEIs) can develop a timetable that is manageable for both candidates and examination staff throughout the exam season in the most cost-effective way. A second, but less common reason for a time limit being imposed on an exam occurs where the speed of the candidates' performance is an intrinsic part of the construct under examination (such as in a typing speed test, for example). In this context the exam will be strictly timed in order to identify the individual differences in speed of responses between candidates and allowing extra time would thus not be appropriate. In exams where speed is not the construct being examined, the 'speededness' of an exam (as determined by the standard time limit) may threaten the exam's validity if the exam is primarily aiming to gauge the critical skills, knowledge and ability of the candidate rather than the speed at which the candidate responds. In exams where speed is construct irrelevant, the granting of extra time to candidates who work more slowly (due to an SpLD, for example) would be an appropriate means of promoting parity and ensuring that the differences in scores between candidates reflect the differences in their subject knowledge, not in their speed.

As a general principle, exam adjustments, such as additional time, are intended to normalise the performance of candidates with SpLD to that of their TD peers and ensure that the differences in scores between candidates reflect the differences in their subject ability and are not the result of their disability. They are not intended to confer an advantage (Bolt, Decker, Lloyd & Morlock, 2011; Portolese, Kreuse & Bonner, 2016). That is to say, granting extra time in exams to candidates with SpLD aims to compensate for the impact of slower processing of information and should level the playing field with their TD peers. Any improvement in performance by the SpLD candidate in receipt of extra time should thus represent a more accurate score that is a valid representation of the student's subject knowledge and skills. It should not simply result in an increase in marks that represents an over-inflated score as this would challenge the validity of the exam results (Bachan, 2017; Sokal & Wilson, 2017; Lewandowski, Cohen, & Lovett, 2013; Pardy, 2016).

In order to identify if the granting of extra time achieves its purpose of achieving equity, rather than according an advantage, four different theoretical frameworks need to be critically examined. Firstly, some researchers propose that, if the granting of extra time in an exam to candidates with SpLD is valid, the extra time should only improve the performance of the students with SpLD and not the performance of their TD peers who are able, it is assumed, to reach their maximum potential in the standard time allocated (this is known as 'the maximum potential thesis') (Cohen, Gregg & Deng, 2005; Crisp, Johnson & Nivakovic 2012; Hill, 1984; Huesman & Frisbie, 2000; Lesaux, Pearson & Siegal, 2006; McKendree & Snowling, 2011; Portolese, Kreuse & Bonner, 2016; Runyan 1991; Runyan & Smith, 1991). Secondly, some researchers propose that, if the granting of extra time in an exam to candidates with SpLD is valid, extra time should provide a differential boost by benefitting students with SpLD to a greater extent than it benefits TD students (Alster, 1997; Goegan, 2014; Kettler et al, 2011; Lindstrom 2010; Lindstrom & Gregg, 2007; Lesaux, Pearson & Siegal, 2006; Ofiesh, Mather & Russell, 2005; Osborne, 2006). Thirdly, it is argued that if extra time improves the test performance of SpLD and TD candidates equally, or if TD candidates improve to a greater extent than SpLD candidates when granted extra time then the adjustment is inappropriate as it is simply inflating scores, rather than enabling more accurate scores by compensating for a disability (Bachan, 2017; Sokal & Wilson, 2017;Lewandowski, Cohen & Lovett, 2013; Lovett & Lewandowski, 2015; Spenceley & Wheeler, 2016). Finally, some researchers argue that the granting of extra time to candidates with SpLD alters the testing conditions for that group and so the exam is no longer administered in a standardised way for all candidates. This is based on the premise that any change to standard conditions may negatively affect the validity of the exam and alter the constructs being assessed in the exam. (Elliott et al., 2004; Pardy, 2016)

Thus, having established that the granting of extra time in exams aims to equalise the opportunity for students with SpLD to demonstrate their subject knowledge under timed conditions, and having identified the four different theoretical frameworks that are used to interpret performance comparability, it is imperative to critically examine the empirical research investigating the practice of granting this exam adjustment in order to identify whether or not it achieves this purpose.

3.2 Quantitative review: Overview of the survey

A survey of the extant quantitative literature[34] identified 17 studies that suggested that the granting of extra time did not confer any advantage to students with SpLD (studies' sample sizes ranged between 32 and 79,963 participants), while 12 studies were identified that suggested that the granting of extra time over-inflated the scores of all candidates (sample sizes between 30 and 119,490 participants). In addition, 4 systematic reviews were identified that concluded that the body of research evidence yields no consensus. These studies can be located in one of three distinct positions regarding the effect of awarding exam adjustments, namely (1) that additional time only improves the performance of students with SpLD, (2) that additional time improves the performance of students with SpLD more than that of their TD peers (albeit both groups benefit to some extent) and (3) that extra time over-inflates scores as both students with SpLD and their TD peers equally improve their performance.

3.2.1 Maximum Potential Thesis (MPT): Additional time only improves exam performance of students with SpLD

Studies that posit that additional time only improves the performance of students with SpLD situate themselves under the umbrella of the Maximum Potential Thesis (MPT). MPT suggests that candidates are unable to perform better than the limit of their knowledge and skills (i.e. exceed their maximum potential) in the exam simply because they are given more time. At the same time, candidates with SpLD are unable to show their full potential without extra time as they process information more slowly and so the standard exam time is not sufficient for these candidates to show the extent of their knowledge and ability (Bolt, 2004; Portolese, Kreuse & Bonner, 2016; Runyan & Smith 1991).

Of the 17 identified studies that suggest that granting extra time fails to confer an advantage, 9 found that additional time improved the exam performance of students with SpLD only, with TD participants failing to improve their marks with extra time[35].

[34] See Appendix 1 for details of the search strategy, databases searched, inclusion and exclusion criteria.
[35] These 9 studies were: Cohen, Gregg & Deng, 2005; Crisp, Johnson & Nivakovic 2012; Hill, 1984; Huesman & Frisbie, 2000; Lesaux, Pearson & Siegal, 2006; McKendree & Snowling, 2011; Portolese, Kreuse & Bonner, 2016; Runyan 1991; Runyan & Smith, 1991.

These 9 studies posit that only students with SpLD improve their performance when granted additional time in exams as TD students are able to reach their potential under standard timed testing conditions.

Runyan (1991), for example, compared the performance in a written comprehension test between University students with SpLD and their TD peers, with both groups taking the test under standard time conditions and also with additional time. This study found that in the 20 minute timed comprehension task, the scores of TD students significantly exceeded those of matched participants with SpLD when both groups were assessed under standard time conditions. However, the scores of the students with SpLD improved materially when awarded additional time to the extent that the gap in performance was closed, with no statistically significant differences then existing between the two groups. Furthermore, the study found that the TD students completed the comprehension task within the 20 minutes normally allocated and failed to show significant improvements in their scores when granted additional time. Runyan (1991) argues that the TD participants were already achieving their ‘maximum potential’ within the standard time allotted and hence did not improve their scores when allowed more time. Conversely, the participants with SpLD failed to complete the comprehension task within the 20 minutes (albeit they answered those questions they attempted correctly). Runyan (1991) concluded that the additional time simply levelled the playing field and normalised the performance of the participants with SpLD to that of their TD peers. The extra time only benefitted those participants who needed it due to their disability.

There are, however, methodological weaknesses in this study that may undermine the reliability of the findings. Firstly, participants were advised that they were taking a timed test but were not informed of the amount of time available in which to complete it (they were simply advised to work as quickly and accurately as possible). When granted extra time, participants were not permitted to change or review the answers that they had made to the test during the first 20 minutes (as these first 20 minutes represented the ‘timed’ conditions). Thus, it could be argued that those participants who worked quickly and completed the test within the standard time were unable to improve their marks when given additional time, even if they had not performed to their potential, as they were not permitted to revisit or revise any answers already given. This may have resulted in a 'ceiling effect' for these participants (most likely to be the TD participants) and so may explain why granting additional time to students with SpLD reduced the gap in performance between SpLD and TD participants in this study.

3.2.3 Differential Boost hypothesis: All students benefit from being granted extra time but SpLD students showed greater benefits than their TD peers.

Supporters of the MPT argue that only individuals with SpLD benefit from extra time (Huesman & Frisbie, 2000; Lesaux, Pearson & Siegal, 2006; McKendree & Snowling, 2011; Portolese, Kreuse & Bonner, 2016). Other studies, however, conflict with these findings and suggest that all students show performance improvements when granted extra time, albeit students with SpLD show significantly greater gains, or a 'differential boost' in comparison with their TD peers (Alster, 1997; Goegan, 2014; Kettler et al, 2011; Lindstrom 2010; Lindstrom & Gregg, 2007; Lesaux, Pearson & Siegal, 2006; Ofiesh, Mather & Russell, 2005; Osborne, 2006). This 'differential boost hypothesis' predicts that there will be greater differences in performance between SpLD and TD candidates under standard time conditions than under additional time, as candidates with SpLD will 'close the gap' when granted extra time and benefit more from the additional time than their TD peers (Goegan, 2014; Sireci, Scarpati & Li, 2005).

Lindstrom (2010) argues that the presence of a differential boost in relation to an exam adjustment acts as a key indicator that the adjustment achieves its aim of leveling the playing field, as the "differential boost is an empirical manifestation that the accommodation speaks to something essential about the disability" (Lindstrom, 2010, p.46). That is to say, any differential in the impact of extra time on performance relates solely to the disability, otherwise the adjustment would impact the performances of the SpLD and TD candidates equally. Lindstrom posits that if TD students also improve their performance when permitted extra time (albeit to a lesser extent than the SpLD group) this does not support the argument that the accommodation is inappropriate on the grounds that it benefits all students, but rather this implies that the standard time conditions are too severe and so insufficient for all candidates.

Lesaux, Pearson & Siegel (2006) investigated the differential boost hypothesis by assessing performance in a reading comprehension test between participants with SpLD and TD peers. Both groups were tested in both timed and extended time conditions and undertook the Nelson Denny Reading Comprehension Test, comprising 7 passages of prose reading followed by 38 multiple-choice questions (Brown, Fishco & Hanna, 1993). Participants were given the first form of the test under the standard timed conditions of 20 minutes and were then given 100% extended time (40 minutes) in which to complete the parallel form of the test. The outcomes were compared to identify any between-group and within-group differences in performance when taking the test under standard time and additional time conditions. Comparisons were made between the number of questions attempted, the number of correct responses achieved, and the number of errors made under both sets of time conditions.

When taking the test under standard time conditions, the TD group attempted significantly more questions on average (34.90 questions out of the possible 38) than the SpLD group (21.64 questions out of 38). When granted 100% additional time, both groups answered all of the questions. Thus, the group with SpLD improved their performance with extra time such that the difference in performance between the two groups was eliminated. Although the TD group also improved their performance in extended time conditions in comparison to standard time conditions, the group with SpLD showed a much greater improvement and benefitted more from the additional time than the TD control group (a 'differential boost'), as the TD group were already close to the test ceiling under standard conditions. Thus, this study found, the participants with SpLD benefitted more from additional time, significantly improving their scores and being able to complete the test. While the TD group improved their performance with extra time by answering more questions this improvement was not statistically significant as this group was already able to perform within the average range under standard time conditions (34.90 questions) and so had little scope to increase their marks before hitting the test ceiling (38 questions).

3.2.4 Extra time over-inflates scores: All students (SpLD and TD students alike) improve their exam performance when granted additional time.

Studies that show that candidates with SpLD benefit more from extra time than their TD peers are, however, contested by studies that conclude that extra time benefits all candidates - over-inflating marks and undermining the validity of the exam results (Amodeo, Marcus, Thornton & Pashley, 2009; Lewandowski, Cohen & Lovett, 2013; Lewandowski, et al., 2007; Lovett & Lewandowski, 2015; Lewandowski, Lovett &

Rogers, 2008). These studies argue that, if all students benefit from being granted additional time, then this adjustment simply facilitates the achievement of higher marks for those who receive the adjustment, rather than representing more accurate marks that compensate for slow processing.

Lewandowski, Cohen & Lovett (2013) examined the effect of granting 50% and 100% extra time to SpLD and TD University students in reading comprehension tests to see whether there was a greater between-group difference in performance under standard time than with extra time. This study posited that, if the granting of extra time is justified, TD participants should achieve higher marks than participants with SpLD when both groups take the test under standard time conditions. However, when both groups are granted extra time there should be no between-group differences in performance. TD candidates should not increase their scores with additional time as it is expected that they reach their full potential under standard time conditions. The two parallel reading comprehension forms of the Nelson Denny Reading Comprehension Test (Brown, Fishco & Hanna, 1993) were combined to produce a single test consisting of 14 passages to read and 76 multiple choice questions to answer. This aimed to circumvent any possible ceiling effect that may arise when extra time was implemented. Each participant's test was marked after 15 minutes (i.e. standard time), again after 22.5 minutes (i.e. with 50% extra time) and a third time after 30 minutes (i.e. with 100% extra time). The numbers of correct responses for each group, as well as the number of questions attempted and the percentage of correct responses, were compared. The results showed that, although neither group was more accurate in their response than the other, the TD participants scored increasingly higher than the SpLD group as time allocation increased. This contradicts the differential boost hypothesis that participants with SpLD will 'catch up' with their TD peers when given additional time. Instead, this study found that all participants improved their scores with extra time, but that TD participants showed greater gains than the participants with SpLD.

Interestingly, when the scores of the participants with SpLD at standard time (15 minutes), 50% extra time (22.5 minutes), and 100% extra time (30 minutes) were compared with the scores achieved by the TD peers under standard time (i.e. their scores after 15 mins), the performance of the SpLD group was lower than the TD group after standard time (15 minutes) but increased after 50% extra time and equalised after

100% extra time. This showed that the SpLD group performed below the TD group under standard conditions but performed similarly to the TD group when granted extra time in comparison to the TD group performance under standard time conditions. Nonetheless, Lewandowski, Cohen & Lovett (2013) posit that the findings overall show that all students improve their performance when they receive additional time, not only students with SpLD, and that this calls into question the validity and appropriateness of allowing extra time as an exam access arrangement.

However, caution should be exercised when generalising the findings of this study to the population as a whole, due to methodological problems that may undermine the reliability of these findings. The Nelson Denny Reading Comprehension subtest was designed (and standardised) as a 20-minute test. Therefore, by doubling the number of test items (through combining the two parallel forms) while restricting the 'standard time' to 15 minutes means that the test is not deemed fully achievable for either SpLD or TD participants within the 'standard time' (as the standard time that the original designers had allocated to this combined activity was 40 minutes). Lewandowski, Cohen & Lovett (2013) argue that this controls for the 'ceiling effect' as the participants will not have reached the ceiling of their ability (as determined by the number of correct answers) within the time permitted but can continue to answer further test items (i.e. time will run out before the number of answers are completed). However, this removal of the test ceiling may account for the finding that TD participants improved their performance when given extra time, as even double extra time (30 minutes) still represented less than the standard time that the test deemed necessary for the completion of both forms concurrently (standard time for both tests concurrently is 40 minutes). Given that the participants with SpLD performed below their TD peers in this study when allowed the 'standard time' of 15 minutes, it is not surprising that they performed below their TD peers when both groups were given 'additional time' of 50% as this was still insufficient time for either group to complete the task. Even when given 100% extra time (30 minutes), this still represented 25% less than the standard time normally allowed by the Nelson Denny test itself (i.e. 40 minutes for 76 questions). When granted 30 minutes to complete the test (i.e. 100% extra time for the purposes of this study) the TD participants completed more of the questions than the SpLD participants. That is to say, the TD group completed, on average, 80% of the test items (SD 9.6) in 30 minutes, while the SpLD group completed, on average, 64% of the test items (SD 12.49) in 30 minutes. This suggests that the TD group should be able to complete all 76 items on the test within the normal standard time of 40 minutes that the Nelson Denny Comprehension test itself sets (given that the TD group had completed 80% of the test items in 75% of the standard time allowed by the test). By contrast, the SpLD group would require extra time to finish the test (as the SpLD group had completed 64% of the test items in 75% of the standard time allowed by the test). It is worth noting that there were no between-group differences in accuracy of response, which suggests that the SpLD participants were slower (but not less accurate) in answering the questions.

Thus, the conclusion of Lewandowski, Cohen & Lovett (2013) that additional time confers an advantage to all students who receive it may only be the case where an exam is not achievable for any candidate in the time allowed. Given the unrealistic time limits placed on the participants in this study, rather than providing evidence to suggest that extra time is not an appropriate adjustment as it improves the scores of all students, not just disabled students, it could be argued that this study provides evidence to suggest that exam designers should calculate the standard exam time carefully to ensure that there is sufficient time for candidates to complete the exam.

3.3 Systematic review

Given that the body of research into the granting of exam access arrangements to students with SpLD appears to be characterised by contradictory findings, Sireci, Scarpati & Li, (2005) conducted a systematic review of eight research studies into the impact of additional time on exams. This aimed to statistically analyse, synthesise and summarise the findings and data from across the range of related studies with the aim of resolving the debate.

Of these 8 studies, 5 were experimental (Elliott and Marquart, 2004; Runyan,1991; Halla, 1988; Hill, 1984; Weaver, 1993) and 3 were quasi-experimental (Huesman and Frisbie, 2000; Alster, 1997; Zurcher and Bryant, 2001). Three of the studies examined the impact of extra time on students with SpLD using the Nelson Denny Reading Comprehension Test (Brown, Fishco & Hanna, 1993) (sample sizes ranged from 30 to 526), two studies used generalised Maths tests (sample sizes ranged from 31 to 124) and three studies used data from national tests, such as the Graduate Record

Examination (GRE), the Scholastic Aptitude Test (SAT) or American College Testing (ACT) (sample size ranged from 34,012 to 52,667). Sireci, Scarpati & Li, (2005) concluded that some studies showed that just the SpLD group benefitted from additional time, some showed that while both groups increased their results, the increase was greater for the SpLD participants than their non-disabled comparators and some showed that both groups benefitted equally. Sireci, Scarpati & Li also identified a range of limitations in the studies, namely that the Nelson-Denny Comprehension Test is a speeded test and thus “all students would be expected to have score gains when extended time is given on a speeded test” (Sireci, Scarpati & Li, 2005, p466). Overall, Sireci, Scarpati & Li concluded that while no clear consensus exists across studies, the outcomes suggest, on balance, that students with SpLD benefit more from additional time than TD students.

Gregg and Nelson (2012) similarly explored the impact of granting extra time in exams to students with SpLD through a meta-analysis that combined and evaluated the findings of 9 studies that had been conducted between 1986 and 2006 (Alster, 1997; Braun, Ragosta & Kaplan, 1986; Cahalan, Mandinach & Camara, 2002; Camara, Copeland & Rothschild, 1998; Cohen, Gregg & Deng, 2005; Lesaux, Pearson & Siegal, 2006; Lindstrom & Gregg, 2007; Ofiesh, Mather & Russell, 2005; Ragosta, Braun & Kaplan, 1991). This meta-analysis compared the test scores of participants with SpLD who were granted extra time with the test scores of TD participants who took the same test under standard conditions across all nine studies. The combined findings of the 9 studies indicated that the TD participants achieved higher scores than the participants with SpLD, despite the latter receiving extra time (mean effect size of d = -.41). In addition, the TD participants achieved higher scores than the participants with SpLD when both groups were granted extra time (mean effect size of d = -.69). Finally, when both groups took the test under standard conditions the TD participants achieved higher scores than the participants with SpLD (mean effect size of d = -.86).

Gregg and Nelson (2012) also made two within-group comparisons. Firstly, the test scores that the SpLD group achieved when granted extra time were compared with the scores they achieved when taking the test under standard conditions. This showed that the SpLD participants achieved higher scores when granted extra time (mean effect size of d = .90). Secondly, the test scores of TD participants taking the test under standard time were compared with their scores when granted extra time. This showed that the TD participants achieved higher scores when granted extra time (mean effect size of *d* = .66).

Overall, Gregg and Nelson (2012) concluded that participants with SpLD underperform in timed tests in comparison to their TD peers. Although this effect was reduced when participants with SpLD were permitted extra time, the extra time did not fully level the playing field or fully compensate for the effect of the disability on exam performance:

> "The most significant finding of our meta-analysis is that…students with LD still underperform academically as compared to their normally achieving peers whether provided extended time or not on standardized tests. While students with LD perform significantly better when provided extended time, the accommodation does not erase the disability" (p136)

Although there are advantages of using a meta-analysis rather than a simple narrative approach to a systematic review, it should be noted that the findings of a meta-analysis can be limited by the difficulties of comparing heterogeneous studies that use a wide variety of different designs and methodologies as well as by the inclusion of a number of studies that have a small sample size, as combining a number of small studies is not the same as a single study with a large sample size. In addition, the inclusion of studies with weak internal, external or construct validity results in the meta-analysis combining studies of different qualities - thereby combining potentially reliable data with potentially unreliable data (Bryman, 2009).

In view of these limitations, Gregg and Nelson (2012) concluded that it would be problematic to generalise the findings of the studies in the field at that time to the population as a whole, due to methodological limitations of the studies involved (especially the lack of ecological validity[36]). This is an observation that is reiterated by the National Centre on Educational Outcomes (NCEO), which conducted a systematic review of 72 studies into the effects of exam access arrangements on students with disabilities between 1999 and 2015. The NCEO identified inconsistency in findings across the studies as well as a 'lack of authenticity of the testing circumstances' (NCEO,

---

[36] Ecological validity refers to the extent to which the study (including its variables and outcomes) reflects the 'real-life' behaviour of the population under study and impacts the extent to which the outcomes of the study can be generalised to that population as a whole in the real world (Kihlstrom, 2021).

2015, p85). The lack of consensus in outcomes as well as poor ecological validity of the extant research identified by NCEO, highlights the need for research into the impact of exam arrangements in the context of the exam itself.

3.4 Limitations of current research

3.4.1 Ecological validity

Studies investigating the impact of granting additional time to students with SpLD in timed tests inform policy around awarding extra time in real-life University exams. However, many of the research studies explored in this literature review adopt methodologies that raise questions relating to ecological validity and the generalisability of the findings to real-life exams. For example, most of these studies investigate the impact of granting extra time in reading comprehension or maths tests. However, the extent to which the findings of these studies can be transferred to the context of a school or university exam is questionable, given that many of the key processes involved in completing a reading comprehension or maths test differ significantly from those required for undertaking a formal exam. An examination testing the candidate's knowledge, understanding, and critical thinking in their taught subject area involves the recall and application of subject knowledge and the ability to develop this into a novel argument. The candidate is required to generate and spell expressive vocabulary, drawing heavily on word retrieval skills. At the same time, presenting a cogent argument in their writing requires the planning, organising and editing of the candidate's ideas. By contrast, a reading comprehension test assesses the candidate's ability to understand a written passage and respond to questions based on the information (either explicit or implied) contained in the passage. Additionally, in a reading comprehension test the vocabulary required is largely given. The candidate is not generally required to generate new information, draw on recall or the interpretation of previously learned concepts, but instead the candidate is required to move or interpret information that is already given in a text to which they can re-refer. Fewer burdens are placed on the candidate's memory and recall, word retrieval, expressive language, planning, structuring and organisational skills; areas that tend to be vulnerable in students with SpLD (Dobson Waters & Torgerson, 2020; Moll et al, 2020; Wagner et al, 2020).

Similarly, it can be argued that many of the processes involved in completing a Maths test differ significantly from those required for undertaking a formal exam. For example, the Maths tests used in the research studies cited above were not testing the participants' knowledge of a recently taught course or curriculum. These were generalised Maths tests that assessed a broad range of pre-existing Maths knowledge and ability. The participants may have had different levels of knowledge of Maths concepts and exposure to Maths courses prior to the test (this potential variability was not considered by the researchers). By contrast, university exams aim to test the subject matter that that has been studied through the year (or course) and assume that the participants have had comparable opportunities to be exposed to the body of knowledge being tested.

The only study identified in this survey of the literature that was conducted within the 'real-life' context of the mandatory high-stakes timed University exams (and so arguably has sufficient ecological validity to generalise findings to actual University exams) was by Duncan & Purcell (2017). This quasi-experimental study examined existing data from the end of year university exams and compared the exam marks of students with SpLD who were granted exam arrangements with the marks of TD peers who took the same exam under standard exam conditions. This study found that students with SpLD who were granted extra time in their exams did not produce a higher word count or achieve higher marks than their TD peers who took the same exam under standard conditions. However, as this study was quasi-experimental (and so based on pre-existing data), it was not possible to compare the performance of the participants with SpLD with and without the extra time and so draw conclusions about the actual impact the extra time had on their performance (except to suggest that this did not indicate an advantage when compared to the performance of their TD peers).

3.4.2 Variability in the amount of extra time granted

A further difficulty in drawing a unified conclusion from the research in this literature review and applying it to the question of extra time in university exams relates to a lack of consistency in the amount of extra time applied across the studies. Reflecting differences in local practices, there is variability across the studies cited above in terms of the amount of extra time granted to the participants and this may contribute to the contradictory findings. For example, Runyan (1991) granted unlimited extra time, Elliott & Marquart (2004) allowed 100% additional time, Lewandowski, Cohen & Lovett (2013) awarded 50% and 100% extra time, while in the study by Duncan & Purcell (2017) the participants were granted 25%. These differences in the amount of extra time awarded in the studies (coupled with differences in inherent 'speededness' of the tests undertaken) hinder any integration of conclusions about the effect of extra time across studies and render it difficult to draw a conclusion from the findings as a whole.

3.4.3 Heterogeneity of sample group and small sample size

One crucial limitation in most of the studies cited in this literature review was a lack of homogeneity in the sample group. Frequently the sample group participants were categorised as LD or 'students with disabilities' and some of the studies included participants with ADHD, where, it could be argued that the cognitive profile differs significantly to the profile of other SpLDs (Elliott & Marquart, 2004; Lewandowski, Cohen & Lovett, 2013; Lewandowski, et al., 2007). This heterogeneity potentially provides a less accurate account of the impact of extra time on students with SpLD specifically and limits the generalisability of the findings to this population (Bryman, 2009).

In addition, a further significant limitation of the studies considered in this literature review relates to small sample size. The majority of the experimental studies included sample sizes that were less than 90 participants in total (Runyan, 1991; Lesaux, Pearson & Siegal, 2006; Alster, 1997; Lewandowski, Cohen & Lovett, 2013; Lewandowski, et al., 2007) with the result that these small sample sizes (and consequent low statistical power) reduce not only the likelihood of detecting a true effect but also reduce the probability that the statistically significant findings actually demonstrated a true effect (Bryman, 2009). This low statistical power coupled with the potential heterogeneity of the participants may undermine the certainty with which these studies' results could be reproduced. In addition, many of the studies in this field were conducted a number of years ago and, arguably, may therefore no longer accurately represent the current situation.

3.5 Conclusion

The rationale that underpins the granting of extra time to candidates with SpLD pivots on the argument that the cognitive deficits associated with SpLD are likely to negatively affect that student's exam performance, resulting in marks that are not representative of the student's subject knowledge and academic skills. Granting extra time to candidates with SpLD reduces the variability in access skills that impact exam marks but are not relevant to the construct that the exam is aiming to measure. However, at the same time, changing the time available for some candidates, but not others, alters the standardisation of the administrative process of the exam and so may impact the validity of the exam marks.

The studies included in this literature review all aimed to investigate the effects of changing the standardised administration of the exam in terms of the time allocation. The objective was to determine whether or not the granting of extra time simply achieves its purpose of compensating for the student's disability and removing the barrier that they would otherwise experience in being able to perform to their potential (and so their marks become a valid representation of their ability) or if it benefits all candidates and so undermines the validity of the exam by simply inflating scores.

It is clear from this review that there is no consensus of opinion as to whether or not extra time benefits all students, or just students with SpLD. This reflects the findings of The National Centre on Educational Outcomes (NCEO), which similarly concluded that 'the findings about the effect of extended time on assessment scores varied widely, yielding no consensus' (NCEO, 2015, p18). The systematic review of literature conducted by NCEO found that all the studies surveyed concurred that extra time has a positive impact on test performance of students with SpLD, although there were varying outcomes in terms of the impact of extra time on assessment scores for TD students. As a result, although there is general agreement in the field that additional time improves outcomes for students with SpLD, there is no unified view about the fairness of this adjustment (i.e. whether it is only students with SpLD who benefit or whether all students benefit similarly from additional time).

Nonetheless, despite the inconsistencies in findings between studies a pattern does emerge across all the studies as a whole, which suggests that the differences in outcomes between studies may be accounted for by the inherent ‘speededness’ of the tests used in each study. For example, those studies that used ‘high speeded’ tests found that all participants improved with extra time. This may be due to the study not allowing sufficient time for any candidate to complete the test under standard time conditions and so all participants improved with extra time due to the removal of any test ceiling. Those studies that were ‘moderately speeded’ (such that the standard time was allocated so that some, but not all, TD candidates were expected to complete the exam without needing extra time) resulted in a differential boost. That is to say, all candidates improved their score with the extra time, but those with SpLD showed greater gains. Those studies that allowed sufficient standard time to enable all TD candidates to complete the exam in standard time showed that only candidates with SpLD increased their scores when granted extra time, suggesting that only those candidates who were not able to complete the exam in standard time benefitted from further time. Figures 3.1 & 3.2 provide a diagrammatical depiction of how the differences in inherent ‘speededness’ of the tests used in the different research studies potentially impact on the outcomes.

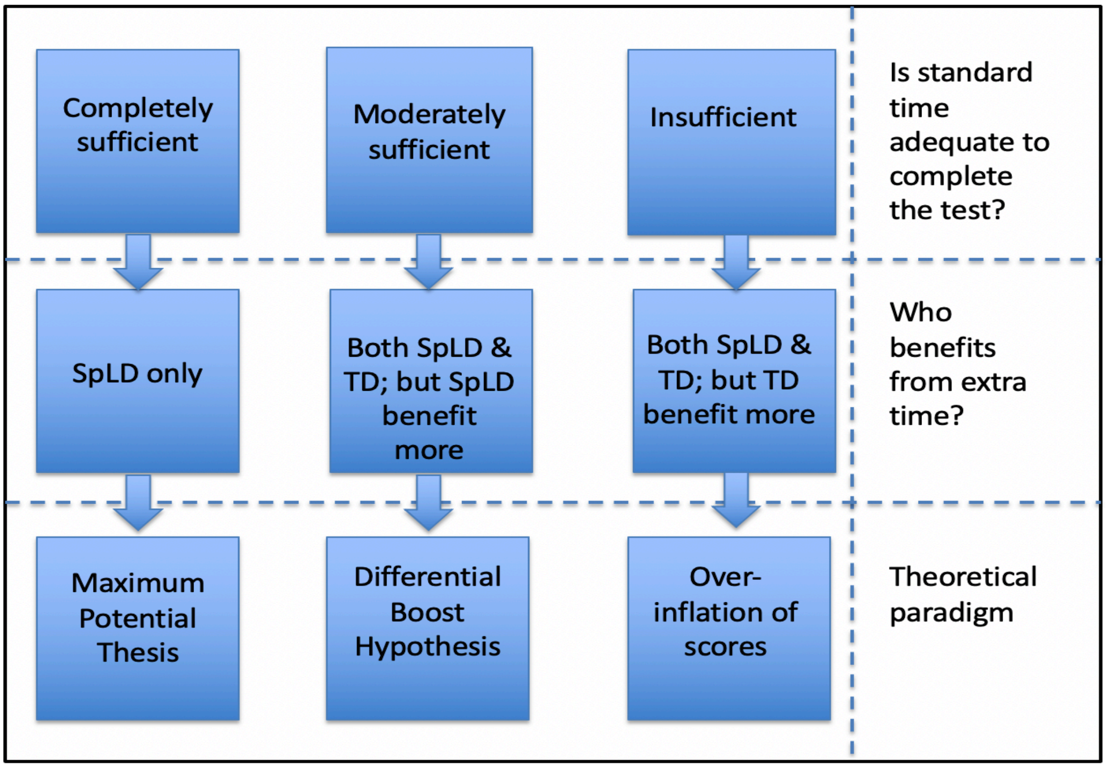


Figure 3.1: Relationship between the amount of standard time allocated in tests and the effect of extra time on participants' scores

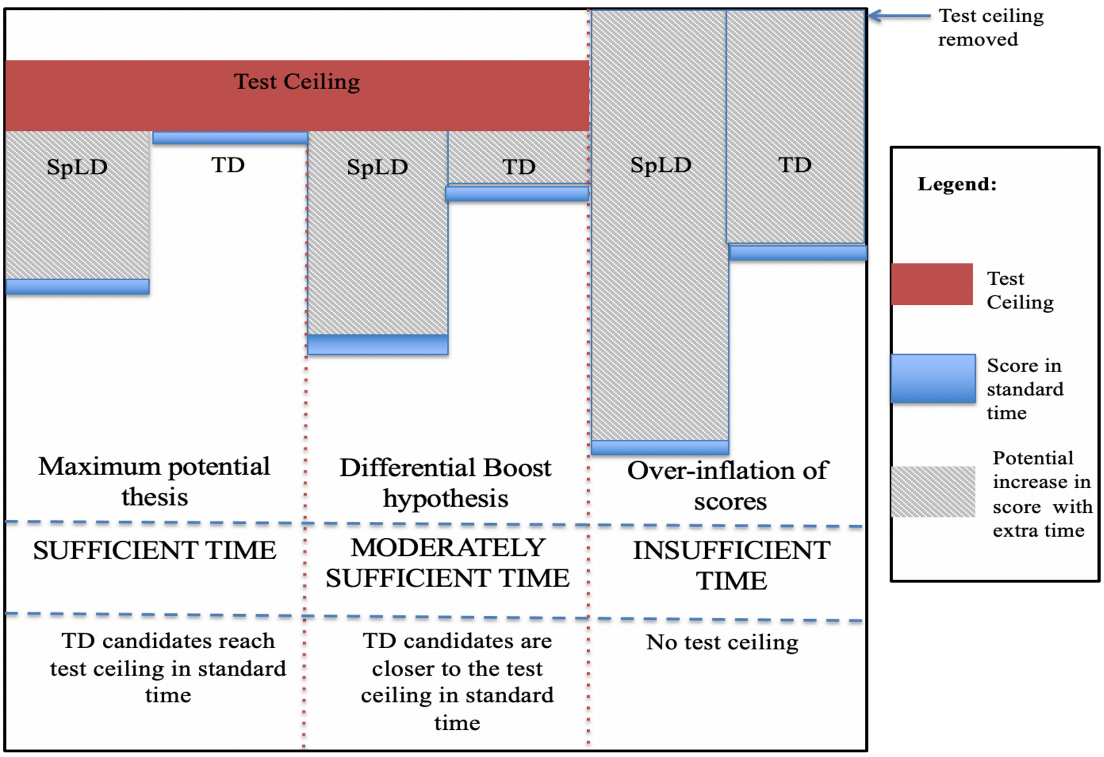


Figure 3.2: Relationship between inherent 'speededness' of the test and the potential increase in score that results from granting extra time.

This explanation is supported by the conclusion of the systematic review of literature conducted by Sireci, Scarpati & Li (2005), which comments that:

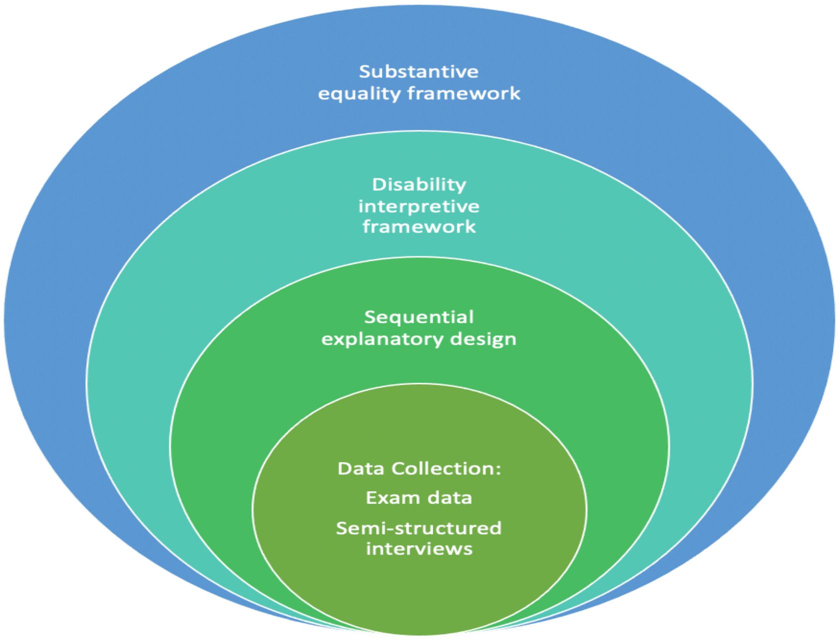
> "Extended-time administrations provide a more precise estimate of students' abilities than standard-time administrations, and differences in scores change are more related to the effect of speed on regular administrations, rather than providing learning disabled students with any advantage due to extended time." (Sireci, Scarpati & Li, 2005, p469)

In summary, there is general agreement in the field that TD students perform better on timed tests under standard time conditions than students with SpLD and that the granting of extra time demonstrates a positive effect on test scores for students with SpLD. There is dissensus, however, over the question of whether TD students also benefit from additional time or whether they are already performing to their potential under standard time conditions. If all students benefit from additional time this raises the possibility that the extra time is simply increasing the score of the candidate in receipt of extra time, rather than achieving a more accurate score that is a valid representation of the student's subject knowledge and skills. Different studies reach different conclusions about this question of the fairness of awarding extra time to candidates with SpLD and assimilating the various findings of the research in this field into a unified conclusion is problematic, given the variety of methodologies, small sample sizes, and wide array of research designs used. As Gregg & Nelson (2012) comment:

> '…the body of research examined is not adequate to address many of the questions pertaining to the comparability of scores with and without [exam] accommodation' and 'the literature is lacking in quantity of studies, restricted in types of design methodologies and under-representative of the diversity of individuals demonstrating the disorder' (Gregg & Nelson, 2012, p142).

As a result of the contradictory research that currently exists, the limitations of these studies and the questions that their findings raise, this literature review demonstrates a clear need for further studies to be conducted so that valid conclusions can be drawn about the comparability of results for HE students with SpLD taking exams with additional time with those of their TD peers who take those same exams under standard conditions in UK universities.

4: METHODOLOGY OVERVIEW



This chapter presents an overview of the research methodology adopted by this research project. It details the research design of the different phases of the research project as well as presenting the overarching design and theoretical perspective that has been applied to the research project as a whole. The chapter also addresses the impact that the researcher's own characteristics (professional and personal) may have on the interpretation and analysis of the qualitative phase of the research project, as well as the impact that the characteristics of the chosen research site may have on the generalisability of the research findings to the field as a whole. This chapter concludes by presenting the research questions that have been developed as a means to explore the question of exam equity (or advantage) in relation to the exam access arrangements granted to HE students with a diagnosis of SpLD.

4.1 Introduction

The design of this current research project has been guided by a disability discrimination and substantive equality theoretical perspective[37]. In other words, it is concerned with equitable outcomes and equal opportunities for students with SpLD in Higher Education. The aim of the research project was to determine whether or not reasonable adjustments in exams (use of a word processor and/or the granting of 25% extra time) achieve their aim of conferring equity to candidates with SpLD in relation to exam performance. Once the effect that exam adjustments have on exam performance had been identified through an analysis of the quantitative data, the research project then aimed to expand this understanding by exploring why and how the phenomenon identified occurs, through an analysis of qualitative data. The ultimate aim of this research project is to provide evidence that will influence the development of educational policy that seeks to promote equality of opportunity and outcome for students with SpLD and to effect material change through the practical application of research evidence on practice.

Given that the research needs to speak to a range of audiences, including policy makers and disability practitioners, a mixed methods research paradigm was adopted as a means to accomplish this. Mertens et al (2010) argue that quantitative data is commonly used to effect policy change, while the qualitative data, as it illuminates the real-life experiences of the population affected by that policy, informs the shape and direction of that policy. In addition, the mixed methods research paradigm used in this research project aimed to foster a more thorough, extensive and integrated portrayal by amalgamating information from complementary data, broaden the ambit of the research, and counteract the weaknesses of using either approach alone (Creswell et al, 2011; Creswell & Clark, 2017; Denscombe, 2008; Denscombe, 2009; Hanson et al., 2005).

[37] Disability discrimination argues that disabled individuals are a marginalised group who experience disadvantage due to societal practices that assume a non-disabled 'norm' (Hackett et al, 2020). Substantive equality is concerned with the equality of opportunity and outcomes for marginalised groups and requires the implementation of specific measures that aim to assist disadvantaged individuals (Fredman, 2016). In the context of students with SpLD in higher education, this requires changes to an environment, such as time-limited exams, that discriminates against students with SpLD and prevents them from enjoying equal opportunities to access their learning and reach their potential. These changes include the awarding of exam access arrangements and course-based adjustments, for example.

The mixed methods approach used in this research project aligns itself with the philosophical outlook of pragmatism[38], which Denscombe (2008) argues 'provides a basis for using mixed methods approaches as a *third alternative*… because good social research will almost inevitably require the use of both quantitative and qualitative research to provide an adequate answer' (Denscombe, 2008, pp273-4). Mixed methods research, underpinned by the assumptions of pragmatism, has the potential to harness the strength of both quantitative and qualitative approaches, answer different questions, and lead to a richer understanding of whether exam adjustments confer equity in exam performance, why and how the outcome occurs, and how the exam adjustments are experienced by students with SpLD (Hanson et al., 2005, p.224).

4.2 Research Design

This research project comprises three phases and employed a sequential explanatory-exploratory strategy. The research project commenced with a largely deductive approach, based on the two-tailed hypothesis that there will be differences in exam marks and length of script between students with SpLD who were granted exam adjustments and their TD peers who sat the same exams under standard conditions. The second quantitative phase compared the marks achieved by students with SpLD in exams taken under standard conditions with the marks achieved by the same students when taking exams with exam adjustments in order to ascertain the magnitude of the impact that exam adjustments have on the marks of candidates with SpLD.

Analysis of the quantitative data aimed to identify any statistical differences to exam mark arising from the granting of exam adjustments and determine whether or not the exam adjustments level the playing field between candidates with SpLD and their TD peers in timed, closed book, written exams. This deductive approach was used to identify and explain any causal relationships between exam adjustments and exam

---

[38] Pragmatism is defined as a problem-oriented approach, based on the premise that the most appropriate research methods to adopt are those that answer the research question most effectively. Pragmatism is located in the view that there are multiple realities and, in order to explore the research question from more than one perspective, the researcher needs to integrate multiple research methods, including both qualitative and quantitative methods, rather than a single methodology. (Creswell, 2017). 'Pragmatism takes the research problem as its starting point and gauges the value of any particular approach or method primarily in terms of how well the outcomes work in practice' (Denscombe, 2009, p.128).

performance (Bryman, 2015; Creswell 2003). Table 4.1 summarises the quantitative research design used in phases one and two.

Table 4.1: Quantitative Research Design used in Phases One and Two

Research Design	Research Strategy
Ontology	Positivism. Universalism. Independent
Epistemology	Causal explanation. Generalisable. Prediction.
Methodology	Quasi-experimental.
Methods	Between group and inter-group comparisons of exam marks, script length and coursework marks.

In the third (final) phase, the research project adopted an inductive, qualitative approach to understand the phenomenon observed, and construct a rich understanding of why and how it occurs (Bryman, 2009). The process of the sequential explanatory-exploratory strategy applied in this phase is shown diagrammatically in Figure 4.1.

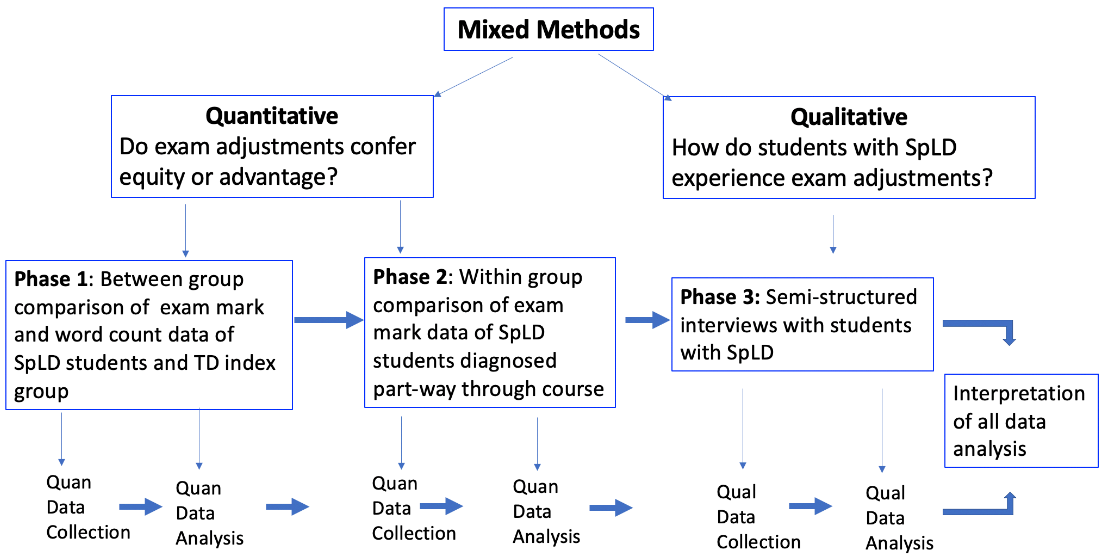


Figure 4.1: The sequential mixed methods strategy process (adapted from Creswell, 2003)

The quantitative data were collected first in this sequential research design and statistical methods were used to select the findings that warranted exploration in the second quantitative phase. The statistical outcomes from the second quantitative phase were then combined with those from the first phase to determine the phenomena of interest to be explored in the final, qualitative, phase of the project. This sequential process enabled the data in each of the phases to expand on and illuminate the results from the previous phases of data collection and analysis, as well as inform decisions about the data collection and appropriate participant selection for the subsequent stages (Creswell & Plano Clark, 2007).

In the qualitative phase, data were collected and analysed to explore why and how the phenomenon identified by the quantitative data occurs. This aimed to provide the basis for understanding the new knowledge revealed by the research project and act as a vital step towards identifying interventions. The qualitative research phase used a constructional ontological perspective that rejects the concept of an objective social reality, 'asserting instead that realities are social constructions of the mind, and that there exist as many such constructions as there are individuals (although clearly many constructions will be shared)' (Guba & Lincoln, 1989, p.43). Applied to this research project, a qualitative approach afforded an opportunity to emphasise the multiplicity of participants' perspectives in understanding how exam accommodations function to confer equity (Glaser, 2007; Strauss & Corbin, 2015), while also supporting the co-construction of meaning between the participants and the researcher (Guba & Lincoln, 1994; Bryman 2009).

In the qualitative phase, grounded theory was used to explore the experiences of students with SpLD who received exam adjustments. Grounded theory may be used to 'construct theory about issues of importance in peoples' lives' (Mills et al, 2006, p2). This inductive approach allowed 'issues of importance' to the research project's participants to emerge from the experiences they disclosed to the researcher during the semi-structured interviews (Guba & Lincoln, 1994). Constructivist grounded theory recognises the influence of the researcher in reconstructing meaning from the participants' data, with the researcher's values and biases inevitably forming part of the outcome (Guba & Lincoln, 1994; Strauss & Corbin, 2015).

This research project, however, diverges from the classic principles of grounded theory as a purely inductive process, towards a more abductive approach, as it locates the grounded theory within a sequential explanatory- exploratory design with the interview questions being shaped by the outcomes of the earlier quantitative data. As the goal of the qualitative research phase in this research project was to provide an understanding and conceptualization of the quantitative findings, the semi-structured questions were formulated within a deductive framework anchored in the quantitative data. Nonetheless, the themes identified from the qualitative data were inductively generated from the participants' responses to the questions and analysed and interpreted using grounded theory tenets to build a conceptual understanding of the experiences of students with SpLD who were granted exam adjustments. Table 4.2 shows the principles of the qualitative research design used within phase three.

Table 4.2: Qualitative Research Design used Within Phase Three

Research Design	**Research Strategy**
Ontology & Epistemology	Relativist; Interpretative paradigm[39]
Theoretical Framework	Substantive equality perspective
Methodology	Constructivist grounded theory
Methods	Semi-structured interviews

As a practitioner in the field, my purpose in conducting this research was to effect policy change. In order to influence policy it was necessary, in the qualitative phase, to have pre-established broad areas of questions to be explored that were located within the phenomenon identified by the quantitative data (i.e. the existence of an awarding gap). Thus, while aiming to align with the inductive nature of grounded theory in terms of content analysis, from a pragmatic point of view (to ensure that this data shapes policy and practice) it was necessary to work within the context of prior knowledge. Therefore, rather than following a purist grounded theory approach, the methodology adopted in the qualitative phase of the project used many key principles of grounded theory as a framework (such as the use of intensive coding that followed the data in order to generate new theory), but combined this with a pragmatic approach, which aims to collect data and develop knowledge to shape practice and effect change. This grounded theory approach, mediated by pragmatism, was also located within an

[39] Bryman (2009) defines a paradigm as a set of beliefs and assumptions about how the research should be conducted and how results should be interpreted for analysis (p. 286).

overarching social justice ontology and a critical theory framework. That is to say, the purpose was to evaluate current assessment practices that assume a non-disabled 'norm' predicated on timed, closed book exams, and influence and change policy towards inclusive and equitable assessment practices. The research questions emerged from this theoretical framework and were intended to effect positive change for students with SpLD by identifying and exploring the reasons for the awarding gap and encouraging pedagogic practices that provide equal educational opportunities for all students.

In summary, the research project used quantitative data to provide statistical evidence of the impact on exam performance of the exam access arrangements. Qualitative methodology (constructivist grounded theory) was then used to explore the experiences of students with SpLD granted exam arrangements with the aim of understanding why and how the phenomenon identified by the quantitative data occurs. This combination of quantitative and qualitative design aimed to answer the research questions while promoting a post-positivist approach and creating a more comprehensive picture of the results. The aim of adopting this mix of methodological and pragmatic approaches is so that this project can present policy makers with both empirical evidence of the awarding gap and an explanation of how and why this exists with a view to influencing positive changes in policy and practice to reduce or eliminate this inequality.

4.3 Researcher reflexivity

Guba & Lincoln (1994), Bryman (2009), and Strauss and Corbin (1998) recognise that the researcher *interprets* the participants' description of their experiences in order to construct theory. In addition, Strauss and Corbin argue that 'it is not possible to be completely free of bias' when interpreting qualitative data (Strauss & Corbin, 1998, p. 97). Thus, in an attempt to promote transparency, the importance of recognising researcher bias is vital.

My position as a researcher undertaking this inquiry is rooted in both an emic perspective (my role as disability adviser supporting university students with SpLD) and an etic perspective (an individual who does not have a diagnosis of SpLD). These identities have the potential to influence my research in multiple ways, from the choice of topic to the coding scheme for the qualitative data. For instance, in my professional role, students with a diagnosis of SpLD regularly share their perceptions of their educational experiences with me (qualitative data), and I also have access to their academic performance results (quantitative data), which influences my understanding of the academic experiences of students with SpLD and, hence, what I consider is important in the data. My 'insider' status potentially also enables me to see links and trends within that data that may not otherwise be obvious. However, my decisions about data selection and its interpretation are also informed by my 'outside observer' perspective (as a TD individual) which, arguably confers a more neutral position than would potentially otherwise be the case.

As a disability practitioner specialising in SpLD, a diagnostic assessor and, previously, a Secondary school teacher, I am intimately concerned with students' experiences, barriers and challenges in higher education and am privy to their direct reports of their perceptions of their experiences (micro view). In my strategic and policy roles (meso-view) I am able to apply that data (along with quantitative data) to inform my own practice as well as in shaping the Neurodiversity service development. I am also well versed in the debates at the university level related to students with SpLD that structure institutional approaches to disability accommodations, both in support and opposition (such as the intensely debated topic of whether or not exam arrangements confer advantage for students diagnosed with SpLD, for example) and so work collaboratively on institutional policy development (meso view). As a Director of the SpLD Assessment Standards Committee (SASC[40]) I take a national policy perspective, collaborating with strategic decision-makers in the field, which furnishes me with a macro-view of how the issues aggregate at a systemic level. These overlapping identities inform my research in explicit ways, as shown diagrammatically in Figure 4.2.

[40] SASC is the national professional association for those working in the field of diagnostic assessment for individuals with SpLD and is responsible for developing and supporting best practice in the assessment of SpLD https://sasc.org.uk/

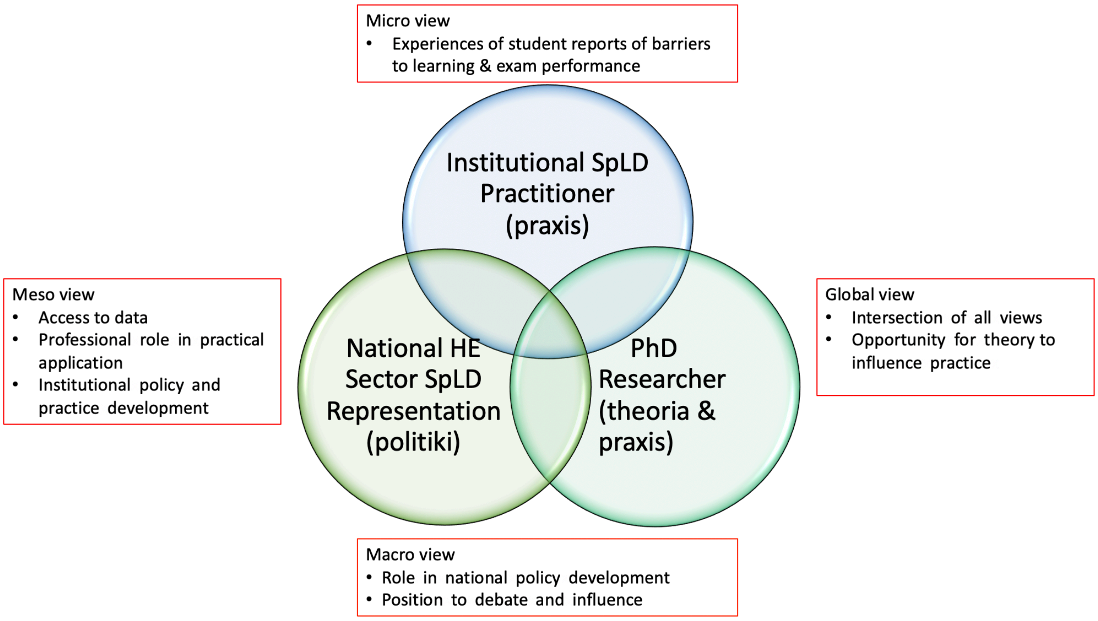


Figure 4.2: Researcher Identities (adapted from Harding, Hofmann, & Alcott, 2021)

My experience as a disability professional delivering services for neurodiverse students and the university more broadly, has fostered in me a deeply rooted commitment to equality of opportunity in education alongside an adherence to affirmative, participatory, student-focused, person-centred, approaches. This informed my decision to engage in research that included the voices of students with SpLD, in addition to collecting and analysing numerical data. I am aware that I am exposed to predominantly negative reports of the academic experiences of students with SpLD as an inevitable corollary of my professional role and so my perception of the lived experience of students with SpLD at University is one beset with systemic disadvantage. However, it is important to strive to ensure that any bias, arising from an expectation of reports of negative experiences, does not unduly affect the inductive coding so that, when developing a conceptual analysis and theoretical interpretation of the participant's accounts, student reports of positive, as well as negative, experiences are represented faithfully and allowed the weighting that the participants intend.

With this in mind, I maintained a reflective log throughout the research project that documented, and made critical reflections on, my research practices and processes. The reflective log assisted in making my biases, and my influences on the data, more visible during the qualitative research process. A key aim of this log was to promote a more transparent and reflexive approach to the qualitative element of this research project by enabling me to more explicitly acknowledge my personal values, assumptions and belief systems and recognise how these subjectivities influenced my interpretation of the data. By documenting these personal and professional assumptions the log provided me with opportunities to reflect on the tension between my aim of collecting and interpreting data neutrally while simultaneously holding preconceived views that students with SpLD experience educational disadvantage occasioned by the teaching, learning and assessment practices at university. This made me reflect on the approaches I use as a researcher, and also in my professional practice, when conducting and interacting during interviews and how I 'hear' and interpret the resultant student stories. As a result, the log provided me with opportunities to explicitly consider how these interactions impact the trustworthiness of the data I present.

In addition, the reflective log enabled me to consider and re-evaluate my methodological decisions. For example, at the outset of this project I assumed that employing mixed methods of data collection across the project would overcome the impact of any researcher biases in the qualitative data collection and interpretation phase as the process of triangulation would cross-reference the data. However, the reflective log encouraged me to question this assumption of validity and prompted me to explore a variety of theoretical readings and competing views on strategies for testing the trustworthiness of the data. These reflections, and the added theoretical readings, prompted me to implement additional measures for testing validity that I had not considered at the outset of the project. For example, I took specific measures in the development and delivery of the interview questions, as well as in the analysis and interpretation of interview data, in order to give due representation to both reports of positive, as well as negative, experiences and ensure that the participants' accounts were authentically presented. The participants in this research project trusted me with their stories and I was acutely conscious of my responsibility to represent their narratives faithfully and in a way that they would recognise. Therefore, I endeavoured to centralize the students' voices by maintaining their words verbatim during the analysis and reporting, using verbatim quotes with only light editing of linguistic 'fillers' (Charmaz, 2000). In this way I strove to 'describe the experiences of others in the most faithful way possible' (Munhall, 2001, p. 540). I also shared with the participants the write up of my interpretation of their data, including the themes identified, and invited their feedback in order to check that my selection and interpretation of their data authentically reflected their experiences and intended meaning. This not only aimed to improve the trustworthiness of the data but also offered participants an opportunity for reciprocity and reconstruction of the data (Creswell & Poth, 2018; Lincoln & Guba, 1985) – an approach that I would not have considered without the influence of the reflective log.

Interestingly, some of the participants responded to say that reading their own stories in this way gave them the courage to be more open about their diagnosis of SpLD and greater confidence to self-advocate. While this project aimed to effect change for students with SpLD in terms of university policy and practices, this was an unexpected outcome that would not have occurred without the reflective log. This has also encouraged me to change my professional practice by now keeping a professional reflective log to help me identify ways that I can promote more emancipatory practices in my professional role.

4.4 Research site

This research was conducted at a Russell Group, collegiate, research intensive HEI. This is a highly competitive environment with a high staff to student ratio. At the same time, the stringent undergraduate admissions requirements[41] suggest that the students at this university have already performed successfully in formal examinations. However, despite the particular nature of Russell Group universities and its cohort, the findings from this research project in relation to the impact that exams and exam arrangements have on students are arguably generalisable to all HEIs where any end of year timed, closed book exams are used, as they explore the particular disadvantages that students with SpLD experience in formal, timed, exams, due to the nature of their disability, and the extent to which exam adjustments are able to compensate for these disadvantages.

[41] Typical offer A*A*A or A*AA, depending on subject (https://www.undergraduate.study.cam.ac.uk)

4.5 Research Questions

Exam access arrangements aim to remove any disadvantage that disabled students experience in exams on account of their disability, but do not aim to confer an unfair advantage. The research questions in this research project were developed to explore whether or not the exam adjustments of extra time (or extra time and the use of a word processor) achieve this aim for students with SpLD and also to better understand the impact that exam adjustments have on this cohort. To interrogate the question of exam equity (or advantage) in relation to access arrangements, the following research questions were developed:

> RQ1: What impact do exam access arrangements have on the exam performance of students with SpLD?
>
> > RQ1a: Do students with SpLD who have access arrangements produce longer exam scripts or achieve higher marks than their TD peers?
> >
> > RQ1b: Do exam access arrangements enable students with SpLD to perform to their potential in exams?
> >
> > RQ1c: How much effect do the exam adjustments have on the exam performance of students with SpLD?
>
> RQ2: What are the perceptions of students with SpLD of the impact that exam access arrangements have on their exam performance?

These questions, and how they are operationalised or explored, are shown in Table 4.3.

Table 4.3: Research questions, adopted methods and output.

Research Question	Method	Quantitative	Qualitative	Output
RQ1a: Do students with SpLD who have access arrangements produce longer exam scripts or achieve higher marks than their TD peers?	Quasi-experimental. Statistical comparison of exam marks and script length of students with SpLD with those of TD index group (SPSS)	Exam mark & script length data		Are the exam grades or script lengths of students with SpLD statistically different to those of their TD peers?
RQ1b: Do exam access arrangements enable students with SpLD to perform to their potential in exams?	Quasi-experimental. Statistical comparison of difference between exam mark and coursework mark of participants with SpLD and those of the TD index group (SPSS)	Exam mark data and coursework mark data		Are students with SpLD achieving marks in exams that reflect their coursework marks and does the relationship between exam marks and coursework marks of students with SpLD mirror those of their TD peers?
RQ1c: How much effect do the exam adjustments have on the exam performance of students with SpLD?	Quasi-experimental. Statistical comparison of exam marks of students with SpLD in exams without exam arrangements with those achieved by the same students in exams with exam arrangements (SPSS)	Exam mark data		Is there evidence that the implementation of exam arrangements alters the exam marks of students with SpLD?
RQ2: What are the perceptions of students with SpLD of the impact of exam access arrangements on their exam performance?	Semi-structured interviews (TA)		Open questions	How do students with SpLD experience exams and how effective do they perceive exam arrangements to be?

Research question 1 is addressed by the two quantitative research phases and research question 2 is addressed by the qualitative research phase.

5: QUANTITATIVE RESEARCH METHODOLOGY

This chapter presents a detailed explanation of the research methodology adopted in the first quantitative phase of the research project. The chapter starts by presenting the overarching hypothesis of this first quantitative phase of the research project, along with the sub-questions that have been developed to operationalise that hypothesis. It then goes on to detail the inclusion criteria used to select participants as well as the sampling process that shaped the data collection process and sample size. Finally, this chapter concludes by presenting the research strategy, data collection procedure, data validation and data analysis techniques that were applied to this first quantitative phase of the project.

5.1 Introduction

The overall aim of this research project was to investigate the impact of exam access arrangements on the exam performance of students with SpLD. This commenced with the first quantitative phase of the project that aimed to identify any causal relationships between variables resulting from the existing policy of changing exam conditions for students with a SpLD through the following two-tailed hypothesis:

H$_1$: As a result of being granted extra time or additional time and the use of a word processor there will be significant differences between the length of exam answers and marks of students with SpLD when compared to their TD peers

This two-tailed hypothesis was operationalised by addressing the following sub-questions:

1. Do students with a SpLD who are granted 25% extra time, or 25% extra time and use of a word processor, produce significantly more words on a Humanities exam paper than their peers who have taken the same exam under standard conditions?

2. Is there a relationship between the number of words produced in a Humanities exam and the quality of exam performance, as indicated by exam grade?

3. Is there a significant difference between the exam marks of students with a SpLD who have been granted 25% extra time, or 25% extra time and use of a word processor, and the exam marks of their peers who have taken the same exam under standard conditions?

This hypothesis assumed that there would be a relationship between the independent variables [SpLD participants with additional time; SpLD participants with additional time and the use of a word processor; index (TD) group] and the dependent variables [length of answers on the exam script as determined by individual word count; number of words per minute produced on the exam scripts; exam mark; classification]. However, a possible outcome of this research was that the data could suggest no relationship exists between the variables cited and, as such, the data would fail to reject the null hypothesis.

5.2 Participants[42]

This research project aimed to generalise its findings to the wider Higher Education (HE) population rather than only apply the findings to the setting of the particular HEI used in this research. Thus, when testing its hypothesis, this research project sought to select a representative participant group in order to promote external validity and allow legitimate

[42] See Appendix 2 for details of data collection protocols and Appendix 3 for ethical considerations

generalisations to be made beyond the participants in the research project (Bryman, 2009). In order to achieve this outcome the following sampling methods were used:

5.2.1 Inclusion criteria

The three sample groups involved consisted of the following participants:

- Students with a diagnosis of SpLD who were granted 25% extra time in their exams and produced a handwritten response

- Students with a diagnosis of SpLD who were granted 25% extra time and used a word processor in their exams (Humanities exams only)

- Students who did not disclose an SpLD and who sat the same exams under standard conditions

Participants included undergraduate students drawn from the Faculties of English, History, Law, Mathematics and Medicine across all three years of the course. This provided a sample population that represented the largest Humanities subject areas (English, History and Law) as well as two key STEM subject (Maths and Medicine) and so comprised a broad student experience. The SpLD participants were matched with an index group (TD students) of similar numbers of candidates on the basis of exam paper characteristics alone (i.e. the paper of a participant with SpLD was matched with a paper from a TD participant who sat the same examination question paper under standard exam conditions). Due to participant anonymity, no other participant characteristics were taken into account.

Given that the choice of University, Faculty and course of study were predetermined, a non-probability (convenience) sampling process was used to obtain the SpLD participant group (Bryman, 2009). In order to achieve a sample size of sufficient magnitude to allow statistically significant inferences to be drawn, all eligible SpLD candidates who agreed to participate were included in the research project. Therefore, participants with SpLD were not randomly selected, as every candidate who disclosed an SpLD (and agreed to take part) from the Faculties of English, History, Law, Maths and Medicine was included.

5.2.2 SpLD group

When selecting participants for the SpLD groups only those candidates who had been formally diagnosed with SpLD (and held a diagnostic assessment report that met national guidelines relating to SpLD assessment[43]) were included in the research project. For the purposes of this research project, students with either dyslexia or dyspraxia (or both) were included. This inclusion criterion for SpLD participants was based on the view that students with dyslexia and those with dyspraxia experience similar challenges with the access skills needed to demonstrate their potential in exams and it is these particular shared characteristics that can render students with these diagnoses unable to complete the exam task in the designated time (Broggi et al, 2019; Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020; Wade & Kazeck 2018; Wilson et al, 2017). This is the rationale that underpins the granting of access arrangements in exams to candidates with SpLD, with the accommodations acting as a means of redressing the disadvantage this group may be subject to in examination situations solely on account of their disability.

5.2.3 Index group

The SpLD participants were paired with an index group (TD students) of similar numbers of candidates on the basis of exam paper characteristics alone (i.e. the paper of a participant with SpLD was matched with a paper from a TD participant who sat the same examination question paper under standard exam conditions).

Participants in the index (TD) group comprised those students who had not disclosed an SpLD or disability[44], who sat the exam under standard conditions and who completed the same exam paper as the SpLD candidates[45]. Given that there was a greater number of eligible TD participants than participants with SpLD, the TD participants included in this research project were selected from the overall group of those eligible through a systematic random sampling approach. Specifically, the first candidate number was used as the starting point, and every 4th

---

[43] National guidelines relating to SpLD referred to were those established by the SpLD Working Group 2005/ DfES guidelines (and subsequent updates).
[44] It should be noted that the TD candidates had not disclosed any disability, but were not screened for any SpLD or other disability. Therefore, it is not possible to assert that this group did not include participants who had an undiagnosed SpLD or any other undiagnosed disability.
[45] Although each candidate answered the same exam paper (compulsory paper) and the same number of questions, as each candidate had a choice of questions on the paper they may not have responded to the same question.

candidate was selected thereafter until sufficient participants had been obtained. While this method does not completely ensure that the sample group obtained is fully representative of the population as a whole, Fox et al (2009) argue that this sampling method enables justifiable conclusions to be drawn about the population being sampled.

5.3 Sampling process

In order to improve the validity of outcomes, the empirical data were collected from the exam papers of candidates who sat the compulsory exam papers in English, History, Law, and Maths exams and from Functional Architecture of the Body (a compulsory topic) in Medicine. As a result, both the SpLD participants and the TD participants (against whom the results from the SpLD paper were compared) sat the same exam paper within each subject area (albeit not necessarily responding to the same questions), thereby allowing a comparison of performance between the groups on the same exam papers. This approach also aimed to promote the ecological validity of the research, as it can be argued that more accurate inferences can be drawn from the sample data taken directly from the exam activity itself compared with collecting data taken from other timed writing situations (such as comprehension tests or free –writing tasks) which are then generalised into inferences about exam performance (Chaytor et al. 2003).

In addition, mark data from the compulsory dissertations from English, History and Law were collected, alongside the marks of the computations project in Maths and from the practical papers from Functional Architecture of the Body in Medicine. This allowed comparisons to be made between each candidate's exam marks and coursework mark and ascertain whether the relationship between exam mark and coursework mark (if any) for the SpLD group mirrored that of the TD group.

5.3.1 Format of Exams[46]

The Humanities exam papers required essay style responses to 3 or 4 questions (which candidates chose from a list of alternatives) and the standard time allowed for the exams was 3 hours. Candidates with SpLD were permitted 3 hours and 45 minutes (i.e. 25% more time than TD candidates). The Maths exams required responses in the form of mathematical

[46] See Appendix 3 for details of the exam questions in each discipline exam paper

calculation and formulae (so all candidates with SpLD handwrote the paper with 25% extra time and none used a word processor) and candidates are required to answer 9 questions within the standard exam time of 3 hours. The assessment in Medicine consisted of 3 written exams and 2 practical exams. Candidates with SpLD were granted 25% extra time in all 3 written exams but sat the practical exam under standard time conditions. The 3 written exams include a 1-hour multiple choice exam paper, a 2-hour short answer exam paper and a 2-hour long answer exam paper (standard time). Both practical exams comprise a 90-minute ‘steeplechase’ paper with 30 stations. A total of 3 questions are set per station, resulting in 90 questions in total. A total of 3 minutes per station is allowed (90 minutes in total).

5.3.2 Sample Size

A total sample group of 714 participants were recruited for this research project, comprising of 357 participants with SpLD and 357 TD participants. Post-hoc power calculations suggested that this sample size should provide 98% power to either reject or fail to reject this research project’s null hypothesis (alpha =.05; two-tailed; effect size d = 0.3)[47], which should be adequate to detect a difference that is actually present and support statistically derived conclusions that can be generalised from the research project to the wider population (Aberson, 2010). This also indicates that the sample size of 714 would be deemed to be sufficient to enable the research project to detect any statistical differences between the exam performances (in terms or word count, mark and degree classification) of students with SpLD who were granted exam arrangements and those of TD students who took the same exam under standard conditions (and so test this research project’s null hypothesis).

The general sample group of 714 participants was categorized into the following 3 sub-groups:

1. 236 participants who disclosed a diagnosis of SpLD and were granted 25% extra time
2. 121 participants who disclosed a diagnosis of SpLD and were granted 25% extra time plus the use of a word-processor
3. 357 TD participants who took the same exam under standard conditions (index group)

The compulsory exam for Maths and Medicine involved more than one exam paper (4 papers in Maths and 3 papers in Medicine) with the result that the 714 candidates represented 1442

[47] GPower analysis for an ANOVA with effect size 0.3, 2 groups, sample size 357 per group.

candidatures. The breakdown of the sample group size between the Faculties is shown in Table 5.1 below:

Table 5.1: Breakdown of sample group size between English, Law, History, Maths & Medicine

Faculty	**SpLD + 25% extra time**	**SpLD + 25% extra time + W.P**	**TD with standard conditions**	**Total No of Participants**	**Total No of papers per participant**	**Total No of Candidatures**
English	34	34	68	136	1	136
Law	22	30	52	104	1	104
History	36	57	93	186	1	186
Maths	76	0	76	152	4	608
Medicine	68	0	68	136	3	408
Total	236	121	357	714		1442

5.3.3 Variables

The independent variables in this research project were (a) student group [students with SpLD; TD students], (b) exam conditions [extra time; extra time and use of a word processor; standard conditions] and (c) faculty [English; History; Law; Maths; Medicine]. The dependent variables in this research project were (a) the number of words on the exam scripts [total mean word count; words per minute] and (b) the participants' scores [mark; classification]. This research hypothesises that the independent variables (i.e. the SpLD group with exam arrangements and the TD group with no exam arrangements in each faculty) influence the dependent variables (i.e. number of words; marks; classification). However, a possible outcome of this research is that the analysis of the data suggests no relationship exists between the variables cited (null hypothesis).

5.4 Design

5.4.1 Research strategy

The aim of this research was to investigate causal relationships between variables resulting from the existing policy of changing exam conditions for students with SpLD in Higher Education (i.e. examining the effects of a change in conditions experienced by the sample groups where that change in conditions was not implemented for the purposes of the research). Therefore, a quasi-experimental design was implemented using a predominantly deductive, quantitative, research approach (Bryman, 2009). This design involved collecting data from the 2016, 2017 & 2018 end of year University examination scripts of students with SpLD who were granted exam arrangements and comparing this with data from the exam scripts of their TD peers who took the same exam under standard conditions. A statistical approach was used to identify and analyse any differences in performance between the groups and explore the relationship between length of exam papers and marks achieved. This comparison aimed to 'establish the causal influence of a particular variable' by comparing the results obtained from the SpLD group with those of the TD (index) group and 'provide lessons of broad applicability and relevance' (Bryman, 2009, p33).

5.4.2 Procedure

The following data was collected and analysed from each participants' exam papers[48]:

- Total number of words on each Humanities exam script (word count)[49]
- Number of words per minute[50] produced by each participant in Humanities papers (expressed as a mean)
- Marks awarded to each paper
- Degree classification conferred

The word count, mean number of words per minute, marks achieved, and distribution of classifications on the papers of the participants with SpLD were compared to the word count, mean number of words per minute, marks achieved, and distribution of classifications on the

[48] The data was collected at one time point, no intervention was implemented and 3 groups were involved.
[49] See appendix 4 for protocol for counting words on the exam scripts
[50] The mean words per minute score was calculated by counting the total number of words on the script and dividing by the exam time permitted.

papers of TD participants who sat the same exams under standard conditions to see if any statistically significant differences were evident.

In order to establish any relationship between the variables cited above, the analysis of the quantitative data included tests for correlations and differences. A Multifactorial Analysis of Variance (ANOVA)[51] was used to explore differences between:

- SpLD candidates with 25% extra time; TD candidates & word count (for essay style questions)
- SpLD candidates with 25% extra time and the use of a Word Processor; TD candidates & word count (for essay style questions)
- SpLD candidates with 25% extra time; TD candidates & mark
- SpLD candidates with 25% extra time and the use of a Word Processor; TD candidates & mark
- SpLD candidates with 25% extra time; TD candidates & classification
- SpLD candidates with 25% extra time and the use of a Word Processor; TD candidates & classification

A Pearson product-moment correlation was used to explore the following relationships:

- Word count & mark (for essay style questions)
- Word count & classification (for essay style questions)

5.4.3 Word counting method

For the purposes of this research project, the method for counting the words on each exam script followed the protocol set out by the Detailed Assessment of Handwriting (DASH) 17+ for counting words (Barnett et al, 2010) [52]. This test instrument was selected as it is a standardised test of hand-writing speed that includes norms for the writing speeds of 400 UK

[51] An ANOVA was used to identify any statistical differences as the data comprised 3 independent (categorical) groups with different participants in each group (SpLD + 25% extra time; SpLD + 25% extra time + use of Word processor; TD). The dependent variables are continuous (interval). Where the data were parametric, approximately normally distributed for each category of independent variable and there was homogeneity of variance, the ANOVA was used. Where these assumptions were not met (non-parametric data) a Kruskall-Wallis H Test was used.

[52] See Appendix 4 for full details of the protocol for counting words

university students (expressed in words per minute) (Barnett et al, 2010)[53]. In addition, this test instrument included a sample group of students with a diagnosis of dyslexia (n= 33), whose free writing performance was compared to age-matched participants without dyslexia[54].

Given this test instrument's high reliability, ecological validity and the inclusion of sample groups comprising HE cohorts (that included students with dyslexia), it was deemed appropriate to mirror the protocol of the DASH 17+ during the word counting data collection phase of this research. The purpose in following an established protocol was to ensure uniformity of word counting approach across all papers and across disciplines, including those papers that included digits, dates, abbreviations, acronyms, hyphenated words and illegible or crossed out words.

Only exam scripts that comprised essay style answers written in continuous prose were included in the evaluation of word count in this research project (i.e. Humanities papers). The exam scripts of the participants in Mathematics and Medicine were excluded from this element of the data collection and analysis. This ensured that the word counts on scripts were directly comparable and did not involve the use of diagrams, calculations, formulae or multiple-choice responses.

5.4.4 Verification of raw data

To promote reliability, an independent research assistant verified the raw data through a second counting of the number of words on a random sample of 54 exam scripts. These 54 sample papers comprised of 2 scripts (selected at random) from each of the 3 participant group in each of the 3 faculty areas in each year. The marks and classification that each paper achieved were provided by the relevant Faculty through the Faculty Mark Book register and were cross-referenced with the University's Student Information Service database.

5.4.5 Data Analysis techniques

The following data was collected from the exam papers of the participants:

---

[53] The DASH-17+ reports high test–retest reliability (test – retest correlations range from .78 to .95), inter-rater agreement (inter-rater coefficient of reliability for free writing = .96) and internal consistency (Cronbach's alpha = .84) (Barnett et al, 2010)

[54] This revealed statistical differences, with the students with dyslexia writing more slowly than their age matched peers ($F(1,64)= 7.15$, $p<.01$) (Barnett et al, 2010).

- the length of answer on the exam paper (as determined by individual word count)
- the mean number of words per minute produced
- exam paper mark (percentage)
- exam paper classification

The empirical data was analysed using graphical models to express basic statistical information, while more advanced statistical information (such as statistical correlations and differences) was obtained by using SPSS version 25. Measures of central tendencies (means) and dispersion (standard deviation, ranges) were calculated. A multi-factor analysis of variance (ANOVA) was used to compare the differences in performance between students with SpLD (with exam arrangements) and TD students (under standard exam conditions). These multiple data analyses were used to summarise and interpret the results of the research project and provide evidence to support a more detailed analysis of the relationship between time and performance in exams, as well as determining whether these findings are valid for the wider population represented by the research project's sample (and so can be generalised) (Bryman, 2009).

6: QUANTITATIVE RESEARCH FINDINGS: HUMANITIES EXAMS

The findings from the first quantitative phase of this research project are split into three chapters; one focusing on the data from Humanities exams, one focusing on the data from Maths exams and one focusing on the data from Medicine Exams. This is the first of those three findings chapters and it presents the outcomes of the analysis of the quantitative data derived from the Humanities exams, with the aim of answering the research question 'do SpLD candidates who have 25% extra time in exams (or 25% extra time and use of a word processor) produce longer scripts and/or achieve higher exam marks than the TD peers who sat the same exams under standard conditions?'

The chapter starts by providing data about the cohort from the School of Humanities as a whole, to provide an overall context, and goes on to present the characteristics and make-up of the sample group drawn from the School of Humanities who participated in this research project. The word count on the exam scripts of the SpLD and TD group are then compared in order to determine whether or not candidates with SpLD who have 25% extra time in Humanities exams (or 25% extra time and use of a word processor) produce longer scripts than their TD comparators (and so potentially receive an advantage). The mean number of words produced per minute by the two groups are also compared to identify whether or not the participants with SpLD write more slowly than their TD peers in exams (and so warrant the granting of extra time).

The chapter goes on to compare the exam marks and degree classifications of the two groups to identify if the participants with SpLD achieved higher marks than their TD peers in exams and so were potentially advantaged by the exam arrangements they received. This chapter also explores the relationship between the word count on participants' exam scripts and the degree classification they achieved, to ascertain whether there is any correlation between script length and mark. This aims to identify whether or not the ability to produce longer script (as facilitated by exam arrangements) is advantaging. Finally, this chapter attempts to identify whether or not the participants with SpLD were under or over-performing in exams by comparing their dissertation results (where they did not receive any adjustments) with their timed, written, closed book exam results (where they received extra time or use of the word processor with extra time) and cross referencing the outcomes with similar data from the TD participants to see if any between group differences exist.

6.1 Introduction

This first quantitative phase of the research project aimed to identify, through an analysis of actual exam data, the extent to which there is evidence to support the commonly-held belief that the granting of exam access arrangements (i.e. use of a word processor and/or extra time) to students with SpLD confers an advantage by enabling them to produce longer scripts than their TD peers and over-inflates their exam marks. Empirical data was collected from the exam scripts of 426 Humanities students from the high-stakes, end of year, examination period in 2016, 2017 and 2018. The exam scripts for all participants in Humanities were drawn from the Faculties of English, Law and History with exam papers from candidates with SpLD being compared to exam papers from candidates who did not disclose a disability and who sat the same exam under standard conditions[55].

[55] Both the SpLD and TD sample groups were matched on the basis of exam paper alone (no other participant attributes were taken into account).

The English, Law and History degree programme at the HEI where this research was conducted comprises a three-year programme and is assessed by exams at the end of each year. Students are also assessed by a written dissertation in addition to exams. The total number of undergraduate candidates who sat exam papers in 2016 - 2018 in the Faculties of English, Law and History was 5180, of which 233 (4.51%) disclosed an SpLD and were granted exam access arrangements (either 25% extra time or the use of a word processor with 25% extra time). Out of the total cohort, 29.07% of candidates (1506) achieved a first class classification (mark of 70% or above); 65.93% of candidates (3415) achieved a 2:1 classification (a mark of 60 - 69%); 3.49% of candidates (181) achieved a 2:2 classification (a mark of 50 – 59%); 1% of all candidates (52) achieved a 3rd class classification (a mark of 40 – 49%) and 0.5% of all candidates (26) failed the exam (a mark below 40%). This is shown in Table 6.1

Table 6. 1: Degree classifications of total cohort in English, Law & History: 2016 - 2018

1st class (>69%)	2:1 (60 – 69%)	2:2 (50 – 59%)	3rd class (40 – 49%)	Fail (<40%)
n=1506 (29.07%)	n= 3415 (65.93%)	n=181 (3.49%)	n=52 (1%)	n=26 (0.5%)

The breakdown of the overall population of all candidates in English, Law and History in 2016 – 2018 can be seen in Table 6.2.

Table 6. 2: Breakdown of candidates in Humanities undergraduate exams: 2016 – 2018

Subject	Number of TD candidates	Number of SpLD candidates	Total candidates	Mean Mark (whole cohort)	Range of marks (whole cohort)
English	1051 (93.68%)	71 (6.32%)	1122 (100%)	67	23 - 88
Law	1719 (96.35%)	65 (3.64%)	1784 (100%)	66	18 - 83
History	2177 (95.73%)	97 (4.27%)	2274 (100%)	67	21 - 84
Total	4947 (95.49%)	233 (4.51%)	5180 (100%)	67	18 - 88

6.2 Sample group

The sample group for this research project was drawn from the Faculties of English, Law and History and comprised 426 students in total (taken from a total population of 5180). The sample consisted of 213 students with SpLD (92 who completed the exam by hand with 25% extra time and 121 who word processed their responses with 25% extra time). These papers were matched with 213 randomly selected TD peers who sat the same exams under standard conditions. The breakdown of the sample group for each year is shown in Table 6.3.

Table 6.3: Breakdown of participants in each year of the research project: 2016 – 2018

Participant group	English: 2016-2018	Law: 2016-2018	History: 2016-2018
SpLD with 25% extra time (N=92)	34 (15.96%)	22 (10.33%)	36 (16.90%)
SpLD with 25% extra time & word processor (N=121)	34 (15.96%)	30 (14.08%)	57 (26.76%)
SpLD whole group (N=213)	68 (31.92)	52 (24.41%)	93 (43.66%)
TD with standard conditions (N=213)	68 (31.92%)	52 (24.41%)	93 (43.66%)

To ensure consistency, the same subject paper was selected from both SpLD and TD participants in each year. Selecting the same paper for SpLD and TD participants aimed to ensure that comparisons could be made from data extracted from the same exam papers, which aims to improve the reliability of the findings. The examination papers selected in each subject were the compulsory papers[56], which are taken by all undergraduate students in each discipline. All papers consist of a written examination lasting three hours (standard time), with SpLD candidates being permitted 3 hours and 45 minutes.

Mean word counts and marks for the participants' exam scripts have been provided in order to identify any between-group, or within-group, differences in word count or marks. In addition, mean marks for participants who completed a dissertation have been provided in order to see if there are any between-group differences in dissertation marks. The dissertations were compiled over the course of 3 months and the word count

[56] See Appendix 1 for examples of the questions on each exam paper.

permitted is between 5,500 words (minimum) and 6,500 words (maximum)[57]. Candidates with SpLD were not granted extra time in the dissertation and so the marks for the dissertation (which all participants took under the same time conditions) have also been compared with the timed, written exams (which the SpLD group took with 25% extra time) to identify any difference between the dissertation mark and the mark achieved in the timed exam. An ANOVA or Kruskall-Wallis *H*[58] test was used to identify any significant difference between the mean marks achieved by each group.

6.3 Comparison of word count between groups

In order to determine whether the SpLD group who were granted the use of a word processor and/or 25% extra time demonstrated an advantage by producing longer scripts, comparisons were made between the length of the exam scripts of candidates with SpLD and those of TD candidates (as determined by individual exam paper word counts). This comparison aimed to determine if the candidates with SpLD who were granted exam access arrangements produced longer exam scripts (and so demonstrate a word count advantage) in comparison with their TD peers who sat the same exam under standard conditions.

Mean word counts for the paper as a whole have been provided in order to explore any between or within-group differences in length of answer across the exam paper. The mean word count has been used as a model for summarising the data with the aim of identifying the ‘typical’ value or central tendency of the data set. While the mean is useful as it includes all the values in the data set, the median is more resistant to outliers and so may more closely reflect the centre of the distribution if the data set includes outliers. Therefore, where the data is not normally distributed or is significantly skewed, the median has been used to provide a better indication of the strength of the central tendency and the dispersion of the data (Field, 2016).

[57] English candidates write a dissertation in year 2 and year 3 of the degree programme. History and Law candidates write a dissertation in year 3 only.
[58] A non-parametric test was used where the data did not meet the assumptions for a parametric test.

6.3.1 Mean word count

The comparison of the mean word count for each group is shown in Table 6.4.

Table 6.4: Mean word count (M) and Standard Deviation (SD) for SpLD & TD participants - Humanities: 2016 – 2018.

Group	Mean word count English		Mean word count Law		Mean word count History		Total mean word count	
	M	SD	M	SD	M	SD	M	SD
SpLD with 25% ET	2862	508.87	3450	530.84	3374	803.55	3203	690.72
SpLD with 25% ET & WP	2891	489.34	3808	641.47	3504	1027.53	3407	882.09
SpLD whole group	2877	495.69	3657	618.07	3454	944.66	3319	809.65
TD	3060	638.17	3811	678.81	3581	762.82	3471	761.58

When the data from all three Humanities subjects was combined, the distribution of mean word count for the SpLD group across all subjects indicated that most of the scripts of the participants with SpLD were between 2,000– 5,000 words in length. Specifically, 66 out of the 213 scripts (30.99%) were between 2,000-3,000 words in length, 93 out of the 213 scripts (43.66%) were between 3,000-4,000 words in length, and 39 out of 213 scripts (18.31%) were between 4,000-5,000 words long. 8 out of the 213 scripts (3.76%) were below 2,000 words in length and 7 of the 213 scripts (3.29%) were above 5,000 words in length. No scripts were above 6,000 words in length.

While this is the outcome for the SpLD group as a whole, the distribution of the data shows some within group differences in the SpLD sample. For example, when the word count data for the three Humanities subjects areas are combined, the mean word count for the SpLD group who handwrote their scripts with 25% extra time (3,203 words) is lower than the mean word count for the SpLD group who word processed their scripts with 25% extra time (3,407 words). A one-way ANOVA [group, word count], shows that this difference is not statistically significant ($F(1, 211) = 3.376$, $p = .068$)[59]. In total, 5 of the 121 scripts (4.13%) of the SpLD participants who word-processed their scripts with extra time were below 2,000 words in length, 38 out of the 121 scripts (31.40%) were between 2,000 – 3,000 words in length, 45 out of the 121 scripts (37.19%) were between 3,000- 4,000 words in length, 28 out of the 121 scripts (23.14%) were between 4,000- 5,000 words in length and 5 scripts (4.13%) were above 5,000 words in length. By comparison, 3 of the 92 scripts (3.26%) of the SpLD participants who handwrote their scripts with 25% extra time were below 2,000 words in length, 28 out of the 92 scripts (30.43%) were between 2,000 – 3,000 words in length, 48 out of the 92 scripts (52.17%) were between 3,000- 4,000 words in length, 11 out of the 92 scripts (11.96%) were between 4,000- 5,000 words in length and 2 of the scripts (2.17%) were above 5,000 words in length.)

Comparing these trends to the data from the TD group who sat the same exams under standard conditions suggests that the TD participants produced longer scripts than the SpLD participants. In total 1 of the 213 scripts (0.47%) of the TD participants was below 2,000 words in length, 55 out of the 213 scripts (25.82%) were between 2,000 – 3,000 words in length, 106 out of the 213 scripts (49.77%) were between 3,000- 4,000 words in length, 44 out of the 213 scripts (20.66%) were between 4,000- 5,000 words in length and 7 of the 213 scripts (3.29%) were above 5,000 words in length). The comparison of the word length of the scripts of the participants with SpLD with those of the TD index group in is shown in Table 6.5.

---

[59] The Shapiro-Wilk test of normality for the word count of the SpLD group who word-processed their scripts in all three Humanities subjects combined indicated that the data are from a normally distributed population (W=.989; p=.429). This was also true for the word count of the SpLD group who handwrote their scripts ($W = .984$; $p = .312$). Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,211) = .975$; p=.293).

Table 6.5: Comparison of script length between SpLD and TD participants

Group	Script length <2000 wds	Script length 2-3000 wds	Script length 3-4000 wds	Script length 4 – 5000 wds	Script length >5000 wds
SpLD + 25% extra time	4.13%	31.40%	37.19%	23.14%	4.13%
SpLD + word-processor + 25% extra time	3.26%	30.43%	52.17%	11.96%	2.17%
TD	0.47%	25.82%	49.77%	20.66%	3.29%

When the data in all three Humanities subject areas were combined, a one way ANOVA [group, word count] revealed significant differences between the mean word count of the SpLD group as a whole and the mean word count of the TD group, with the SpLD group producing significantly fewer words on the exam scripts ($F(1,424) = 3.981$; $p = .047$[60]) The effect size was very small ($r = .096$; $r^2 = .009$; Cohen's $d = .193$). The comparison of the word count between the TD and SpLD group as a whole is shown in Figure 6.1.

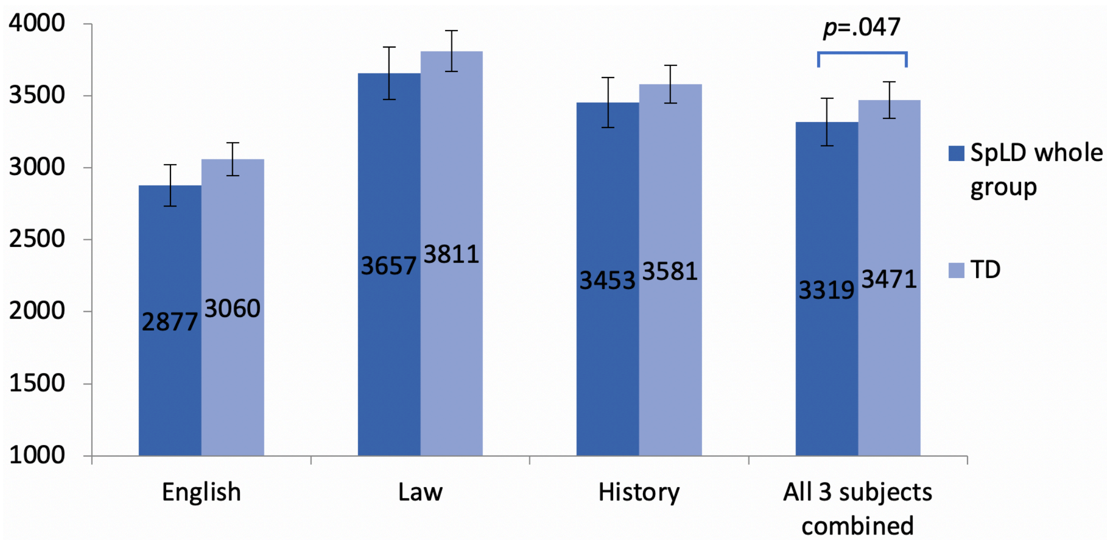


Figure 6.1: Total mean word count (with standard error) SpLD whole group and TD group: 2016-18.

In addition to the between-group differences that were identified between the word

[60] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,184) = .471$; $p = .493$). The Shapiro-Wilk test of normality for the word count of the SpLD group ($W = .981$; $p = .78$) and the TD group ($W = .979$; $p = .061$) indicated that the data were statistically normal.

count of the papers of the SpLD whole group and the TD group, it is useful to compare the word counts of the sub-groups to identify if there were any within-group differences or differences between any of the SpLD sub-groups and the TD group. Therefore, the data between the following sub-groups was compared:

- Mean word count of the SpLD group who handwrote their scripts with 25% extra time and mean word count of the TD group
- Mean word count of the SpLD group who word processed their scripts and mean word count of the TD group
- Mean word count of the SpLD group who handwrote their scripts and mean word count of the SpLD group who word-processed their scripts.

When the word count data from all three Humanities subject areas were combined, no statistical differences were identified between the SpLD group who word-processed their scripts with 25% extra time and the TD group, or the SpLD group who handwrote their scripts with 25% extra time and the SpLD group who word-processed their scripts with 25% extra time. However, statistically significant differences were identified between the SpLD group who handwrote their scripts with 25% extra time and the TD group, with the SpLD group producing significantly fewer words on the script than the TD group, despite the additional time. The effect size was small (r = .181; r^2 = .033; Hedges' g = 0.360; Cohen's d = 0.368). The outcomes of the one-way ANOVA comparing the word count of the sub-groups are shown in Table 6.6.

Table 6.6: Comparison of difference in total mean word count for SpLD and TD participants: All three Humanities subjects combined 2016 – 2018

Group comparison	*df*	*F*	*Sig.*
SpLD 25% extra time (handwritten) group compared with TD group[61]	1, 303	8.417	.004

[61]Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,303)$ =2.592; p=.108)

SpLD word-processor & 25% extra time group[62] compared with TD group	1, 332	.478	.490
SpLD handwritten group[63] compared with SpLD word processed group	1, 211	3.376	.068

The comparison of the total mean word count across all three participant groups in English, Law and History is shown in Figure 6.2

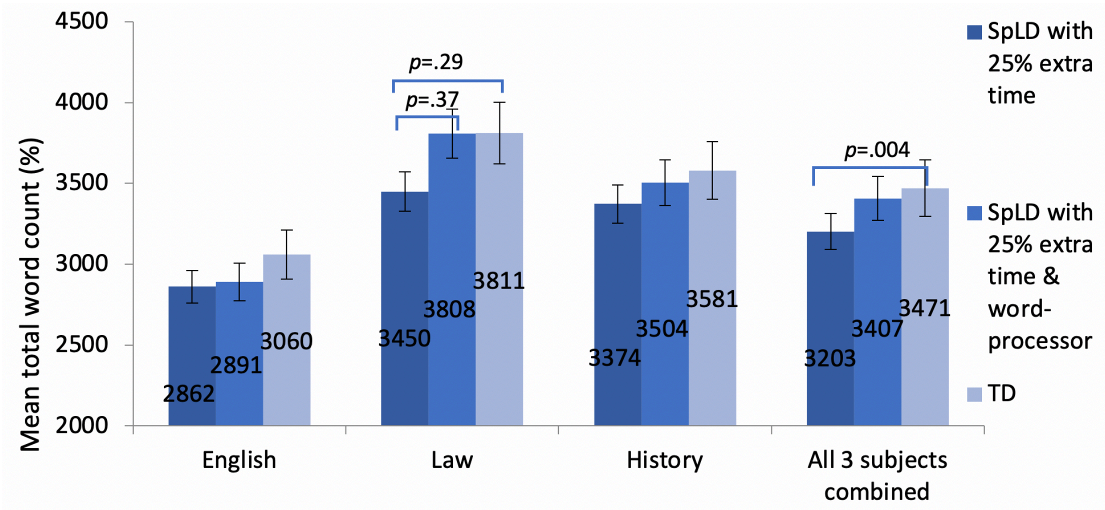


Figure 6.2: Total mean word count (with standard error) SpLD sub-groups and TD group: Humanities 2016 - 2018.

Overall, this analysis of the data has failed to identify that the SpLD group produce longer scripts than the TD participants. Instead, the reverse is the case with the SpLD candidates producing, on average, shorter scripts in Humanities written exams than the TD group despite the extra time. This suggests that the exam access arrangements granted to the SpLD group has not conferred any advantage in terms of an increase in word count and may, especially in the case of the SpLD participants who did not use a word-processor in addition to the extra time, fail to level the playing field in terms of quantity of output.

[62] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,332)=4.037$; $p=.054$)
[63] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,211)=3.140$; $p=.086$)

6.3.2 Comparison of number of words per minute on scripts

While an analysis of the total word count data aimed to show the effect that the granting of additional time may have on the length of the scripts, it is useful to determine if the students with SpLD produced their ideas more slowly in exams than was the case for their TD peers. The theory that individuals with SpLD process information more slowly is a key principle underpinning the rationale for granting extra time in the exams (Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020; Warren, 2017). If it is the case that individuals with SpLD process information more slowly in exams, it would be expected that the number of words produced each minute by the participants with SpLD on the exam scripts would be lower than the number of words produced each minute by the TD participants (Warren, 2017). This comparison removes any mitigating effect that additional time may have on the numbers of words produced by the SpLD group on the script as a whole. In order to determine if this was the case, the mean number of words per minute produced by the participants with SpLD and the TD participants were compared. This facilitated a more direct comparison of writing speed between the two sample groups and the outcomes are shown in Table 6.7.

Table 6.7: Comparison of the Mean number of words per minute produced by the SpLD and TD groups: Humanities 2016 - 2018.

Group	Total mean words per minute: English	Total mean words per minute: Law	Total mean words per minute: History	Total mean words per minute: All 3 subjects combined
SpLD ET	12.87 (SD=2.10)	14.77 (SD= 3.76)	14.99 (SD=3.57)	14.16 (SD=3.71)
SpLD WP& ET	12.69 (SD=2.33)	16.92 (SD=2.85)	15.57 (SD=4.57)	15.10 (SD=3.28)
SpLD whole group	12.78 (SD=2.20)	16.01 (SD=3.41)	15.35 (SD=4.20)	14.69 (SD=3.91)
TD	16.95 (SD=3.49)	21.15 (SD=3.75)	19.89 (SD=4.24	19.26 (SD=4.22)

When the data from all three Humanities subject areas were combined, a one-way ANOVA [group, words per minute] confirmed that the difference in the mean number of words per minute produced by the TD group (19.26 mean words per minute) in comparison with the SpLD group as a whole (14.69 mean words per minute) in the exams was significant ($F(1,424) = 141.059$; $p < .001$). The effect size was very large ($r = .46$; $r^2 = .21$; Hedges' $g = 1.021$; Cohen's $d = 1.022$). This revealed that the SpLD group produced significantly fewer words per minute than the TD group in the Humanities exams. The comparison of the mean words per minute produced by the SpLD whole group and the TD group in each of the three subject areas is shown graphically in Figure 6.3.

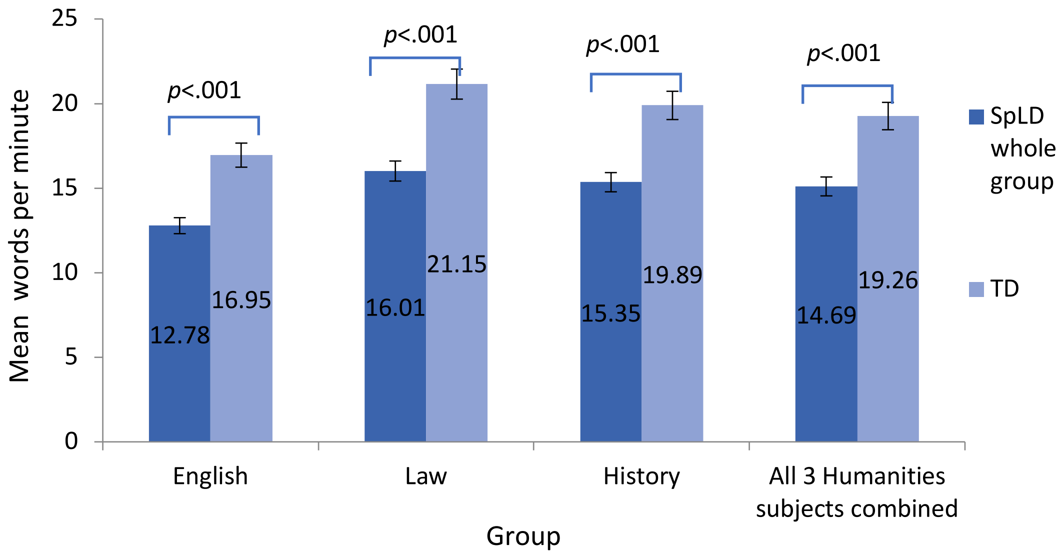


Figure 6.3: Comparison of mean words per minute (with standard error): Humanities 2016-18

The data further shows that the SpLD group who handwrote their scripts produced fewer words per minute (14.16 mean words per minute) than the SpLD group who word-processed their scripts (15.10 mean words per minute), albeit a one-way ANOVA [group, words per minute] confirmed that the difference was not significant ($F(1,211) = 3.404$; $p = .066$). The comparison of the mean words per minute produced by the SpLD sub-groups and the TD group in each of the subject areas is shown graphically in Figure 6.4.

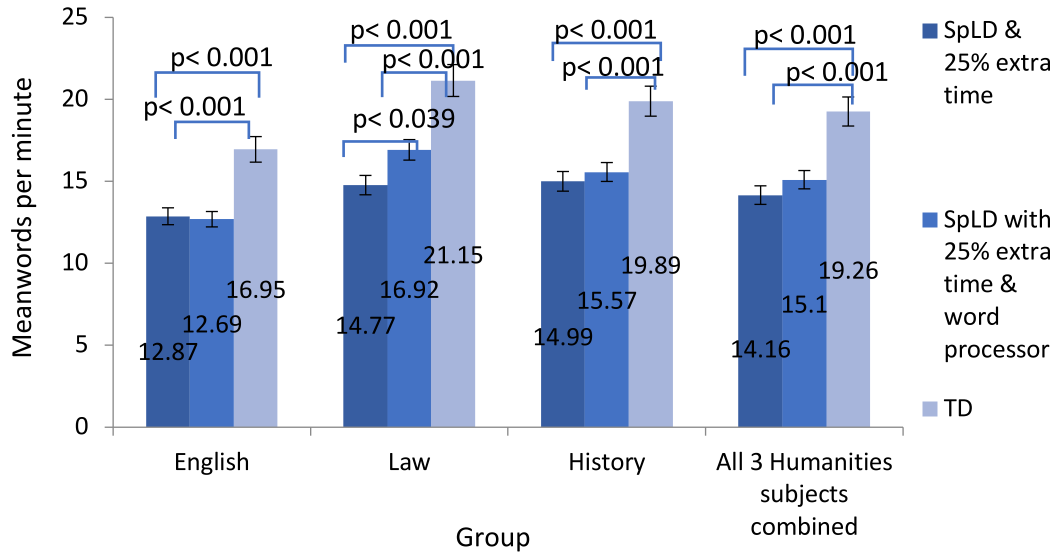


Figure 6.4: Mean words per minute (with standard error) SpLD sub-groups & TD group: Humanities 2016 - 2018

In summary, when analysing the data in relation to the average number of words that each participant group produced in the exam, it can be seen that there were statistically significant differences between the participants with SpLD and the TD participants in the number of words each group produced per minute, with the SpLD group producing fewer words per minute than the TD group. This finding that SpLD candidates write more slowly than their TD peers in Humanities exams supports the rationale for granting extra time to students with SpLD, in order to compensate for the slower production of work. When the overall numbers of words on the script were analysed, significant differences remained in overall word count between the SpLD group and the TD group, with the SpLD group who handwrote their scripts producing significantly fewer words on their exam scripts, despite the additional time. This outcome suggests that the implementation of exam arrangements does not advantage the SpLD group in terms of length of script and may, especially in the case of SpLD candidates who handwrite their scripts, fail to successfully level the playing field in terms of quantity of written output.

6.3.3 Crossed out words

The protocol for counting the number of words on the script followed the guidelines of a standardised test of handwriting, The Dash 17+ (Barnett, et al. 2010) stipulates that the total number of words counted should include any crossed out or illegible words. This is based on the rationale that the candidate has produced the word in the time allocation, even if it has been crossed out. However, it could be argued that crossed out or illegible words do not contribute towards the mark achieved and so it is valuable to compare the percentage of crossed out/illegible words between groups to identify any between group differences. This comparison applies only to the SpLD group who handwrote their scripts and the TD group, as there were no crossed out/illegible words in the typed scripts produced by the SpLD group who were granted the use of a word-processor. The comparison of the mean numbers of crossed out/illegible words between the SpLD (handwritten) and TD groups is shown graphically in Figure 6.5.

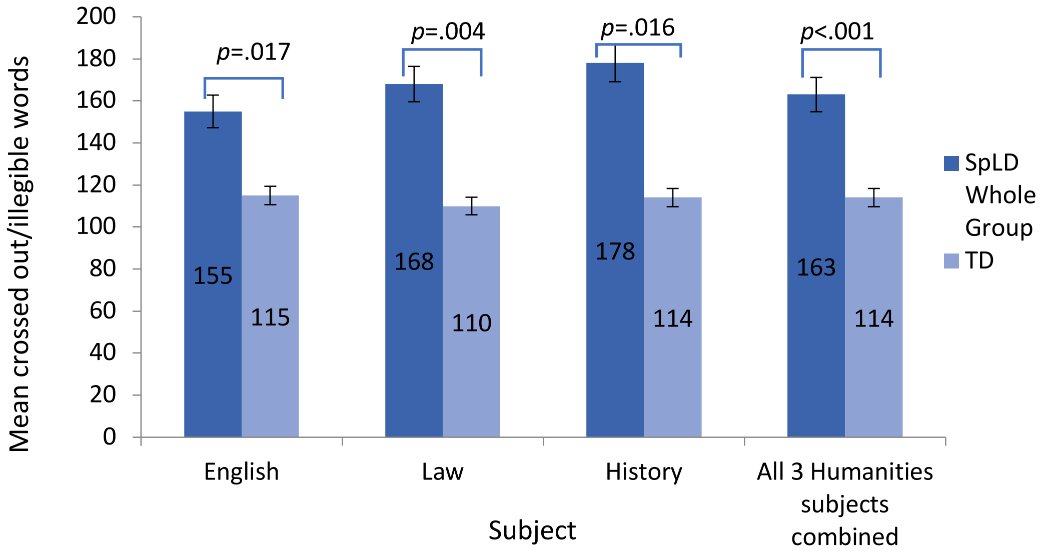


Figure 6.5: Mean number of crossed out/illegible words per question (with standard error): SpLD group with 25% extra time and TD group: Humanities 2016 - 2018

When the data from all three Humanities subjects is combined, the outcome shows that the SpLD group produced, on average, a higher number of crossed out/illegible words in their exam scripts than the TD group. A Kruskal-Wallis *H* test[64] [group, number of crossed out/illegible words] showed that the between-group differences in the number of crossed out/illegible words was significant, with the SpLD group crossing out (or

[64] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the number of crossed out/illegible words of the SpLD group who handwrote their scripts with 25% extra time were not normally distributed (W=.816, p< .001) and the number of crossed out/illegible words of the TD group were also not normally distributed (W=.862, p< .001).

writing illegibly) significantly more words than the TD group ($\chi^2 = 22.705$, p <.001, with a mean rank of 140.95 for the SpLD group and 190.73 for the TD group). The effect size was medium ($r = .22$; $r^2 = .05$; Hedges'g[65] = .57). The finding that the SpLD group cross out their work (or write illegibly) significantly more frequently than their TD peers in the same exam, reflects the findings of studies that suggest that students with SpLD experience more difficulties with structuring ideas, grammar, spelling and legibility of handwriting in essay writing than their TD peers, resulting in increased tendency to cross out words (Callens & Brysbaert, 2019; Callens, Tops & Brysaert, 2012; Jones, Snowling & Moll, 2016; Snowling et al, 2020;).

6.4 Comparison of marks

While the analysis of word count aimed to determine whether or not the SpLD group who were granted exam arrangements in the form of the use of the word processor and/or 25% extra time were advantaged in terms of increased written output, a more relevant test of whether these exam arrangements advantaged one particular group over another is, arguably, a comparison of exam performance as indicated by exam mark. That is to say, if the data supports the commonly held view that the use of a word processor and/or 25% extra time advantages students with SpLD in comparison to their TD peers, then it would be expected that the students with SpLD who were granted these exam arrangements will achieve higher marks than their peers taking the same papers under standard conditions. In order to evaluate this hypothesis, the mean marks of all participants were compared, and the results are shown in Table 6.8.

[65] Hedges' g has been used due to calculate effect size to the uneven size of the sample groups (Sawilowsky, 2009).

Table 6.8: Mean exam marks achieved by SpLD & TD participants: Humanities 2016-18

Participant group	Exam mark: English 2016- 18		Exam mark: Law 2016-18		Exam mark: History 2016-18		Exam mark: All three subjects 2016-18	
	M	SD	M	SD	M	SD	M	SD
SpLD ET	64	4.39	65	4.82	66	4.11	65	4.43
SpLD WP& ET	63	5.18	64	5.40	66	4.49	65	5.05
SpLD whole group	63	4.79	64	5.17	66	4.33	65	4.77
TD	67	4.55	66	3.83	68	3.98	67	4.22

WP = word processor, ET = extra time

When all three Humanities subjects are combined, a one-way ANOVA[66] [group, mark] confirmed that the difference in mean marks achieved in the Humanities exams in 2016-18 by the TD group (67%) compared to the SpLD group who handwrote their script with 25% extra time (65%) was significant, with the SpLD group achieving significantly lower marks ($F(1,303) = 15.986$; $p < .001$). The effect size was small to medium ($r = .22$; $r^2 = .05$; Hedges'g[67] = .45). The comparison of the mean marks of the SpLD group who handwrote their scripts with 25% extra time and the TD group is shown graphically in Figure 6.6.

[66] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,1303) = .512$; $p = .475$). The Shapiro-Wilk test of normality for the mark of the SpLD group who handwrote their script with 25% extra time indicated that the data were statistically normal ($W = .991$; $p = .198$). Similarly, the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were statistically normal ($W = .981$; $p = .196$).

[67] Hedges' g has been used to calculate effect size due to the unequal sample size (Sawilowsky, 2009).

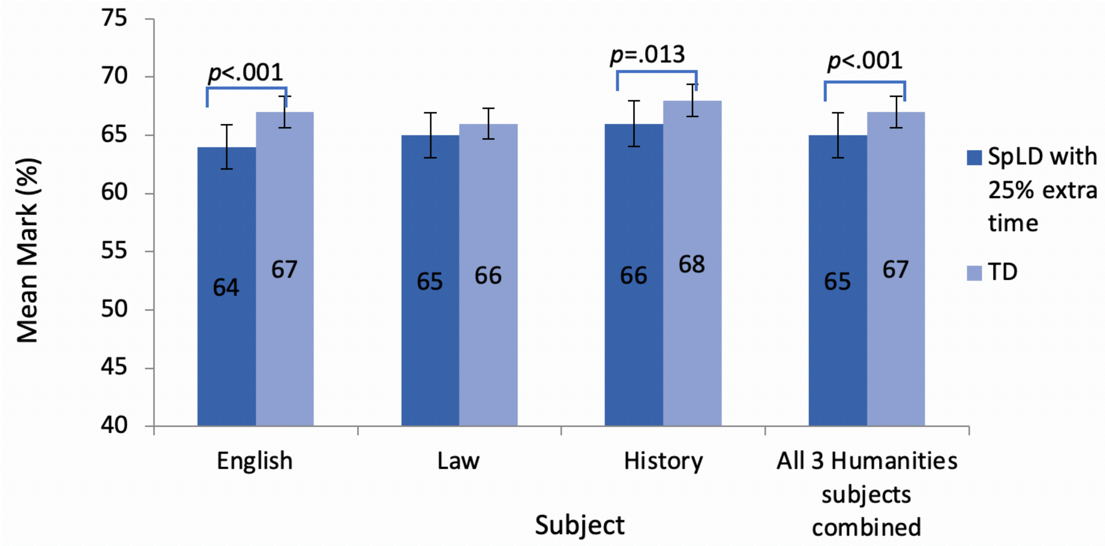


Figure 6.6: Comparison of mean percentage mark per year (with standard error): SpLD group with 25% extra time (handwritten scripts) and TD group: Humanities 2016-18.

In addition, a one-way ANOVA[68] [group, mark] confirmed that the difference in mean marks achieved in the Humanities exams in 2016-18 by the TD group (67%) compared to the SpLD group who word-processed their scripts with 25% extra time (65%) was significant, with the SpLD group achieving significantly lower marks ($F(1,332)$ = 25.455; p <.001). The effect size was small to medium (r = .22; r^2 = .05; Hedges'g[69] = .50). A comparison of the mean marks of the SpLD group who word-processed their scripts with 25% extra time and the TD group is shown graphically in Figure 6.7.

[68] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,1332)$ =1.570; p=.211). The Shapiro-Wilk test of normality for the mark of the SpLD group who word processed their scripts with 25% extra time indicated that the data were statistically normal (W=.991; p=.198).

[69] Hedges' g has been used to calculate effect size due to the unequal sample size (Sawilowsky, 2009).

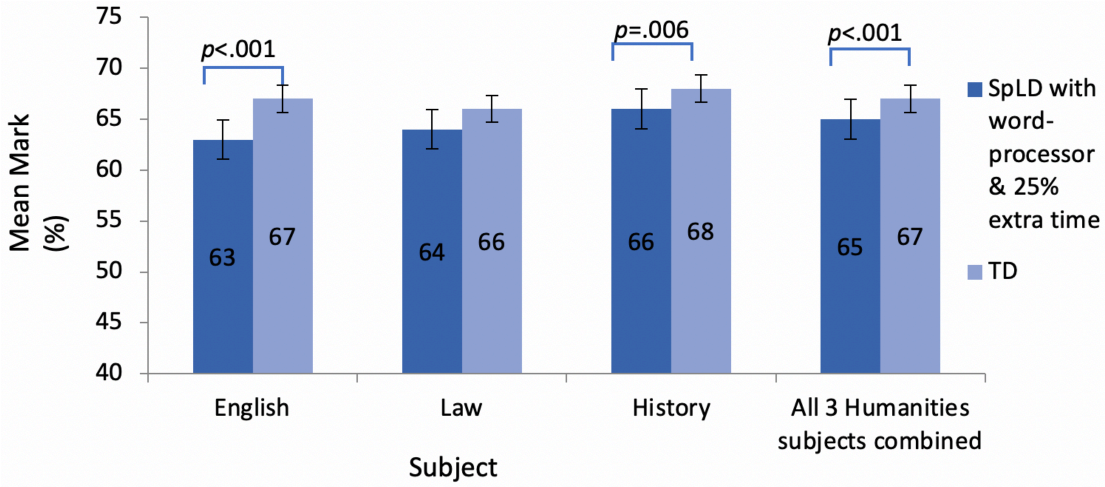


Figure 6.7: Comparison of mean percentage mark per year (with standard error): SpLD group using a word processor with 25% extra time and TD group: Humanities 2016 – 2018.

When the mean marks of the SpLD group as a whole were compared with the mean marks of the TD group in the Humanities exams, a one-way ANOVA[70] [group, mark], confirmed that the difference in marks achieved by the TD group (67%) compared to the SpLD whole group (65%) was significant, with the SpLD group as a whole achieving lower marks (F(1, 1,424) = 30.209; p <.001). The effect size was small to medium (r = .22; r^2 = .05; Hedges'g = .44; Cohen's d=.44). A comparison of the mean percentage mark per year between the SpLD whole group and TD group is shown graphically in Figure 6.8.

[70] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups (F(1,424) =1.531; p=.217). The Shapiro-Wilk test of normality for the mark of the SpLD whole group indicated that the data were statistically normal (W=.978; p=.061).

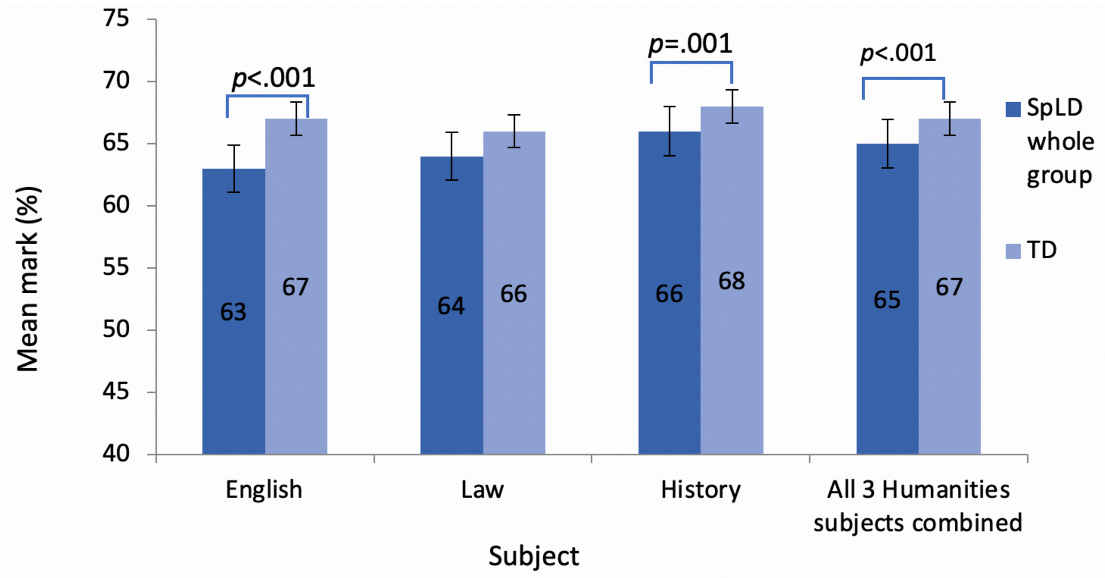


Figure 6.8: Comparison of mean percentage mark per year (with standard error): SpLD whole group and TD group: Humanities 2016 - 2018

In summary, an analysis of the data using one-way ANOVAs [group, mark] revealed that the SpLD candidates achieved statistically significantly lower mean marks than their TD peers in Humanities exams in 2016-18, despite receiving extra time (or the use of a word processor with extra time). This difference in marks between the two groups suggests that, rather than conferring an advantage by inflating marks, the exam adjustments granted to the SpLD candidates may fail to fully compensate for the impact of having an SpLD on construct-irrelevant skills in exams.

6.5 Comparison of classifications

When comparing the performance of SpLD participants with the performance of TD participants, it could be argued that a comparison of marks fails to provide the whole picture. Marks are categorised into degree classifications in order to distinguish the position of the candidate within the cohort as a whole. Given that the classifications pivot on set grade boundaries, any identified differences in marks may not result in similar differences in classification (depending on whether or not the difference in marks crosses a classification boundary). In the institution where the research was conducted, a mark of 70% or above results in a classification of a First-class honours degree (1st); a mark between 60 - 69% equates to Second-class honours, upper division (2:1); a mark between 50 - 59% equates to Second-class honours, lower division (2:2);

and a mark between 40 - 49% results in a Third-class honours degree (3rd). Marks below 40% are classified as a fail.

When the marks from the exam scripts collected in this research project were categorised into classification the results show that the participants with SpLD achieved fewer first-class classifications than the TD participants, a similar number of 2:1 classifications, and a greater number of 2:2 and 3rd degree classifications. Thus, the SpLD group overall achieved fewer higher classifications than the TD group and a greater number of lower classifications. Specifically, 15.96% of SpLD candidates achieved the highest classification (first-class) compared with 30.05% of the TD participants. 71.36% of the SpLD candidates achieved a 2:1 classification compared with 66.67% of the TD participants. 11.27% of the participants with SpLD achieved a 2:2 classification in comparison with 3.29% of their TD comparators and 1.41% of the candidates with SpLD achieved a 3rd class classification whereas no TD participants received this grade. A Kruskal-Wallis *H* test[71] [group, classification] confirmed that these differences were significant, with the SpLD group achieving significantly lower degree classifications ($\chi^2 = 19.786$, $p < .001$, with a mean rank of 191.96 for SpLD and 235.04 for TD).[72] The comparison of degree classifications between SpLD and TD groups in Humanities exams during 2016–18 are shown in Figure 6.9.

[71] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the degree classification data for the TD group were not normally distributed (W=.405, p< .001) and the degree classification data for the SpLD were not normally distributed (W=.392, p< .001). Thus the data did not meet the assumptions required to conduct an ANOVA.

[72] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,134)$ =1.199; p=.275)

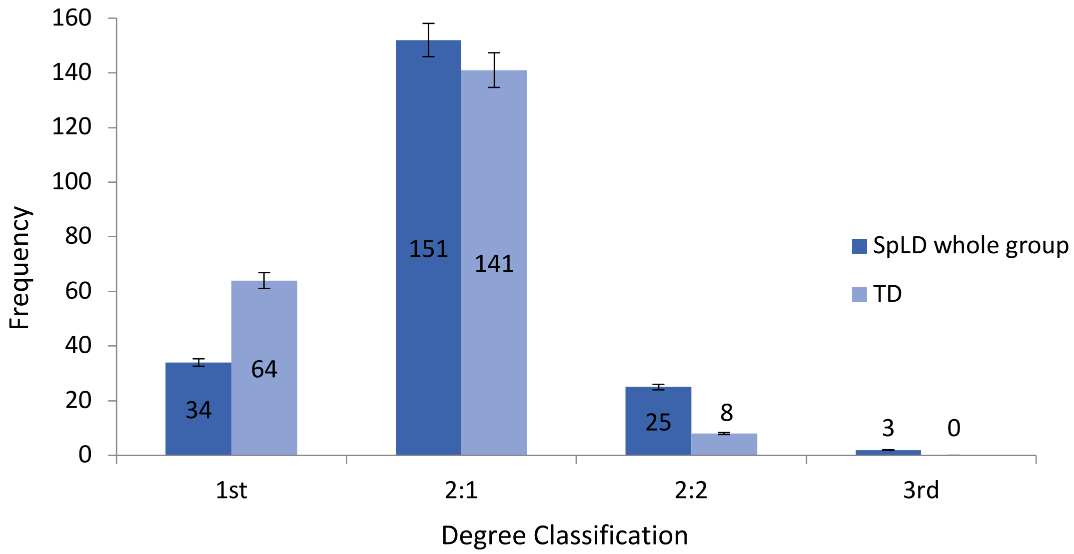


Figure 6.9: Comparison of classification (with standard error) between SpLD and TD groups: All three Humanities subjects combined 2016-18.

The data from this research project shows that, not only are the TD candidates who sat Humanities exam under standard conditions achieving, on average, higher mean percentage marks in the exam than the SpLD participants, the TD group are also being awarded higher degree classifications than the SpLD participants. Not only is this the outcome when the classifications of the TD group are compared with those of the SpLD group as a whole, but this is also the case when the classifications of the TD group are compared with each of the SpLD sub-groups. That is to say, the statistical comparison between the classifications of the SpLD participants who handwrote their scripts with 25% and the TD group who sat the same exams under standard conditions showed that the SpLD participants who used a word-processor with 25% extra time achieved lower classifications than the TD group (*p* <.001)[73]. Similarly, the SpLD group who word-processed their scripts with 25% achieved statistically significantly lower classifications than the TD group (*p* <.001)[74].

---

[73] A Kruskal-Wallis *H* test [group, classification] comparing the classifications achieved by the SpLD participants who handwrote their scripts with 25% extra time with those of the TD participants revealed a significant between group difference (χ^2 = 11.61, *p* = .001, with a mean rank of 174.42 for TD and 143.75 for SpLD).

[74] A Kruskal-Wallis *H* test [group, classification] comparing the classifications achieved by the SpLD participants who word-processed their scripts with 25% extra time with those of the TD participants revealed a significant between group difference (χ^2 = 14.365, *p* <.001, with a mean rank of 189.12 for TD and 155.22 for SpLD (word-processed sub-group))

The data, therefore, suggests that when exam marks are categorised into classifications, the SpLD group are achieving significantly lower classifications than the TD group, even with exam adjustments. The classifications of TD participants more closely reflect the classifications of the cohort as a whole than is the case for the SpLD participants; given that 29% of the whole Humanities cohort achieved a 1st classification, 66% achieved a 2:1, 4% achieved a 2:2, 0.7% achieved a 3rd and 0.3% failed the exam[75]. This highlights a clear awarding gap between the exam achievements of the SpLD students and the cohort as a whole in Humanities exams. Thus, contrary to the belief that the exam arrangements of 25% extra time (or 25% extra time and use of the word-processor) advantage participants with SpLD in terms of over-inflating their performance, this data suggests that candidates with SpLD are potentially underperforming in the exams (as measured by classification as well as mark) and that the exam arrangements fail to fully level the playing field.

6.6 Comparison between mark and word count

The view that additional time, especially when used in conjunction with a word processor, advantages a student pivots on the assumption that these adjustments enable the student to produce longer answers, which in turn, correlates with higher marks. In order for the data to support this view, the evidence should suggest that length of answers (as measured by the number of words on the script) should correlate closely with marks (i.e. long answers should result in high marks and shorter answers should result in lower marks). In order to investigate this hypothesis, the average number of words on the scripts was compared with the classification the script achieved to ascertain if a higher number of words resulted in a higher classification and if shorter answers correlated with lower classifications. The comparison between the mean word count and the classification achieved in Humanities exams 2016-18 is shown in Table 6.9.

[75] Exam classification data for whole cohort Humanities exams was provided by Office of Student Statistics at the research site Institution

Table 6.9: A comparison between mean word count and degree classification: Humanities (all 3 subjects combined) 2016 - 2018.

	1st	2:1	2:2	3rd
Mean word count per paper for candidates with SpLD & 25%Extra time	3028	3123	3372	n/a
Mean word count per paper for candidates with SpLD, word processor & 25% extra time	3338	3321	3322	3390
Mean word count per paper for candidates with SpLD, whole group	3267	3297	3364	3390
Mean word count per paper for TD candidates under standard time conditions	3141	3591	4156	n/a

A Pearson product-moment correlation coefficient was computed for the SpLD participants as a whole to assess the relationship between the number of words on a script and the classification achieved. When the data from all three Humanities subject areas was combined a weak, non-significant negative relationship was identified (r_p = -.029, N=213, p =.679). A Pearson product-moment correlation coefficient was also computed for the TD participants to assess the relationship between the number of words on a script and the classification achieved and a weak, significant negative relationship was identified (r_p = -.313, N=213, p < .001), suggesting that as word count increases, classification decreases.

However, it should be noted that when all three Humanities subjects are combined any relationship may be masked. For example, there were only 2 scripts that resulted in a 3rd classification and of these two scripts, one produced a very low word count (1778) and one produced a very high word count (5002). It can be argued that very low word count results in a low mark due to insufficient quantity of argument or detail and also a very high word count potentially results in a low mark due to the foregrounding of excessive detail over argument. However, when the very high and very low word count are averaged, the resulting figure (3390) is not representative of the data.

Nonetheless, having identified that a high word count on a script fails to correlate strongly with a high classification, a comparison between the word count and the percentage mark was evaluated to identify any relationship between the mark and the word count on the scripts. This comparison of the relationship between word count and mark for the SpLD group is represented in Figure 6.10.

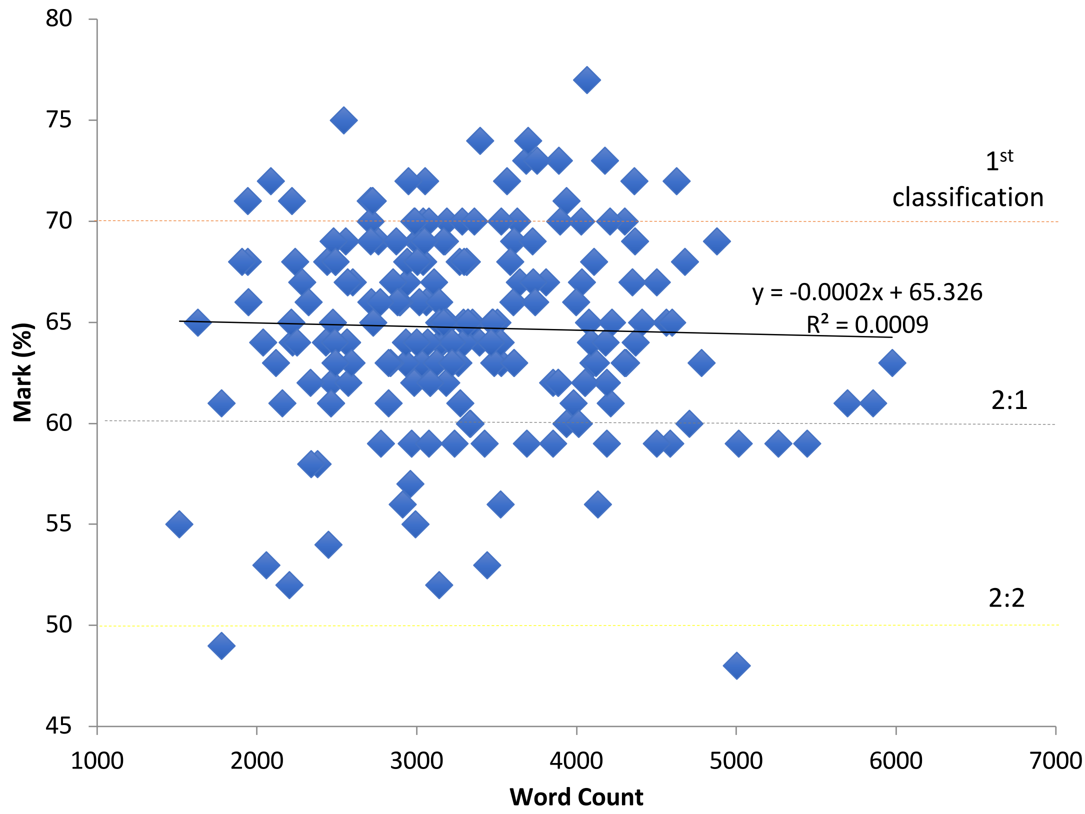


Figure 6.10: Comparison of word count and mark for SpLD whole group: All three Humanities subjects combined 2016 – 2018

A Pearson product-moment correlation coefficient was computed for the participants with SpLD (whole group) to assess the relationship between the number of words on a script and the exam percentage mark achieved when all three Humanities subject areas are combined and a non-significant negative relationship was identified (r_p = -.030, N=213 p = .661). A simple linear regression was calculated to predict exam mark based on word count and a non-significant regression equation was found (F(1,211) = .193, p = .661) with an R^2 of .001[76].

[76] The data met the assumptions of homogeneity of variance and linearity and are approximately normally distributed.

A Pearson product-moment correlation coefficient was then computed for TD students who sat the exam under standard conditions to assess the relationship between the number of words on a script and the exam mark achieved and a significant moderate negative relationship was found (r_p = -.401, N=213, p < .001). A simple linear regression was calculated to predict mark based on word count and a significant regression equation was found (F(1,211) = 40.514, p < .001) with an R^2 of .161 (so 1.61% of the variation in mark can be explained by the model). Participants predicted mark is equal to 74.86 - .0006 (mark) words when words are measured in units of 1. The slope coefficient for mark was -.0022, so participants' mean marks decreased by .0022 for each additional word produced in the exam script[77]. The relationship between word count and exam mark for the TD participants in Humanities exams is shown in Figure 6.11.

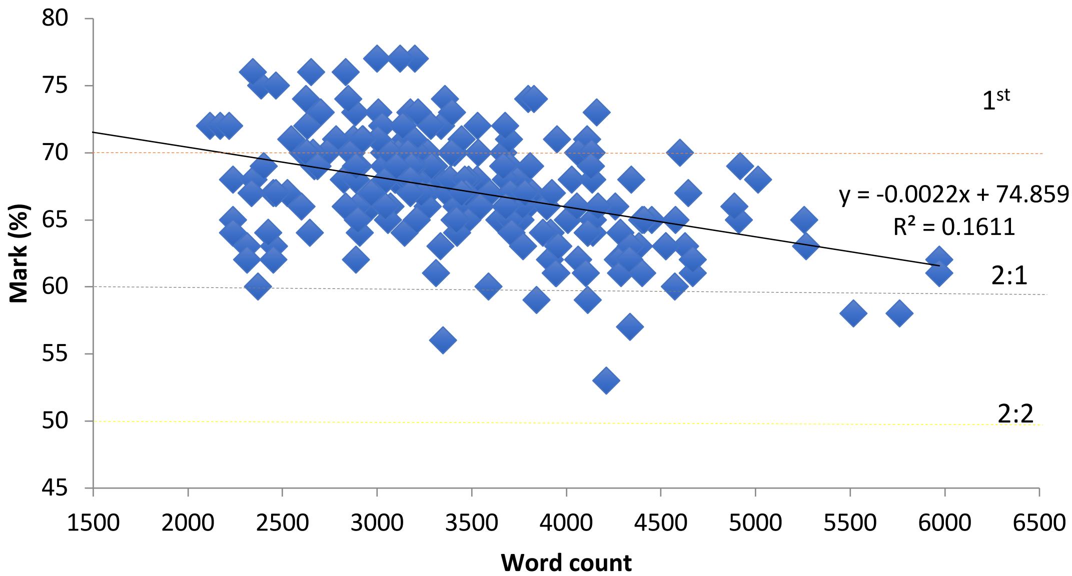


Figure 6.11: Relationship between word count and mark for TD group: All three Humanities subjects combined 2016- 2018

In order to ascertain if differences in the relationship between word count and mark exist between the TD group and the SpLD group as a whole, a multi-factor ANOVA[78] exploring between group differences for the effect of word count on mark was

[77] The data met the assumptions of homogeneity of variance and linearity and are approximately normally distributed.
[78] Levene's Test of Equality of Error Variances confirmed that variances are equal across the two participant groups (F(1,424) =1.531; p=.217)

conducted. This revealed a significant between group difference in the relationship between marks and word count, with a medium effect size ($F(1,424) = 3.981$, $p = .047$; Wilks' $\Lambda = .914$, partial $\eta^2 = .086$). This outcome suggests that any increase in word count produced by the TD group in Humanities exam has a greater negative effect on their marks than any increase in word count does for the SpLD group.

To identify any potential differences between the SpLD sub-groups, a multi-factor ANOVA[79] exploring within-group differences for the effect of word count on mark was conducted. The multi-factor ANOVA revealed no significant between group differences in the relationship between marks and word count between the SpLD group who handwrote their scripts and the SpLD group who word processed their scripts ($F(1,212) = .505$, $p = .478$; Wilks' $\Lambda = .982$, partial $\eta^2 = .002$). This suggests that the effect on mark of any increase in word count is similar for those SpLD participants who handwrite their scripts as it is for those who word process their scripts.

6.7 Outlier data

It should be borne in mind that the data above is reported as averages, which may mask individual differences. It is therefore worth evaluating outlier data. This will include examining the highest and lowest word counts on individual papers and their corresponding marks and classification.

6.7.1 Highest word counts

The highest word count in an English exam paper was 4,667 words. This paper was produced by a TD candidate under standard time conditions and received a mark of 61% (2:1 classification). In the Law exams, the highest word count was 5,268. This paper was produced by a TD candidate under standard time conditions and received a mark of 63% (2:1 classification). The highest word count in the History exam papers was 5975. This paper was produced by a candidate with SpLD who used a word processer with 25% extra time in the exam and the paper achieved a mark of 63% (2:1 classification)

---

[79] Levene's Test of Equality of Error Variances confirmed that variances are equal across the two participant groups ($F(1,211) = .184$; $p = .669$)

When all three subjects are combined, from the whole group of 426 participants, only 14 participants produced a paper with a word count above 5,000 (7 TD participants and 7 SpLD participants - of whom 5 used a word processor with extra time and 2 handwrote their scripts with extra time). The mean mark of these 14 candidates was 60%, which suggests that high word count is not closely associated with a high mark. This scrutiny of the exam scripts with the highest word counts suggests that unusually long answers are not commonly produced and that TD candidates produce a similar number of these unusually long papers as SpLD candidates.

6.7.2 Lowest word counts

The lowest word count on an English exam paper was 1,316 words. This paper was produced by a TD candidate and received a mark of 57% (2:2 classification). The lowest word count on a Law exam paper was 2,239 words. This paper was produced by a TD candidate and received a mark of 64% (2:1 classification). In History, the lowest word count on the exam paper was 1631, which was produced by a candidate with SpLD who handwrote the scripts with 25% extra time and received a mark of 65% (2:1 classification).

When all three subjects are combined, from the whole group of 426 participants only 8 candidates produced papers with a word count below 2,000 (1 TD participant and 7 SpLD participants - of whom 5 used a word processor with extra time and 2 handwrote their scripts with extra time). The mean mark of these 8 candidates was 62%. This data suggests that unusually short answers are rarely produced and that candidates with SpLD produce more of these unusually short scripts than TD candidates. A breakdown of the word count (in groups of 1,000 words) for each participant group is shown in Table 6.10.

Table 6.10: Breakdown of word counts of each participant group: All three Humanities subjects.

	Under 2,000 words	2,000–3,000	3,000–4,000	4,000-5000	Above 5,000 words
SpLD with 25% extra time (n=92)	2 (2.17%)	28 (30.43%)	49 (53.26%)	11 (11.96%)	2 (2.17%)
SpLD with word processor & 25% extra time (n=121)	5 (4.13%)	38 (31.40%)	45 (37.19%)	28 (23.14%)	5 (4.13%)
SpLD whole group (n= 213)	7 (3.29%)	66 (30.99%)	94 (44.13%)	39 (18.31%)	7 (3.29%)
TD candidates with standard conditions (n=213)	1 (0.47%)	54 (25.35%)	107 (50.23%)	44 (20.66%)	7 (3.29%)
Total (n=426)	8 (1.88%)	120 (28.17%)	201 (47.18%)	83 (19.48%)	14 (3.29%)

The word count data indicates that 94.83% of the candidates in the Humanities exams produced papers between 3,000 – 5,000 words in length, while only 3.29% of the candidates produced papers over 5,000 words long and only 1.88% of the candidates produced papers less than 3,000 words in length. More candidates produced scripts that were between 3,000 – 4,000 words in length (47.18%) than any other category.

6.8 Dissertation

Out of the 426 participants in this research project, 302 participants were assessed by dissertation in addition to the exam. This consisted of 136 of the English participants, 38 of the Law participants and 128 of the History participants. All students who produce the dissertation do so in word-processed form, no additional time is granted (all candidates are permitted up to 3 months to write the dissertation) and the maximum permitted word count is 6,500 (minimum = 5,500). Therefore, the students with SpLD were assessed under the same conditions as TD students. Given that university students with SpLD generally report that their coursework marks are higher than their exam marks (Cameron, 2016; Camilleri et al, 2019; Chetcuti, 2019; Fullarton & Duquette, 2016), it is worthwhile comparing the marks of the SpLD students in their exams with the marks they achieved in their dissertations to see if any differences exist, on the basis that coursework may be a more representative assessment of their academic potential. In order to establish if any differences between exam marks and dissertation marks are specific to students with SpLD, the exam marks and dissertation marks of TD students have also been compared and the outcomes for these two groups have been cross referenced.

When the dissertation marks from all three Humanities subjects are combined the mean dissertation mark for the 151 SpLD participants was 69% (SD = 4.78) and the mean dissertation mark for the 151 TD participants was 68% (SD = 4.20). A one-way ANOVA[80] [group, dissertation mark] confirmed that the between-group difference in dissertation mark was not significant ($F(1,300) = 2.487$, $p = .116$). 3 of the 151 dissertations (1.99%) of the SpLD participants achieved a mark between 50-59% (a 2:2 classification), 83 out of the 151 dissertations (54.97%) achieved a mark between 60-69% (2:1 classification) and 65 of the 151 dissertations (43.05%) achieved a mark above 70% (1st class classification). Similarly, 3 of the 151 dissertations (1.99%) of the TD participants achieved a mark between 50-59% (a 2:2 classification), 90 out of the 151 dissertations (59.6%) achieved a mark between 60-69% (2:1 classification) and 58 of the 151 dissertations (38.41%) achieved a mark above 70% (1st class classification). The comparison of the dissertation marks achieved by the SpLD and TD groups in Humanities subjects in 2016-18 is shown in in Table 6.11.

[80] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,300)$ =1.958; p=.163). The Shapiro-Wilk test of normality indicated that the dissertation marks of the SpLD group were normally distributed (W=.993; p=.641) and the dissertation marks of the TD group were normally distributed (W=.988, p=.229).

Table 6.11: Comparison of the dissertation marks achieved by the SpLD and TD groups in Humanities subjects in 2016-18.

Group	N	50 – 59% (2:2)	60 – 69% (2:1)	>69% (1st)
SpLD Whole Group	151	3 (1.99%)	83 (54.97%)	65 (43.05%)
TD	151	3 (1.99%)	90 (59.60%)	58 (38.41%)

This reveals that, unlike their performance in the timed, closed book, written exams, the performance of the SpLD and TD groups in coursework assessments is statistically similar, and suggests that the disparity in assessment performance that exists in the timed, unseen, closed book, written exams cannot be accounted for by a difference in ability between the two groups.

When the exam marks and dissertation marks of the participants were compared, a one-way ANOVA[81] [exam mark, dissertation mark] confirmed that the difference between the mean dissertation mark (68%) and the mean exam mark (67%) for the TD group was not significant ($F\,(1{,}150) = 1.552, p = .568$). However, a one-way ANOVA[82] [exam mark; dissertation mark] confirmed that the difference between the mean dissertation mark (69%) and the mean exam mark (65%) for the SpLD group, was significant (F $(1{,}150) = 3.063$, $p > .001$), with the SpLD group achieving significantly lower marks in their exams than they received for their dissertations. This was a large effect size ($r = .39$; $r^2 = .15$; Hedges' g[83] $= .835$, Cohen's $d = .836$) and suggests that the participants with SpLD underperform in timed, closed book, written exams.

In order to ascertain if differences in the relationship between dissertation mark and exam mark exist between participant groups, a multi-factor ANOVA exploring between group differences for the effect of exam mark on dissertation mark was conducted.

[81] The Shapiro-Wilk test of normality indicated that the dissertation marks of the TD group were normally distributed (W=.991, p=.198) and the exam marks of the TD group were also normally distributed (W=.988, p=.229). Levene's Test of Homogeneity of Variances confirmed that variances are equal (F19,127) =1.522; p=.088).

[82] The Shapiro-Wilk test of normality indicated that the dissertation marks of the SpLD group were normally distributed (W=.993, p=.641) and the exam marks of the SpLD group were also normally distributed (W=.973, p=.596). Levene's Test of Homogeneity of Variances confirmed that variances are equal (F(19,127) =1.135; p=.325).

[83] Hedges' g has been used due to calculate effect size to the unequal sample sizes (Sawilowsky, 2009).

When all three Humanities subjects are combined, the multi-factor ANOVA[84] revealed a significant between group difference in the relationship between dissertation marks and exam marks, with the SpLD group showing a larger difference between their exam mark and dissertation mark than the TD group. The effect size was large ($F(1,300) = 24.265$, $p < .001$; Wilks' $\Lambda = .857$, partial $\eta^2 = .143$). The between group comparison of the difference between the mean exam mark and mean dissertation mark for the SpLD and TD groups in Humanities subjects areas is shown graphically in Figure 6.12.

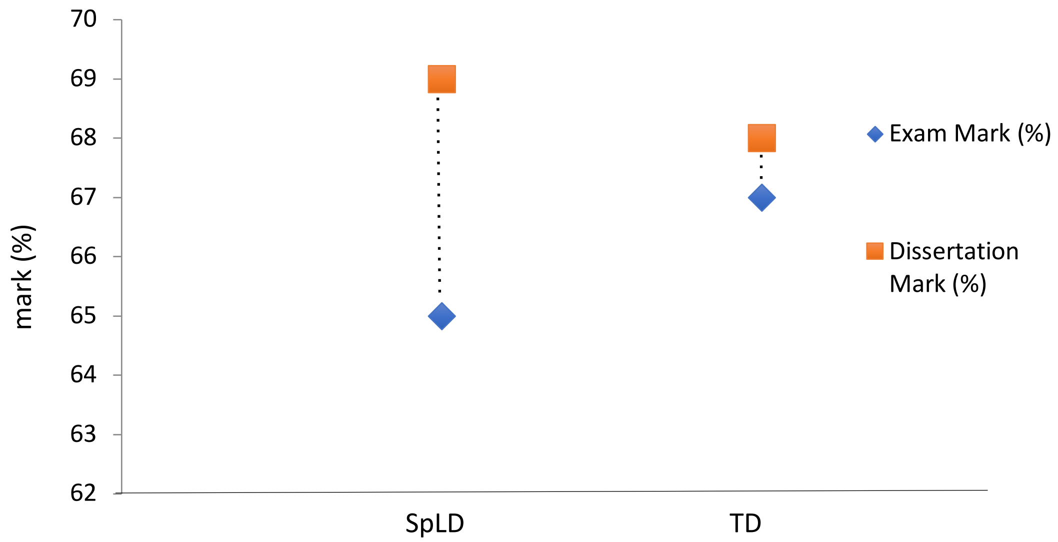


Figure 6.12: Between group comparison between exam mark and dissertation mark for TD & SpLD groups: All three subjects combined 2016-18.

In addition to the between-group differences in the relationship between exam marks and dissertation marks, potential between-group differences in the relationship between exam classification and dissertation classification were analysed. When all three Humanities subjects were combined a total of 43.05% (65 out of 151) of the SpLD participants who wrote a dissertation achieved a First-class classification for the dissertation, while only 15.96% (34 out of 213) of the SpLD participants who sat the exam achieved a First-class classification in the exam. In comparison, 38.41% (58 out of 151) of the TD participants who wrote a dissertation achieved a First-class classification in the dissertation and 30.05% (64 out of 213) of the TD participants who

[84] Levene's Test of Equality of Error Variances confirmed that variances are equal ($F(1,300) = .347$; $p = .556$)

sat the exam achieved a First-class classification in the exam. A multifactorial ANOVA[85] comparing the relationship between exam classification and dissertation classification across the two groups showed significant between-group difference, with the SpLD group showing a greater difference between dissertation classification and exam classification than the TD group. The effect size was medium ($F(1,300) = 19.708$, $p <.001$; Wilks' $\Lambda = .742$, partial $\eta^2 = .062$).

6.9 Conclusion

Overall, the data identified significant differences between the total word count produced by the SpLD and TD groups in Humanities exams in 2016-18, with the SpLD group producing fewer words overall on a script. The effect size was small and related only to the SpLD group who handwrote their scripts with 25% extra time – no statistically significant differences were observed between the length of scripts produced by the SpLD group who word processed their scripts with extra time and those of their TD peers who sat the same exam under standard conditions. In addition, statistically significant differences were observed between the number of words that the SpLD group produced per minute of the exam and the number of words that the TD group produced in each minute of the exam, with the SpLD group producing fewer words per minute (this was the case both for the SpLD group who handwrote their scripts and for the SpLD group who word-processed their scripts). The effect size was large. This finding supports the rationale for granting extra time in exams to candidates with SpLD simply on the grounds of leveling the playing field in terms of quantity of output.

When performance in terms of exam mark and degree classification was analysed, significant between-group differences emerged, with the SpLD participants achieving significantly lower exam marks and lower exam classifications than their TD peers who sat the same exams under standard conditions. The classifications and the marks of the TD participants in this research project more closely reflect the results achieved by the

[85] Levene's Test of Equality of Error Variances confirmed that variances are equal across the two participant groups ($F(1,300) =1.802$; $p=.181$).

whole University cohort of students in English, Law and History exams than was the case for the SpLD participants. Specifically, both the TD group and the whole University cohort achieved a mean mark of 67% in Humanities exams in 2016-18, while the SpLD participants achieved a mean mark of 65%. The comparison of classifications between the participants in this research project and the whole University cohort is shown in Table 6.12.

Table 6.12: Comparison of classifications between the whole of the University cohort in English, History and Law combined and the participants in this research project: 2016-18.

Group	1st class classification	2:1 classification	2:2 classification	3rd class classification
Whole University Cohort	28.14%	66.81%	3.9%	0.74%
TD study Participants	30.05%	66.67%	3.29%	0%
SpLD study participants	15.96%	71.36%	11.27%	1.41%

Interestingly, when the marks achieved in the dissertations by both the TD and SpLD groups in this research project were analysed, no statistical differences between the two groups were identified. However, differences emerged in the between-group relationship between exam mark and dissertation mark, with the gap between the exam mark and dissertation mark being greater for the SpLD group than the TD group. This suggests that the SpLD group are underperforming in the exams (as measured by exam mark) by comparison to their potential (as measured by dissertation mark), despite the exam adjustments. Thus, the findings from this phase of the research project suggest that the exam access arrangements of 25% extra time or 25% extra time and the use of a word processor do not advantage the SpLD group in terms of increased word count, higher marks or higher classifications than their TD peers in timed, written, exams. In fact the opposite is the case, with the exam adjustments failing to fully compensate for the impact of having an SpLD on construct-irrelevant skills in exams, leaving an awarding gap that is not accounted for by differences in ability.

7: QUANTITATIVE RESEARCH FINDINGS: MATHS EXAMS

7.1 Introduction

7.2 Sample Group

7.3. Comparison of marks

7.4 Comparison of Classifications

7.5 Assessed coursework (Computational) project

7.6 Comparison of exam mark and coursework mark

7.7 Conclusion

This is the second of the three findings chapters from the first quantitative phase of this research project that aims to address the research question of whether or not exam arrangements granted to SpLD candidates confer an advantage or promote equity. This chapter presents the analysis and outcomes of three years of data relating to the exam marks achieved in Maths exams by candidates with SpLD and compares this with similar exam data from their TD peers who sat the same exams under standard conditions.

The chapter starts by providing data about the cohort from the Department of Mathematics as a whole, to provide an overall context, and goes on to present the characteristics and make-up of the sample group drawn from the Department of Mathematics who participated in this research project. The chapter then compares the exam marks and degree classifications of the participants with SpLD with those of the TD participants to identify if the participants with SpLD achieved higher marks or degree classifications than their TD peers in exams and so were potentially advantaged by the exam arrangements they received. Finally, this chapter attempts to identify whether or not the participants with SpLD were under or over-performing in exams by comparing their assessed coursework results (where they did not receive any adjustments) with their timed, written, closed book, exam results (where they received extra time) and cross referencing the outcomes with similar data from the TD participants to explore any between group differences.

7.1 Introduction

The Mathematics degree programme at the HEI where this research was conducted comprises a three year programme and is assessed by exams at the end of each year. Students in the second year and third year are also assessed by a written, computational project (coursework) in addition to the exams. There is no assessed coursework in the first year. The total number of undergraduate students who sat the Maths exams in the years 2016-18 combined was 2084 of which 93 (4.46%) disclosed an SpLD and were granted 25% additional time in the exam. In the 2016-18 summer exams, 29.99% of candidates (625) achieved a first class classification (mark of 70% or above); 40.02% of candidates (834) achieved a 2:1 classification (a mark of 60–69%); 21.59% of candidates (450) achieved a 2:2 classification (a mark of 50–59%); 7.2% of all candidates (150) achieved a 3rd class classification (a mark of 40–50%) and 1.2% of all candidates (25) failed the exam (a mark below 40%). The breakdown of the overall population of 2084 candidates can be seen in Table 7.1.

Table 7.1: Breakdown of candidates in Maths undergraduate exams: 2016-18

Exam year	Number of TD candidates	Number of SpLD candidates	Number of candidates in whole cohort	Mean Mark (whole cohort)	Range of marks[86] (whole cohort)
2016	658 (96.20%)	26 (3.80%)	684	65	20 - 82
2017	675 (96.15%)	27 (3.85%)	702	66	22 – 91
2018	675 (96.70%)	23 (3.30%)	698	66	21 - 89
2016-18	2008 (96.35%)	76 (3.65%)	2084	66	21- 91

7.2 Sample Group

The total sample group drawn from the Department of Mathematics for this research project comprised 152 participants in total (selected from a total population of 2,084). The sample group consisted of 76 participants with SpLD, all of whom wrote by hand

[86] Lowest - highest score (data published by the students' statistics office)

with 25% extra time. These participants were matched with 76 randomly selected TD peers who sat the same exams under standard conditions. Due to the nature of the exam, which consisted of writing mathematical formulas and calculations rather than prose, none of the candidates used a word processor.

A total of 608 papers were drawn from year 1, year 2 and year 3 of the Maths undergraduate degree in 2016 - 18. This represented 4 papers sat by each of the 152 participants (76 SpLD and 76 TD participants). To ensure consistency, the same subject papers[87] were selected from both SpLD and TD participants in each year[88]. The standard time allowed for the Maths exam was 3 hours and SpLD candidates who are granted 25% extra time were allowed 3 hours and 45 minutes for the exam.

The mean exam marks achieved by each sample group has been compared in order to identify any between-participant differences in exam performance. The mean marks for the coursework (computational) project have also been provided to identify any between-group or within-group differences between the coursework project mark and the mark achieved in the timed exam. Extra time was not granted to candidates with SpLD for the coursework project and so both groups undertook the assessed coursework under the same time conditions.

7.3 Comparison of marks

Given that the Maths exams require responses in the form of mathematical calculation and formulae, word count is not a valid measure of the effect of extra time on exam scripts. Therefore, no analysis of word count was undertaken on the Maths papers and only data relating to exam mark was analysed. If the data supports the commonly held view that awarding 25% extra time advantages students with SpLD in comparison to their TD peers, then it would be expected that the students with SpLD who were granted extra time will achieve higher marks than their TD peers taking the same papers under standard conditions. This would suggest that the extra time over-inflated the

[87] The 4 papers were: 1) Vectors and Matrices; 2) Differential Equations; 3) Groups and Vector Calculus; 4) Numbers and Sets; Dynamics and Relativity.
[88] See Appendix 1 for examples of questions on each paper.

performance of the candidates with SpLD. In order to evaluate this hypothesis, the scores of all participants were compared and the results are shown in Table 7.2.

Table 7.2: Mean exam marks achieved by SpLD & TD participants: Maths 2016-18

Participant group	Exam mark: 2016		Exam mark: 2017		Exam mark: 2018		Exam mark: 2016-18 combined	
	M	SD	M	SD	M	SD	M	SD
SpLD with 25% extra time	60	10.58	62	10.99	61	8.89	61	10.30
TD	63	12.68	67	10.87	67	10.09	65	11.34

When the data from all three years were combined, a Kruskal-Wallis *H* test[89] [group, mark] confirmed that the difference in marks achieved in the Maths exams in 2016-18 by the TD group (65%) compared to the SpLD group who were granted 25% extra time (61%) was significant, with the SpLD group achieving significantly lower marks (χ^2 = 3.558, p = .049, with a mean rank of 69.77 for SpLD and 82.23 for TD). This was a small effect size (r = .18; r^2 = .03; Cohen's d = .37). These findings suggests that the exam access arrangements of 25% extra time granted to SpLD candidates in Maths exams do not advantage the SpLD group in terms of over-inflating their marks. The comparison of the mean marks of the SpLD group who were granted 25% extra time and the TD group is shown graphically in Figure 7.1.

[89] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups (F(1,150) = .063; p = .802). The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were *not* statistically normal (W = .994; p = .002). However, the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were statistically normal (W = .985; p = .483). As the assumptions for an ANOVA were not met, a parametric test was used.

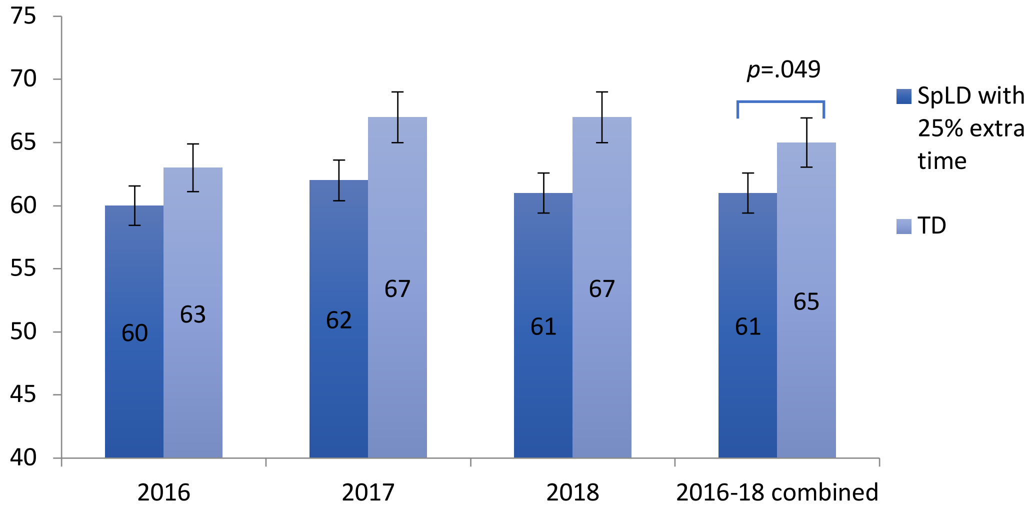


Figure 7.1: Comparison of mean percentage mark per year (with standard error): SpLD group with 25% extra time and TD group: Maths 2016-18.

7.4 Comparison of Classifications

When the marks were categorised into classification the results showed that the participants with SpLD achieved fewer first-class classifications than the TD participants, a greater number of 2:1 and 2:2 classifications and a similar number of 3rd classifications. A Kruskal-Wallis *H* test[90] [group, classification] confirmed that the classifications awarded in the Maths exams in 2016-18 to the TD group were significantly higher than those awarded to the SpLD group who were granted 25% extra time ($\chi^2 = 4.989$, $p = .026$, with a mean rank of 68.93 for SpLD and 84.07 for TD). The effect size was small ($r = .008$; $r^2 < .001$; Cohen's $d = .02$). The comparison of classifications achieved by each group can be seen in Figure 7.2.

[90] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,150) = .754$; $p = .387$). However, the Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were *not* statistically normal ($W = .869$; $p < .001$) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were also not statistically normal ($W = .849$; $p < .001$). As the assumptions for an ANOVA were not met, a parametric test was used.

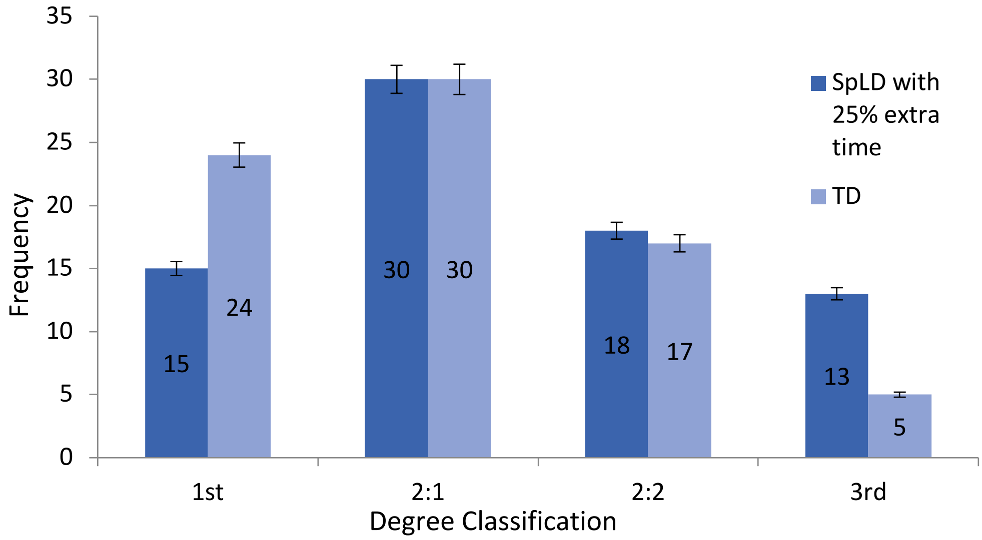


Figure 7.2: Comparison of classifications (with standard error) between SpLD and TD groups (Maths 2016-18)

The data from this research project showed that the SpLD participants who sat the Maths exam with 25% extra time and the TD candidates who sat the same exam under standard conditions were not only achieving, on average, lower marks, but also lower classifications. The data also showed that the marks and the classifications of the TD group were more closely aligned to the marks and classifications of the cohort as a whole than is the case for the SpLD group. The comparison of the exam classifications of the whole Maths cohort in the years 2016-18 and those of the participants in this research project over the same time period is shown in Table 7.3.

Table 7.3: Comparison of exam classifications between whole cohort and research project participants: Maths 2016-18

Group	1st	2:1	2:2	3rd	Fail/unclassified
Whole cohort (n=2084)	31.99%	40.02%	21.77%	6.2%	0.2%
SpLD (n=76)	19.73%	39.47%	23.68%	17.11%	0%
TD (n=76)	31.58%	39.47%	22.37%	6.58%	0%

In conclusion, the analysis of the data shows that the SpLD participants who were granted 25% extra time in Maths exams 2016-18 achieved lower exam marks and lower degree classifications than the TD participants who sat the same exam under standard conditions. This suggests that the exam arrangement of 25% extra time granted to the participants with SpLD in Maths 2016-18 exams did not advantage participants with SpLD in terms of over-inflating their performance as measured by degree classification or exam mark and, to the contrary, failed to fully level the playing field.

7.5 Assessed coursework (Computational) project

Out of this research project's sample group of 152 participants, a total of 102 participants (51 participants with SpLD and 51 TD participants) were assessed by coursework (a computational project). No additional time is granted for the computational project, so, unlike in the timed exams, the participants with SpLD were assessed under the same conditions as TD participants for their coursework. The breakdown of computational project marks between the SpLD and TD groups in Maths 2016-18 is shown in Table 7.4

Table 7.4: Comparison of mean and median computational project marks between SpLD and TD group: Maths 2016 – 18.

Group	Number of candidates	Mean project Mark (percentage)	Median project mark (percentage)	Standard deviation
SpLD participants	51	66	68	13.02
TD participants	51	65	65	13.05

A comparison between the computational project marks for the SpLD and TD groups in Maths 2016-18 is shown graphically in Figure 7.3.

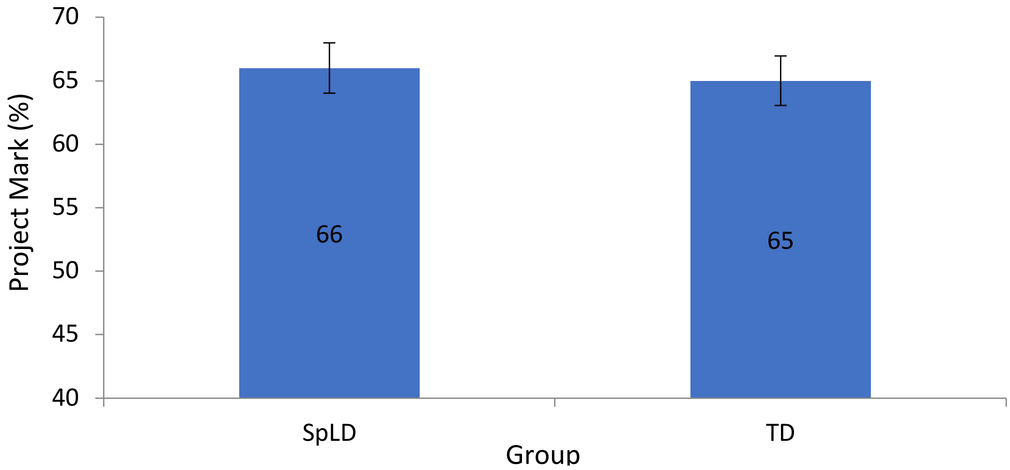


Figure 7.3: Comparison of project marks (%) (with standard error) between SpLD and TD groups: Maths 2016-18

A Kruskal-Wallis *H* test[91] [group, project mark] comparing the project marks achieved by the SpLD participants and those of the TD participants revealed no significant between group differences (χ^2 = .362, p = 548; mean rank mark of 46.14 for SpLD and 42.86 for TD).

[91] Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,86)$ = .553; p = .459). However, the Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were *not* statistically normal (W = .906; p = .002) . The Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were statistically normal (W = .976; p = .477). As the assumptions for an ANOVA were not met, a parametric test was used.

7.6 Comparison of exam mark and coursework mark

The marks of the candidates with SpLD in their computational project are un-affected by any adjustments as extra time was not granted in the assessed coursework and both the SpLD and TD groups were assessed under similar conditions. If it can be reasoned that comparing the project marks of the SpLD group with the project marks of the TD group provides a direct comparison of each group's performance in the subject (as this assessment was conducted under similar conditions for both groups) a between group comparison of any differential between the exam mark and the project mark should also identify if the exam mark of the SpLD group has been over-inflated by the exam adjustments. That is to say, any differential between the exam mark and project mark of the TD group should mirror the differential between the exam mark and project mark of the SpLD group, with the coursework mark acting as a base-line comparator of the subject knowledge and critical skills that the exams are aiming to assess.

In order to explore any between group differences in the differential between exam mark and project mark, a comparison of the exam mark and project mark was conducted for each group[92]. The comparison between the exam mark and project marks for both participant groups are shown in Table 7.5 and Figure 7.4.

Table 7.5: Comparison of exam mark and project mark between SpLD and TD group

Group	Mean Exam mark (%)	Mean Project mark (%)
SpLD (n=51)	61 (SD=9.80)	66 (SD=13.02)
TD (n=51)	66 (SD= 12.39)	65 (SD=13.05)

[92] The exam marks of the participants who were examined in their first year of the Maths course were excluded from this analysis as they did not submit coursework. Only the exam marks of those participants who submitted a computational project only were included.

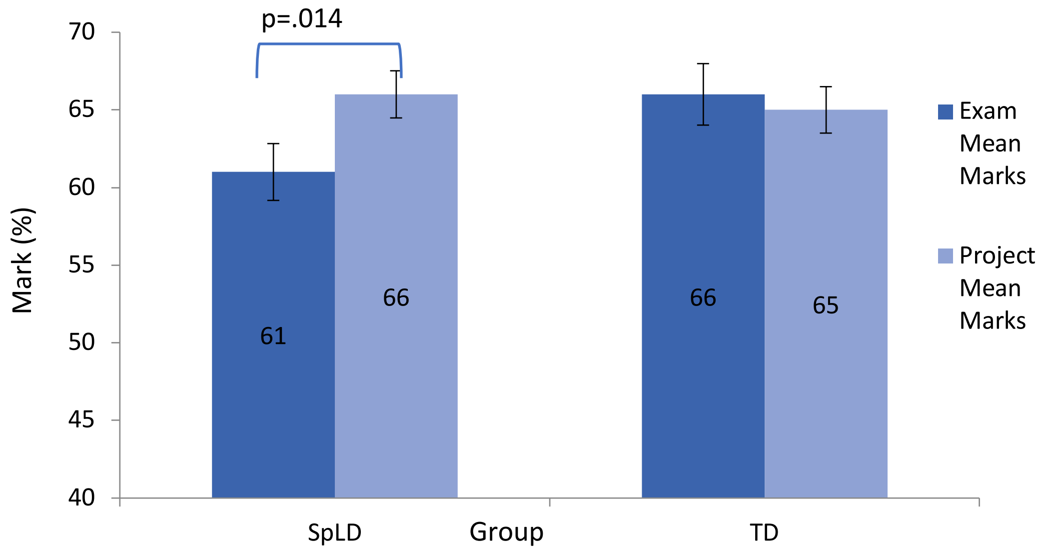


Figure 7.4: Comparison of mean exam mark and mean project mark (with standard error) between SpLD and TD group

A Kruskal-Wallis *H* test [exam mark, project mark] confirmed that the difference between the mean project mark (65%) and the mean exam mark (66%) for the TD group was not significant ($\chi^2 = 36.406$, $p = .231$). However, a Kruskal-Wallis *H* test [exam mark, project mark] confirmed that the difference between the mean project mark (66%) and the mean exam mark (61%) for the SpLD group was significant ($\chi^2 = 58.451$, $p = .014$), with a small effect size ($r = .053$; $r^2 = 0.03$; Hedges' $g = .106$; Cohens' $d = .106$).

In addition to the difference in mean marks between the exam and project for the SpLD group, a difference in classifications was also evident for this group. For example the SpLD group achieved more 1st class classifications in their assessed coursework than they achieved in the exam, and fewer 2:2 and 3rd classifications. A comparison of the classifications achieved in the exam and those achieved in the assessed coursework project by the SpLD group is shown in Table 7.6.

Table 7.6: Comparison of exam and project classifications: SpLD group Maths 2016-18.

Assessment	1st	2:1	2:2	3rd	Fail/unclassified
Exam	19.73%	39.47%	23.68%	17.11%	0%
Project	45.10%	39.22%	5.88%	7.84%	1.96%

A Kruskal-Wallis *H* test [exam classification, project classification] confirmed that the difference between the coursework project classifications and the exam classifications for the SpLD group was significant ($\chi^2 = 10.512$, $p = .033$). This combination of a lower mean mark in exams in comparison to the mean coursework marks for participants with SpLD coupled with a statistically lower classification achieved in the exam in comparison to the coursework project not only suggests that the extra time in the exams does not appear to over-inflate the scores of the SpLD candidates but instead indicates that the SpLD group are underachieving in the timed exam.

By comparison, less difference was observed between the project marks and the exam marks for the TD group. This is shown in Table 7.7.

Table 7.7: Comparison of exam and project classifications: TD group Maths 2016-18

Assessment	1st	2:1	2:2	3rd	Fail/unclassified
Exam	31.58%	39.47%	22.37%	6.58%	0%
Project	37.25%	29.41%	25.49%	5.88%	1.96%

A Kruskal-Wallis *H* test [exam classification, project classification] confirmed that the difference between the project classifications and the exam classifications for the TD group was not significant ($\chi^2 = 36.406$, $p = .231$). This suggests that, in contrast to the SpLD group, the TD group perform similarly in their coursework as they do in their exams.

Having examined the differences between the exam mark and the coursework project mark for each group, the data was then compared to ascertain if there were any between-group differences in the differential identified between each group's exam mark and their project mark. A multi-factorial ANOVA exploring between-group differences for the effect of exam mark on project mark was conducted. This revealed a significant between-group difference in the relationship between coursework marks and exam marks ($F(1,300) = 42.32$, $p = .002$; Wilks' $\Lambda = .25$, partial $\eta^2 = .15$). The differences between the mean exam mark and the mean coursework project mark for the two participant groups is shown in Figure 7.5.

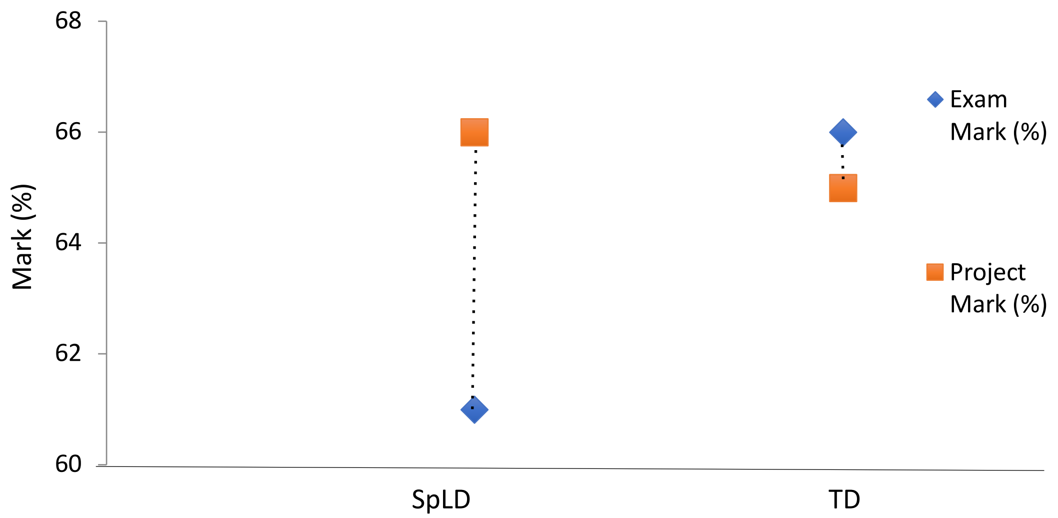


Figure 7.5: Between group comparison between exam mark and project mark for TD & SpLD groups: Maths 2016-18

In addition to the between-group difference in the relationship between coursework marks and exam marks, between-group differences were also identified in the relationship between exam classification and coursework classification. In total, 23 out of the 51 SpLD participants (45.10%) who submitted coursework achieved a First-class classification for their coursework project, yet only 15 out of the 76 SpLD participants (20%) who sat the exams achieved a First-class classification in their timed exam. By comparison, 19 out of the 51 TD participants (37.25%) who submitted coursework achieved a First-class classification in their coursework project and 24 of the 76 (31.58%) TD participants who sat the exam achieved a First-class classification in their exam. A multi-factorial ANOVA comparing the differential between exam classification and project classification between the two groups showed that there was a significant difference, with the SpLD group showing a greater differential between exam classification and project classification than the TD group ($F(1,300) = 50.94$, $p < .001$; Wilks' $\Lambda = .38$, partial $\eta^2 = .16$). If the project work is assumed to be a more accurate representation of an SpLD candidate's subject knowledge and ability than a timed exam (Keogh et al, 2002), this further suggests that the SpLD group are underachieving in the exam.

7.7 Conclusion

Overall, significant differences between the SpLD groups and the TD group were identified in exam mark and exam classification in the Maths 2016-18 exams, with the SpLD group achieving lower marks and lower degree classifications than their TD peers who sat the same exam under standard time conditions. However, no differences between the SpLD group and the TD group were identified in the coursework project mark (where no exam arrangements were granted to the SpLD group). At the same time, significant between-group differences were identified in the relationship between the exam mark and assessed coursework mark of the two groups, with the SpLD group showing a larger gap between their exam marks and their coursework marks than was the case for their TD comparators. Furthermore, the SpLD group demonstrated statistical differences between their exam classifications and their coursework classifications (with exam classification being significantly lower than the coursework classification). This differential suggests that rather than over-inflating the exam scores of the SpLD group, granting 25% extra fails to fully level the playing field with the result that the SpLD group may be underperforming in the timed, closed book, written exam (as measured by exam mark and classification) in comparison to their potential (as measured by coursework project mark and classification), despite the additional time.

8: QUANTITATIVE RESEARCH FINDINGS: MEDICINE EXAMS

This is the third of the three findings chapters that combine to form the first quantitative phase of this research project that aims to address the research question of whether or not exam arrangements granted to SpLD candidates confer an advantage or promote equity. This chapter presents the analysis and outcomes of three years of data relating to the exam marks achieved in Medicine exams by candidates with SpLD and compares this with similar exam data from their TD peers who sat the same exams under standard conditions.

The chapter starts by providing data about the Medical School cohort as a whole, to provide an overall context, and it then goes on to present the characteristics and make-up of the sample group from the School of Medicine who participated in this research project. The chapter then compares the exam marks and degree classifications of the participants with SpLD with those of the TD participants to identify if the participants with SpLD achieved higher marks than their TD peers in exams and so were potentially advantaged by the exam arrangements they received. Finally, this chapter attempts to identify whether or not the participants with SpLD were under or over-performing in exams by comparing their practical exam results (where they did not receive any adjustments) with their timed, written, closed book, exam results (where they received extra time) and cross referencing the outcomes with similar data from the TD participants to explore any between group differences

8.1 Introduction

The medicine programme at the HEI where this research was conducted comprises a 6 year course of which the first 3 years represent pre-clinical (undergraduate) studies, covering basic medical science. Assessment in the first 3 years is by formal, timed, written exams and practical exams at the end of each year. For the purposes of this research project, data has been drawn from the exams for years 1 & 2 as the format, content and marking rubric of the exams are similar in both years. The format, content and marking rubric of the third year exams differ considerably and so have not been included.

The total number of students who sat the Medical science exams in 2016-18 was 1,960 of which 76 (3.88%) disclosed an SpLD and were granted 25% additional time in the exam. In the 2016-18 summer exams, out of the total cohort, 18.21% of candidates (357) achieved a first-class classification (mark of 70% or above); 49.03% of candidates (961) achieved a 2:1 classification (a mark of 60 – 69%); 31.48% of candidates (617) achieved a 2:2 classification (a mark of 50 – 59%); 1.07% of all candidates (21) achieved a 3rd class classification (a mark of 40 – 50%) and 0.2% of all candidates (4) failed the exam (a mark below 40%). The breakdown of the overall population of all 1,960 candidates can be seen in Table 8.1.

Table 8.1: Breakdown of candidates in Medical Science undergraduate 2016-18.

Year	Number of TD candidates	Number of SpLD candidates	Total candidates	Mean Mark (whole cohort)	Range of marks (whole cohort)
2016	637 (96.52%)	23 (3.48%)	660	62	36 – 97
2017	622 (95.99%)	26 (4.01%)	648	63	38 – 98
2018	625 (95.86%)	27 (4.14%)	652	62	34 - 96
Total	1,884 (96.12%)	76 (3.88%)	1,960	62	34 – 98

8.2 Sample Group

The total sample group drawn from the Department of Medicine for this research project comprised 136 participants (selected from a total population of 1,960) of which 68 participants disclosed an SpLD and received exam arrangements and 68 participants were randomly selected TD peers who sat the same exams under standard conditions. Due to the nature of the exam, which consisted of writing answers that included chemical formulas, graphs and diagrams rather than exclusively continuous prose, no candidates used a word processor.

Data from a total of 408 written papers, drawn from Functional Architecture of the Body (FAB)[93], was included. This represented 3 papers sat by each of the 68 participants (68 SpLD and 68 TD participants). In addition, 272 practical papers from Functional Architecture of the Body (FAB) were included (2 papers for each candidate). Candidates with SpLD were not granted extra time in the practical exam (as this exam involved specimen identification and oral work rather than reading and/or writing) and so the outcomes of the practical exams (which all participants took under standard conditions) will be compared with the timed, written exams, for which the SpLD group received 25% extra time.

The assessment for FAB consists of 3 written exams and 2 practical exams[94]. Candidates with SpLD were granted 25% extra time in all 3 written exams but sat the practical exam under standard time conditions. The 3 written exams comprised a 1 hour multiple choice exam paper, a 2 hour short answer exam paper and a 2 hour long answer exam paper (standard time). The practical exams comprised a 90-minute ‘steeplechase’ paper with 30 stations. Mean marks for each of the three papers, as well as for the exam as a whole, have been provided in order to explore any within-participant differences in marks between the exam papers. The mark for the practical exam has also been provided to explore any within or between-group differences between the practical exam mark and the mark achieved in the written exam. ANOVAs were used to test any significance between the mean marks[95].

[93] Functional Architecture of the Body is a compulsory topic
[94] See Appendix 1 for examples of the exam questions.
[95] Equivalent non-parametric tests were used where the data did not meet the assumptions for an ANOVA

8.3 Comparison of exam results

Given that the Medicine exams require responses in the form of chemical formulae, graphs and diagrams as well as written explanation, word count is not a relevant measure of the effect of extra time on exam scripts. Therefore, no analysis of word count has been undertaken on Medicine exam papers. Instead, a comparison of performance as determined by exam mark and degree classification awarded has been undertaken in order to identify whether the granting of extra time advantaged the SpLD group over their TD peers in the exams.

8.3.1 Comparison of marks from FAB written exams 2016-18

The mean marks for the SpLD and TD groups were compared in each of the three FAB papers, namely the multiple-choice paper (MCQ), the short written answer paper and the long written answer (essay style) paper. The results are shown in Tables 8.2 & 8.3.

Table 8.2: Mean percentage exam marks achieved by SpLD participants (with standard deviation): FAB written exams in Medical Science 2016 - 18.

Participant group/year	% score Paper 1 (MCQ)		% score Paper 2 (Short answer)		% score Paper 3 (Long answer)		% score Total Exam Paper	
	M	SD	M	SD	M	SD	M	SD
SpLD 2016 (n = 20)	68	10.57	63	6.54	59	4.88	61	4.86
SpLD 2017 (n=23)	69	10.49	63	8.37	55	4.85	60	5.06
SpLD 2018 (n=25)	68	8.78	63	6.43	56	5.45	60	4.21
SpLD 2016-18 (n=68)	68	9.87	63	6.50	57	5.21	60	4.75

Table 8.3: Mean percentage exam marks achieved by TD participants (with standard deviation): FAB 2016-18 written exams in Medical Science.

Participant group/year	% score Paper 1 (MCQ)		% score Paper 2 (Short answer)		% score Paper 3 (Long answer)		% score Total Exam Paper	
	M	SD	M	SD	M	SD	M	SD
TD 2016 (n=20)	70	11.23	61	7.16	60	6.83	63	7.56
TD 2017 (n=23)	65	11.01	63	8.40	56	7.25	61	7.31
TD 2018 (n=25)	68	11.12	61	8.01	58	7.15	60	6.45
TD 2016-18 (n=68)	67	11.20	61	8.17	58	7.38	61	6.80

A Kruskal-Wallis *H* test [mark, group] confirmed that the difference in mean scores achieved on each paper by the TD group compared to the SpLD group in the written Medicine (FAB) exams was non-significant. This can be seen in Table 8.4.

Table 8.4: Comparison of difference between % score for each paper between SpLD participants with 25% extra time and TD participants: Medicine 2016 - 18.

Paper	χ^2	p	Mean rank mark for SpLD	Mean rank mark for TD
Fab paper 1[96] (mcq)	1.320	.251	72.38	64.63

[96] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal ($W = .957$; $p = .019$) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal ($W = .941$; $p = .003$).

Fab paper 2[97] (short written answers)	1.523	.217	72.65	64.35
Fab paper 3[98] (long written answers)	.216	.642	66.93	70.07
Fab exam as a whole[99]	.010	.920	68.16	68.84

Although the SpLD group achieved slightly lower mean results in the FAB exam as a whole (60%) than the TD group (61%), a Kruskal-Wallis *H* test [% mark, group] confirmed that this difference was not significant (χ^2 = .010, *p* = .920; mean rank mark of 68.16 for SpLD and 68.84 for TD). There were some variations in the marks achieved by each group at the individual paper level, but these differences were also not significant. For example, the SpLD group achieved a mean score of 68% in the MCQ paper, in comparison to the TD group who achieved a mean mark of 67%, but a Kruskal-Wallis *H* test [% mark, group] confirmed that this was not significant (χ^2 = 1.320, *p* = .251; mean rank mark of 72.38 for SpLD and 64.63 for TD). In the FAB short written answer paper, the SpLD group achieved a higher mean score (63%) than the TD group (61%), but again a Kruskal-Wallis *H* test [% mark, group] confirmed that this difference was not significant (χ^2 = 1.523, *p* = .217; mean rank mark of 72.65 for SpLD and 64.35 for TD). In the FAB long answer paper the SpLD group achieved a lower mean mark (57%) than the TD group (58%), and a Kruskal-Wallis *H* test [% mark, group] again confirmed that this difference was not significant (χ^2 = .216, *p* = .642; mean rank mark of 66.93 for SpLD and 70.07 for TD). This outcome can be seen in Figure 8.1.

[97] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal (*W* = .906; *p*< .001) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal (*W* = .940; *p* = .003).

[98] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal (*W* = .952; *p* = .010) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal (*W* = .915; *p*< .001).

[99] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were statistically normal (*W* = .973; *p* = .147) but the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal (*W* = .933; *p* = .001).

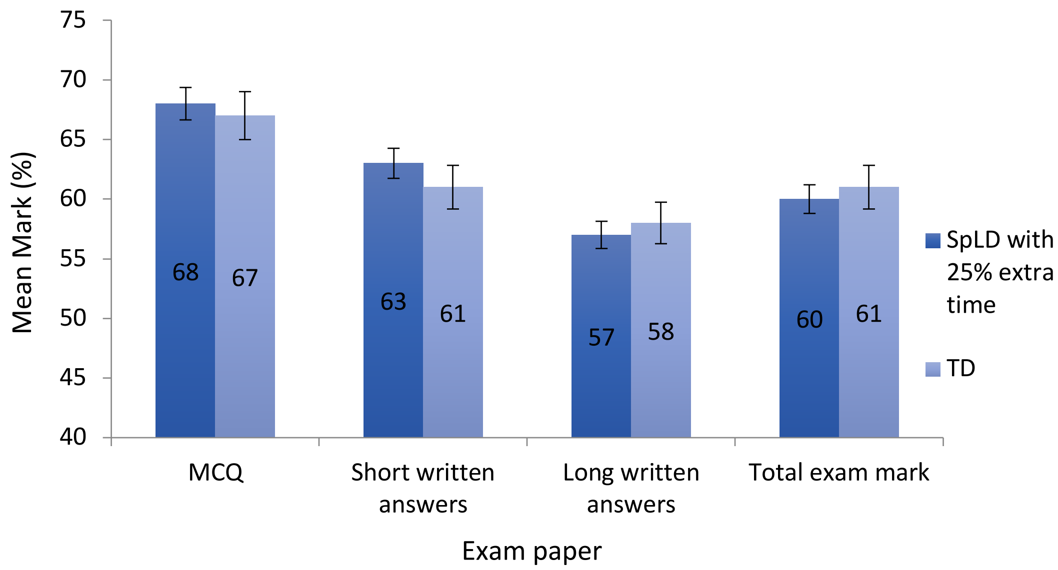


Figure 8 1: Mean marks (with Standard Error) for SpLD and TD participants: Medicine 2016-18

Overall, these outcomes suggest that there were no significant differences between the mean marks that the SpLD group achieved in the 2016-18 Medicine (FAB) exams with extra time in comparison with the TD group who sat the same exams under standard time conditions. This suggests that the extra time did not confer any advantage.

8.3.2 Comparison of Classifications

When the marks are categorised into classifications the results show that the participants with SpLD achieved fewer first-class classifications than the TD participants, a greater number of 2:1 and 2:2 classifications and a similar number of third-class classifications. A comparison of the classifications of the two participant groups can be seen in Figure 8.2.

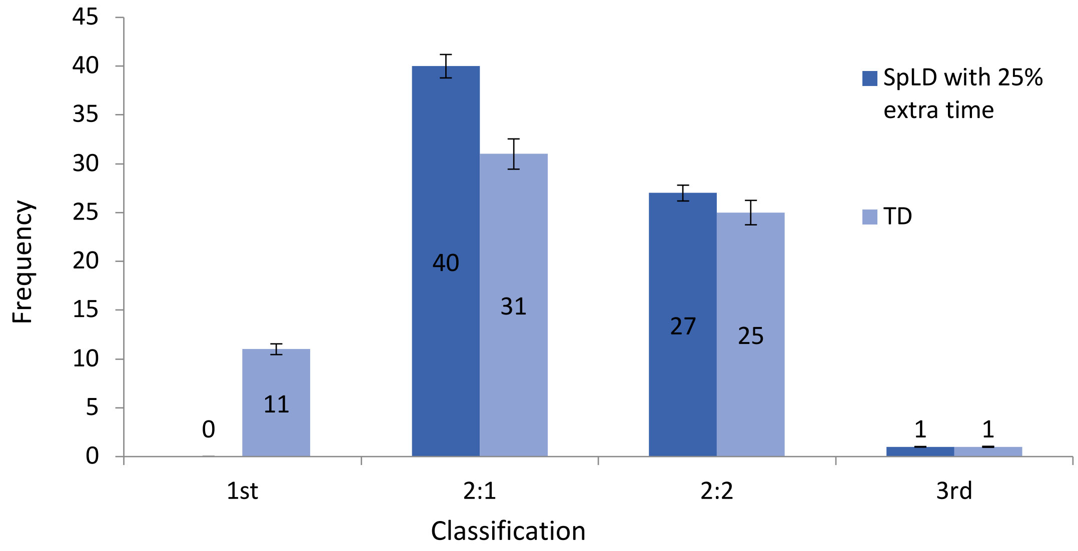


Figure 8.2: Comparison of classifications between SpLD and TD groups (with Standard Error): Medicine (FAB) 2016-18.

When comparing the exam classification of the SpLD and TD group to the classifications of the cohort as a whole, it was evident that the classifications of the TD group were more closely aligned to the classifications of the cohort as a whole compared to those of the SpLD group. That is to say, none of the SpLD group achieved marks above 70% (1^{st} class classification), 58.82% of the SpLD group (N=40) achieved marks in the 60 – 70% range (2:1 classification), 39.71% of the SpLD group (N=27) achieved marks between 50 – 60% (2:2 classification), and 1.47% (N=1) scored below 50% (a 3^{rd} class classification). In contrast, 16.18% of the TD group (N=11) achieved marks above 70% (1^{st} class classification), 45.58% of the TD group (N=31) achieved marks in the 60 – 70% range (2:1 classification), 36.76% of the TD group (N=25) achieved marks between 50 – 60% (2:2 classification), and 1.47% (N=1) scored below 50% (a 3^{rd} class classification). This comparison can be seen in Table 8.5.

Table 8.5: Comparison of exam classifications between whole cohort and research project participants: Medicine (FAB) 2016-18.

Group	1st Class	2:1	2:2	3rd	Fail/unclassified
Whole cohort (n=1960)	17.21%	46.03%	35.48%	1.07%	.2%
SpLD (n=68)	0%	58.82%	39.71%	1.47%	0%
TD (n=68)	16.18%	45.58%	36.76%	1.47%	0%

In conclusion, a Kruskal-Wallis *H* test [% mark, group] comparing the percentage marks achieved by the SpLD participants who were granted 25% extra time in the written FAB exams in Medicine 2016-18 and those of the TD participants who sat the same exam under standard conditions revealed no significant between group differences ($\chi^2 = .010$, $p = .920$; mean rank mark of 68.16 for SpLD and 68.84 for TD). However, some differences were evident in classifications awarded, with the SpLD groups achieving fewer 1st degree classifications, but more 2:1 and 2:2 classifications than the TD group. Both groups achieved a similar number of 3rd degree classifications. A Kruskal-Wallis *H* test[100] [written exam classification, group] confirmed that the difference in written exam classifications between the two groups was significant ($\chi^2 = 3.88$, $p = .048$; mean rank mark of 61.28 for SpLD and 74.72 for TD). Thus, there was no indication that the extra time advantages participants with SpLD by over-inflating their performance when this is measured by exam mark or degree classification and an awarding gap in degree classification persists, despite the exam adjustments.

8.4 Practical exams

The assessment for FAB also includes two practical exams. Both practical exams

[100] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal ($W = .750$; $p < .001$) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal ($W = .754$; $p < .001$).

comprise a 90-minute 'steeplechase' paper consisting of 90 questions in total. As this is not a written paper, candidates with SpLD are not granted additional time and so candidates with SpLD and TD candidates take these exams under the same time conditions. Comparing the marks of the SpLD students in their written exams with the marks they achieved in their practical exams may indicate whether or not a candidate with SpLD is either over or under performing in the written exam in comparison to his/her potential as suggested by practical exam marks. In order to establish if any differences between written exam marks and practical exam marks are specific to students with SpLD, the written exam marks and practical exam marks of TD students have also been compared and the outcomes for these two groups are cross referenced. The breakdown of the practical exam marks between the SpLD and TD groups in each year is shown in Table 8.6.

Table 8.6: Comparison of practical exam marks between SpLD and TD group in Medicine (FAB) 2016 - 18.

Group	Mean practical exam mark (%)	Standard deviation
SpLD participants (n=68)	70	11.57
TD participants (n=68)	71	12.63

A Kruskal-Wallis *H* test[101] [group, practical exam mark] comparing the practical exam marks achieved by the SpLD participants and those of the TD participants in the practical FAB exams revealed no significant between group differences ($\chi^2 = .005$, $p = .946$) with a mean rank mark of 20.38 for SpLD and 20.63 for TD candidates. A comparison between the practical FAB exam marks for the SpLD and TD groups in Medicine 2016-18 is shown graphically in Figure 8.3.

[101] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal ($W = .964$; $p = .044$), albeit the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were statistically normal ($W = .971$; $p = .119$).

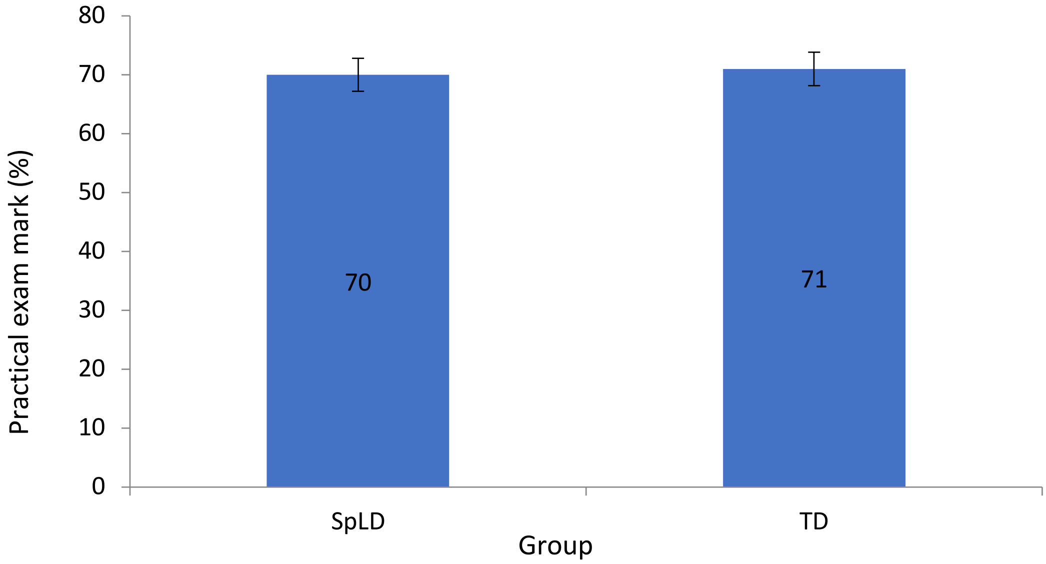


Figure 8.3: Comparison of practical exam marks (%) (with Standard Error) between SpLD and TD groups: Medicine (FAB) 2016-18.

8.4.1 Comparison of Practical Exam Classifications

When the marks from the practical exams are categorised into classifications the results show that the participants with SpLD achieve slightly fewer 1st and 2:1 classifications than the TD participants, more 2:2 classifications but fewer 3rd degree classifications than the TD participants. However, a Kruskal-Wallis *H* test[102] [practical exam classification, group] confirmed that the difference in practical exam classifications between the two groups was not significant (χ^2 = .175, p = .675; mean rank of 67.21 for SpLD and 69.79 for TD). A comparison of the classifications of the two participant groups can be seen in Figure 8.4.

[102] A non-parametric test has been used as the data did not meet the assumptions for an ANOVA. The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were not statistically normal (W = .750; p < .001) and the Shapiro-Wilk test of normality for the mark of the TD group indicated that the data were not statistically normal (W = .754; p < .001).

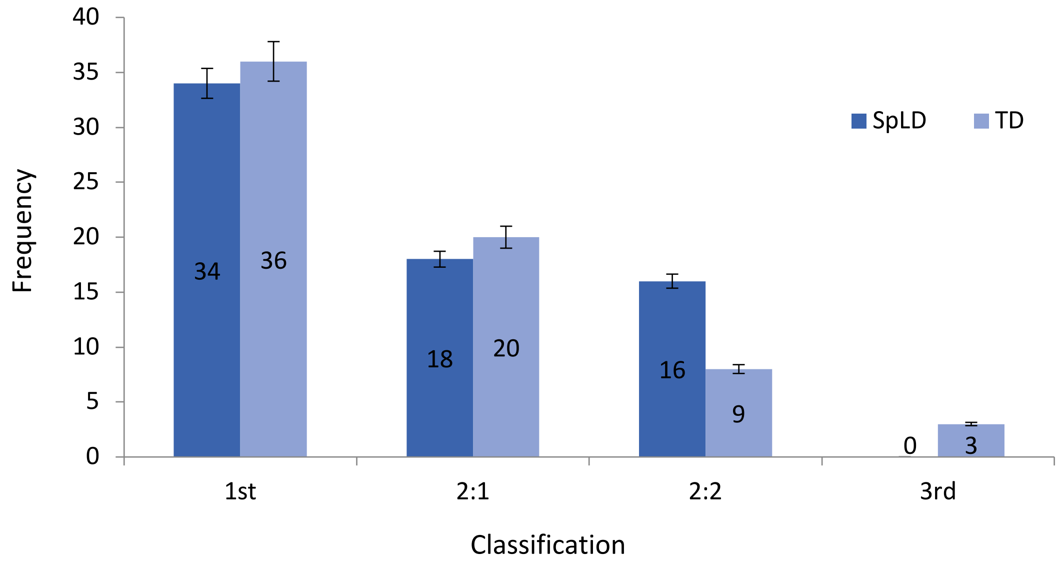


Figure 8.4: Comparison of classifications between SpLD and TD groups (with Standard Error): Medicine (FAB) 2016-18.

8.5 Comparison of written exam mark and practical exam mark

If it can be reasoned that comparing the practical exam marks of the SpLD group with the practical exam marks of the TD group provides a direct comparison of each group's performance in the subject (as this assessment was conducted under the same conditions for both groups) a between group comparison of any differential between the written exam mark and the practical exam mark should also identify if the written exam mark of the SpLD group has been over-inflated through the awarding of additional time. That is to say, any differential between the written exam mark and practical exam mark of the TD group should mirror the differential between the exam mark and practical exam mark of the SpLD group. Thus, a comparison of the exam mark and the practical exam mark was conducted for each group and the outcomes are shown in Table 8.7.

Table 8.7: Comparison of written exam mark and practical exam mark between SpLD and TD group: Medicine (FAB) 2016-18

Group	Mean MCQ exam mark (%)	Mean short answer exam mark (%)	Mean long answer exam mark (%)	Total mean written exam mark (%)	Mean practical exam mark (%)
SpLD (n=68)	68 (SD=9.87)	63 (SD=6.50)	57 (SD=5.21)	60 (SD=4.75)	70 (SD=11.57)
TD (n=68)	67 (SD=11.20)	61 (SD=8.17)	58 (SD=7.38)	61 (SD=6.80)	71 (SD=12.63)

A Kruskal-Wallis *H* test [total written exam mark, practical exam mark] confirmed that the difference between the mean written exam mark (60%) and the mean practical exam mark (70%) for the SpLD group was significant (χ^2 = 31.035, p = .020; with a mean rank of 36.83 for the written exam and 56.25 for the practical exam). This represented a large effect size (r = -0.49, r^2 = 0.24; Cohen's d= -1.130). In addition, a Kruskal-Wallis *H* test [written exam mark, practical exam mark] confirmed that the difference between the mean written exam mark (61%) and the mean practical exam mark (71%) for the TD group was also significant (χ^2 = 39.873, p = .001; with a mean rank of 30 for the written exam and 58.50 for the practical exam). This was also a large effect size (r = -0.44, r^2 = 0.20; Cohen's d= -0.99). The comparison of the written exam mark and the practical exam mark between the two participant groups can be seen in Figure 8.5.

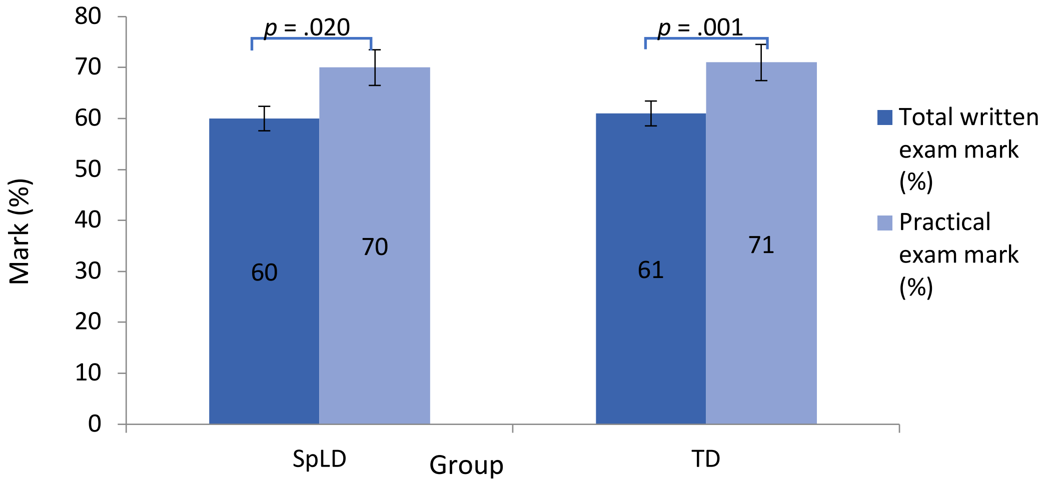


Figure 8.5: Comparison of written exam mark (%) and practical exam mark (%) between SpLD and TD group (with Standard Error): Medicine (FAB) 2016-18.

A Pearson product-moment correlation coefficient was conducted for the SpLD participants to assess the relationship between the marks achieved on the written exam and the marks achieved in the practical exam and a moderate, significant positive relationship was found (r_p = .536, N=68, p = <.001), suggesting that those participants who achieved higher scores in the written exam, relative to the cohort, similarly achieve higher scores in the practical exam (relative to the cohort), while those participants who achieved lower scores in the written exam, relative to the cohort, similarly achieved lower scores in the practical exam (relative to the cohort). This is shown in Figure 8.6.

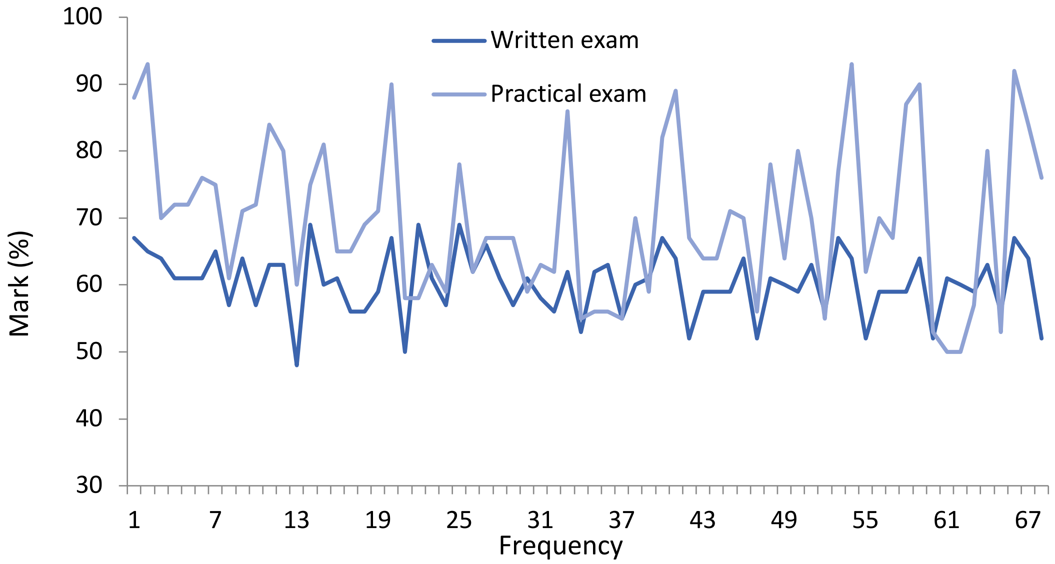


Figure 8.6: Comparison of written exam mark (%) and practical exam mark (%) for SpLD group: Medicine (FAB) 2016-18

A Pearson product-moment correlation coefficient was also conducted for the TD participants to assess the relationship between the marks achieved on the written exam and the marks achieved in the practical exam for the TD group. Reflecting the outcome of the SpLD group data, a moderate, significant positive relationship was found (r_p = .434, N=68, p = <.001), suggesting that those participants who achieved higher scores in the written exam, relative to the cohort, similarly achieve higher scores in the practical exam (relative to the cohort), while those participants who achieved lower scores in the written exam, relative to the cohort, similarly achieved lower scores in the practical exam (relative to the cohort). This outcome is shown in Figure 8.7.

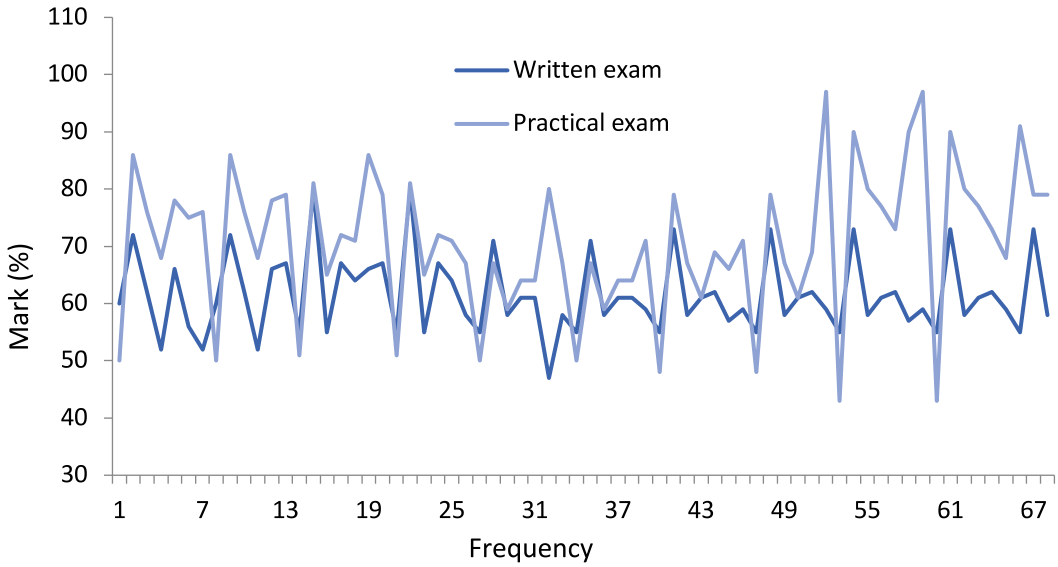


Figure 8.7: Comparison of written exam mark (%) and practical exam mark (%) for TD group: Medicine (FAB) 2016-18.

Having examined differences between the written exam mark and the practical exam mark for each group, the data was then compared to ascertain if there were any between group differences in the differential identified between each group's written exam mark and their practical exam mark. A multifactorial ANOVA[103] exploring between group differences in the relationship between the practical exam marks and written exam marks revealed a non-significant between group difference ($F(1,135) = .152$, $p = .698$; Wilks' $\Lambda = .023$, partial $\eta^2 = .023$). The differences between the mean written exam mark and the mean practical exam mark for the two participant groups is shown in Figure 8.8.

[103] To conduct parametric tests that require normally distributed data (such as an ANOVA), the written exam total score for the TD group was transformed. Following a reciprocal transformation the Shapiro-Wilk test of normality indicated that the data are statistically normal (W=.935; p=.510). In addition, the practical exam score for the SpLD group was transformed. Following a reciprocal transformation the Shapiro-Wilk test of normality indicated that the data are statistically normal (W=.968; p=.073). The Shapiro-Wilk test of normality indicated that the written exam data of the SpLD are statistically normal (W=.973; p=.147) and the practical exam data of the TD group are statistically normal (W=.971; p=.119). Box's Test of Equality of Covariance Matrices confirmed that variances are equal across the groups (Box's M=7.877; F =2.583; p=.051).

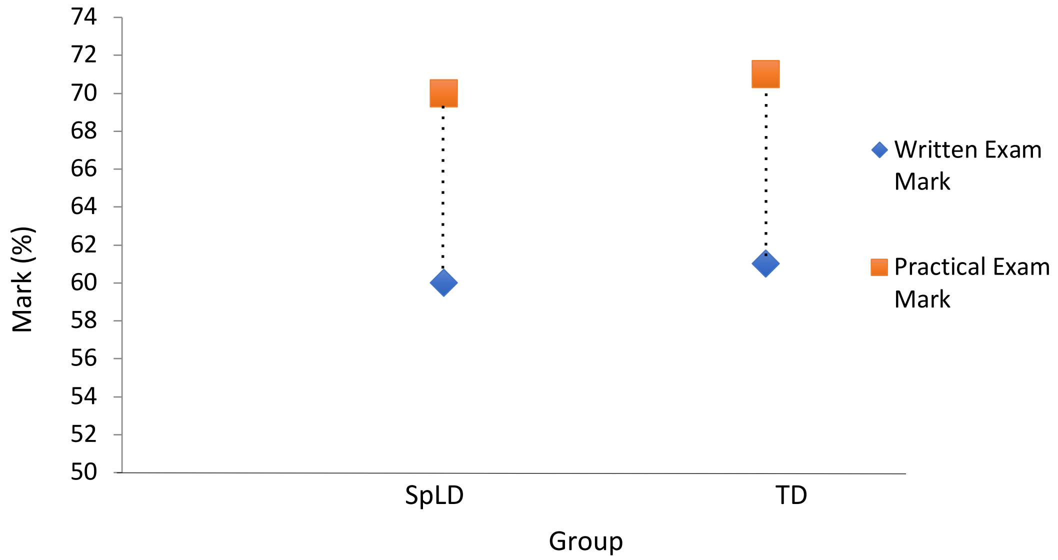


Figure 8.8: Between group comparison between written exam mark and practical exam mark for TD & SpLD groups: Medicine (FAB) 2016-18

Although no between-group differences in the relationship between the written exam marks and the practical marks were identified, some between group differences existed in the relationship between written exam classification and practical exam classification. For example, 34 out of the 68 SpLD participants (50%) achieved a first-class classification in their practical exam, yet none achieved a first-class classification in their written exam. In comparison, 36 out of the 68 TD participants (52.94%) achieved a first-class classification in their practical exams and 11 of the TD participants (16.18%) achieved a first-class classification in their exam. A multifactorial ANOVA comparing the differential between exam classification and practical exam classification between the two groups showed that there was a significant difference, with the SpLD group showing a greater differential between the written exam classification and the practical exam classification than the TD group ($F(1,135) = 6.85$, $p = .015$; Wilks' $\Lambda = .28$, partial $\eta^2 = .11$). This outcome suggests that the adjustment of 25% extra time does not fully compensate for the impact of construct irrelevant skills on candidates with SpLD in written Medicine exams when results are categorised into classifications.

8.6 Conclusion

Overall, no significant differences between the SpLD groups and the TD group were identified in the mean written exam mark in the Medicine 2016-18 exams, albeit differences were identified in the classifications awarded to each group (with the SpLD group achieving lower classifications than their TD peers). At the same time, no differences between the SpLD group and the TD group were identified in the mean practical exam mark or in the relationship between the mean written exam percentage mark and the mean practical exam percentage mark between the two groups. However, there were between-group differences in the relationship between the written exam *classification* and the practical exam *classification*, with the SpLD group achieving lower written exam classifications in comparison to the practical exam classification than the TD group. Thus, the data from the Medicine exams in 2016-18 suggest that granting 25% extra time in exams does not advantage the SpLD group and an awarding gap is still evidence when this is measured by degree classification.

9: QUANTITATIVE RESEARCH FINDINGS: STUDENTS DIAGNOSED DURING THE COURSE

This chapter presents the findings from the second quantitative phase of this research project. It focuses on participants with SpLD who were diagnosed part-way through their course and so sat some exams without any exam adjustments and later exams with exam arrangements. The purpose of this chapter is to further explore Research Question 1: ‘What impact do exam access arrangements have on the exam performance of students with SpLD’?

This chapter commences by presenting the methodology adopted in this second quantitative phase of the research project, including an explanation of the research design, data collection, and data analysis procedures used, as well as outlining the composition of the sample group. Confounding factors, such as the impact of study skills sessions, as well as other limitations are also addressed. The chapter then compares the marks that the participants with SpLD achieved before they received exam arrangements (and so sat their exams under standard conditions) with the marks they achieved in later exams where they received exam adjustments. These exam marks, covering the full three years of the participants' degree course, are also compared to the marks of matched TD comparators over the same period to establish if any changes that occur following the implementation of exam arrangements are significantly different to the 'normal' increase in exam results that would be expected as the degree course progresses.

This chapter then goes on to compare the exam results of participants with SpLD who received 25% extra time only with those who received 25% extra time and the use of a word processor, in order to identify whether the implementation of a word processor in conjunction with extra time in exams contributes to a greater increase in marks than is the case for extra time alone. The chapter also compares the data from participants with a diagnosis of dyslexia with the data from participants with a diagnosis of dyspraxia (and also those with a dual diagnosis) to identify if any differences in outcomes exist between these two categories of SpLD in exam situations.

Following this, the chapter considers the impact that engaging with study skills sessions may have had on the exam results of the participants with SpLD, on the premise that accessing study skills sessions aim to improve learning outcomes (Saeidi et al, 2019). This part of the chapter details the methodology employed to evaluate the potential effect of study skills sessions on exam performance and presents the findings of an analysis of the relationship between the number of study skills sessions accessed and the exam mark achieved. It also presents a comparison of the exam marks achieved by participants who used study skills sessions and those who did not, in order to explore the potential efficacy of study skills sessions in terms of improved exam performance.

9.1 Introduction

The analysis of the data from the first quantitative phase of the project addressed the question of the impact of exam arrangements on students with SpLD by comparing the exam results of participants with SpLD who sat the exam with adjustments with those of their TD peers who sat the same exam under standard conditions to identify any statistical differences in marks or length of scripts between these two groups (on the basis that any statistical differences arise from the different exam conditions that the two groups experienced). This aimed to address Research Question 1: 'What impact do exam access arrangements have on the exam performance of students with SpLD?' by interrogating whether or not students with SpLD who have access arrangements produce longer exam scripts or achieve higher marks than their TD peers.

The data from this first phase of the research project not only showed that higher word count failed to correlate with high marks, but also found that the use of a word processor and/or 25% extra time did not advantage the SpLD group in terms of increased word count, higher marks or higher classifications than their TD peers in timed, written, exams. In fact, the data suggests that the exam arrangements granted to the participants with SpLD failed to fully level the playing field as the SpLD participants produced fewer words on the exam scripts and achieved significantly lower exam marks and degree classifications than their TD peers who sat the same exams under standard conditions. At the same time, a comparison of the assessed coursework marks between the two groups showed that the participants with SpLD achieved similar marks in assessed coursework to their TD peers; it was only in the timed, closed book exams that the differences existed. This, therefore, suggests that the differences in exam performance is not a question of differences in ability between the two groups but a question of the barriers that exist in the exam that are disadvantaging the candidates with SpLD; barriers that are not being fully overcome by the exam arrangements granted. The question then arises of whether or not the exam adjustments actually made any difference to the participants with SpLD in their exams. That is to say, if the awarding gap between SpLD and TD students still exists despite the use of exam adjustment, do the access arrangements help at all? Would the candidates with SpLD have achieved the same exam marks whether or not they had not received any exam adjustments?

In order to explore this question, a second phase of the project was conducted which focused on identifying any changes in the exam results of undergraduate students with SpLD who are granted exam adjustments part-way through their university course. This aimed to further explore Research Question 1: 'What impact do exam access arrangements have on the exam performance of students with SpLD'? by quantifying the effect that implementing exam adjustments had on the exam performance of students with SpLD.

Specifically, the students' exam results obtained before they received a diagnosis of SpLD (and so sat their exams under standard conditions) were compared to the same students' exam results after they received the diagnosis of SpLD and were granted exam adjustments. This comparison aimed to establish whether or not the participants with SpLD achieved higher exam marks after they received exam adjustments than they achieved when they sat their exams under standard conditions, thereby exploring the question of whether or not the exam adjustments had any positive effect on the exam results of the participants with SpLD. As detailed in the literature review chapter, a broad range of research has been conducted into the impact that extra time has on the exam performance of individuals with SpLD (albeit with contradictory outcomes). However, no research has been identified that compares the exam performance of students before they were diagnosed with SpLD with the exam performance of the same individuals after they were diagnosed with SpLD and received exam arrangements. This second phase of the project aims to focus on this gap in current research.

It should be pointed out that it is anticipated that the marks that all students achieve in exams are likely to increase year on year during their undergraduate degree, due to the cumulative impact of teaching and learning resulting in mounting improvements in learning outcomes as a course progresses (Paolini, 2015). To minimize the possibility that any identified change in exam performance of the participants with SpLD after they receive exam adjustments were simply the result of this 'expected maturation' effect (or, for that matter, caused by changes in exam difficulty in different years of study) the first year grades of the participants with SpLD were compared to the first year grades of the same number of TD participants; the second year grades of the participants with SpLD were compared with the second year grades of the same number of TD participants, and so on. Participants were matched on the basis of gender, age, course of study and exam time point. Any changes in the exam marks of the SpLD participants were compared with changes in exam marks of the matched TD peers over an equivalent period of time. This comparison aimed to minimise confounding variables by identifying whether the exam adjustments granted affected the exam performance of the participants with SpLD to a greater extent than would be predicted by the 'normal' changes demonstrated by the index group.

This second phase of the project also explored the impact of study skills sessions on academic performance by identifying whether or not the participants with SpLD who received additional 1:1 specialist study skills support following their diagnosis improved their exam performance to a greater extent than those who did not access the sessions and also whether there was any correlation between the number of study skills accessed and exam outcomes.

9.2 Methodology

9.2.1 Research Design: Procedure

Data was collected from undergraduate students who were diagnosed with SpLD part-way through their three-year university course[104]. The examination marks of the participants before they were diagnosed with SpLD were compared with the performance of the same participants after they were diagnosed with SpLD and received extra time (or extra time and the use of a word processor) to identify any changes in marks that occurred following the awarding of exam adjustments. This data was cross-referenced with any changes in exam marks of their TD peers who sat the same exams under standard conditions over the same period of time, to identify if any changes in marks simply represent the 'normal maturation' of student grades across the lifetime of their degree, or if any change in exam marks of the participants with SpLD following the implementation of exam adjustments is of a magnitude to be statistically

[104] Unlike in Phase 1 of the study, the participants in phase 2 were drawn from the full range of Humanities and STEM discipline areas studied at the University.

significant. In addition, data regarding the number of 1:1 study skills sessions undertaken by the participants with SpLD was analysed and compared with their exam marks, with the aim of exploring the impact (if any) of study skills sessions on exam performance. The data from this phase has been analysed using SPSS (version 25) to identify any statistical differences in the exam results between groups and identify any relationship between exam arrangements and grades or study skills sessions and grades (using a correlation analysis).

9.2.2 Sample group

The sample group consisted of 160[105] participants in total. This included 80 participants with SpLD who had been diagnosed part-way through their course (30 of whom were diagnosed after the first year of their course and 50 of whom were diagnosed after the second year of their course) and an index group of 80 TD participants who sat the same exams under standard exam conditions. The SpLD participants were matched with TD students on the basis of gender, age, course of study and exam time point. The overall exam grade achieved by a participant with SpLD at three time points across their course of study (at the end of the 1st, 2nd and 3rd year) was compared with the overall exam grade of a TD participant at the same time points (the TD participants selected sat the same exams, but under standard exam conditions). The sample groups consisted of the following participants:

- 30 students who received a diagnosis of SpLD in the second year of their course and who were granted 25% extra time (or 25% extra time and the use of a word processor) in their exams in the second and third year, but who sat their first-year exams under standard conditions.
- 50 students who received a diagnosis of a SpLD in the third year of their course and who were granted 25% extra time (or 25% extra time and the use of a word processor) in their exams in the third year, but who sat their first- and second-year exams under standard conditions.

---

[105] A power and sample size (PSS) analysis conducted through STATA, using the mean and standard deviations resulting from the data collected, confirmed that a total sample size of 160 participants provided this research project with a power of 90%

- 80 TD students who sat the same exams under standard conditions in all three years of their course.

Participants with SpLD were recruited from the research site University's Disability Database, which holds information on all SpLD students' assessment dates (data is held for over 1000 current SpLD students who would be eligible to participate in this research project). Participants with SpLD were selected via simple random selection using an online random number generator until sufficient participant numbers were achieved. TD participants were recruited using the research site University's main student database, which provides information about the course and year of study of all students. Eligible students were emailed with details of the project to ask if they would be interested in participating in the research project. The exam mark data was collected from the research site (University) database, which holds data on the exam marks achieved by each individual student in each year of study.

9.2.3 Variables

The independent variables in this research project were (a) student group [students with SpLD; TD students], (b) exam conditions [exam adjustments; standard conditions] and (c) year of adjustments [year 2; year 3; no adjustments]. The dependent variable in this research project was the participants' exam marks. This research project hypothesises that receiving exam adjustments influences exam marks. However, a possible outcome of this research project was that the analysis of the data could suggest no relationship exists between the variables.

9.2.4 Data Collection and analysis techniques

The following data was collected and analysed from each participant:

- SpLD participants: Exam grades in each year of study. Comparisons were made between the grades achieved in exams taken with and without exam arrangements to see if the exam grades improved following diagnosis and the implementation of exam arrangements.
- TD participants: Exam grades in each year of study. Comparisons were made between the grades achieved in each year's exams to identify any normal maturation or increase in exam grades across the years of the course.

- SpLD participants: The number of study skills sessions were identified for each SpLD participant. The exam results of SpLD participants who used study skills sessions were compared to the exam results of SpLD participants who did not use study skills sessions to identify any relationship between change in exam grade and study skills usage.

The overall exam grade achieved by participants with SpLD at three time points across their course of study was compared with the overall exam grade of TD participants at the same time points (the TD participants selected had taken the same exams, but under standard exam conditions).

9.2.5 Impact of study skills sessions

This second quantitative phase of the research project also sought to evaluate the impact that study skills sessions may have on the exam performance of students with SpLD by exploring any causal relationships between variables [exam marks; study skills sessions] through the following two-tailed hypothesis:

> H_1: there will be significant differences between the exam marks of students with SpLD who accessed study skills sessions and the exam mark of students with SpLD who did not access study skills sessions

This two-tailed hypothesis was operationalised by addressing the following sub-questions:

1. Do students with a SpLD who access study skills sessions achieve higher exam marks than their SpLD peers who do not access study skills sessions?

2. Is there a relationship between the number of study skills sessions accessed and exam grade?

Out of the 80 participants with SpLD in this phase of the research project, 32 participants accessed study skills sessions. Data was collected in relation to the number of study skills sessions each of the 32 participants accessed and the change in mark each of these participants achieved. A Pearson product-moment correlation coefficient was undertaken to identify any relationship between the number of study skills sessions used and exam performance (i.e. do those participants who use more study skills sessions achieve higher results than those who use fewer sessions). A comparison was also made between any change in marks achieved by the SpLD participants who accessed study skills sessions and any change in mark achieved by the SpLD participants who did not access study skills sessions to identify whether or not the study skills sessions provided made any significant contribution to boosting exam marks. It should be pointed out that, as only 32 participants accessed study skills sessions, post-hoc power calculations suggested that this sample size only has a 40% power of detecting a 'true' effect (alpha =.05; two-tailed; effect size $d = 0.4$).

9.2.6 Limitations

Several factors relating to the research design and methodology of the quantitative phases of the research project may limit the inferences that can be drawn from the data collected. Firstly, as the first quantitative phase used a quasi-experimental approach it was not possible to investigate the effect of the exam arrangements on the performance of students with SpLD by testing these same participants under standard exam conditions (i.e. without granting the accommodations to this sample group) and evaluating any differences. While this limitation was largely addressed in this second quantitative phase of the project, as this was also quasi-experimental it was not possible to test participants performance with and without adjustments in the same exam at the same time-point (albeit this was tested across similar exams in the three different years of the course).

A second limitation resulting from the research design relates to the potential impact on the dependent variables [marks, classification] of the mix of sample group characteristics. Harris et al (2020) argue that gender, race and social factors (including previous educational experience) can affect exam performance. Thus, confounding variables such as gender, age, ethnicity and type of previous schooling potentially pose a threat to the validity of the inferences the research project makes about cause and effect as these factors have the potential to affect exam results.

To protect the anonymity of the participants in the first phase of the research project, exam scripts were identified by candidate number only. Thus it was not possible to identify which (if any) of the students had received study skills tuition or had previous experience in the use of access arrangements. In addition, information relating to age, gender, ethnicity, socio-economic background and previous educational experience of participants was unavailable (due to anonymity) and it was therefore not possible to take these variables into account when making comparisons or to analyse any effect that these variables may have had on exam performance. However, in phase 2 of the research project, the identities of the participants were known to the researcher. To improve the internal validity of this phase of the research project, these variables have been accounted for through the design of phase 2 of the research project before the data was gathered (by restriction and matching). As far as possible, only TD participants sharing similar characteristics to those of the SpLD participants in terms of course, gender, ethnicity and nationality, were selected in order to reduce the impact of confounding variables. While every attempt was made to also match participants on the grounds of age and type of previous schooling, minor differences between the two groups exist in the areas of age (with SpLD tending to be slightly older on admission than was the case for the TD participants), type of previous schooling (a higher percentage of participants with SpLD attended independent schools than was the case for the TD participants), and mean UCAS tariff points on admission (the SpLD group achieved lower mean UCAS tariff points on admission). However, no between group differences existed between the mix of gender, nationality, ethnicity or course of study[106].

9.3 Comparison of marks before and after diagnosis

The participants' exam results obtained before they received a diagnosis of SpLD (where the participants sat their exams under standard conditions) were compared to the same participants' exam results after they received the diagnosis of SpLD and were granted exam adjustments to identify any statistical differences between the exam marks achieved when they sat the exams under standard conditions and the marks they achieved when they sat the exams with adjustments.

---

[106] See Appendix 6 for full details of the course of study, gender, ethnicity, nationality, age, previous schooling and UCAS points on entry for the participants

9.3.1 SpLD participants who were diagnosed after the first year of their course.

The mean exam marks of the participants who were diagnosed with SpLD after their first-year exams (and so sat their first year exams under standard conditions and received exam arrangements in the second and third year exams) increased by a mean of 4 marks between the first and second year (i.e. following the implementation of exam adjustments). Between the second and third year (both exams being taken with exam arrangements) the exam marks of the SpLD group increased by a mean of 1 mark. A Wilcoxon matched pairs test[107] was conducted to compare the marks achieved when the exams were taken with exam arrangements and the marks achieved when the exams were taken under standard exam conditions. There was a significant difference in the scores for exam marks with exam arrangements (Mdn=66.5%) and exam marks under standard exam conditions (Mdn=61%); z = -4.169, p <.001 (W+ = 15.44, W- = 6.67). This indicates that the SpLD group who received a diagnosis after the first year of their course increased their exam scores significantly following the granting of exam adjustments in their second and third years. The effect size was medium to large (r = .35; r^2 = .12; Hedges' g[108] = .737, Cohen's d = .746).

In comparison, the exam marks of the matched TD group increased by a mean of 1 mark in each year of their course. A paired-samples t-test[109] was conducted to compare the marks achieved in the third year of the course with the marks achieved in the first year of the course, to identify if the increase in marks over the lifetime of the course was significant. There was a non-significant difference between the exam marks achieved in the first year (M=66%, SD=3.10) and exam marks achieved in the third year (M=68%, SD=2.84); $t(29)$= .128, p = 899. This indicates that the increase in the scores of the TD group across the lifetime of the course does not represent a significant increase.

---

107 The Shapiro-Wilk test of normality for the mark of the SpLD group indicated that the data were *not* statistically normal (W = .871; p = .002). As the assumptions for a parametric test were not met, a non-parametric test was used.

108 Hedges' g has been used to calculate effect size due to the unequal sample sizes (Sawilowsky, 2009).

109 The Shapiro-Wilk test of normality for the mean mark of the TD group in year 1 (W = .957; p = .257) and the mean mark in year 3 (W = .954; p = .210) indicated that the data were statistically normal. The data met the assumptions for a parametric test.

The identified change in exam results of the SpLD participants who received exam adjustments in their second and third year has been compared to any changes in the exam results of the matched TD participants who took the same exams over the same period of time, but with no exam adjustments. This aims to explore whether or not the changes identified in the exam performance of the SpLD group were significantly different to those of the matched TD group at the same time point. A Kruskal-Wallis *H* test[110] [group, change in exam marks between year 1 and year 2] identified a significant between-group difference in the change of marks between year 1 and year 2 (i.e. before and after the group with SpLD were granted their exam arrangements) ($\chi^2 = 12.545$, $p < .001$, with a mean rank of 22.55 for the TD group and 38.45 for the SpLD group). The effect size was large ($r = .38$; $r^2 = .14$; Hedges' $g = .80$, Cohen's d = .81). However, a Kruskal-Wallis *H* test[111] [group, change in exam marks between year 2 and year 3] identified no significant between-group difference in the change of marks between year 2 and year 3 ($\chi^2 = .008$, $p = .929$, with a mean rank of 30.30 for the TD group and 30.70 for the SpLD group). This suggests that the granting of exam arrangements resulted in a significant boost in marks for the SpLD group, over and above that expected by the normal increase in marks as demonstrated by the TD group. However, in the subsequent years, when the exam adjustments were ongoing, the increase in marks achieved by the SpLD group mirrored the 'normal' increase demonstrated by the TD group. A comparison of the increase in the marks of both groups is shown in Table 9.1 and Figure 9.1

[110] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the TD group were not normally distributed (W=.895, p= .006) and the mark difference data for the SpLD were not normally distributed ($W = .911$, $p = .016$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

[111] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the TD group were not normally distributed ($W = .930$, $p = .006$) and the mark difference data for the SpLD were not normally distributed ($W = .955$, $p = .055$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

Table 9.1: Comparison of mean marks achieved by each group in three years of the course: SpLD group diagnosed after year 1

Group	Mean Mark Year 1 (%)	Mean Mark Year 2 (%)	Mean Mark Year 3 (%)
SpLD (n=30)	61	65	66
TD (n=30)	66	67	68

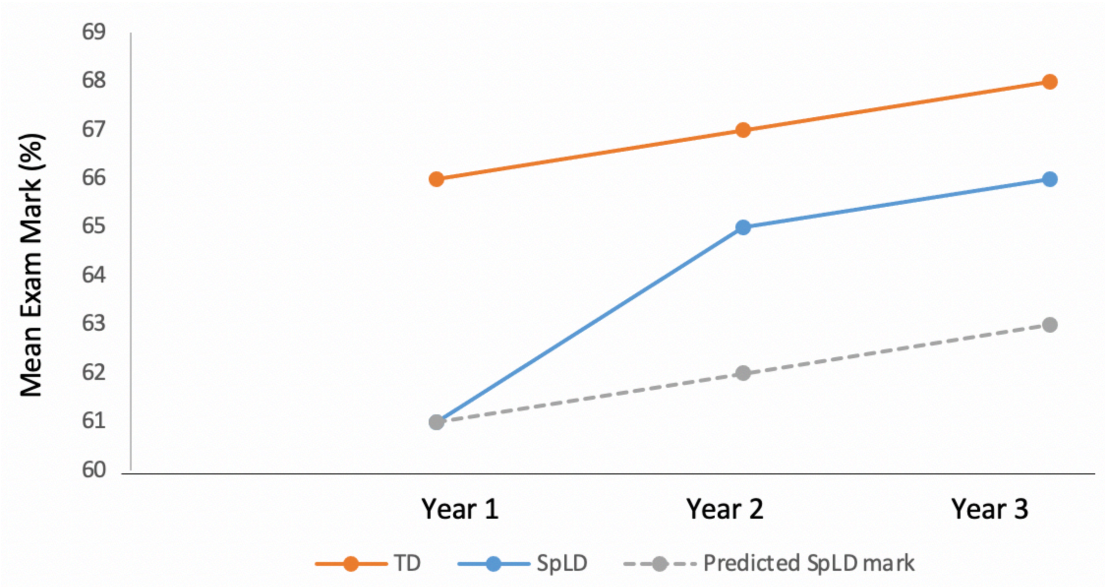


Figure 9.1: Comparison between exam marks of TD group and exam marks of SpLD group who received exam arrangements in 2nd and 3rd year and predicted SpLD exam marks.

Having analysed the differences between the SpLD and TD groups in terms of the change in marks across each year, it is also useful to compare the differences in marks between the two groups in each year to see if the mark differences between the groups that have been identified are of a magnitude that is statistically significant (i.e. does an awarding gap exist between the two groups and, if so, does the awarding gap close after the exam arrangements are granted). Therefore, a Kruskal-Wallis *H* test[112] [group,

[112] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that although the exam mark data for the TD group were normally distributed ($W = .945$, $p = .122$), the exam mark data for the SpLD were not normally distributed ($W = .871$, $p = .002$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

mark] comparing the exam marks achieved by the SpLD participants in year one (i.e. before exam adjustments were implemented) with those of the TD participants was conducted. This confirmed that the difference in mean marks achieved by the TD group in year 1 (66%) compared to the SpLD group (61%) was significant, with the SpLD group achieving significantly lower marks ($\chi^2 = 12.789$, $p < .001$, with a mean rank of 22.47 for SpLD and 38.53 for TD). The effect size was large ($r = .38$; $r^2 = .15$; Cohen's $d = .83$). In addition, a Kruskal-Wallis *H* test[113] [group, mark] comparing the exam marks achieved by the SpLD participants in year two (i.e. after exam adjustments were implemented) with those of the TD participants confirmed that the difference in mean marks achieved by the TD group (67%) compared to the SpLD group (65%) was significant ($\chi^2 = 5.198$, $p = .02$, with a mean rank of 26.52 for SpLD and 34.48 for TD), albeit the effect size was small ($r = .18$; $r^2 = .03$; Cohen's $d = .37$). This was also the case in year 3, where the marks of the SpLD group (66%) were significantly lower than those of the TD participants (68%) ($\chi^2 = 4.989$, $p = .024$, with a mean rank of 25.93 for SpLD and 35.07 for TD[114]), and, again, a small effect size ($r = .13$; $r^2 = .02$; Cohen's $d = .28$). This outcome suggests that the awarding of exam access arrangements to exam candidates with SpLD narrows the awarding gap between SpLD candidates and their TD peers (shifting the difference in exam marks between the two groups from a large effect size to a small effect size), but the adjustments fail to fully close the gap.

9.3.2 SpLD participants who were diagnosed after year 2

The participants diagnosed with SpLD after their second-year exams sat their first and second year exams under standard conditions and only received exam arrangements in their third year exams. The exam marks of this SpLD group increased by a mean of 4 marks after the implementation of exam adjustments in the third-year exams. By contrast, the exam marks of the same SpLD group increased by a mean of 1 mark between the first and second year of their exams (both of which were taken under

[113] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the exam mark data for the TD group were not normally distributed ($W = .916$, $p = .021$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

[114] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the exam mark data for the TD group were not normally distributed ($W = .864$, $p = .001$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

standard conditions). A Wilcoxon matched pairs test[115] was conducted to compare the marks achieved by the SpLD participants when the exams were taken with exam arrangements with the marks achieved by the same SpLD participants when the exams were taken under standard exam conditions. There was a significant difference in the scores for exam marks following the implementation of exam arrangements (Mdn=66%) and exam marks under standard exam conditions (Mdn=61%); $z = -3.861$, $p <.001$ (W+ = 16.15, W- = 11.25). This indicates that the SpLD group who received a diagnosis after the second year of their course increased their exam scores significantly following the granting of exam adjustments in their third year. The effect size was medium ($r = .30$; $r^2 = .09$; Hedges' *g*[116] = .628, Cohen's $d = .636$).

At the same time, the increase in scores of the SpLD group between the first and second year exams (both taken under standard conditions) was not significant. A Wilcoxon matched pairs test[117] was conducted to compare the marks achieved in year one (when the exams were taken under standard conditions) with the marks achieved in year two (when the exams were also taken under standard exam conditions) and this revealed a non-significant difference ($z = -1.194$, $p = .233$ (W+ = 18.17, W- = 11.61). This indicates that the SpLD group did not increase their exam scores significantly during the years in which they sat the exams under standard conditions; the significant increase in marks only occurred once the exam adjustments were granted in their third year.

In comparison, the exam marks of the matched TD group increased by a mean of 1 mark in each year of their course. A paired-samples t-test was conducted to compare the marks achieved by the TD participants in the third year of the course with the marks they achieved in the first year of the course, and this revealed that the increase in marks over the lifetime of the course was non-significant for this group ($t(30)= .228$, $p = 821$).

---

[115] The Shapiro-Wilk test of normality for the mark of the SpLD group pre-test indicated that the data were *not* statistically normal ($W = .871$; $p = .002$). As the assumptions for a parametric test were not met, a non-parametric test was used.

[116] Hedges' g has been used due to calculate effect size due to the unequal sample sizes (Sawilowsky, 2009).

[117] The Shapiro-Wilk test of normality for the mark of the SpLD group in year 1 indicated that the data were *not* statistically normal ($W = .863$; $p = .001$) and the mark of the SpLD group in year 2 were also not normally distributed ($W = .899$; $p = .007$). As the assumptions for a parametric test were not met, a non-parametric test was used.

The identified change in exam results of the SpLD participants who received exam adjustments only in their third year has been compared to any changes in the exam results of the matched TD peers who took at the same exams over the same period of time, but with no exam adjustments. This aims to identify if the changes identified in the exam performance of the SpLD group were significantly different to those of the matched TD group at the same time point. A Kruskal-Wallis *H* test[118] [group, change in exam marks between year 1 and year 2] identified no significant between group difference in the change of marks between year 1 and year 2 (i.e. when both groups sat their exams under standard arrangements) ($\chi^2 = .143$, $p = .705$, with a mean rank of 51.59 for the TD group and 49.41 for the SpLD group). This indicates that the change in marks achieved by the SpLD group between years 1 and 2 (the exams in both of these years were taken were under standard conditions) mirrored the change in marks achieved by the TD group in years 1 and 2 (who also sat the exams under standard conditions). However, A Kruskal-Wallis *H* test[119] [group, change in exam marks between year 2 and year 3] identified a significant between-group difference in the change of marks between year 2 and year 3 (i.e. when the SpLD group sat their exams with exam adjustments) ($\chi^2 = 6.481$, $p = .011$, with a mean rank of 43.15 for the TD group and 57.85 for the SpLD group). The effect size was medium ($r = .30$; $r^2 = .09$; Hedges' $g = .623$, Cohen's d = .631). A comparison of the increase in the marks of both groups is shown in Table 9.2 and Figure 9.2.

[118] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the TD group were not normally distributed ($W = .950$, $p = .034$) and the mark difference data for the SpLD were not normally distributed ($W = .908$, $p = .001$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

[119] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the TD group were not normally distributed ($W = .930$, $p = .006$) and the mark difference data for the SpLD were not normally distributed ($W = .955$, $p = .055$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

Table 9.2: Comparison of mean marks achieved by each group in three years of the course: SpLD group diagnosed after year 2.

Group	Mean Mark Year 1 (%)	Mean Mark Year 2 (%)	Mean Mark Year3 (%)
SpLD (n=50)	61	62	66
TD (n=50)	65	67	68

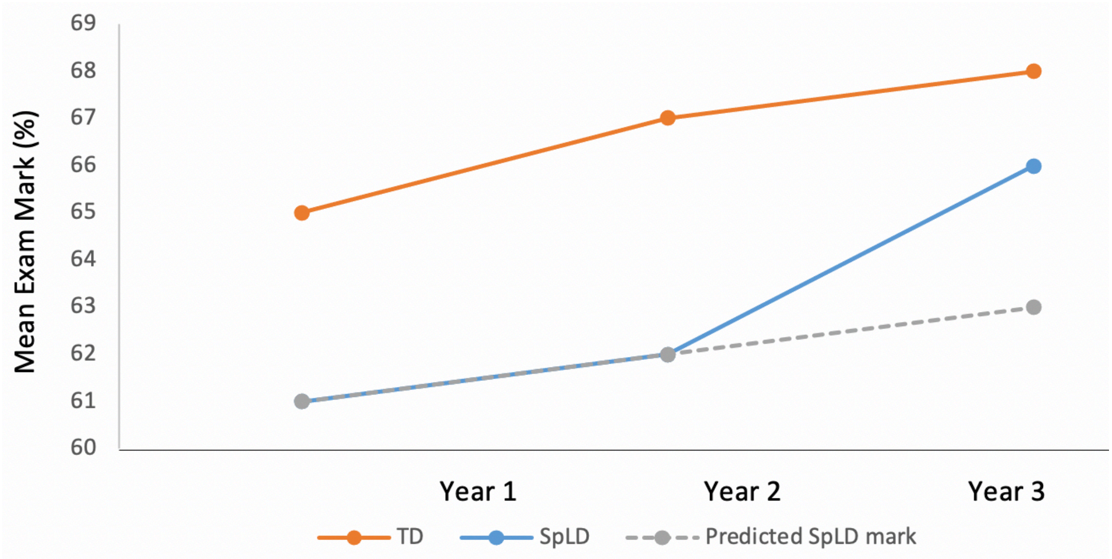


Figure 9.2: Comparison between exam marks of TD group and exam marks of SpLD group who received exam arrangements in 3rd year and predicted SpLD marks.

Having analysed the differences between the SpLD and TD groups in terms of the change in marks across each year, it is also useful to compare the differences in marks between the two groups in each year to see if the mark differences between the groups that have been identified are of a magnitude that is statistically significant (i.e. does an awarding gap exist between the two groups when both groups take exams under standard conditions and, if so, does the awarding gap close after the exam arrangements are granted?). Therefore, a Kruskal-Wallis *H* test[120] [group, mark] comparing the exam

[120] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that although the exam mark data for the TD group were normally distributed ($W = .989$, $p = .204$), the exam mark data for the SpLD were not normally distributed ($W = .882$, $p < .001$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

marks achieved by the SpLD participants in year one (i.e. before exam adjustments were implemented) with those of the TD participants was conducted. This confirmed that the difference in mean marks achieved by the TD group in year 1 (65%) compared to the SpLD group (61%) was significant, with the SpLD group achieving significantly lower marks ($\chi^2 = 21.566$, $p < .001$, with a mean rank of 37.07 for SpLD and 63.93 for TD). The effect size was large ($r = .33$; $r^2 = .11$; Cohen's $d = .71$). In addition, a Kruskal-Wallis H test[121] [group, mark] comparing the exam marks achieved by the SpLD participants in year two (also before exam adjustments were implemented) with those of the TD participants confirmed that the difference in mean marks achieved by the TD group (67%) compared to the SpLD group (62%) was significant ($\chi^2 = 22.251$, $p < .001$, with a mean rank of 36.84 for SpLD and 64.16 for TD), again the effect size was large ($r = .41$; $r^2 = .17$; Cohen's $d = .89$). Despite the awarding of exam adjustments in year three, the differences between the two groups remained significant, with the SpLD group achieving lower marks (66%) than their TD peers (68%) ($\chi^2 = 4.474$ $p = .034$, with a mean rank of 44.38 for SpLD and 56.62 for TD[122]). However, this difference represented a small effect size ($r = .18$; $r^2 = .03$; Cohen's $d = .36$). As with the participants who were diagnosed after the first year of their course, this outcome suggests that the awarding of exam access arrangements to exam candidates with SpLD narrows the awarding gap between SpLD candidates and their TD peers (shifting the difference in exam marks between the two groups from a large effect size to a small effect size), but the adjustments fail to fully close the gap.

[121] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that although the exam mark data for the TD group were normally distributed ($W = .972$, $p = .270$), the exam mark data for the SpLD were not normally distributed ($W = .918$, $p = .002$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

[122] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that although the exam mark data for the TD group were normally distributed ($W = .966$, $p = .151$), the exam mark data for the SpLD were not normally distributed ($W = .945$, $p = .048$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

9.3.3 Within group differences: a comparison of the change in marks of the SpLD subgroup who handwrote their exam scripts with those who used a word processor.

Out of the 80 SpLD participants in this research project, 29 handwrote their exam scripts with 25% extra time and 59 used a word processor with 25% extra time. Given that the SpLD participants experienced two different exam conditions, a comparison between the marks achieved by these two SpLD sub-groups was made to explore whether the use of a word processor contributed to a greater increase in marks than was the case for SpLD participants who handwrote their scripts. A Kruskal-Wallis *H* test[123] [group, change in exam marks after receiving exam arrangements] confirmed that there was a non-significant between group difference in the change in exam marks following the implementation of exam arrangements for the SpLD group who handwrote their exam script with 25% extra time and the SpLD group who word processed their exam scripts with 25% extra time (χ^2 = 1.926, p= .165, with a mean rank of 45.28 for the for the SpLD group who handwrote their exam script with 25% extra time and 37.78 for the SpLD group who word processed their exam scripts with 25% extra time). This indicates that, following the implementation of exam arrangements, the change in marks achieved by the SpLD group who word processed their exam scripts with 25% extra time was statistically similar to the change in marks achieved by the SpLD group who word processed their exam scripts with 25% extra time. Thus, there is no evidence to suggest that using a word processor with extra time boosts exam marks more than is the case when writing by hand with extra time in exams. This outcome contradicts the study by Duncan & Purcell (2017), which found that SpLD Humanities exam candidates achieved higher marks when using a word processor than when writing by hand. However, the findings of this second phase of the project mirrors the outcome of the first phase of this current research project that showed that both the SpLD group who word processed their script with 25% extra time and the SpLD group who handwrote their script with 25% extra time achieve statistically similar exam marks

[123] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mean mark difference data for the SpLD group who word processed their scripts were not normally distributed (W = .951, p = .035). Thus, the data did not meet the assumptions required to conduct an ANOVA.

9.4 Differences between SpLD diagnostic categories

As the data from the first phase of the research project was collected anonymously, one of the limitations of the first phase of the research project was the inability to determine if any differences exist in the outcomes for participants with dyslexia in comparison with participants with dyspraxia (or participants with a dual diagnosis of dyslexia and dyspraxia). Due to the construct-irrelevant deficits that are attributable to both dyslexia and dyspraxia (and given that both groups receive similar exam arrangements), it was hypothesised that the characteristics shared by these groups that are relevant to the context of formal, timed exams, were sufficiently similar to achieve reliable and valid findings in relation to exam outcomes, even if these two sub-groups were combined into one category of ‘SpLD’. Nonetheless, it could be argued that the differences between the diagnostic profiles of individuals with dyslexia and those with dyspraxia may result in potential differences in outcomes between these two categories of SpLD in exam situations.

The data from the participants in this second phase of the research project, by contrast, was collected confidentially, rather than anonymously, making it possible to identify the diagnosis disclosed by each participant. Thus, the data from each of the SpLD diagnostic categories could be compared with each other, identifying whether or not any differences are evident between the outcomes of the participants with dyslexia and the outcomes of the participants with dyspraxia (or participants with dual diagnosis). This analysis may also help to clarify whether the hypothesis of the first phase of the research project in relation to the inclusion criteria for the SpLD group was justified.

9.4.1 Breakdown of sub-groups of SpLD

Out of the 80 participants with SpLD, 33 disclosed a diagnosis of dyslexia, 36 disclosed a diagnosis of dyspraxia and 11 disclosed a diagnosis of both dyslexia and dyspraxia. The breakdown in terms of the year of diagnosis is shown in Table 9.3.

Table 9.3: Breakdown of SpLD participants' disclosed diagnoses.

Diagnostic category	Year of diagnosis: year 2 (n=30)	Year of diagnosis: year 3 (n=50)
Dyslexia (n=33)	12	21
Dyspraxia (n=36)	14	22
Dyslexia and Dyspraxia (n=11)	4	7

9.4.2 Comparison of exam performance between the SpLD sub-groups: participants diagnosed with SpLD after year 1

The exam marks achieved before their diagnosis by those participants who were diagnosed with SpLD after the first year of their course (and so sat the first year exams under standard conditions and the second and third year exams with exam arrangements) were compared across the three SpLD diagnostic categories to see if there were any differences in exam performance between participants with dyslexia and participants with dyspraxia (or participants with both dyslexia and dyspraxia) when they sat the exams under standard conditions. A Kruskal-Wallis *H* test[124] [diagnosis, mean marks before exam arrangements] confirmed that there was no significant between group differences ($\chi^2 = 1.450$, $p = .484$, with a mean rank of 13.92 for the participants with dyslexia, a mean rank of 15.57 for the participants with dyspraxia and a mean rank of 16.52 for the participants with a dual diagnosis). The exam marks achieved post diagnosis (i.e. after receiving exam adjustments) by those participants who were diagnosed with SpLD after the first year of their course were compared across the three SpLD diagnostic categories to explore whether any differences exist in exam performance between participants with dyslexia and participants with dyspraxia (or participants with both dyslexia and dyspraxia) when they sat the exams with exam adjustments. A Kruskal-Wallis *H* test [diagnosis, mean marks after exam arrangements] confirmed that there was no significant between group differences ($\chi^2 = .757$, $p = .685$, with a mean rank of 15.79 for the participants with dyslexia, a mean rank of 16.50 for

[124] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the data for the participants who had a diagnosis of dyslexia and dyspraxia were not normally distributed (W=.781, p=.006). Thus, the data did not meet the assumptions required to conduct an ANOVA.

the participants with dyspraxia and a mean rank score of 16.68 for the participants with a dual diagnosis). A Kruskal-Wallis *H* test[125] [diagnosis, change in exam marks after receiving exam arrangements] also confirmed that there was no significant between group difference in the change itself in exam marks following the granting of exam arrangements between the sub-categories of SpLD ($\chi^2 = 1.497$, $p = .477$, with a mean rank score of 15.38 for the participants with dyslexia, a mean rank score of 16.93 for the participants with dyspraxia and a mean rank of 10.88 for the participants with a dual diagnosis[126]). The breakdown of exam marks achieved by each diagnostic sub-group before and after receiving exam arrangements is shown in Table 9.4.

Table 9.4: Breakdown of exam marks achieved by each diagnostic sub-group before and after receiving exam arrangements

Diagnosis	Mean exam mark under standard conditions (SD)	Mean exam mark with exam arrangements (SD)	Change in mean mark following exam arrangements (SD)
Dyslexia (n=12)	59.21 (5.81)	65.17 (9.95)	5.96 (6.46)
Dyspraxia (n=14)	61.00 (6.16)	66.04 (6.03)	5.04 (4.61)
Dyslexia & dyspraxia (n=4)	64.50 (3.08)	67.00 (4.04)	2.50 (3.76)

The comparison of the mean exam marks achieved before and after the granting of exam access arrangements between the participants with dyslexia, dyspraxia and those with a dual diagnosis of dyslexia and dyspraxia is shown graphically in Figure 9.3.

[125] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the data for the participants who had a diagnosis of dyslexia and dyspraxia were not normally distributed ($W = .729$, $p = .024$). Thus, the data did not meet the assumptions required to conduct an ANOVA.

[126] It should be born in mind that only 4 participants disclosed both dyslexia and dyspraxia, which is not considered a sufficiently large sample size to detect an effect.

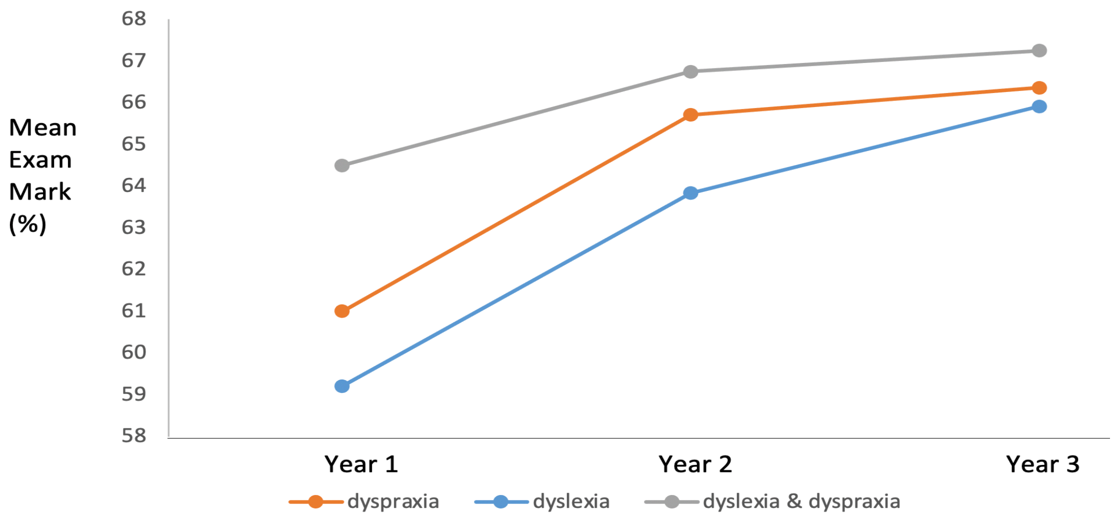


Figure 9.3: Comparison between the exam marks of the SpLD sub-categories for participants who received exam arrangements in the 2nd and 3rd year of their course.

9.4.3 Comparison of exam performance between the SpLD sub-groups: participants diagnosed with SpLD after year 2

A similar outcome was observed in the exam data of the SpLD participants who were diagnosed after the second year of their course (and so sat the first *and* second year exams under standard conditions and the third year exams with exam arrangements). The exam marks achieved when the exams were taken under standard conditions (i.e. without any exam arrangements) by those participants who were diagnosed with SpLD after the second year of their course were compared across the three SpLD diagnostic categories. A Kruskal-Wallis *H* test[127] [diagnosis, mean marks before exam arrangements] confirmed that there was no significant between group differences (χ^2 = .401, p = .818, with a mean rank of 25.12 for the participants with dyslexia, a mean rank of 24.84 for the participants with dyspraxia and a mean rank of 28,71 for the participants with a dual diagnosis). The exam marks achieved post diagnosis (i.e. after receiving exam adjustments) by those participants who were diagnosed with SpLD after the second year of their course were compared across the three SpLD diagnostic categories and a Kruskal-Wallis *H* test [diagnosis, mean marks after exam

[127] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the group with dyslexia were not normally distributed (W = .866, p = .008). Thus, the data did not meet the assumptions required to conduct an ANOVA.

arrangements] confirmed that there was no significant between group differences (χ^2 = .442, p = .802, with a mean rank of 24.43 for the participants with dyslexia, a mean rank of 25.52 for the participants with dyspraxia and a mean rank of 27.57 for the participants with a dual diagnosis).

A Kruskal-Wallis H test[128] [diagnosis, change in exam marks after receiving exam arrangements] also confirmed that there was no significant between-group difference in the change itself in exam marks after receiving exam adjustments between the thee sub-categories of SpLD (χ^2 = .200, p = .905, with a mean rank of 24.74 for the participants with dyslexia, a mean rank of 25.57 for the participants with dyspraxia and a mean rank of 27.57 for the participants with a dual diagnosis[129]). The breakdown of exam marks achieved by each diagnostic sub-group before and after receiving exam arrangements is shown in Table 9.5.

Table 9.5: Comparison of mean exam marks (with standard deviation) between the three SpLD diagnostic categories.

Diagnosis	Mean exam mark under standard conditions (SD)	Mean exam mark with exam arrangements (SD)	Change in mean mark following exam arrangements (SD)
Dyslexia (n=21)	61.52 (6.62)	65.71 (8.21)	4.19 (5.71)
Dyspraxia (n=22)	61.50 (5.72)	65.86 (3.96)	4.36 (6.44)
Dyslexia & dyspraxia (n=7)	61.71 (4.23)	66.57 (1.72)	4.86 (4.05)

The comparison of mean marks before and after the granting of exam access arrangements between the participants with dyslexia, dyspraxia and those with a dual diagnosis of dyslexia and dyspraxia is shown graphically in Figure 9.4.

[128] A non-parametric test was used as the Shapiro-Wilk test of normality indicated that the mark difference data for the group with dyspraxia were not normally distributed (W = .763, p = .000). Thus, the data did not meet the assumptions required to conduct an ANOVA.
[129] It should be born in mind that only 7 participants disclosed both dyslexia and dyspraxia, which is not considered a sufficiently large sample size to detect an effect.

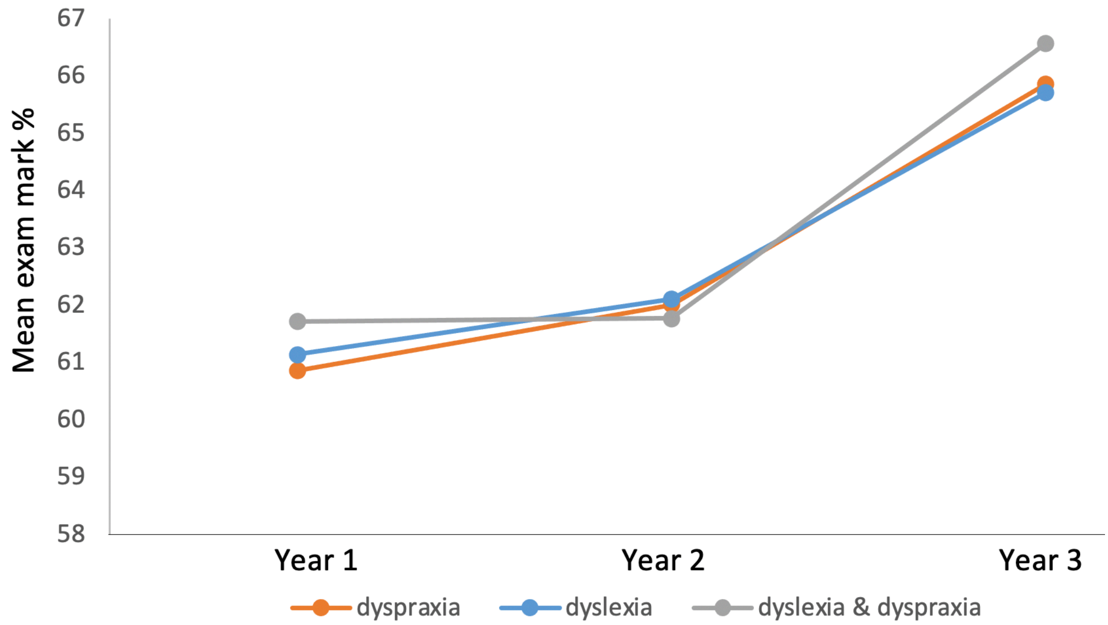


Figure 9.4: Comparison between the exam marks of the SpLD sub-categories for participants who received exam arrangements in the 3rd year of their course only.

9.4.4 Conclusion

When the data from this phase of the research project is separated out into the three different sub-categories of SpLD (namely, dyslexia, dyspraxia and a dual diagnosis of dyslexia and dyspraxia) the analysis shows that there are no significant differences in exam performance between the three SpLD diagnostic categories. That is to say, participants with dyslexia performed similarly to those with dyspraxia (and also similarly to those with a dual diagnosis) both when taking the exams under standard conditions as well as when taking the exams with adjustments. The outcome in relation to participants with a comorbid diagnosis of dyslexia and dyspraxia fails to support extant studies in the field that suggest that having more than one diagnosis has a greater negative impact on learning outcomes than having one diagnosis as the simultaneous presence of different specific learning difficulties result in multiple difficulties (Kirby, 2012; Cappa, Muzio & Giulivi, 2012).

However, the outcome of this current research project may be explained by the small number of participants in this sample group who have disclosed a dual diagnosis, which may limit the ability of the research project to reliably detect an effect. Interestingly, this low number of participants disclosing dual diagnosis of both dyslexia and dyspraxia also fails to support research studies that suggest that co-occurrence is the norm rather than the exception (Kirby, 2012; Kirby & Sugden 2007; Kaplan 1998). One explanation of the low numbers of participants with dual diagnosis in this phase of the research project could be that, if co-occurrence is the norm, coupled with the fact that having more than one SpLD has a greater negative impact than having one SpLD alone, students with co-occurring SpLDs are arguably more likely to have sought a diagnosis earlier in their academic career (with their SpLD being identified at school rather than part way through their degree course). The participants in this research project had received late diagnoses, and so it could be argued that these late diagnoses are potentially the result of more subtle 'endophenotype' of SpLD, where the characteristics have not been sufficiently problematic prior to studying at degree level to cause the student to seek a diagnostic assessment at an earlier stage (which is less likely to be the case for individuals who have co-occurring SpLDs). This potential explanation is supported by Callens & Brysbaert (2019) who suggest that 'as HE is not compulsory, individuals with dyslexia entering this sector of education are likely to be a specific, highly motivated subgroup with possibly less severe symptoms' (p.194). Perhaps it could be argued that those individuals with SpLD for whom the impact is more significant are potentially more likely to seek a diagnosis earlier in their academic career than is the case for those for whom the academic impact is less apparent.

Nonetheless, despite these contradictions, the lack of any statistically significant differences in the exam marks between the three SpLD sub-groups in this current phase of the research project suggests that the exam performance of the participants with dyslexia, both with and without exam arrangements, are sufficiently similar to that of the participants with dyspraxia (and also to those with co-occurring diagnoses) to justify amalgamating these sub-groups into a single SpLD group for the purposes of this research project's data analysis.

9.5 Study skills

An extensive range of literature exists that suggests that explicitly taught study skills techniques, which confer on students a greater understanding of how to study successfully, can positively influence the academic achievement of university students and contribute to improved learning outcomes (Barrable et al 2018; Hassanbeigi et al., 2011; Rahim & Meon, 2012; Saeidi et al, 2019; Wibrowski et al., 2017). While the literature suggests that this is the case for students generally in higher education, students with SpLD experience particular challenges with a wide range of skills involved in effective academic study, due to the underpinning deficits associated with the profiles of dyslexia and dyspraxia. These difficulties may include, among others, problems with note-taking in lectures, summarising from texts, planning, organising and structuring ideas in essays, proofreading, fluency of written expression, reading speed and the need to re-read frequently to extract meaning, spelling and reading accuracy, recalling information, processing information at speed, general organisation, time management and concentrating for sustained periods (Broggi et al, 2019; Callens & Brysbaert, 2019; Falzon, 2020; Sayeski, 2019; Snowling et al, 2020; Wade & Kazeck 2018; Wilson et al, 2017). While this list of challenges is not exhaustive, it does illustrate the range of difficulties that students with SpLD are likely to experience when studying in higher education and the obstacles these students may face in achieving their potential in the HE environment.

In recognition of the barriers to learning that HE students with SpLD encounter due to the nature of their disability, 1:1 study skills tuition with a tutor specialising in teaching students with SpLD is a common academic support provision that is offered to university students diagnosed with SpLD (Rodger et al, 2015; Saeidi et al, 2019; Wibrowski et al., 2017). This support aims to help students to develop strategies to address the challenges that their SpLD may present in accessing their course. The current body of literature concerned with the efficacy of specialist 1:1 study skills tuition suggests that University students with SpLD identify number of key benefits of the sessions, including; the development of independent learning skills; a person-centred, individualised, development of study strategies; the development of critical analysis and problem-solving techniques; increased confidence and improved academic self-esteem; greater self-understanding and self-advocacy skills; increase in motivation and achievement; and the ability to transfer strategies learned to other aspects of life, including the workplace (Kirwan and Leather, 2011; Rahim & Meon, 2012; Saeidi et al, 2019; Wibrowski et al., 2017). It should be borne in mind, however, that the evidence for the efficacy of the study skills tuition identified in these studies (including the reports of increased achievement), was based on students' self-reported improvements, rather than using quasi-experimental data identifying a quantitative change in marks. This current project aims to fill that gap.

9.5.1 Methodology

Out of the 80 participants with SpLD in this second phase of the research project, only 40% (32 participants) accessed study skills sessions, with 60% (48 participants) electing not to access any of the study skills sessions to which they were entitled. Out of those participants who accessed study skills sessions, the median number of sessions used was 3[130].

9.5.2 Relationship between the number of study skills sessions used and the change in mark.

A Pearson product-moment correlation coefficient was computed for the SpLD participants who accessed study skills sessions to identify any relationship between the number of study skills sessions used and exam performance. That is to say, among those students who accessed study skills sessions, did those who used a higher number of sessions achieve a greater increase in exam mark than those who used fewer sessions (as measured by the change between the mark achieved before the participant was diagnosed with SpLD and the mark achieved after the participant was diagnosed with SpLD and accessed study skills sessions)? A very weak, non-significant positive relationship was identified (r_p = .122, N=32, p = .507), suggesting no statistical relationship between the number of study skills sessions used and the improvement in exam performance. The relationship between number of study skills sessions accessed and exam marks attained is shown in Figure 9.5.

---

[130] The Shapiro-Wilk test of normality for the number of study skills sessions used by the SpLD (W = .531; p < .001) indicated that the data were not statistically normal. Therefore, the median is a more representative figure to describe this data than the mean.

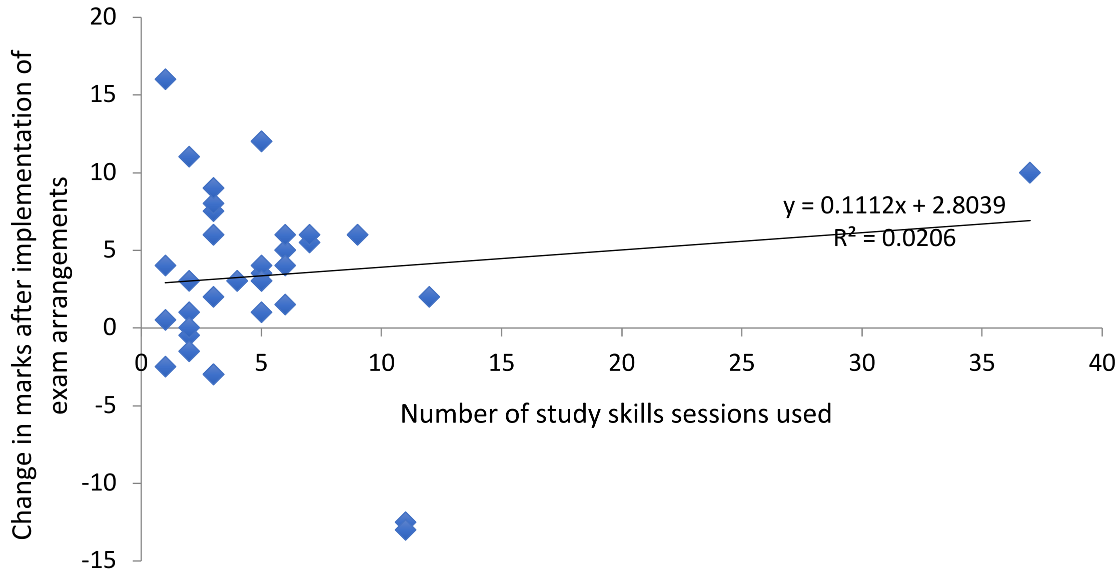


Figure 9.5: The relationship between number of study skills sessions accessed and change in exam marks.

9.5.2 Comparison between SpLD participants who accessed study skills sessions and SpLD participants who did not

The participants with SpLD who accessed study skills sessions increased their marks by a mean of 3.34 points (SD 5.16) in exams after being diagnosed with SpLD and receiving both study skills sessions and exam access arrangements. By comparison, the SpLD participants who did not access study skills sessions achieved a mean mark increase of 5.26 points (SD 5.77) in exams after being diagnosed and receiving only exam access arrangements (but not study skills sessions). A one-way ANOVA [group, increase in mark], showed that this difference between the two groups was not statistically significant ($F(1, 78) = 2.301$, $p = .133$)[131]. Therefore, this comparison of the increase in marks between the SpLD participants who did and did not access study skills sessions fails to suggest that the study skills sessions improved the performance of the SpLD participants who used them to a greater extent than was the case for the SpLD participants who did not use study skills sessions.

---

[131] The Shapiro-Wilk test of normality for the increase in mark of the SpLD group who accessed study skills sessions indicated that the data are from a normally distributed population ($W = .953$; $p = .053$). This was also true for the increase in mark of the SpLD group who did not access study skills sessions ($W = .960$; $p = .281$). Levene's Test of Homogeneity of Variances confirmed that variances are equal across the two participant groups ($F(1,78) = .299$; $p = .586$).

9.6 Conclusion

In summary, this second phase of the research project suggests that the implementation of exam access arrangements makes a significant positive difference to the exam marks of candidates with SpLD and significantly reduces the awarding gap that exists between the SpLD and TD participants. However, despite the granting of exam access arrangements (which aim to level the playing field), the awarding gap between the two groups is reduced in timed, closed book, written exams, but not fully closed.

The data from this research project shows that the marks of the TD participants most closely reflect those of the entire University cohort than is the case for the participants with SpLD. This suggests the existence of a university wide awarding gap, with students with SpLD achieving significantly lower exam marks than their TD peers across the University as a whole, despite the implementation of exam adjustments. The increase in marks of the whole University cohort from 2016 to 2018 is shown in Table 9.6.

Table 9.6: Increase in marks of the whole University cohort across three years (2016-18)

Total number of undergraduate students	Exam year	Mean mark (%)
12, 163	2018	68
11,800	2017	67
11,786	2016	65

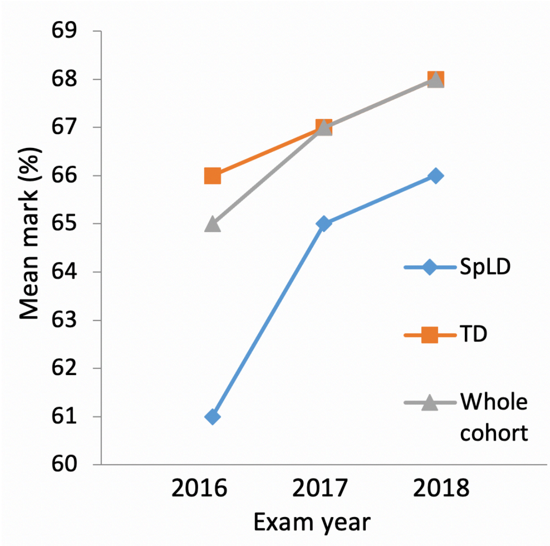


Figure 9.7: Comparison between exam marks of Whole University cohort, TD group and SpLD group who received exam arrangements in 2nd year.

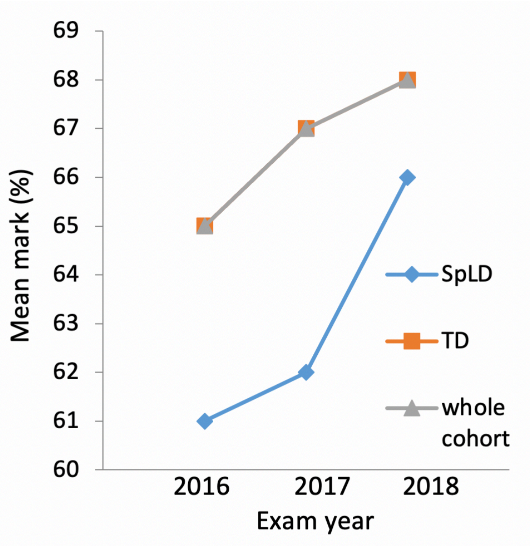


Figure 9.6: Comparison between exam marks of Whole University cohort, TD group and SpLD group who received exam arrangements in 3rd year.

These outcomes, that reveal both an awarding gap between the SpLD and TD participants in this project, and also a University-wide awarding gap between University students with SpLD and their TD peers in timed, closed book exams, indicate that timed, written, closed book exams comprise explicitly disabling processes that disadvantage candidates with SpLD. The data analysis above suggests that these disabling processes are *intrinsic* to the environment of timed, written, closed book exams as the disadvantage persists despite the granting of exam arrangements.

The second strand of this phase of the project focused on the efficacy of study skills sessions for students with SpLD in terms of improvements in exam performance. A key aim of study skills sessions is to improve the academic performance of students with SpLD and one method of measuring academic performance is through the marks that are achieved in formal exams. A study by Shetty and Srinivasan (2014), exploring the effectiveness of study skills techniques on the academic performance of dental students, found a positive correlation between the number of study skills techniques and academic performance. Similarly, in this current research project, the data revealed that the greater the number of study skills sessions used, the greater the increase in marks, however this relationship was not statistically significant. In addition, when the change in marks achieved by the SpLD participants who accessed study skills sessions was compared with the change in mark achieved by the SpLD participants who did not access study skills sessions, no significant differences were identified. This suggests that, overall, for the participants in this research project, the study skills sessions provided made no significant contribution to boosting exam marks.

This outcome conflicts with the findings of Shetty & Srinivasan (2014), but the failure of this current research project to identify a positive relationship between study skills sessions and exam performance may be accounted for by the fact that only 32 participants accessed study skills sessions, which provided this part of the research project with a low power and limits the likelihood of detecting a 'true' effect. Nonetheless, the lack of correlation between study skills sessions and improved exam performance may suggest that the study skills sessions being delivered are not successfully providing students with the strategies that they need in an exam situation. However, an alternative explanation for this outcome is that study skills strategies can take time to become effective and embedded (Mortimore & Crozier, 2006; Saeidi et al, 2019; Wibrowski et al., 2017) and the participants in this research project only accessed study skills sessions in the second or third years their degree. The low number of study skills sessions used by those participants who took them up (only 3 study skills sessions on average were taken up by each of the 32 participants who accessed this support) may also account for the lack of any significant improvement in mark for those who used study skills sessions in comparison to those who did not use the sessions (i.e. 'too little too late'). That is to say, while the evidence from the literature in the field clearly identifies improved learning outcomes for University students with SpLD who access study skills sessions, the failure of this research project to fully corroborate these findings may arise from the insufficient number of study skills sessions used by the participants (i.e. not enough study skills sessions were accessed to make a difference to the exam marks), accessing the study skills sessions relatively late in their academic career, as well as by the low statistical power of the research project associated with the small sample size. The findings of this research project so far in relation to study skills sessions, as well as the impact of exam arrangements, on exam performance will be explored in more detail in the following, qualitative, phase of this research project through semi-structured interviews with the students themselves.

10: LITERATURE REVIEW: THE STUDENTS' VOICES

This chapter reviews the current body of research within which the qualitative phase of this research project is conceptually located, and identifies the gap that this qualitative phase of the research project aims to address: namely, how do HE students with SpLD experience exam access arrangements and why do the adjustments of 25% extra time (or the use of a word processor with 25% extra time) in exams fail to close the awarding gap? This chapter interrogates the extant research that sheds light on the barriers to learning that students with SpLD experience in Higher Education as well as the barriers that they experience in formal, timed, written, closed book exams.

10.1 Introduction

The literature review in chapter 3 revealed the existence of a significant number of quantitative studies investigating the impact of granting exam adjustments to students with SpLD. These studies employed a statistical approach to the question of whether or not exam arrangements confer equity or an advantage and suggest that exam adjustments boost the performance of students with SpLD. However, the contradictory findings of the quantitative studies, and lack of consensus about whether this boost creates a level playing field or confers an advantage, suggests that the effect of the current practice of exam adjustments is highly nuanced and complex.

Furthermore, the quantitative body of literature tends to focus on the impact of granting exam adjustments to students with SpLD from an essentially 'transactional' perspective. That is to say the studies concentrate on the effect of awarding exam adjustments on exam marks for students with SpLD while overlooking the consequences that these adjustments, and how they are acquired, have on the individual student's learning experience. In order to explore the more nuanced and complex facets of the issues suggested by the contradictory nature of the quantitative studies, and identify why or how the phenomena observed occurs, it is, therefore, critical to 'give voice' to the students' perspectives. Exploring the qualitative literature in this field aims to provide deeper insights into the impact of exam arrangements on students with SpLD, through a greater understanding of their lived experience and the interrelationship between the issues. By reflecting on the student's experiences and perspectives and integrating these with the findings of the quantitative studies, this review of the qualitative research in the field of exam adjustments aims to complement the quantitative literature review and act as a step towards 'making the world visible' (Denzin & Lincoln, 2018, p3)[132].

The qualitative literature reveals the authentic human dimension of studying at university with a diagnosis of SpLD. Participants describe how they simply aspire to be facilitated to be students like any other, to achieve their potential in their course, and be allowed an equal opportunity to succeed (Camilleri et al., 2020; Chetcuti et al, 2019; Shaw, Anderson & Grant, 2018; Yssel, Pak, and Beilke, 2016). The extent to which these aspirations can be realised pivot on the availability and nature of support and exam adjustments, as well as the inclusivity of teaching and institutional practices (Camilleri et al., 2020; Chetcuti et al, 2019; Peterson, 2016; Timmerman & Mulvihill, 2015; Yssel, Pak, and Beilke, 2016;). The qualitative studies illuminate the nuanced ways in which students with SpLD experience University exams and unveil the unintended consequences that assessments have on this cohort. High dependency on timed, closed book, written exams remains the prevailing custom and cultural norm, with limited assessment alternatives and a discernible lack of diversification (Chetcuti

[132] See Appendix 1 for full details of the literature review search strategy, including the databases searched, the inclusion and exclusion criteria and the screening rubric.

et al, 2019; Newman, 2019; Pino & Mortari, 2014; Shaw et al, 2016). These traditional assessment practices are reported to have a disproportionate impact on students with SpLD as evidenced by the enduring existence of an awarding gap and the lower completion rates of students with SpLD, as well as by the need to implement adjustments to the standard assessment conditions to make them accessible for candidates with SpLD (Camilleri et al., 2020; Chetcuti et al, 2019; Falzon, 2020; Lightfoot et al, 2018; Monagle, 2015; Pino & Mortari, 2014). At the same time, the studies reveal that learning and assessment are inextricably interwoven. That is to say, while the difficulties that students with SpLD experience are made visible in the exams, this underperformance does not occur in isolation in the assessment task (i.e. not just in the exam itself), but is the cumulative consequence of University policies, practices, structures and barriers to learning that pervade the curriculum as a whole, working in tandem with barriers within assessment practices (Camilleri et al., 2020; Chetcuti et al, 2019; Lightfoot et al, 2018; Oloffson, Taube, Ahl, 2015; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018; Zambrano, 2016).

Overall, the key themes arising from the analysis of the qualitative studies exploring the experiences of University students with SpLD using exam access arrangements fall into two main categories. Firstly, students report that, in the exams themselves, they experience the after-effect of the difficulties and barriers they encounter in accessing the course throughout the year. That is to say, the barriers that pervade their teaching and learning result in less exposure to learning throughout the year, coupled with less effective revision, with the outcome that students with SpLD are less prepared than their TD peers when they enter the exams and so are at a disadvantage (Camilleri et al., 2020; Chetcuti et al, 2019; Lightfoot et al, 2018; Newman, 2019; Pino and Mortari, 2014; Shaw, Anderson & Grant, 2018). Secondly, participants with SpLD report a range of barriers they experience in the exams themselves, due to the construct irrelevant skills that are needed to access the exams that represent specific deficits associated with SpLD, and which the exam adjustments fail to fully overcome (Camilleri et al., 2020; Chetcuti et al, 2019; Cameron, 2016; Dobson Waters and Torgerson, 2020; Pino & Mortari, 2014; Lightfoot et al., 2019; O'Byrne, et al., 2019). This paints a picture of multiple and various disadvantaging factors woven throughout the very fabric of teaching, learning, and assessment for university students with SpLD, culminating in underperformance in exams and a resultant awarding gap. This cumulative effect, described by the reviewed studies, is shown diagrammatically in Figure 10.1.

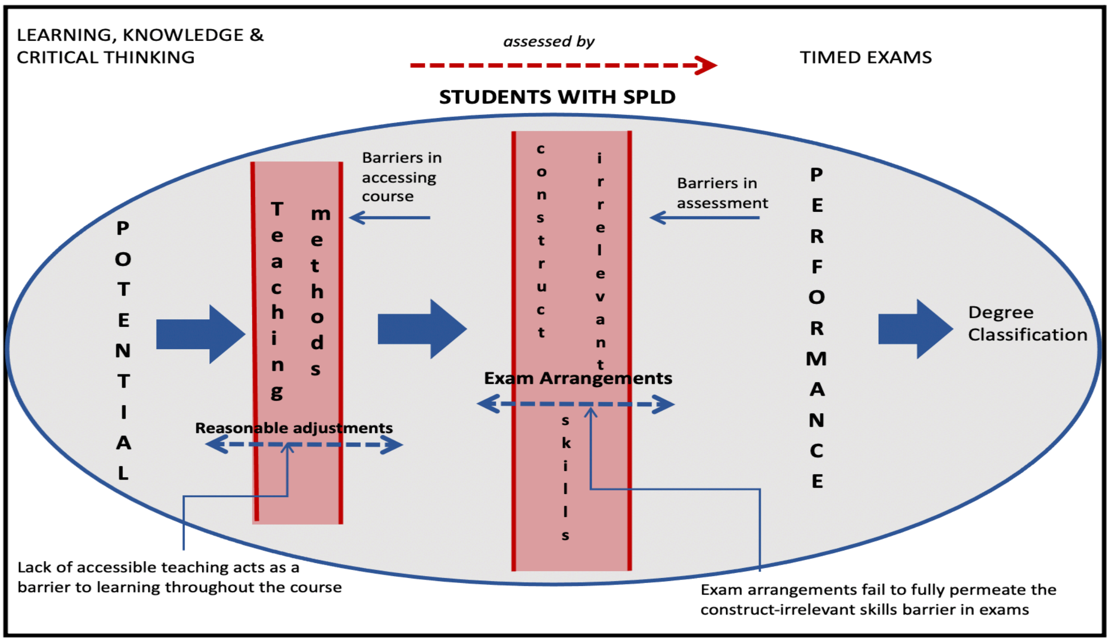


Figure 10.1: The barriers to learning and exam performance that exist for students with SpLD.

10.2 Barriers to learning

10.2.1 Course accessibility

Participants with SpLD in a number of studies identified barriers to their learning throughout their course that engendered a 'back-wash' effect on exam performance. The barriers meant that students with SpLD lacked equal opportunities to the same learning as their TD peers throughout their degree and in preparation for the exams, which negatively played out in their exam performance. Participants with SpLD in these studies described conventional teaching and learning methods as inaccessible due to being heavily dependent on literacy skills, with an emphasis on acquiring knowledge from reading texts, retaining that knowledge through making notes, and demonstrating their knowledge, understanding and critical thinking via written essays and assignments (Camilleri et al., 2020; Chetcuti et al, 2019; Falzon, 2020; MacCullagh, Bosanquet, & Badcock, 2016; Oloffson, Taube & Ahl, 2015; Pino and Mortari, 2014). As a result, the pre-eminent need for competent and fluent literacy skills acted as a significant barrier to their learning:

‘I have witnessed many of my non-dyslexic peers writing entire assignments just hours before deadlines. I would work weeks ahead, carefully articulating my thoughts and transcribing them into the essay. Even with this painstaking approach, I would usually get lower grades.’ (Shaw et al, 2016, p 2043)

Research suggests that while no differences between students with SpLD and their TD peers are identifiable in higher order thinking skills (Rodriquez-Goncalves et al, 2021), students with SpLD present with particular difficulties with the fluency of literacy skills required for the range of text analysis and production that underpins their learning in their discipline (Chetcuti et al, 2019; Falzon, 2020; Fullarton & Duquette, 2016; Marinkovich et al, 2016; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018). The relevant areas of deficit in this context include difficulties with reading speed and accuracy, the ability to identify and succinctly summarise the key idea in texts for the purposes of note-taking, spelling and grammar accuracy, writing fluency (including word retrieval, vocabulary choice, and fluency of written expression), organising and structuring ideas across an essay, and editing and proofreading skills (Camilleri et al., 2020; Chetcuti et al, 2019; Fullarton & Duquette, 2016; Kearns & Whaley, 2019; Kimel & Ahissar, 2020; Oloffson, Taube & Ahl, 2015; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018; Spear-Swerling, 2019). Competency in these aspects of literacy is critical for students to cope successfully with the demands of academic study in tertiary education, placing students with SpLD at a significant disadvantage (Camilleri et al., 2020; Chetcuti et al, 2019; Falzon, 2020; Fullarton & Duquette, 2016; Marinkovich, 2016; O’Byrne, 2019).

In addition to the difficulties encountered in showing their potential in writing, participants with SpLD in many of the reviewed studies also reported problems resulting from slow reading and the need to re-read frequently in order to absorb material (arising from working memory issues), which reduced the amount of teaching and learning material they were exposed to across the duration of the course; locating them at a disadvantage prior to entering the exam (Camilleri et al., 2020; Couzens et al, 2015; Fullarton & Duquette, 2016; Kafle, 2015; MacCullagh, Bosanquet, & Badcock, 2016; Oloffson, Taube & Ahl, 2015):

‘My weaknesses can be summarised within four domains: slow reading speed; slow writing speed; difficulty in assimilating large chunks of verbal information; and an inability to retain written information if attempting to read

> quickly. If I read something slowly I'll take it in, but it's *insanely* slow' (Shaw et al, 2016, p2044)

Students described utilising a more strategic approach to their learning as a compensatory strategy for their inability to read the same quantity as their peers, such as employing a greater degree of deep learning, alongside creative and innovate critical thinking, in order to exploit to its maximum potential the limited material that they actually read (Camilleri et al., 2020; Chetcuti et al, 2019; Kafle, 2015; Olofsson, Ahl, & Taube, 2012; Oloffson, Taube & Ahl, 2015; Pino and Mortari, 2014; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018). Nonetheless, despite the use of these strategies, students with SpLD continued to feel disadvantaged by their inability to cover the same breadth and depth of text-based learning as their peers (even with assistive technology), with the gaps in learning manifesting themselves in exam underperformance (Camilleri et al., 2020; Chetcuti et al, 2019; Kafle, 2015; Newman, 2019; Olofsson, Ahl, & Taube, 2012; Oloffson, Taube & Ahl, 2015; Pino & Mortari, 2016; Shaw et al, 2016; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018).

10.2.2 Study Skills tuition

A number of studies identified the benefit of 1:1 study skills sessions in developing compensatory strategies and offsetting the impact of weak literacy skills, difficulties with essay writing, revision, and recall on learning (Couzens et al., 2015; Dobson Waters & Torgerson, 2020; Nelson & Reynolds, 2015; Peterson, 2016; Saeidi et al, 2019; Sumner, Crane & Hill, 2020; Degirneci, Baglama, & Yucesoy, 2020). However, others noted that study skills tutors had insufficient training in, or knowledge of, the range and the use of assistive technology with the effect that, although students were familiar with the functionality of the software, their study skills tutors lacked the skills to show them how to apply the technology efficiently in the context of their studies (Peterson, 2016). In addition, the generic nature of study skills support, and a lack of sufficiently detailed understanding of the course by the support tutor, may render the sessions irrelevant, particularly in STEM disciplines (Lyman et al., 2016; Monagle, 2015; Shaw, Anderson & Grant, 2018):

> ' "dyslexia tutors" … because they did not know anything about our degree, really couldn't help ….the provision of dyslexic support equipment [isn't

> helpful] as is not medicine-specific – for example dictionary software access.' (Shaw et al., 2016, p 2048).

Many studies also noted the additional time cost of accessing 1:1 study support sessions, which inhibited students with SpLD from accessing the additional support, even though they recognised that it could be materially helpful (Lyman et al, 2016; Newlands, et al, 2015; Nelson & Reynolds, 2015; Peterson, 2016; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018).

10.2.3 Assistive Technology

In a number of studies assistive technology was generally seen as beneficial as it helped with reading, spelling, essay planning and getting ideas down on paper (Dawson et al, 2019; Horn & Huber, 2020; Peterson, 2016; Ruhkamp, 2015; Timmerman & Mulvihill, 2015; Williams, 2015), could be used to make materials more accessible (Sumner, Crane & Hill, 2020; Degirneci, Baglama, & Yucesoy, 2020) or to customise the presentation of the texts (such as changing the typeface, font size, background colour, line spacing, etc.) (Rauschenberger, Baeza-Yates & Rello 2019; Dawson et al, 2019). However, participants reported that they were generally not permitted to access assistive technology in the exam itself, and even where they were granted the use of the word processor, the grammar checker was disabled (Couzens et al., 2015; MacCullagh, Bosenquet & Badcock, 2016; Nelson & Reynolds, 2015; Newlands, et al, 2015).

10.2.4 Time cost

A common theme that emerged across the literature was the significant time cost associated with the profile of SpLD, with the result that students with SpLD spend considerably more time engaging with their learning and assessment activities than is the case for their TD peers, but with less return (Camilleri et al., 2020; Chetcuti et al, 2019; Fullarton & Duquette, 2016; Lovett and Leja, 2015; Lyman et al, 2016; MacCullagh, Bosanquet, & Badcock, 2016; Newlands, et al, 2015; O'Byrne et al, 2019; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018; Timmerman & Mulvihill, 2015; Zambrano, 2016). For example, in a series of semi-structured interviews with 13 University students with SpLD and 20 TD peers, investigating the factors underpinning differences in academic success between the two groups, MacCullagh, Bosanquet, & Badcock (2016) found that the participants with SpLD spent considerably more time reading, re-reading and overlearning material than was the case for their TD peers. Due to their slow reading speed, and the need to re-read frequently, it took the SpLD participants much longer to cover the set material and engage with independent research. They also spent more time producing notes from lectures (recording the sessions and replaying the recordings to fill in numerous gaps, while their TD peers did not report listening to the lecture again) with the result that there was less time available for the remaining academic tasks, such as reading the recommended texts:

> "I use the recorded lectures to re-listen to the face-to-face lectures… It may take me 2 hours to go through a 1-hour lecture online, or longer, depending on what's in it." (MacCullagh, Bosanquet, & Badcock, 2016, p10).

Students with SpLD also reported taking much longer than peers to complete essay writing tasks:

> 'I will mull over it for like, 14-odd hours or whatever. Whereas a normal person—they'll write an essay easily within 1 or 2 hours' (Shaw, Anderson & Grant, 2018, p 228)

While the act of spending more time engaging with learning, and overlearning, may have advantages in terms of deeper learning, the ramification for exams are that students with SpLD have less time to research topics to the same breadth or depth as their TD peers and so cover less material and are exposed to less learning, disadvantaging them prior to entering the exam room (Camilleri et al., 2020; Chetcuti et al, 2019; Fullarton & Duquette, 2016; Lyman et al, 2016; MacCullagh, Bosanquet, & Badcock, 2016; O'Byrne et al, 2019; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018).

In addition to the extra time that students with SpLD devote to their learning during their course of study, they also experience a greater time cost when engaging with exam preparation than is the case for their TD peers, due to the impact of their SpLD on revision efficacy (Camilleri et al., 2020; Chetcuti et al., 2019; Fullarton & Duquette, 2016; Lyman et al., 2016; MacCullagh, Bosanquet, & Badcock, 2016; O'Byrne, et al., 2019; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018; Timmerman & Mulvihill, 2015; Zambrano, 2016). O'Byrne, et al. (2019), for example, identified that students with SpLD adopt a highly time-consuming approach to revision in order to compensate for recall and memorisation difficulties in the exam. Participants described adopting the strategy of rote learning and committing information to memory verbatim in preparation for the exam in order to avoid losing time trying to formulate the answer from scratch in the exam itself:

> ‘I had to learn things off by heart … because in an exam I knew that I’d mess up my time if I tried to oh what about this what about that’(O’Byrne, et al, 2019, p1037).

At the same time, participants describe the highly inefficient and time-consuming act of revision, where they review material copious times, but without being able to commit it to memory:

> ‘…you’re constantly revising the same material and you don't feel like you're getting anywhere’ (Shaw, Anderson & Grant, 2018, p 228).

Not only do students with SpLD experience significant time cost when engaged with their learning, to the detriment of their breadth of study, but also the administrative burden of negotiating and applying for exam adjustments adds a further time cost that is not encountered by their TD peers (Lightfoot et al, 2018; Pinot & Mortari, 2014; Shaw, Anderson & Grant, 2018). The application procedures for requesting exam adjustments commonly involve cross-institutional communication and time consuming, administratively heavy, processes. Negotiating exam adjustments requires efficient organisation skills, the dissemination of diagnostic evidence, and the ability to engage in explanations justifying the specific adjustments being sought. These activities are time consuming and place an additional burden on those areas of specific deficit associated with the cognitive profile of SpLD, such as executive functioning, literacy skills and processing speed, with the result that these additional activities further reduce the time available for students with SpLD to engage with their learning and preparation for assessment (Cole & Cawthon, 2015; Lightfoot et al, 2018).

10.2.5 Othering

The sense of being different to others, less able, less suitable for academic environments than their peers, and an ‘imposter’, was described by participants with SpLD in a number of studies, due to the perceived deficiencies associated with having a profile of SpLD (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al., 2019; Lightfoot, 2018; Monagle, 2015; Majer, 2018; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). This tendency for students with SpLD to compare themselves with their peers (in order to understand where they are located in the hierarchy of academic worth) and finding themselves wanting is exacerbated by the processes of exam arrangements, which are seen as ‘outing’ and so further add to the feeling of being different and ‘lesser’ (Chetcuti et al., 2019; Cameron, 2016; Lyman et al, 2016; Lightfoot, 2018; Monagle, 2015). Participants described how taking exams separately to others made their SpLD visible and placed them in a position where they had to explain their disability to their peers:

> ‘you feel like a bit of an outcast being like, oh well yeah, I didn’t see you in the exam, oh well yeah, I was in my own special room doing it with a scribe. That can be embarrassing’ (Cameron, 2016, p235)

This vignette highlights how the sense of being a ‘misfit’, who is conspicuously set apart, is brought sharply into focus for students with SpLD by the physical location of a separate exam venue and the provision of exam conditions that are different to those received by their TD peers. This insidiously feeds their perception of being less academically capable than their peers (Cameron, 2016; Camilleri et al., 2020; Lyman et al, 2016; Lightfoot, 2018 Monagle, 2015). Given that a systematic review conducted by Lightfoot (2018) of 37 qualitative studies involving students with SpLD revealed that positive self-perception and a confident outlook on their own learning were crucial to the academic success of students with SpLD, the contribution that separate exam conditions makes to fostering this pre-existing narrative of the ‘misfit’ (who is less academically able), adds yet another layer of experiences that act to inadvertently undermine academic self-efficacy and, thus, the academic performance of students with SpLD (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al, 2019; Lyman et al, 2016; Lightfoot, 2018 Monagle, 2015):

> ‘I hated being separated from the rest of the class…. isolating me from the rest of the class changed the way I looked at myself. It also affected my social interactions with [peers]. The embarrassment these experiences caused are something I can never forget and have led me to constantly struggle with my self-esteem’ (Chetcuti et al, 2019, p18)

Not only did students with SpLD report feeling othered and ‘outed’ by the awarding of the exam adjustments (as their ‘deficiencies’ were made visible by sitting exams in venues apart from their peers), but they also felt uncomfortably exposed by the need to disclose their diagnosis to academic and administrative staff in order to request course based adjustments and exam arrangements in the first place (Cole & Cawthon, 2015; Lightfoot, 2018; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). Participants described feeling ambivalent about this need to actively disclose a diagnosis of SpLD in order to be able to access their education. They felt that the diagnosis was perceived negatively and stigmatised them, but that they were forced into a situation where they needed to explicitly identify with the diagnosis in order to be able to maximize their educational opportunities (Cole & Cawthon, 2015; Lightfoot, 2018). Some felt that the negative associations of SpLD outweighed any benefits (Camilleri et al., 2020; Cole & Cawthon, 2015; Lightfoot, 2018), while others felt that the support and adjustments that they received as a result of the diagnosis allowed them to perform more closely to their potential and so was worth the negative repercussions of the diagnosis (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al., 2019; Cole & Cawthon, 2015; MacCullagh, Bosanquet, & Badcock, 2016; Shaw, et al 2016; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018).

10.2.6 Perception of advantage

Most of the reviewed studies found that exam adjustments were commonly perceived by academic staff and TD peers as advantaging students with SpLD and this further added to the sense of being lesser and othered described by participants with SpLD (Byrne, 2018; Cameron, 2016; Camilleri et al., 2020; Chetcuti et al, 2019; Lightfoot, 2018; Lyman et al, 2016; Shaw et al, 2016; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). The perception of students with SpLD themselves who use exam access arrangements is that they do not confer an advantage, but are implemented to create equity by offsetting the disadvantage that candidates with SpLD experience due to the impact of the deficits associated with SpLD (Lyman et al, 2016; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018; Zambrano, 2016):

> ‘If I need to wear glasses I am allowed to wear them in the exam. This does not give me an advantage over someone who does not need glasses. Similarly, for my dyslexia, if you give me…some more time you are not giving me something extra. You are just giving me what I need to start off at par with others who do not have dyslexia’ Camilleri et al., 2020, p18)

However, this view was not reflected in their experiences of the attitudes of others (both staff and peers) who expressed concerns that the exam adjustments potentially advantaged students with SpLD (O'Byrne, 2018; Camilleri et al., 2020; Chetcuti et al., 2019; Pinot and Mortari, 2014; Lyman et al, 2016) and suggested it was 'unfair that… just because I am having a struggle in my life that I should get something that other people don't get" (Lyman et al, 2016, p129).

10.3 Barriers in exams

The literature above reveals that students with SpLD encounter a range of barriers in accessing teaching and learning that result in less exposure to learning throughout the year and so they are less prepared than their TD peers when they enter the exams. In addition to the impact on exam performance of inaccessible teaching and learning practices, the literature also suggests that students with SpLD experience barriers in the exam itself which are injurious to their exam performance, adding a further disadvantage and fostering an awarding gap.

10.3.1 Generic nature of exam arrangements

Participants in the reviewed studies felt that the range of exam adjustments available were not sufficient for the range and context of the exams they encountered (Lightfoot et al, 2018; Rogers et al 2019; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). In particular, it was felt that there was a limited 'set menu' of exam adjustments available to a student with SpLD (namely; extra time, use of a reader/amanuensis and the use of a word processor) and these were not always the most appropriate adjustments or applicable in all exam contexts. For example, in practical medical clinical exams (OSCEs) where candidates are required to respond to a 'real-life' case scenario via a timed station, candidates with SpLD reported that the ability to write down the conversation, rather than having extra time, would have been a helpful adjustment, but this was not an available option:

> 'I just don't remember what the patient tells me…. And, I asked for a pen and paper before the station. But they said I wasn't allowed it….. They just said no. So I had to accept that' (Shaw, Anderson & Grant, 2018, p229).

Nonetheless, participants reported that exam adjustments were critical to their academic success and that without these accommodations they were likely to underperform in the exam to an even greater extent (Fullarton & Duquette, 2016; Lightfoot et al, 2018;

Lyman et al, 2016; MacCullagh, Bosanquet, & Badcock, 2016; Shaw et al., 2016; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018).

10.3.2 Anxiety

Camilleri et al. (2020), in a qualitative case study into the examination experiences of eight students with SpLD, argue that while exams are stressful experiences for any candidate, due to the high-stakes nature of the activity, those with SpLD are disproportionately affected as they 'enter the examinations race at a disadvantage' and this triggers particularly high levels of anxiety (Camilleri et al., 2020, p16). High incidences of anxiety among students with SpLD, that serve to impede candidates' ability to perform to their potential in exams, are also described in numerous other studies (Dobson Waters & Torgerson, 2020; Chetcuti et al, 2019; O'Byrne, et al, 2019; Livingston, Siegel, & Ribary, 2018; Stagg, Eaton & Sjoblom, 2018; Boyes, et al, 2016). Nelson, Lindstrom & Foels (2015) for example, in a study of 50 University students with SpLD and 50 TD peers, found that clinically significant anxiety was 5 times more prevalent in the participants with SpLD than their TD peers. They argue that, while for TD participants some exam anxiety was desirable to promote exam performance, the extremely high levels of anxiety experienced by participants with SpLD significantly inhibited their exam performance and so acted as a 'construct-irrelevant factor that reduce[d] the validity of the test scores' (Nelson, Lindstrom & Foels, 2015, p 422):

> 'Examinations make me very anxious …When I am preparing for an examination … I cry because I think that I'm not going to pass. I don't sleep at night and wake up thinking that I will blank out during the examinations. Every time I think about examinations I get the shivers…I know that I am intelligent and creative, but somehow this never shows up in my examination results' (Chetcuti et al., 2019, p 38)

For university students with SpLD, anxiety, it is argued, is fuelled by low academic self-esteem and a past history of underperformance in exams (due to educational assessment tasks that rely heavily on automatic literacy skills, efficient working memory and robust executive functioning) as well as the 'back-wash' effect of difficulties accessing the course. It is further inflamed by self-doubt about their academic ability in a culture that conflates literacy skills with intelligence and thus results in anxiety-based thought processes focusing on the likelihood of failure (Camilleri et al., 2020; Cameron, 2016; Chetcuti et al, 2019; Fullarton & Duquette, 2016; O'Byrne, et al, 2019). This concatenation of events, where previous underperformances reinforce anxiety, serves to undermine future exam performance and detrimentally impacts exam achievement (Camilleri et al., 2020; Cameron, 2016; Fullarton & Duquette, 2016; Kafle, 2015; Nelson, Lindstrom & Foels, 2015; O'Byrne, et al, 2019; Stagg, Eaton & Sjoblom, 2018).

10.3.3 Construct irrelevant skills

The reviewed studies also reveal that students with SpLD experience a range of challenges associated with construct-irrelevant skills when undertaking written, time-bounded, closed book exams. These difficulties include insufficient time, problems with memorisation and recall of information, inadequate literacy skills, word retrieval deficits, and a lack of fluent and concise written expression (Camilleri et al., 2020; Chetcuti et al, 2019; MacCullagh, et al., 2016; Ofiesh, Moniz, & Bisagno, 2015; Pino & Mortari, 2014; Shaw et al, 2016; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018). Participants also commonly reported misreading the question in the exam and so losing significant marks:

> [in the exam] 'I lost 20 marks simply because I had misread a question. The question should have read 'not using this theory' and I had read it 'using this theory'. What I had written down was perfect, but it didn't answer this question' (Chetcuti et al, 2019, p 49)

In exams where essay type answers act as a vehicle for presenting the candidates understanding, participants with SpLD reported having significant problems showing their true ability due to difficulties with essay planning, organisation, getting ideas down on paper, structuring those ideas and with written expression, word finding, grammar, punctuation and spelling. In addition, participants reported significant problems deconstructing the question and identifying the expectations of the essay, all of which affected the quality of the response (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al, 2019; Lightfoot et al., 2019; O'Byrne, et al., 2019; Pino & Mortari, 2014):

> 'I don't understand what the examiners expect of me. It's not that I wouldn't know the answer; it's just that I wouldn't have understood what was being asked. To make matters worse, sometimes the questions are so long and winding that it takes me forever to read them through. This affects my understanding' (Chetcuti et al, 2019, p 47)

This was described as less of a problem with coursework-based assignments, with the result that clear discrepancies between achievement in coursework and exams were reported, with the coursework more closely reflecting the students' potential (Camilleri et al., 2020; Chetcuti et al, 2019; MacCullagh, et al., 2016; Pino & Mortari, 2014; Shaw et al, 2016; Shaw and Anderson, 2017; Shaw, Anderson & Grant, 2018):

> 'The overall degree he got (an upper second) was not a reflection of his abilities at all as I came to know him. In his exams he was getting a 2.2 and all of his coursework was first class. So even the fact that he had extra time, I don't think that was adequate compensation. So I felt really strongly after that – that here we are assessing a student within a system of assessment that is obviously not giving him a fair deal because he can't really demonstrate what he knows and what he is capable of.' (Riddell & Weedon, 2006, p66)

Participants in the reviewed studies reported that these difficulties were not adequately addressed or offset by the awarding of extra time with the result that they felt significantly disadvantaged by the policy of using written, time-bounded, exams as the prime assessment modality, as the construct-irrelevant barriers in the exams prevented them from genuinely demonstrating their subject knowledge, understanding, and critical thinking skills. They felt that the emphasis on time constrained, closed book, written exams privileged memorization and recitation over understanding, even though critical thinking, deep understanding and the ability to apply their subject knowledge aligned more closely with the stated purpose of higher education (Camilleri et al., 2020; Chetcuti et al, 2019; Pinot & Mortari, 2014; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). As a result of these perceived disadvantages, the participants with SpLD reported a strong aversion to this assessment modality, which, they felt, failed to equitably assess their knowledge and skills as it inherently discriminated against candidates with SpLD (Camilleri et al., 2020; Chetcuti et al, 2019; MacCullagh et al., 2016; Newlands et al., 2015; Pino & Mortari, 2014; Shaw et al, 2016; Shaw, Anderson & Grant, 2018; Shrewsbury, 2016):

> 'Even with all of the support in place, I still do not feel able to compete with 'normal' students within the current assessment methods. I have excelled on many occasions where we were assessed practically or verbally. However…these assessments have all been formative, therefore not graded…Assessments that contribute towards my grades are still mainly examinations based on reading and writing. So, however much I have shone in the working environment, I continue to feel that I let myself down with assessment grades' (Shaw et al, 2016, p2049 - 50).

Although the use of closed book, timed, written exams was almost universally disliked by students with SpLD, this antipathy was more evident in exams requiring essay style responses than was the case for exam formats that inherently provided scaffolding and emphasised the application of knowledge over memorisation or recitation (Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018, Shaw et al, 2016). For example, participants in the interpretive phenomenological study of the experiences of medical students with dyslexia by Shaw, Anderson & Grant (2018) identified that many of the exam formats used in Medicine assessments, particularly multiple-choice exams, were less discriminatory against candidates with SpLD. These were seen as testing applied knowledge, rather than the ability to structure an answer or recall abstract information. Thus, students with SpLD perceived the format of medical exams to be more accessible to candidates with SpLD than essay format exams and medical students in the reported studies felt less disadvantaged in exams than SpLD participants who sat exams requiring essay style responses:

> '… They were MCQ[133] questions, true or false questions and single best answer questions. So, maybe that's, possibly that's why I never experienced problems throughout medical school—because I never really had to write anything' (Shaw, Anderson & Grant, 2018, p228).

10.4 Conclusion

This review of the qualitative literature reveals a picture of multiple systemic barriers that coalesce to disadvantage students with SpLD by permeating all aspects of their teaching and learning experience and culminate in exam underperformance and an awarding gap. Students with SpLD report spending considerably more time engaging with their learning and assessment activities than is the case for their TD peers, leaving insufficient time to cover the same material to the same breadth or depth as their TD peers. At the same time, they report barriers to learning arising from inaccessible teaching and course materials. Key activities demanded by a degree course revolve around reading textbooks and journal articles, making notes from reading, taking notes in lectures, and writing essays and assignments. These conventional teaching and

[133] Multiple Choice Questions

learning methods rely heavily on fluent and effective literacy skills, efficient processing speed and working memory and robust executive functioning – all of which represent construct irrelevant skills that act as a barrier to academic performance for students with SpLD due to the nature of their cognitive profile.

The barriers identified above are woven into the very fabric of the teaching and learning of students with SpLD and disadvantage them throughout the academic year, having a 'back-wash' effect in the exams by combining to hinder their preparedness for the exam. In addition to this, students with SpLD then experience further barriers in the exams themselves that impede them from performing to their potential in the exam moment. These barriers include the exam's emphasis on construct-irrelevant skills needed to access the exam (such as accurate and efficient literacy skills, memorisation and recall, and processing speed) as well as the impact of low academic self-esteem and high exam anxiety, fueled by poor past exam performance, which impedes exam success for students with SpLD. The menu of exam adjustments available, such as extra time, were reported as helpful but not varied enough or sufficient to offset the full impact of the range of disadvantages a student with SpLD experiences in the exam situation. The impact of these multiple and various disadvantaging factors are cumulative and combine to trigger exam underperformance for students with SpLD, while not impacting their TD peers, thereby resulting in an awarding gap.

It should be pointed out that the extant studies investigating the lived experiences of university students with SpLD do not present an entirely negative picture. Participants in the studies also describe positive experiences of support, including individual supportive and empathetic tutors who foster self-confidence and success, and the benefit of the deeper engagement with learning activities that is a consequence of the need to overlearn and re-read material (MacCullagh, 2014; MacCullagh et al., 2017). These positive outcomes reportedly co-occur alongside the negative ones. However, despite the positive experiences reported, the emphasis on closed book, time bounded, written, high-stakes exams was universally described in negative terms by the participants with SpLD in the reviewed studies, with participants feeling disadvantaged in the preparation for the exam, both in term of difficulties accessing the course, as well as in the exam proper.

In conclusion, the experiences described by the participants in these studies highlight the structural lack of inclusion inherent in HEIs, where disabling experiences arise from the barriers in the environment, undermine the educational experience and academic self-efficacy of students with SpLD, and exacerbate the impact of the cognitive difficulties of students with SpLD on their academic performance. That is to say, the socially produced disadvantages of having SpLD in University work in tandem with the inherent cognitive profile of SpLD to negatively impact their learning experience and self-efficacy (and, thus, academic performance) (Cameron, 2016; Lyman et al, 2016; Lightfoot, 2018: Monagle, 2015). Thus, future research is warranted into the ways in which the assessment of students in HE could be made more equitable for those with SpLD, to ensure that they are tested in a way that exposes what they know, rather than exposes their disability.

11: QUALITATIVE RESEARCH METHODOLOGY

This chapter presents the research methodology that was applied to the qualitative phase of this project. The chapter commences with an overview and explanation of the project's decision making in relation to the qualitative methodology applied and how this meshes with the project as a whole. The chapter then elaborates on the research strategy, including how the semi structured interview questions were developed and implemented, as well as details of the participants in this phase of the project, the procedure and the data analysis techniques used.

11.1 Introduction

While the first two, quantitative, phases of the research project involved empirical testing of the hypotheses, this third, qualitative, element of the research project used a systematic, inductive approach to understanding the real-life experiences of student with SpLD who use exam access arrangements. This qualitative element is located within the framework of grounded theory and used an interpretative data collection approach (Charmaz & Thornberg, 2020). Corbin & Strauss posit that the purpose of grounded theory is to 'construct theory grounded in the data' (Corbin & Strauss, 2015, p3). The purpose of this qualitative phase of the project was to explore the experiences of students with SpLD who use exam arrangements and, by following the data that emerged, to construct theory (rather than just a description) around how the experience of the exam environment (including the experience of using exam arrangements) influenced the exam behaviour and performance of students with SpLD. In other words, this qualitative phase aimed to build new theories directly from the data collected from the participants with SpLD, using inductive analysis, rather than using pre-existing theoretical frameworks (Corbin & Strauss, 2015).

11.2 Methodology

This phase of the research project used qualitative data drawn from semi-structured interviews[134] conducted in 2020 with 10 students with SpLD who received exam arrangements in the same subject areas as the participants in phase 1 of the research project (namely, English, History, Law, Maths and Medicine). Participants were asked to describe their experiences of using exam arrangements in order to obtain the lived experience of receiving and using exam arrangements with the aim that the ensuing data may provide insights into the barriers to exam performance that are both overcome by exam arrangements, and also the barriers that remain despite the granting of exam access arrangements. Thus, the qualitative phase of the research project aimed to enhance the understanding of the outcomes of the first two, quantitative, phases of the research project by exploring the reasons for the awarding gap, building a basis for triangulation through a conversation with the quantitative data from the first two phases of the project as well as from the literature in the field, and acting as a significant step towards the identification of interventions that would promote future good practice.

Importantly, the quantitative data from the first two phases of the research project have explored the hypothesis that exam arrangements either advantage students with SpLD or promote equity and have identified that exam adjustments neither confer an advantage, nor create a level playing field in exams with their TD peers. Instead, the quantitative findings suggest that an awarding gap between students with SpLD and their TD peers persists in exams, despite the implementation of exam adjustments, albeit these adjustments do help to narrow that gap. This new theoretical position (i.e. students with SpLD underperform in exams, despite the granting of exam adjustments) raises new questions, such as 'what disability related disadvantages do exam adjustments help overcome and what barriers to exam performance do they fail to help with'? It is in addressing these new questions that the qualitative data is solicited and applied. Grounded theory, used as a framework to collect and interpret the qualitative

[134] See Appendix 7 for details of the semi-structured interview questions

data through purposive sampling and inductive analysis, both illuminates the key findings of the quantitative data and generates new theory. The themes identified foster a deeper understanding of the effects that exam arrangements have on students with SpLD who use them by uncovering the ways in which exam arrangements help to overcome barriers in exams that students with SpLD experience and also what barriers to exam performance are not overcome by the exam arrangements. This assists in developing an explanation of how and why the awarding gap persists despite the implementation of exam adjustments.

11.2.2 Research strategy

A predominantly inductive, qualitative research approach was adopted in this third phase of the research project, through a series of semi-structured interviews, using questions shaped by the findings of the first two phases of the research project. The interviews were deemed *semi-structured* as, although they were structured around a series of pre-formulated questions in a set order, the questions were open-ended and provided opportunities for the participants to share additional information by going into more detail where they felt appropriate, to steer the conversation towards experiences they wished to discuss that were, perhaps, tangential to the set questions, and to respond in a way that reflected their own contextualisation of their experiences. Thus, the semi-structured interview approach enabled the participants to share data beyond the reach of the trigger questions as well as allowing the researcher to probe further with additional prompts or supplementary questions as necessary (Drever, 2003). This approach provided the potential to increase the richness and authenticity of the data gathered (Bryman, 2015; Drever, 2003; Kvale, 1996; Rabionet, 2011).

Although this project came to the qualitative phase inductively (applying grounded theory principles of using intensive coding to follow the data and allowing concepts to emerge and generate theory) in practice, this project adopted a pragmatic version, rather than a purist version, of grounded theory (Charmaz, 2014; Morgan, 2020). In order to be fully inductive, the process of developing the research questions should occur when the data is being collected, such that the questions are developed and refined reflexively during the data collection phase. That is to say, the analysis of the initial data should inform the collection of subsequent data, and so on in a continuous cycle, until a well-integrated, substantiated theory can be constructed (Agee, 2008; Corbin & Strausss, 2015). However, in contrast to these principles of grounded theory, this project did not develop ongoing iterations of questions that 'changed during the process of research to reflect an increased understanding of the problem' (Creswell & Creswell, 2018), or in response to the researcher's reflections on their own perspectives and position in the data interpretation process (Corbin & Strausss, 2015). Instead, in order to develop knowledge that can be directly applied to assessment practices in Universities (i.e. a pragmatic imperative), the research questions were developed prior to the data collection phase as a result of the outcomes of the quantitative phase of the project, with the quantitative findings determining the primary focus and parameters of the project (that focus being the factors driving the awarding gap experienced by students with SpLD in exams). Nonetheless, by being open ended, the questions used in this project were designed to be broad, flexible, and sufficiently 'discovery oriented' to elicit unexpected data, enable participants to take the discussion in their direction of choice and lead to new theories (Creswell & Creswell, 2018), albeit within the pre-set parameters of the topic of interest (the awarding gap).

11.2.3 Semi-Structured Interview Schedule

A collaborative approach to the development of the questions for the semi-structured interviews was adopted. This emanates from a belief in a methodological approach that is defined by the active involvement of the stakeholders, with the aim of collectively identifying solutions to the problems, and also speaks to the broader responsibility of the researcher to include elements of participation in the research process, where possible, as opposed to conducting research 'on' or 'about' the participants (Walker & Loots 2017, p.168). In order to foster a collaborative relationship with the population under study, the questions in the semi-structured interview schedule were developed in concert with 2 sabbatical officers from the University's Student Union, who were undergraduate students between 2016-19 and who were granted exam arrangements (25% extra time and use of word processor) in their exams on account of a diagnosis of SpLD.

These sabbatical officers[135] assisted in developing the questions for the semi-structured interview schedule, such that the questions reflect the key topic areas that they, as the population being studied, wished to be probed further in the light of the research project's findings so far. The sabbatical officers had already undertaken substantive work related to reducing the awarding gap under the umbrella of the current Access and Participation Plan, which obligates the University to reduce the awarding gap of disabled students by 2025. As such, these sabbatical officers were already conversant with the issue of underachievement in exams experienced by students with SpLD and, in their role as active officers of the Student's Union, were committed to the work of identifying the factors driving the awarding gap and effective interventions. Thus, a co-design approach to the qualitative phase of this project, in terms of question creating and framing, was deemed an appropriate approach to ensure collaboration between the researcher and research population and the opportunity for the population under study to influence outcomes and strategy.

The findings of the first two phases of this research project were used as a frame of reference in the design of the interview questions for this qualitative phase of the project, with the broad areas for discussion having been generated from the outcomes of the research findings so far alongside the project's research objectives. By exploring these broad topics through the development of semi-structured interview questions, this qualitative phase of the research project aimed to elucidate and provide a deeper understanding of the factors that influence exam performance for students with SpLD and has the potential to highlight hitherto unidentified, yet salient, issues (such as student' priorities, key concerns and preferred approaches to interventions) (Cohen, Manion & Morrison, 2018, (Bryman, 2015; Drever, 2003). The findings of the first two phases of the research project were shared with the two sabbatical officers and, through a group discussion, a series of relevant questions were mapped onto the key research objectives. Thus, questions were created that aimed to generate deeper insights into the impact of exam access arrangements for students with SpLD, particularly in relation to

[135] The participating sabbatical officers, who are receiving a full-time salary for their role representing University students, were not paid or otherwise incentivised for their participation in this project.

their perceptions of the purpose of the exam arrangements, whether or not the adjustments meet this purpose, the particular barriers that exist in exams for students with SpLD and what may be potential solutions to overcoming these barriers. Two questions exploring the impact of study skills sessions were also included in the interview questions, in light of the ambiguous findings of the second phase of the research project relating to the efficacy of study skills tuition.

The semi-structured interview schedule was piloted with 3 participants initially and then reviewed to ensure that the questions were appropriate and relevant as well as adequate for eliciting useful data (Cohen, Manion & Morrison, 2018). The pilot interviewees did not ask for clarification during the interview, or respond in such a way as to suggest that the questions needed refining and so no revisions to the questions were made in light of the results of the pilot. All of the 10 semi-structured interviews were conducted as 1:1, face-to-face (online), interviews and by the same interviewer (this research project's researcher) in order to minimize differences that may arise from different interviewers asking the questions or probing differently, which can result in differences in the amount of detail elicited by each interview and variations in the interpretation of the questions by the participant (Cohen, Manion & Morrison, 2018, (Bryman, 2015; Drever, 2003).

The findings of the first two phases of the research project were not shared with the participants who were interviewed in this third phase of the research project for ethical reasons. Basic tenets of research, according to Oliver & Barnes (2012), require the researcher to ensure that the participants being studied are not disempowered by the research itself. The first two phases of the research project identified that, despite the exam access arrangements granted to students with SpLD, an awarding gap remains. In addition, the impact of 1:1 study skills tuition is ambiguous. Sharing this information with students who would be sitting future exams under these conditions, and who may be accessing study skills sessions could, potentially, have a negative psychological impact, including affecting self-esteem, motivation and engagement with support. In addition, revealing the outcomes of the first two phases of the research project may inadvertently influence the responses of the participants during the interviews. In order to avoid these risks, the outcomes of the quantitative phase of the research project were not divulged to the participants, however, the participants will have access to the outcomes of this phase of the research project following completion.

11.2.4 Population and sampling

As the first phase of this research project comprised participants who were undergraduate students studying English, History, Law, Maths and Medicine, the qualitative phase of the research project involved 10 participants with a diagnosis of SpLD who were recruited from the same Faculties. While in the first quantitative phase of the research project it was not possible to determine the subcategories of SpLD that the participants identified with (as the data was supplied anonymised), in the second quantitative phase of the research project 41% of the participants had a diagnosis of dyslexia, 45% had a diagnosis of dyspraxia and 14% had a diagnosis of both dyslexia and dyspraxia. Reflecting these participant attributes, the qualitative phase of the research project recruited 4 participants (40%) with a diagnosis of dyslexia, 4 participants (40%) with a diagnosis of dyspraxia, and 2 participants (20%) with a diagnosis of both dyslexia and dyspraxia[136].

In order to further reflect the attributes of the quantitative phase of the research project, the qualitative phase of the research project recruited 5 participants who were diagnosed with SpLD prior to starting at university and 5 who were diagnosed part-way through their course (and so sat some exams under standard conditions and some exams with exam adjustments). In addition, in the first quantitative phase of the research project, 66% of the participants were granted 25% extra time and 34% were granted the use of a word processor with 25% extra time. In the second quantitative phase of the research project, 64% of the participants were granted 25% extra time and 36% were granted the use of a word processor with 25% extra time. To reflect these attributes, in the qualitative phase of the research project, 70% of the participants recruited had handwritten their exams with extra time (N=7) and 30% used a word-processor in addition to extra time (N=3). Finally, in the second quantitative phase of the research project, 40% of the participants engaged with specialist study skills tuition, and so four

[136] 50% of the sample group were male participants and 50% female.

of the participants (40%) selected to participate in the qualitative phase of the research project were chosen as they had received study skills sessions. A comparison of the participant attributes across all 3 phases of the research project are shown in Table 11.1.

Table 11.1: A comparison of the participant attributes across all 3 phases of the research project

Participant attributes	Phase 1 (quantitative) N=357	Phase 2 (quantitative) N=80	Phase 3 (qualitative) N=10
Dyslexia	% unknown	41%	40%
Dyspraxia	% unknown	45%	40%
Dyslexia + dyspraxia	% unknown	14%	20%
Extra time only	66%	64%	70%
Extra time + word processor	34%	36%	30%
Male	% unknown	47.5%	50%
Female	% unknown	52.5%	50%
% accessed study skills tuition	% unknown	40%	40%

This variation of participant characteristics aimed to broadly represent the array of characteristics of the group in the previous samples, such that the range of perceptions and experiences of the different groups included in both quantitative phases of this research project is reflected. The participants were recruited from the research site University's Disability Database, which holds information on all students with SpLD (data is held for over 1500 current students with SpLD who would be eligible to participate in this research project). Eligible students[137] were emailed with details of the research project to ask if they would be interested in participating. In recognition of the value of the contribution that the participants made to this research project and their time, all participants were offered a £10 supermarket, bookstore or multi-retailer gift voucher (selected by participant from the list of options).

[137] 10 eligible students were contacted initially and all 10 responded agreeing to participant.

11.2.5 Procedure

In total, 10 participants (5 male, 5 female) from 5 departments in the University (English, History, Law, Maths and Medicine) took part in the research project [138] (see Table 2 for detailed participant information). All of the participants reported either single (N=8) or dual (N=2) diagnoses. The participants' characteristics are shown in Table 11.2

Table 11.2: Participant characteristics

Department	Name (pseudonym)	Gender	SpLD diagnosis	Exam arrangements received	Accessed study skills?
English	James	M	dyspraxia	25% ET	Y
English	Bella	F	dyslexia + dyspraxia	25% ET & WP	N
Law	Alex	M	dyslexia	25% ET & WP	N
Law	Anna	F	dyspraxia	25% ET	N
History	Daniel	M	dyspraxia	25% ET & WP	N
History	Juliet	F	dyslexia	25% ET	Y
Maths	Graham	M	dyslexia + dyspraxia	25% ET	Y
Maths	Laura	F	dyslexia	25% ET	N
Medicine	Robert	M	dyspraxia	25% ET	Y
Medicine	Joanna	F	dyslexia	25% ET	N

All participants were white/British, and the mean age of participants was 20 years (SD=.67). Participants focused on their experiences as university students, but they also offered insights into their experiences in compulsory education. The number and characteristics of the selected participants aimed to provide a range of insights that

[138] Kvale & Brinkmann (2009) suggest that 15 ± 10 qualitative interviews should be sufficient to reach saturation point in qualitative interview design.

illuminated the research objectives until saturation point has been achieved. After 8 interviews had been conducted no new insights were being generated. Two further interviews were conducted to ensure that saturation and redundancy had been reached (making 10 interviews in total) and, as no additional insights were elicited, no further interviews were therefore deemed necessary (Creswell & Clark, 2017).

The interviews were carried out remotely (in accordance with public health guidance on physical distancing[139]) via Microsoft Teams, in rooms where only the interviewer and the interviewee were present and were recorded using the Teams recording facility. The interviews were transcribed verbatim with light editing of linguistic 'fillers'. The length of the interviews ranged from 60 min to 90 mins (M = 77), amounting to 13 hours and 37 minutes of interview time in total.

11.2.6 Data Analysis techniques

The salient characteristics of the data were analysed and systematically coded via thematic analysis techniques using NVivo12.0 qualitative research software package (QSR International, 2018). Thematic analysis offers a strategic way of organising, analysing and interpreting qualitative data according to 'patterns' (themes) that both respond to the research questions and reflect the content of the data (Braun & Clarke, 2006) and enables researchers to critically examine the dialectical relationship between the research questions and what the data presents in order to acquire an in-depth appreciation of the complexity of the data (Strivastava & Hopwood, 2009). A thematic analysis strategy can also offer opportunities to highlight similarities across data, as well as differences, thereby enabling diversity in the data set to be managed and unexpected insights to be highlighted (Braun & Clarke, 2006)

Firstly, to identify the patterns of topics in the data, each transcript was read and re-read thoroughly, and initial impressions were noted down. Then, each line of the transcript was labelled with a short phrase to sum up the meaning as initial, preliminary

[139] The interview stage of this research project coincided with the first national 'lock-down' occasioned by the COVID-19 pandemic (March 2020). As a result, in-person interviews were not permitted by public health guidelines and so these interviews were conducted remotely via the video call facility of MS Teams. A revised ethics application was submitted in relation to this alteration in the interviewing process and interviewing was delayed until the revised ethics proposal was agreed.

coding. In order to then shift this descriptive analysis approach to a more conceptual analysis, the codes[140] were surveyed for commonalities and grouped into initial broad themes[141] and categories. This was repeated for each transcript and the data set as a whole were compared. The themes were then reviewed against the coded extracts, as well as the data set as a whole, creating a thematic 'map' of the patterns identified in the data. The specifics of each theme were refined by reflecting on the commonalities between the themes (i.e., where they converged and where they diverged), and the overall narrative the analysis produced, creating definitions and labels for each theme. The themes identified were repeatedly referred back to the transcripts and the data re-examined for any overlapping concepts or informational redundancy and to re-confirm or refute them (Braun & Clarke, 2006; Creswell, 2015). This aimed to ensure that the data grouped within each theme was internally consistent and so formed a coherent theme and, at the same time, each theme was individually distinctive and different to the others (Braun & Clarke, 2006). Once saturation point was achieved, and no new information arose from the data, no further themes were developed. As part of the final analysis, illustrative examples were selected and linked back to the research question and literature. The transcripts were further reviewed by repeatedly re-playing the recorded interviews and focusing on the content furnished by the participants to ensure that the quotations selected accurately represented the identified themes. It should be noted that similarities between the experiences of students with dyslexia and dyspraxia (and those with dual diagnoses) were evident across all the themes, rather than being specific to each SpLD sub-group.

The theoretical standpoint of the qualitative phase of the research project is that an inequality in exam performance exists between students with SpLD and their TD peers arising from barriers in the exam environment that specifically disadvantage individuals with SpLD. It was necessary for the researcher to be explicitly conscious of this theoretical position when interpreting the data to refrain, insofar as possible, from interpreting the participants' responses from the point of view of this theoretical standpoint, rather than that of the participants. In addition, as a disability adviser, the

[140] Creswell (2015) defines codes as 'labels used to describe a segment of text' (Creswell, 2015, p. 243).
[141] Creswell (2015) defines themes as 'similar codes aggregated together to form a major idea in the database' (Creswell, 2015, p. 244).

researcher's professional experience is characterised by anecdotal reports of the barriers experienced by students with SpLD in their academic study. This personal experience inevitably formed a filter through which the data was examined and it was therefore necessary to also maintain a critical awareness of this potential bias, which may cause the researcher to assign greater weight to negative participant reports than to positive ones when interpreting the data (i.e. privileging data that fits in with the researcher's own belief system over data that challenges those beliefs). While it is not possible to fully eradicate these biases, viewpoints and filters, by explicitly maintaining a critical awareness of the potential impact of the theoretical standpoint of the research alongside the researcher's biases, the researcher aimed to minimize the effect of the researcher on the interpretation of the data and protect the rigour of this phase of the research project (Johnson et al, 2020).

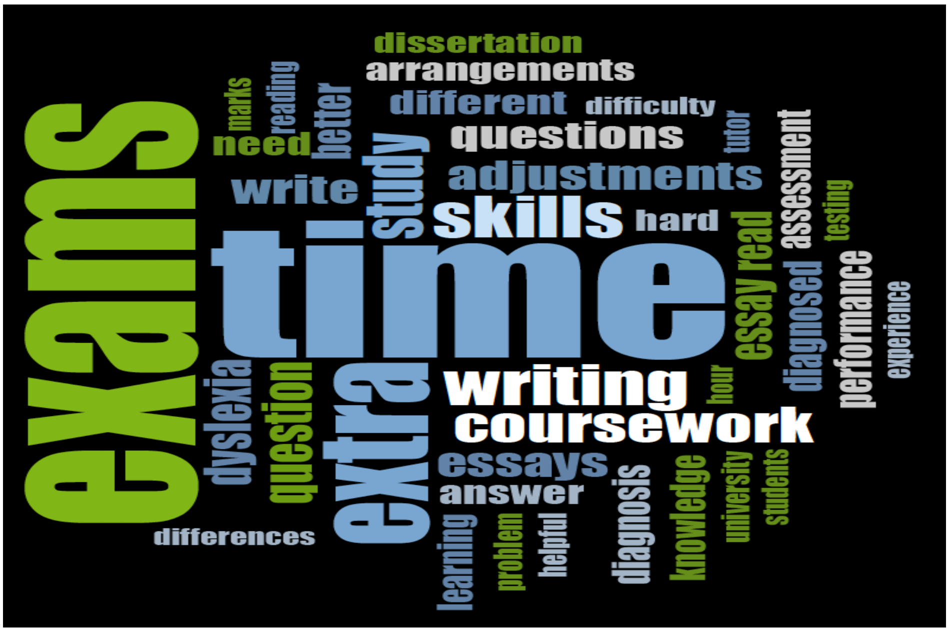
In order to further minimise the impact of the researcher's bias on the interpretation of the data and the themes identified, the 'trustworthiness' and credibility of the interpretation of the data was checked with the participants from whom the data was collected (Creswell & Poth, 2018; Lincoln & Guba, 1985). Each of the participants was emailed a summary of the themes identified in their transcript and invited to verify accuracy or correct any misinterpretations or omissions. This aimed to ensure the authenticity with which the themes identified represent the meanings intended by the participants. The feedback confirmed the participants' satisfaction in the trustworthiness of the interpretation and categorisation of the data.

11.3 Conclusion

In keeping with the principles of grounded theory, the qualitative research phase did not commence with a pre-identified set of concepts or themes, instead, these emerged during the analysis of the data. The participants in this qualitative phase of the project described their experiences in timed, closed book exams and recounted the different ways in which the exam environment presented barriers to their performance. They also described the strategies they used to compensate for these barriers as well as identifying those barriers that were not overcome (or sufficiently reduced) by their compensatory strategies. They identified possible approaches to assessment that could potentially remove or reduce the barriers sufficiently to foster a level playing field. Using intensive coding, these descriptions were then categorised into themes, links were made between categories and the categories were integrated into the overarching concept that there are barriers inherent to timed, closed book, exam environments that inhibit students with SpLD from performing to their potential. This main, overarching theme (the core category) allowed all the other emerging themes to integrate and thus form a structure that facilitated a comprehensive theoretical explanation of how these barriers in exams occur and how they are experienced by the students with SpLD (Corbin & Strauss, 2015; Creswell & Creswell, 2018)

Thus, the data was collected through purposive sampling and semi-structured, open ended interview techniques and, rather than testing the hypothesis that students with SpLD who have exam arrangements perform differently to their TD peers in exams (as measured by word count and mark), this project used the data to identify emerging themes that provide an understanding of how the phenomenon of the research project (inequality) occurs.

12: QUALITATIVE PHASE 3: FINDINGS



12.1 Introduction

12.2 Themes

12.2.1 Course accessibility
12.2.2 Levelling the playing field
12.2.3 How the adjustments help: Time to complete the paper
12.2.4 How the adjustments help: Reducing anxiety
12.2.5 Quantifying the impact of extra time
12.2.6 Barriers that are not overcome by exam adjustments
12.2.7 Generic nature of exam arrangements
12.2.8 Construct irrelevant skills: Lack of automaticity of Literacy skills
12.2.9 Construct irrelevant skills: Slow processing
12.2.10 Construct irrelevant skills: Working memory deficits
12.2.11 Interaction between slow processing and working memory deficits
12.2.12 Recall difficulties
12.2.13 Fatigue from the extra time
12.2.14 Impact of SpLD on revision

12.3 Differences between performance in coursework and performance in Exams

12.4 Efficacy of study skills support

12.5 Barriers to accessing study skills

11.5.1 Time cost
11.5.2 Othering

12.6 Over-arching theory

12.7 Solutions

11.6.1 Co-design of assessments
11.6.2 More applied, meaningful, and relevant assessment
11.6.3 Diversified assessment
11.6.4 Clarifying expectations
11.6.5 Use of assistive technology
11.6.6 Training for examiners and test designers
11.6.7 Increased take up of study skills sessions

12.8 Conclusion

This chapter presents the findings from the qualitative phase of this research project and answers research question 2: "What are the perceptions of students with SpLD of the impact that exam access arrangements have on their exam performance?" The chapter draws on qualitative data from semi-structured interviews that interrogate the experiences of students with SpLD who use exam access arrangements in order to identify how effective they perceive these arrangements to be.

The earlier findings chapters focused on the outcomes of the quantitative data comparing the exam performance of students with SpLD with their TD peers and revealed that the exam arrangements that students with SpLD are granted neither confer an advantage nor level the playing field in exams and an awarding gap persists despite the implementation of exam adjustments (albeit the adjustments narrow the awarding gap). The identification of an awarding gap between the SpLD and TD participants in this project, despite the granting of exam arrangements, suggests that explicitly disabling processes are intrinsic to timed, written, closed book exams and disadvantage candidates with SpLD. This chapter builds on those findings and aims to uncover the nature of these disabling processes and how candidates with SpLD experience the resultant disadvantage. In this chapter, the participants, through semi-structured interviews, discuss the barriers to performance that they experience in the exams, as well as in the course itself, and detail the ways in which the exam adjustments do and do not overcome those barriers. The participants also present their views on possible solutions to those barriers; data which may assist in determining possible interventions and policy changes aimed at resolving the awarding gap that was identified in the quantitative phase of this research project.

12.1 Introduction

This research project aimed to investigate whether students diagnosed with specific learning difficulties (SpLD) who are permitted 25% additional time, or the use of a word processor with 25% additional time, demonstrate an unfair advantage in timed, formal University exams compared to their non-disabled peers or whether the exam access arrangements have simply ensured parity by relieving the students with SpLD of 'any substantial disadvantage that might arise as a result of their disability' (Equality Act 2010).

The first phase of the research project demonstrated that, far from conferring any advantage to students with SpLD, the exam adjustments failed to even level the playing field in terms of exam marks and/or exam classifications. Despite receiving exam adjustments a significant awarding gap remained, with participants with SpLD achieving lower exam marks and lower degree classifications than those awarded to their TD peers. At the same time, the second phase of the research project showed that the granting of exam arrangements had a significant positive impact on the exam marks of the participants with SpLD and narrowed (albeit failed to close) the awarding gap between SpLD and TD participants. Thus, the data from the first two phases of this research project identified the presence of an awarding gap between students with SpLD and their TD peers when they are assessed though time-constrained, closed book, written exams, despite the implementation of exam adjustments whose purpose is to level the playing field. While the quantitative data in the first two phases of this research project has been able to show *what* is happening (i.e. the continued existence of an awarding gap despite the exam arrangements), it is unable to explain *why* this is occurring (i.e. why students with SpLD demonstrate poorer performance in exams compared to their TD peers, despite performing similarly to their TD peers in assessed coursework, and why the exam adjustments fail to overcome this disadvantage in exams).

Therefore, following the first two phases of this research project, a third, qualitative, phase of the research project was conducted to augment the quantitative research by investigating students' own experiences of using exam access arrangements in order to construct an understanding of why the awarding gap continues to exist despite the awarding of exam adjustments. A detailed analysis of these various personal accounts aimed to centralise the students' voices in the research project and contribute the following:

- a broader explanation of the findings of the quantitative data.
- an interpretation of the ambiguous findings of the quantitative data (particularly in relation to the efficacy of study skills sessions and any differences in quantitative outcomes between disciplines).
- an understanding of those aspects of the research project that are not quantifiable.
- a broader perspective than can be achieved from one data collection method alone.
- a basis for the identifications of potential intervention (Bryman, 2009).

Thus, an analysis of the lived experiences of students with SpLD using exam arrangements, through the collection of qualitative data, aims to develop an understanding of the barriers that impede the performance of SpLD students in exams and inhibit them from achieving their full potential. This understanding, it is hoped, may acting as a step towards the identification of approaches and policies that may mitigate or navigate these barriers.

12.2 Themes

The themes that were identified in the literature review chapter were similarly reflected in the experiences described by the participants in this research project. In the semi-structured interviews the participants all described how the adjustments and support they received on account of their diagnosis of SpLD helped to overcome some, but not all, the barriers that prevented them performing to their potential in their timed, closed book, written exams. They identified the tension between the benefits of course-based adjustments (such as handouts in advance, modified reading lists, permission to record in lectures and tutorials, study skills tuition) which were designed to make the course accessible, and the disadvantages of the inaccessible teaching approaches commonly adopted. Similarly they identified the tension between the benefits of exam access arrangements on the one hand (which was universally seen as a significant factor in allowing them to perform more closely to their potential in the exam), and the construct irrelevant access skills demanded by the exams, on the other hand, that pressurise core deficits associated with SpLD and impede them from showing their ability. On balance, the participants felt that the negative impacts of inaccessible course delivery and exam formats were not sufficiently offset by the course-based adjustments and exam accommodations; hence the persistence of the awarding gap.

12.2.1 Course accessibility

Similarly to the participants in the studies evaluated in the previous literature review chapter, the participants in this phase of the research project did not present an entirely negative picture. Participants commented on the benefits of course based adjustments during their studies, which enabled them to access their course more fully, revise more effectively, and improve their essay writing techniques - thereby contributing to their exam performance:

> 'Arriving at uni already with the diagnosis meant I could record in lectures and have power-points before the lecture. I got more help with the reading lists and also had them earlier and so I wasn't struggling from the get-go.' (Juliet).

However, at the same time, the participants described the greater impact of the barriers they experienced due to inaccessible pedagogic practices during the year, such as lecture style delivery of information which relies on being able to listen and take notes simultaneously at speed, which had a 'back wash' effect on their exam performance:

> 'The problem isn't just in the exam itself, it was in teaching before that, and so I wasn't prepared for the exams. For example, I can't keep up with getting notes down in lectures. I can either listen and understand what's being said, but then I can't remember the information when the lecture is over, or I can just take notes and not try to understand the material, kind of 'lecture capture' I guess. So you're not just disadvantaged in the exam itself, you're also disadvantaged before the exams by the inaccessible teaching.' (James)

The participants felt that the adjustments they received during the course, while helpful, were insufficient to offset the disadvantages they experience due to SpLD, as conventional teaching and assessment approaches adopted at university engender a learning environment that is fundamentally antagonistic to their cognitive profile:

> 'Attending a university course in the traditional way that its structured is counter-intuitive for someone like me with dyslexia. If I was given options, I'd *never* imagine myself landing on 'you'll have one-hour blocks where you'll sit and listen to a lecture and write notes. And then you write up about it once a week and then at the end of the year you're examined on it'. That is the last way I'd ever choose to learn or be tested and its deeply inaccessible' (Anna).

12.2.2 Levelling the playing field

While there are conflicting views about the impact of granting exams arrangements to students with SpLD, there is a consensus in the field that the purpose of the exam arrangements is to level the playing field (Camilleri et al., 2020; Chetcuti et al., 2019; Lightfoot et al, 2018; MacCullagh et al, 2016). In this current research project all the participants, similarly to those in the extant literature, expressed the belief that the purpose of the exam arrangements was to compensate for the barriers that the exam presented to candidates with SpLD and thereby confer equity across the cohort as a whole. They opined that the exam arrangements simply helped to overcome construct irrelevant barriers (insufficient time), but did not artificially enhance their ability by boosting their subject knowledge or critical thinking skills (i.e. the constructs that the exam is testing):

> 'The idea is to level the playing field, so that everyone has the opportunity to access the questions and respond in a way that reflects what they've been learning….if someone could write a thousand words and I could only write 600 words, say, in the same time, that extra time would allow me to write the same amount compared to someone who can write faster and who can organise their thoughts faster.' (Daniel)

12.2.3 How the adjustments help: Time to complete the paper

The participants explained that, rather than conferring an unfair advantage, the exam adjustments enabled them to improve their exam performance by allowing them to complete more of the paper, take the time they needed to plan and organise their ideas, implement strategies to help compensate for problems with memory and factual recall and thus reduce anxiety and exam stress. Hjarn (2020) argues that granting extra time to candidates with SpLD enables the candidate to complete the paper on the basis that ‘while the time is adequate for others, extra time will help test-takers with dyslexia to overcome a construct-irrelevant barrier (lack of time)’ (p2). This view was mirrored by the participants in this current research project, who commented that the extra time was necessary to enable them to read the questions and formulate written responses and so complete the paper - at the same time identifying that completing the paper in the standard time allowed was not similarly problematic for their TD peers:

> ‘Before I had extra time I didn’t finish exams at all, I often had 3 or 4 questions left. And that’s quite a lot of the marks in the exam lost. When the time was up I’d still have a whole page to go and I’d be like ‘ohh, there’s still another 3 questions’. But my friends all managed to finish in time, it was just me that was slow.’ (Robert)

The participants commented on the difficulty they experienced in producing essay style responses in exam conditions and how, in addition to providing additional time for reading the questions, the extra time allowed them more time to execute the various stages of essay writing, including planning and organising their ideas:

> ‘The extra time lets me get my thoughts in order and to go through my thinking and re-organise my ideas so that I have the narrative that was the one that I wanted to produce. Essentially the adjustments allow me to go through the stages that I need to go through in order to produce an essay.’ (Alex)

While the difficulties in compositional writing are, perhaps, more closely aligned with the longer, essay style, questions associated with Humanities subject, for the participants studying STEM subjects the extra time allowed them to implement strategies they had developed to compensate for difficulties they experienced in the exam with factual recall; strategies that were more time consuming:

> I often rely on superior comprehension to get through the [Maths] exam - I learn the gist of the formulae and then spend time in the exam deducing the likely formulae by reverse-engineering the question, and I then can proceed. I find that the extra time gives me time to go through the process of reverse-engineering the question. But it would be a lot quicker if I could just remember the formulae in the first place.’ (Laura).

Deficits in the retrieval of factual information, as well as the learning and recall of serial order information, is commonly reported in individuals with SpLD (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al, 2019; MacCullagh, et al, 2016; Shaw & Anderson, 2018; Snowling et al, 2016).This may be explained by underlying deficits in working memory and the role that working memory plays in developing secure phonological representations in long term memory (as shown in Baddeley's revised model of working memory, Figure 12.1). Thus, in individuals with SpLD, there appears to be a deficit in the ability to retain a sequence of speech sounds in working memory and then transmit that sequence to long-term memory (Norris, Page & Hall, 2018).

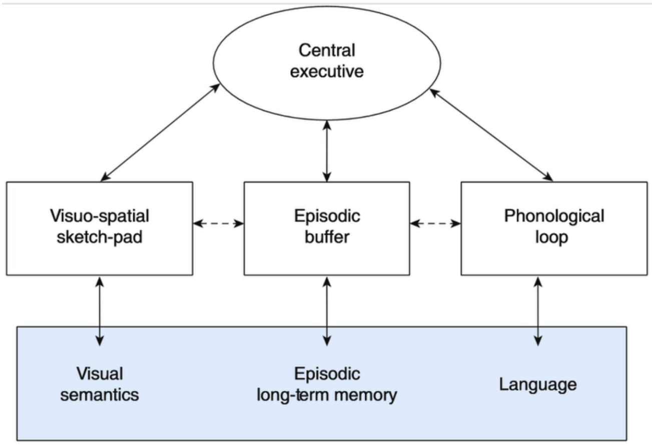


Figure 12.1: The revised model of working memory (Baddeley, 2000)

12.2.4 How the adjustments help: Reducing anxiety

As identified in the previous literature review, a number of studies have found that HE students with SpLD are at significantly greater risk of experiencing high levels of exam anxiety and stress compared to their TD peers. This is attributed to the problems candidates with SpLD experience with literacy skills, processing speed and working memory, which are construct irrelevant access skills required by formal timed examinations and are exacerbated by the time-pressured, high stakes, nature of the exam (Dobson Waters & Torgerson, 2020; Camilleri et al., 2020; Chetcuti et al, 2019; Livingston, Siegel, & Ribary, 2018; Nelson, Lindstrom & Foels 2015; O'Byrne, et al, 2019; Stagg, Eaton & Sjoblom, 2018; Boyes, et al, 2016). The participants in this current research project similarly identified that receiving extra exam time helped to reduce anxiety and thus promoted improved processing and clearer thinking; thereby enhancing performance:

> Having the exam adjustments definitely reduced anxiety and therefore that meant that I was clearer with my thoughts and more able to access knowledge as I was less anxious and stressed. It helps me to feel less rushed and less panicked where I can't think straight and I start confusing information. (Graham)

These findings are also congruent with the findings of Lewandowski et al (2014), O'Byrne, et al (2019), Shaw et al. (2016), Shaw & Anderson (2017) and Shaw, Anderson & Grant (2018), all of whom found that the extra time granted to students with SpLD had the impact of reducing anxiety and exam stress, with participants describing how they felt more relaxed, less frustrated and therefore able to think coherently during the exam and, as such, more likely to perform to their potential.

12.2.5 Quantifying the impact of extra time

These descriptions of the value of extra time in allowing candidates with SpLD to go through the stages needed to complete the exam paper, and implement the strategies that compensate for working memory deficits, support the findings of the second quantitative phase of this research project that showed a significant difference in exam performance of the participants with SpLD pre and post their diagnosis (i.e. with and without the extra time). Similarly, the participants in this qualitative phase of the research project described an underperformance in their exams prior to their diagnosis (in the absence of exam adjustments) and associated their improved exam performance following diagnosis with the implementation of exam accommodations; commenting on the significant mark difference that the extra time produced:

> 'In the previous exams [without extra time] what I did was to start writing as soon as the time started as when I have planned I haven't had time to get down everything that I wanted to write. When I didn't have extra time or the word processor the quality was lower because I didn't plan, well I wouldn't even finish the essay' (Alex)

However, despite recognising that receiving 25% extra time in exams helped candidates with SpLD to complete the exam, many participants commented that the 25% extra time was not sufficient, especially for longer essay style questions:

> 'I feel like it's a lot better with the extra time, but I still don't have enough time to show properly what I know. I don't do introductions or conclusions because I don't have enough time, so I just have to hope I can get away without them provided I've got the content right in the rest of the essay.' (Anna)

The qualitative studies reported in the previous literature review chapter similarly revealed that participants with SpLD consider the extra time they receive in exams to be insufficient, albeit more extra time is fatiguing, and so counterproductive (Camilleri et al., 2020; Chetcuti et al, 2019; Pinot & Mortari, 2014; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018; Timmerman & Mulvihill, 2015). At the same time, a wider survey of the extant literature exploring exam accommodations reveals a lack of any empirical validation of the amount of extra time that is demonstrably sufficient to level the playing field in university exams (Holmes & Silvestri, 2019; Lewandowski et al, 2013).

12.2.6 Barriers that are not overcome by exam adjustments

A systematic review by Dobson Waters and Torgerson (2020) that explored the efficacy of support provided to HE students with SpLD identified that the deficits associated with SpLD that represent construct irrelevant exam access skills, such as processing speed, working memory and recall, automaticity of literacy skills and the ability to organise and structure ideas in essay format, are exacerbated by time-bounded, closed book, written exam situations. Dobson Waters and Torgerson conclude that the exam adjustments, while helpful, are not able to fully offset the impact of the high-pressure situation of an exam, as they are not as effective as they would be in a less pressurised environment. Similarly, in this current research project, while the participants felt that the exam adjustments helped significantly, they were still conscious that the adjustments failed to fully level the playing field**:**

> 'Having the extra time and laptop in the exam makes a whole grade difference. However there are still things that it doesn't help with and so I still don't show my real ability in exam. I can't fully express what I know in the way that the examiners want to see it in the exams. I can do that in my coursework, but not in the exams. So I always get higher grades for my coursework.' (Bella)

This also mirrors the findings of the second phase of this research project that showed that the extra time significantly improved marks for the candidates with SpLD, narrowing the awarding gap but not fully closing it. Also, the participants' observation of the material difference in performance between exams and coursework supports the findings of the first phase of the research project that revealed a significant difference between exam marks and coursework marks for students with SpLD, while no significant differences in performance between these two modes of assessment were identified for their TD peers.

The participants provided a number of explanations for the persisting awarding gap, despite the exam arrangements. Firstly, they commented that they needed more than 25% extra time in order to be able to perform to their potential, as 25% extra time did not permit them to allocate time strategically, such as allowing additional time for planning at the start of the exam, or leave time for proofreading their work at the end of the exam. In addition, the participants identified a range of barriers that acted as drivers of the awarding gap that were not directly related to the time allowance of the exams and so were not overcome by the provision of extra time. All of the participants identified that the adjustments failed to fully level the playing field where success in the exams is heavily dependent on effective literacy skills, working memory, memorisation and the ability to inherently understand the expectations of a written, essay style, response:

> 'Asking us to respond in an essay style format is testing competencies that aren't related to the subject, such as spelling skills, being able to remember names and dates accurately, and being able to write in a linear way under time pressure.' (Juliet)

12.2.7 Generic nature of exam arrangements

The impact of the lack of a range of bespoke, context sensitive, exam adjustments was identified as a factor in the awarding gap in a number of the studies evaluated in the previous literature review chapter (Lightfoot et al, 2018; Rogers et al 2019; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018), with Dobson Waters and Torgersen (2020) arguing that 'a one-size solution will not fit all' (para. 51). Similarly, in this current research project, the participants commented that, although they had a very detailed and lengthy diagnostic assessment that identified their individual cognitive profile, the exam arrangements were generically 'extra time and the option to use a word processor'. The exam recommendations were perceived as being the same regardless of the individual's particular cognitive profile or the distinct nature of the exam that they would be sitting:

> 'The adjustments are blanket and not tailored to the specific thing that I'm going to struggle with. 'Cos that thing can vary from exam to exam, and from person to person, so having a range of adjustments based on what I need and how I work would be more helpful.' (Robert)

Participants also explained that they often relied on assistive technology during their course (such as a thesaurus to help with word retrieval, speech-to-text software to help with spelling and writing fluency, text-to-speech software to help with reading accuracy and proofreading, and mind mapping software for organising ideas) but these tools were not available to them as adjustments in the exams. Thus, in the exam, they were not able to use the strategies they developed for essay writing during the year, leaving them at a disadvantage.

This experience was similarly identified in the extant literature, which also found that the main strategies used by students with SpLD to overcome difficulties with the structuring and expression of ideas in prose (such as using proofreading services, text-to-speech software, vocalising ideas, mind-mapping extensively) were strategies that were only available to students during the production of coursework and not in an exam situation (Couzens et al., 2015; MacCullagh, Bosenquet & Badcock, 2016; Nelson & Reynolds, 2015; Newlands, et al, 2015; Pinot & Mortari, 2014). This demonstrates how the interaction between deficits in the various literacy skills of spelling (inhibiting the use of expressive language), grammar, sentence structure and text structure, impacts negatively on academic writing, resulting in the potential for marks to be lost for candidates with SpLD that cannot be offset in the exams where the assistive technology that is commonly used to overcome these deficits is not generally permitted.

12.2.8 Construct irrelevant skills: Lack of automaticity of literacy skills

Lack of automaticity of literacy skills was identified in the literature review as a particular barrier to students with SpLD in exams (Camilleri et al., 2020; Chetcuti et al, 2019; Pinot & Mortari, 2014; Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018). Similarly, the participants in this third phase of the research project described difficulties performing to their potential in exams where the format of the questions privilege literacy skills, rather than subject knowledge:

> 'The problems that I have with spelling, grammar and structuring my ideas prevents me from being able to show my real knowledge in the exams. So I'm not showing what I can do just because I can't do the spelling and the structuring. I end up dropping marks even though I know the material, it's just that I'm not able to express it so well.' (Anna)

> ‘People are judgmental when it comes to spelling, grammar, sentence structure. And even though they are meant to be judging you on your ideas they are not seeing that. They are using spelling and grammar as a proxy for intelligence… spelling, grammar, and how you phrase your ideas, rather than your knowledge and thinking, is influencing the examiners and affecting your mark.’(James)

This concern that examiners perceive literacy skills as synonymous with intelligence or ability, influencing both the construction of the exam format (i.e. ‘discuss’) as well as the marking, is supported by a study by Majer in 2018 into commonly held perceptions of SpLD. Majer identifies that ‘SEND policy whilst formulated within an inclusive social model ... perpetuates the notion of the deficient student and affirms socio-historic connections between literacy and intelligence’ (Majer, 2018, p26). Thus, Majer argues, there are entrenched societal perceptions of a direct link between literacy skills and intelligence, unconsciously accentuated by SEN policy and practice in educational institutions. This unconscious bias appears to be unwittingly influencing both the format of exam questions and the examiners approach to marking, and thus inadvertently disadvantaging candidates with SpLD in exams.

Not only do difficulties with spelling, grammar, and written expression impede the exam performance of students with SpLD, but inaccuracy with reading also can have negative consequences. The tendency of individuals with SpLD to misread words is inevitably problematic in exam situations where time pressure and anxiety may exacerbate the tendency to misread (and where any misreading has the potential to cost marks, as it may lead the candidate to misunderstand the question being asked (Camilleri et al., 2020; Chetcuti et al, 2019; Ellis, 2016; O’Brien et al, 2012; Shaw et al, 2016). Congruent with the studies identified in the literature review, the participants in this current research project reported a tendency to misread in exams. The extra time provided the opportunity to check the accuracy of reading, which the participants felt helped to some extent:

> ‘It wasn’t the reading itself that was the problem, it was more the misreading of words and missing words out that was a problem. I read ‘impossible’ as ‘possible’ and then thought ‘I don’t know what this answer is’, but because I had time I then went back, and I realised what I’d done.’ (Alex)

Nonetheless, despite the extra time, misreading remained an issue that had the potential to cost marks. This was equally a problem in STEM exams as it was in Humanities exams:

> 'When I came out of the [Maths] exam I said to my friend 'that was a giveaway question on the NOT diagram". He said, "it didn't say 'NOT', it said 'NOR'. Misreading that one letter meant I got zero for that question as I'd answered a different question.' (Laura)

In addition to difficulties with reading accuracy, spelling and grammar, the participants also identified difficulties with the organisation and structuring of ideas across the essay responses in their exams that were not sufficiently compensated for by the exam arrangements:

> 'Because of my dyspraxia I have problems organising my thoughts and making sentences and paragraphs flow as well, and the extra time and the computer didn't help much with that flow.' (Daniel).

Studies suggest that difficulties with the organisation of ideas in the written work of students with SpLD are exacerbated by the closed book, time pressured environment of written University exams, which the exam adjustments are unlikely to be sufficient to offset (Dobson Waters and Torgerson 2020; Fullarton & Duquette, 2016). In addition, the extant literature suggests that not only do students with SpLD experience difficulties in essay writing due to deficits in literacy skills and problems with planning, structuring, organising and editing ideas, but that they also experience problems with the contextual features of essay writing, such as identifying and responding to the expectations of the essay (Camilleri et al., 2020; Chetcuti et al, 2019; Kearns & Whaley, 2019; Kimel & Ahissar, 2020; O'Byrne, et al. 2019). Reflecting this, the participants with SpLD in this current research project commented that they found it particularly hard to work out the expectations and 'ingredients' of an exam essay response, which was something that the extra time and use of a word processor was unable to address:

> 'I don't really know what the examiners want, because in the exam it just isn't really clear what we're supposed to do, so I just write for half an hour about things that I know about in the hope that that is what they were looking for.' (James)

12.2.9 Construct irrelevant skills: Slow processing

There is general consensus in the field that individuals with SpLD present with deficits in processing speed (Camilleri et al., 2020; Gooch, et al, 2016; Georgiou, Ghazyani &

Parrila, 2018; Moll et al, 2016; Snowling et al, 2020). Slow processing has the potential to negatively impact on the performance of university students with SpLD in timed exams, as the underlying processes required for the efficient execution of the access skills needed in exams relies heavily on processing speed (Fullarton & Duquette, 2016). In this current research project, participants commented that barriers arising from slow processing compromised both their learning during the year as well as their performance in time bounded exams. The participants described how slow processing speeds caused all of their study activities to take longer, including reading, making notes, and revision. This exposed the students with SpLD to less learning than their TD peers across the year and placed them at a disadvantage in terms of preparation for the exams; compounding the disadvantage they experience as a result of reading and responding to the questions more slowly in the exam itself:

> 'Inevitably I'm slower than other people, so that's an issue all during the year. Reading and taking notes takes me longer so I don't get as much learning as my peers. I get to cover less during my revision because I can't do that as fast, and I'm slower at getting my ideas down in the exams. So I am at a disadvantage all the way through.' (Graham)

12.2.10 Construct irrelevant skills: Working memory deficits

Working memory can be conceptualised as the ability to retain and manipulate information over a short period of time (managed and coordinated by the central executive) and is used to plan and execute actions in concert with short term memory (with the aid of the attention-related and integration processes of the central executive) (Baddeley et al, 2019; Broggi et al, 2019; Liebel & Nelson, 2017). Thus, working memory is used to perform most tasks that require focused concentration and multi-step problem solving (Baddeley et al, 2019; Brydges et al, 2018). Deficits in working memory represent a core diagnostic feature of SpLD and can negatively impact reading comprehension and prose writing activities (Camilleri et al, 2020; Hatcher, Snowling and Griffiths; 2010; MacCullagh, Bosanquet, & Badcock, 2016; Oloffson, Taube & Ahl, 2015). Reflecting the extant research, working memory deficits were described by the participants in this research project as having a detrimental impact on their reading comprehension and production of prose in the exams, which the extra time was unable to fully compensate for:

> 'I have problems with keeping my ideas in my head long enough to structure and write them down. They just kind of dissolve. And the extra time doesn't help a whole lot with that.' (James)

12.2.11 Interaction between slow processing and working memory deficits

All the participants identified the 'multiple deficit' arising from the interaction between slow processing and deficits in working memory (i.e. how slow processing overburdens the, already weak, working memory) and felt doubly disadvantaged as a result. This is congruent with research studies identifying that slow processing places an additional burden on the working memory of students with SpLD, and further aggravates problems with the many activities involved in university study that demand efficient working memory (Camilleri et al, 2020; Fullarton & Duquette, 2016; MacCullagh, Bosanquet, & Badcock, 2016). For the participants in this current research project, this had an impact both on their study in preparation for the exams (such as listening and taking notes in lectures, recalling what had been read, and revision) as well as in the exams themselves (such as retaining ideas in working memory sufficiently long to structure the argument and formulate written responses):

> 'Everything to do with my study take me longer because I'm slow at reading and processing ideas. And then, on top, I have problems with memory. That's a double whammy as I have to hold my thoughts in my head for longer before I can get them down on paper because I am slow. And then, because I forget quickly, they evaporate. I can't hold them long enough. I have to keep going over and over it to make it stick and that means that it takes me much more time and effort to get my thinking down. That's a problem all through the year and I don't get through as much as others do, so I'm not as prepared for the exams as my friends are. And then the exam puts a lot of burden on those things too. So being slow and having a poor memory puts me at a double disadvantage.' (Daniel)

12.2.12 Recall difficulties

In addition to working memory deficits, students with SpLD experience impairments in long term memory capacity (arising from the deficits in working memory, which then impacts on efficient associative search mechanisms), which affects factual recall and the retrieval of information (Baddeley et al, 2019; Gooch et al, 2016; Moll et al, 2016; Snowling et al, 2020). Similarly, the participants in this current research project identified difficulties recalling factual information and described how this was a barrier in exams that privileged information recall over applied knowledge and reasoning; a difficulty that could not be overcome by additional exam time:

> 'The hardest thing in the exams was remembering the names of the people you're meant to be talking about. They shouldn't be marking you on your ability to remember the name, they should be marking you on your awareness of the theory and how to apply it. But unfortunately it's a necessary requirement to know the name. So it's a barrier.' (Daniel).

The impact in exams of difficulties with recall appears not to be confined to the context of Humanities subjects, but similarly affects students in STEM subjects. A range of studies show that individuals with SpLD experience deficits in their capacity to recall number facts, despite not presenting with any difficulty in their Maths reasoning ability, due to deficits in the phonological storage and rehearsal loop mechanism which is required to effectively memorise mathematical formulae in correct serial order (Baddeley, Hitch & Allen, 2018; Boedigheimer et al, 2015; Chinn, 2019; Gomez et al, 2015; Zoccolotti et al, 2020). These studies identified that students with SpLD typically use reconstructive tactics to compensate for their deficits in factual recall and the participants in this current research project similarly described using a reconstructive approach to compensate for difficulties recalling number facts:

> 'In Maths there are pretty bad exam costs associated with not being able to memorise effectively. I'm expected to commit to memory formulae and theorems, and then reproduce these in exams. Often, this will be the first part of a long question, and if I can't recall the precise formula, I can't do the rest of the question. Extra time often can't help with this if I simply haven't memorised the appropriate formula.' (Graham)

12.2.13 Fatigue from the extra time

Research by Lyman et al. (2016), MacCullagh, Bosanquet, & Badcock (2016), Shaw & Anderson (2017) and Shaw, Anderson & Grant (2018) concur that students provided with extra time in exams report cognitive fatigue, which has the potential to result in decreased exam performance. Similarly, the participants in this current research project reported experiencing mental fatigue as a result of the extra time and that this caused a reduction in later exam performance:

> 'There are disadvantages of having extra time because you get tired and lose concentration because the exam is long, especially if it's the second one in a day. Other people have only had to have 6 hours of exams in the day and, with the extra time, I've had to focus intensely for seven and a half hours of exam time and that's exhausting, especially if you have exams the next day too. There isn't any recovery time and then you don't do very well in the later exams as you're too tired, more tired than other people who didn't have extra time'. (Robert)

12.2.14 Impact of SpLD on revision

Participants commented that the deficits associated with SpLD not only impacted on the construct irrelevant skills needed for the exams, but also negatively affected their revision in the run up to exams. As a result, participants described how they had to commit to more time to study and revision than was the case for their peers, but were still unable to cover as much material as their peers – resulting in a disadvantage when going into the exams:

> 'I think it's not just the exam, it's the pre-exam element too, its the whole revision process. I find having dyspraxia means that I can't prepare for my exams like my friends can, as I don't have enough time. I was working in the library all day and all night, and my friends weren't but they still read more than I did.' (Daniel).

The participants also explained that, as they were unable to cover the same breadth of material as their peers during revision, this limited their choice of questions on the exam paper:

> 'Because I can't revise as quickly as others, I have to pick and choose what to revise. But that means that I can't just choose what questions I want to answer on the paper, I am restricted to those topics that I managed to revise. They may not be the best questions for me to answer to get the highest marks, but I don't have a choice. I have to select questions relating to topics that I have revised and that's not as much as others.' (James)

This was also described as a particular problem for Maths students, as, while in some disciplines it is possible to perform adequately in the exam despite being unable to revise all of the topics, in Maths exams the inability to revise all of the topics from the year prevented the participants from being able to answer mandatory questions:

> ‘The maths exams are notorious for examining every bit of the course. They give you a question on almost everything that you have covered during the year, so if you’re missing one bit there will be questions that you can’t answer. I remember in my first-year maths exam I was missing a lot to do with matrices and there was a matrices question in the statistics paper and no amount of extra time could have helped with that ‘cos I just hadn’t covered it in time.’ (Graham)

The inability of Maths students to be able to avoid any topic areas in the exam that they had insufficient time to revise for may account for the findings of the first phase of this project, where the difference in the awarding gap between SpLD and TD candidates in Maths was greater than the difference in the awarding gap between SpLD and TD candidates in the Humanities subjects.

12.3 Differences between performance in coursework and performance in exams

The data from the first phase of this research showed a significant gap between the exam mark and coursework mark for candidates with SpLD, and this gap was significantly larger for the SpLD group than the TD index group. This finding is supported by the extant literature, which similarly identifies that students with SpLD perform closer to their potential in coursework than is the case in exams (Dobson Waters & Torgeson 2020; Shaw et al, 2016; Shaw & Anderson 2017). Reflecting the body of research, all participants in this qualitative phase of the project stated that they underperformed in the exams by comparison to the marks and feedback from their coursework:

> ‘I was doing really well in my coursework essays; they were high First’s and then in the exam I got a 2:2 in that same paper. Definitely there’s a difference for me between exams and coursework. My dissertation went really well, but my exam results were much lower.’ (Juliet)

The participants gave a variety of reasons to explain why they were able to perform more closely to their potential in coursework than was the case in exams. Firstly, they unanimously agreed that the coursework provided them with adequate time to plan, write fluently, revise, and edit their writing, and overcome the difficulties with literacy skills and organisation of ideas, with the result that they could communicate their ideas more cogently and with greater precision in coursework than was the case in the exams (even with the extra time):

> ‘Yeah, coursework is good, because you can take time to make sure you get your ideas written clearly and it takes time, *days*, for me to do that. In the dissertation you have time to think about your ideas. A lot of the time you’re

> spending in exams on just working out how to write the sentence and paragraph and you don't have the time to really think about the ideas.' (James)

Richardson (2015) suggests that the purpose of exams is 'to test rather than to teach, whereas assessment by coursework achieves both outcomes' (p.439). Similarly in this project, the participants highlighted the greater pedagogical value of coursework, with its benefits to learning arising from feedback opportunities and its integration into the course of study. They contrasted this to the exam experience:

> 'My dyslexia didn't really impact on my coursework. I wrote a dissertation so I could express my ideas, and the feedback from my supervisor really helped, and I did much better in that than in any of my exams. Coursework helped me to improve my learning because it means that you have to research and so you become more knowledgeable in your subject. That doesn't happen in exams (Juliet)

There is a range of research that suggests that coursework enables candidates to showcase a more extensive array of skills than is the case in time-constrained exams, including research skills and innovative thinking, and so has the potential to deepen understanding (Chetcuti et al, 2019, Shaw & Anderson, 2017; Shaw, Anderson & Grant, 2018; Shaw et al, 2016). Reflecting these views, the participants in this current project remarked that coursework allowed them to show a broader range of skills than was the case in exams:

> 'There are multiple skills that I don't think I get to show in the exams including creative approaches, flair in approach, finding multiple approaches to problem solving. I was able to show all that in my coursework. You could take an angle that was quite unique, and you could choose something that you could become an expert in.' (Anna)

The participants identified that these differences in the character of coursework and timed exams contributed to a full classification difference between the exam and coursework grade. This is consistent with the findings of the quantitative phase of this project that showed that the participants with SpLD achieved, on average, lower classifications in their exams than was the case in their assessed coursework.

12.4 Efficacy of study skills support

Most HEIs provide additional support to students with SpLD in the form of 1:1 study skills support (Dobson, 2018; Lyman et al, 2016; Peterson, 2016). However, the earlier qualitative literature review chapter revealed some variability in the reported efficacy of 1:1 study skills tuition, with some students with SpLD finding the support helpful with improving study techniques, while others found the study skills tutors' lack of course knowledge limited the value of the support they could offer (Couzens et al., 2015; Dobson Waters & Torgerson, 2020; Nelson & Reynolds, 2015; Peterson, 2016; Saeidi et al, 2019; Sumner, Crane & Hill, 2020; Degirneci, Baglama, & Yucesoy, 2020). Reflecting this research, the participants in this current research project commonly sought study skills sessions to develop more effective writing strategies, including strategies for organising and structuring their written work as well as improving written expression. The participants reported that the bespoke approach to the support (based on where they were and what they needed to do), rather than a generic, predetermined, 'set menu' of skills tuition, was most helpful:

> 'The tutor took me back to basics about how to write a sentence, because, to me, I just couldn't work it out. For me, it wasn't innate. Also it was based on where I was and what I needed to do, it wasn't some pre-set programme, which was hugely helpful.' (James)

Many participants also used the sessions for exam preparation and to develop strategies for using their extra time in the exam:

> 'The study skills sessions helped with organising my thoughts and ideas in the exams and making use of the extra time. It gave me a method on how to approach the exams. An essay is a linear way of putting things, I spark off with lots of fantastic ideas, but these need to be wrestled into a linear argument and someone who thinks in a linear way, like the study skills tutor, can show you the formula of how to do this. They bridge that way of thinking, of how to frame your ideas in that linear way that the exam demands of you.' (Juliet).

The participants studying STEM subjects described the importance of working with a study skills tutor who shared a discipline specific background:

> 'I had a problem to start off with because my study skills tutor didn't have a background in Maths and that didn't translate into improvements in my performance as a lot of the skills I developed were just not effective. Then I was given a tutor with a Maths background and that was really helpful. She could give me general advice on how to study and revise Maths, figure presentation and that kind of thing. That helped me improve my course performance '' (Graham)

While the importance of a study skills tutor with a shared subject background was identified particularly by participants in this research project who were studying STEM subjects, this may be a reflection of the fact that most of the study skills tutors had a background in Arts and Humanities, and so already shared a common discipline background with those participants studying Arts and Humanities. As identified in the literature review chapter, this outcome is congruent with findings by Lyman et al, 2016, Monagle, 2015 and Shaw et al, 2016, who similarly found a higher take up of study skills tuition by students with SpLD studying Arts and Humanities than was the case for those studying STEM, which, they suggest, arises from an insufficiently detailed understanding of STEM subjects by study skills tutors.

12.5 Barriers to accessing study skills

Studies by Barrable et al (2018); Hassanbeigi et al (2011); Kirwin & Leather (2011); Rahm & Meon (2012); Saeidi et al, 2019; Shetty & Srinivasan (2014) all identify some lack of take up of study skills sessions by students with SpLD who are eligible to access this support. Similarly, only 40% of the participants with SpLD in the second quantitative phase of this current research project took up their study skills provision[142]. Participants in this third, qualitative phase, of the research project (who similarly had a 40% take up of study skills sessions) provided insights into the reasons for this lack of take-up of additional support.

12.5.1 Time cost

The most common reason cited by the participants for not accessing study skills sessions was the time cost, given that students with SpLD are already 'time poor', and the risk that, despite the time investment, it may not help:

> 'There never seemed time to have the sessions. Also, not knowing if they would be effective meant that I didn't feel that I could dedicate time to it. Every moment of my time is so valuable as I take so much longer to get through my work than my peers, if the study skills sessions didn't help in a really significant way I would have lost more time and I couldn't afford to take that chance.' (Joanna)

[142] Participants with SpLD in this qualitative phase of the project were selected based on similar characteristics to those who participated in the second qualitative phase of this project, including a 40% take up of study skills tuition.

12.5.2 Othering

The participants in this current research project also commented that asking for additional support, such as study skills tuition, compounded feelings of inadequacy and 'othering', discouraging them from accessing the sessions:

> 'There's the stigma around asking for extra help in form of study skills help. You don't want anything different to others. Being singled out as needing study skills support made me feel exposed and made me feel less able. So that's put me off taking it up.' (Anna)

This participant articulates the latent vulnerability of her academic self-esteem and how this can be triggered by institutional policies that 'other' the student; in this case, receiving different treatment through 1:1 specialist study skills support. Extant studies suggest that the impact on self-esteem of the stigma and 'othering' associated with SpLD is deeply entrenched, as negative societal perceptions of SpLD are internalised by the individual from an early age with the result that the vulnerability to low academic self-esteem remains lifelong and is easily triggered by subsequent experiences (Cameron, 2016; Camilleri et al., 2020; Chetcuti et al., 2019; Lightfoot, 2018; Majer, 2018). As articulated by the participants in this research project, these societal perceptions, in conjunction with time costs, potentially act as barriers to students accessing valuable study skills support that could otherwise improve their learning outcomes and exam performance.

12.6 Over-arching theory

As indicated in the previous qualitative methodology chapter, the decision to use a grounded theory methodology for the qualitative data collection, interpretation, synthesis and analysis phase of this project arose from a key aim of this project to generate new theory grounded in the data (Corbin & Strauss, 2015). The over-arching theory yielded by the major themes that emerged from the qualitative data, is that timed, closed book, written, exams (and also, to a large extent, the teaching and learning processes and practices adopted throughout the degree course) are predicated on a non-disabled norm that is intrinsically disadvantaging for students with SpLD. The systematic analysis and interpretation of the qualitative data and themes described above reveals that timed, closed book, written, exams are viewed as inherently 'ableist' and deeply inaccessible to candidates with SpLD. While adding adjustments on to a fundamentally discriminatory, ableist exam model reduces some of the difficulties candidates with SpLD experience in accessing the exam on a level playing field with their TD peers, these ‘bolt on’ changes fail to alter the inherently ableist structure sufficiently to overcome the disadvantages that are intimately interwoven into the very fabric of the exam environment. Thus, even with the provision of exam adjustments, timed, written, closed book exams inevitably continue to represent a fundamentally ableist structure that is disabling and disadvantaging for students with SpLD.

In addition, the data reveals that the ‘bolt on’ nature of the adjustments bring the ableist structure of exams sharply into focus by adding in an additional layer of othering, stigmatisation and discrimination. That is to say, the exam arrangements themselves identify the student as ‘other’ and this inevitably raises debates around advantage where the testing conditions for one particular group (in this case, student with SpLD) are visibly different to the testing conditions that are considered the norm and applied to non-disabled students (Bachan, 2017; Sokal & Wilson, 2017; Spenceley & Wheeler, 2016).

12.7 Solutions

Having identified a range of barriers that exist in current assessment practices, and which contribute to an awarding gap, the participants in this research project suggested a number of potential remedies that they felt would improve the inclusivity and accessibility of assessment and embody principles of equality, inclusion and agency - thereby reducing the awarding gap. These solutions included:

- Co-design of assessments
- Authentic assessment
- Diversified assessment; to include greater coursework, oral assessment and authentic assessment
- Shorter exams, testing knowledge at end of each module (rather than long exams at the end of the year, testing the full year’s knowledge), with a greater amount of extra time
- Clearer guidelines on expectations from exam responses
- The facility to use assistive technology in the exam

- Opportunities for skill development integrated into assessment and locating exams into a pedagogic framework of assessment *as* learning, rather than assessment *of* learning.
- Staff training in inclusive assessment and inclusive curriculum design

In particular, participants felt that assessments should foreground and emphasise the application of knowledge and critical thinking, rather than information recall, and reduce the reliance on construct irrelevant skills, such as processing speed and efficient literacy skills.

12.7.1 Co-design of assessments

The participants suggested that by being actively and collaboratively involved in assessment design, they could suggest adjustments to the design of assessment that can reduce unintended barriers, such as the provision of 'in exam' reference materials to reduce the burden on memory and recall and foreground subject understanding and critical thinking:

> 'Having a formulae or theorem book that you could refer to in the [Maths] exam, or including required material in question prompts, would solve the memorising problem and allow me to show comprehension and what I understand, which is where I can excel.' (Laura)

In addition, for those disciplines, such as Medicine, that already provide candidates with reference material in the exams, participants suggested that employing a co-design approach to assessment could help to avoid the unintended consequences of the current inaccessible location of those reference materials:

> 'We could help with redesigning the booklet. Even simple things such as not having the reference materials printed at the back of the booklet, being able to have that separate so you're not constantly flicking backwards and forwards and then realising that you've forgotten what the reference material said' (Robert)

The participants revealed how the location of resources and reference materials can inadvertently adversely affect the performance of students with SpLD and that providing the reference materials on a separate sheet, rather than appended to the end of the booklet, would improve accessibility**.** Previous studies have similarly found that students with SpLD achieve lower marks than their peers in exams when reference material is printed on the back page, due to working memory deficits that provoke the need to repeatedly turn between the question and the source material (Crisp, Johnson and Novakovic, 2012). Printing reference material on a separate sheet represents an easily achievable, low cost, and non-controversial solution to this barrier. Thus, a co-design approach to assessment, with its ability to capture students' unique insights into the experience of sitting exams, as well as their understanding of their own requirements, has the potential to enable construct irrelevant barriers to be more readily identified and resolved.

12.7.2 More applied, meaningful, and relevant assessment

Arguably, the purpose of traditional exams is to infer from the candidate's performance in the exam the extent to which that candidate has acquired the competencies required by that discipline (Rogers et al 2019; Lightfoot et al, 2018). Ullah (2020), however, suggests that traditional exams 'tend to reveal only whether the student can recognize, recall or "plug in" what was learned out of context'(Wiggins, 1990, p.1), and posits that assessments that reflect real-life situations (i.e. authentic assessment), by contrast, require the student to be able to apply their conceptual understanding in context. Thus authentic assessment, Ullah (2020) argues, has the potential to deepen the students' understanding of their subject and also allows them to meaningfully rehearse the skills required by their professional lives. Reflecting these arguments, the participants in this current research project similarly suggested that assessments should be more authentic and related to the 'real-life' situations they would encounter in their professional life:

> 'Exams should be more related to the real-life tasks that you would have to do in your vocation. In my subject [Law] they could show you a video of a situation or a debate between two people and you could discuss it. Any exam should be testing a range of different competencies.' (Alex)

This concern about the lack of authenticity of assessment practices, and the role this potentially plays in the skills gap, is not new. As early as 2000, Bowden et al identified the gap between the skills that graduates had developed during their study and the skills that their profession required. Yet this debate remains largely unresolved. Tyszko & Sheets (2019), for example, similarly identified the skills gap evident when students enter employment and that, while HEIs express the importance of developing key employability skills in the graduate body, there remains a lacuna between the recognition of the importance of developing these skills and the teaching and learning processes that promote them. Similarly to the participants in this current research, Tyszko & Sheets (2019) argue for the inclusion of authentic assessment opportunities in the curriculum and suggest a co-design approach (between students, HEIs and employers) to re-align assessment and learning towards addressing this skills gap.

12.7.3 Diversified assessment

A number of studies suggest that using a variety of assessment methods (rather than just a timed, closed book, written exam) provides opportunities to develop a more inclusive and authentic assessment design and promote academic learning (Chetcuti et al, 2019; He, Wei & Lu, 2018; Li et al, 2017; Newman, 2019). Similarly, the participants in this project identified the practical and pedagogical value of diversified assessment and the scope this offers for more equitable, inclusive and authentic assessment:

> 'More non exam assessment would help, such as open book exams, more coursework and assignment based assessment, or a portfolio of your best essays across the year..… something with a more creative focus that can use the skill set that I have, that I can really deploy in my subject.' (Daniel)

This recommendation for increasing assessment by coursework as a means to reducing the awarding gap mirrors the outcomes of the quantitative element of this project, which found that participants with SpLD achieve the same levels as their TD peers in coursework, whereas they underachieve, in comparison to their TD peers, in the closed book, timed, written exams. However, diversifying assessment can include a broader range of assessment methods than simply a choice between coursework or timed, closed book, written exams. Boedigheimer et al. (2015), Gaudet, (2015) and Hazen (2020), for example, identify the value of oral exams in facilitating a student in demonstrating their depth of understanding as 'the probing of the oral format can force students to show that they are able to wrestle with the nuances of a question' (Hazen, 2020, p3). Similarly, in this current research project, the participants commented that their oral ability was superior to their written expression and that an oral style exam, where the examiner entered into a dialogue to prompt answers, had the potential to enable them to show their learning more readily than is the case with a written assessment:

> ‘The style that I do best with is when people are asking me questions, oral exams, cos I can explain my ideas well, its just when I write them down its hard to express them in a way that represents the quality of my ideas. ’ (Robert)

12.7.4 Clarifying expectations

The participants in this research project also highlighted difficulties that they experienced in pinpointing the expectations of the exam essay. The solutions they suggested included receiving clearer criteria, both of the content being assessed and also the format of expected responses, as well as feedback and exemplars of responses. This, they felt, would help them to improve their grasp of the ‘ingredients’ of an exam essay and enable them to produce responses that met the criteria more closely:

> ‘It would help to have good guidance for the expectations of the exams, including model answers of what a good exam response looks like. Feedback on your exam answers would be hugely helpful, so you can look at your scripts and see where to improve and how you met, or didn’t meet, the criteria. You could learn from that.’ (James)

12.7.5 Use of assistive technology

As highlighted in the literature review chapter, students with SpLD commonly use a range of assistive technology during their studies to make their academic materials more accessible (Sumner, Crane & Hill, 2020; Degirneci, Baglama, & Yucesoy, 2020). This may include using digital tools to customise the format and presentation of the texts, as well as using speech-to-text software and spelling and grammar checkers to compose written work, text-to-speech software to read texts, and software to convert texts to audio form in order to absorb research and proofread (Dawson et al, 2019; Horn & Huber, 2020; Rauschenberger, Baeza-Yates & Rello 2019). Similarly, participants in this current research project reported that the use of assistive technology represents a key learning strategy that they develop and rely on throughout the duration of their course and were sensitive to the impact of not being allowed access to this strategy during the exams themselves:

> ‘Using text-to-speech software would be really good to have in the exam as I use it in my coursework and its really good for alerting me to typos that I don’t spot when I read my work back myself.’ (Alex)

12.7.6 Training for examiners and test designers

Crisp, Johnson and Novakivic (2011) suggest that the format, style and construction of exam questions can have a disproportionately negative impact on students with SpLD and conclude that ‘those involved in examination design have a responsibility to ensure students [with SpLD] are not disadvantaged’ (p814). Similarly, the participants in this current research project felt it would be helpful for text setters and examiners to receive training in SpLD so that the features of the exam questions and marking approaches do not inadvertently disadvantage candidates with SpLD:

> ‘I think there’s a lot of unconscious bias in marking and the way the questions are set. Examiners having some training about what dyslexia and dyspraxia really means and the way that affects people would be helpful.’ (Juliet)

12.7.7 Increased take up of study skills sessions

In addition to the suggestions around making the course and the exams more accessible, the participants in this research project also suggested inclusive solutions to the lack of take up of 1:1 study skills tuition. The participants identified the perceived stigmatisation associated with the 1:1 specialist study skills support provided to students with SpLD, and the sense of ‘othering’, as one factor in the decision not to access 1:1 study skills sessions, along with the time cost of attending these separate, additional sessions. They commented on where skills provision should ‘sit’ and how students and staff could collaboratively take responsibility for skills development by embedding skill development into the curriculum for all students:

> ‘It would be good if academic skills development could be part of the curriculum, rather than having to attend separate study support sessions which make you feel different. It would just be seen as a normal part of being a student and then it doesn’t matter whether or not you have a disability, you’ll still get to develop those skills that you need to study.’ (Robert)

12.8 Conclusion

The quantitative elements of this project identified a significant gap between the academic performance in timed, closed book exams of students with SpLD and their TD peers. The data showed that this awarding gap was reduced, but not closed, by the exam adjustments of the use of a word processor and/or 25% extra time. At the same time, no awarding gap existed between the coursework performance of SpLD and their TD peers. Despite not being privy to this data, the participants in the qualitative element of this project described similar experiences. That is to say, they commented that exam arrangements were beneficial and improved their exam performance but were insufficient to fully compensate for the construct irrelevant barriers that candidates with SpLD experience in the exam solely due to their neurodiverse profile. They also identified the difference between their coursework marks and timed exams marks and that, unlike the timed exams, the coursework allowed them to more closely perform to their potential.

The participants volunteered a range of reasons clarifying why the exam arrangements help to narrow the awarding gap. For example, they explained that the adjustments allowed them:

- To complete more of the exam.
- To plan and organise ideas better (and thereby improve the essay structure and coherence of their responses)
- To implement strategies that compensate for poor factual recall.
- To reduce anxiety and exam stress (which improved task-focused thinking)

The participants also identified the benefits to their performance of course-based adjustments and study skills support, but still commented that exam adjustments were the single most important factor in boosting their exam performance and reducing the awarding gap. Although the participants unanimously agreed that the exam adjustments made a significant positive difference to their exam grade, they still recognised that the awarding gap persisted, along with a discrepancy between their exam performance and their performance in coursework. They identified the following reasons for this:

- 25% extra time, whilst helpful, is insufficient additional time. At the same time, the extra time is fatiguing, especially where candidates have consecutive exams on the same day and/or on subsequent days.
- Students with SpLD work more slowly across the year and have more difficulty accessing their teaching and learning than their TD peers. Thus, they are exposed to less material than their TD peers, including less revision material and hence are disadvantaged going into the exams.

- Candidates with SpLD were unable to use their assistive technology in exams (such as text-to-speech software, planning software, speech-to-text software, etc) and so the strategies they developed during the year for essay writing were denied them in exams.
- The construct irrelevant skills that are needed to access the exams act as a significant barrier. These include accurate and automatic literacy skills, working memory, memorisation and recall, processing speed and the ability to inherently understand the expectations of a written, essay style, response. Figure 12.2 shows how the underlying cognitive profile of SpLD links to the construct irrelevant skills needed to access exams, resulting in the barriers described by the participants and leading to underperformance in exams (lower marks).

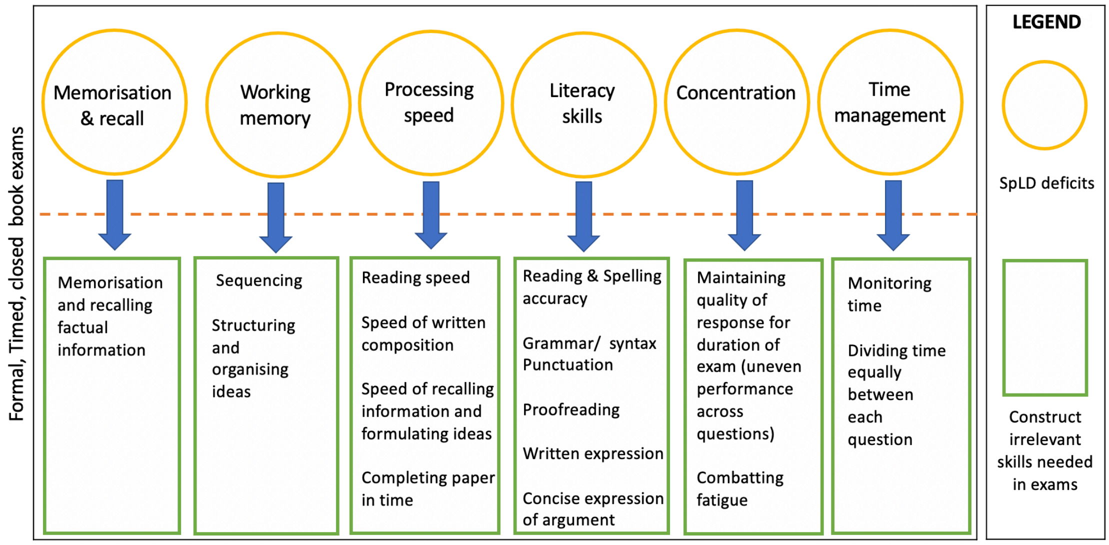


Figure 12.2: Deficits underpinning SpLD and associated exam access skills

Identifying that timed, closed book, written, exams are predicated on a non-disabled norm that is intrinsically disadvantaging for students with SpLD, the participants suggested shifting from a reliance on timed, closed book, written exams as the predominant assessment approach towards an increased focus on improving inclusivity and accessibility of assessment. They argued that closing the awarding gap requires the development of an environment that is accessible to all, and should include the following:

- Co-designed/co-developed assessments.
- Inclusive, universal design for teaching and learning as well as for assessment (addressing problems in the inaccessible teaching and learning environment, as well as in exams).

- Diversified assessments (a shift away from the 'one size fits all' model).
- Ability to use a full range of assistive technology in all assessments, including exams.
- Feedback from exam scripts, continual practice with writing exam responses (with feedback) and clear, written, explicit, guidance on the expectations of an assessment or exam response.
- Questions that focus on understanding, applied knowledge and critical thinking, rather than memorisation, processing speed and literacy skills.
- Authentic, real life, meaningful assessment (assessment *as* learning, rather than assessment *of* learning).

Overall, the first two quantitative stages of the project identified that, far from acting as an advantage, the exam adjustments of 25% extra time and use of a word processor fail to fully level the playing field, albeit they are effective in narrowing the awarding gap. This was similarly described by the participants in this qualitative phase of the project. The inclination from this data is to focus on the exam format to see what alterations can be made to reduce the inequality (such as universal design for assessment and diversified assessments). However, the outcomes of the qualitative phase of the project suggests that this is not the full picture. While considering changes to the format and conditions of the exam is undoubtedly useful, participants have also pointed to the impact of the lack of understanding of the expectations of the exam as well as the 'back-wash' effect of teaching practices that are not fully accessible. This suggests that work around developing inclusive teaching and learning environments, clear learning outcomes from the assessment, explicitly identifying the competence standards being tested in the exam, feedback from the exam scripts to help improve the understanding of the expectations and practice in how to produce an exam essay/response, as well as training for examiners, should all be considered if a level playing field is to be achieved. The accumulated impact from both the assessment process itself and also the teaching and learning that precedes it are experienced simultaneously and cumulatively to have a disproportionately disadvantaging effect, resulting in a less accurate assessment of the subject knowledge, understanding and critical thinking skills of candidates with SpLD.

13: DISCUSSION

This is the discussion chapter of this thesis and it presents the central arguments of the research project. This chapter focuses on analysing, interpreting and explaining the findings that have emerged from the research project as a whole, and relates the findings to the existing literature as well as to the thesis' hypothesis and research questions.

13.1 Introduction

The aim of this research was to investigate whether students diagnosed with SpLD who are permitted 25% additional time, or the use of a word processor with 25% additional time, demonstrate an unfair advantage in timed, written, closed book, university exams compared to their TD peers, or whether the exam access arrangements have simply ensured parity by relieving the students with SpLD of any substantial disadvantage in the exams that might arise as a result of their disability.

This research was provoked by the tension that currently exists in the field of SpLD between the requirement to ensure equal opportunities on the one hand and the imperative to uphold academic standards on the other. Altering the testing conditions of exams for a particular cohort (i.e. through the implementation of exam adjustments) undermines the principle of assessing all candidates in a consistent way (Elliott et al, 2004; Pardy, 2016) and thus raises the concern that exam arrangements potentially advantage one group over another (Bachan, 2017; Sokal & Wilson, 2017; Lovett & Lewandowski, 2015). At the same time, standard time conditions in exams discriminate against students with SpLD due to the impact on access skills[143] of the cognitive deficits associated with SpLD (MacCullagh, Bosanquet, Badcock, 2017; Snowling et al, 2020). Universities have a statutory duty to relieve a student of 'any substantial disadvantage that might arise as a result of their disability' in the exam (Equality Act 2010) and exam adjustments are intended to normalise the performance of candidates with SpLD to that of their TD peers in formal, timed, closed book examinations (Lyman, 2016; McFarlane, 2019). A model of the way in which exam access arrangements are designed to bridge the barriers that students with SpLD experience in exams on account of the deficits in construct irrelevant skills, and enable them to show their academic potential when tested under timed exam conditions, is shown in Figure 13.1

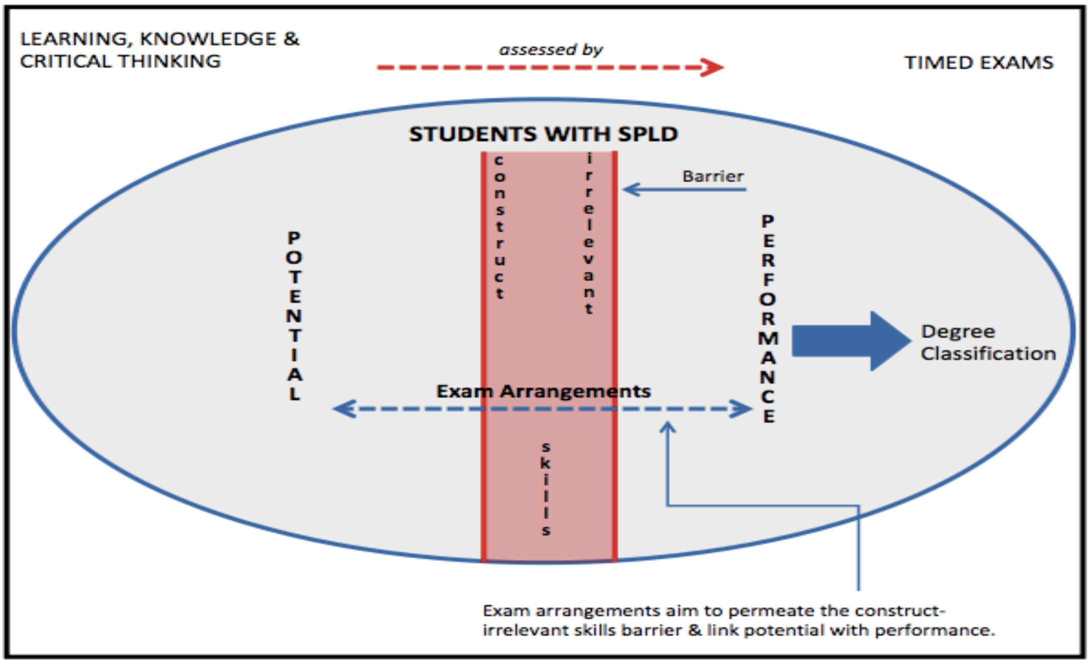


Figure 13.1: A diagrammatical depiction of the underpinning rationale for granting exam arrangements to candidates with SpLD.

[143] These are not the skills being tested by the exam, but deficits in these skills negatively affect the ability of the student to perform to their potential in the exam situation.

This current research project aimed to explore the debate about the fairness of granting exam adjustments by comparing the exam performance of students with SpLD who were granted exam access arrangements with that of their TD peers who took the same exams under standard conditions. This was operationalised by testing the following, two-tailed, hypothesis:

> H_1: As a result of being granted extra time, or additional time and the use of a word processor, in exams there will be significant differences between the length of exam answers and marks of students with SpLD when compared to their typically developing peers.

The first two quantitative phases of this research project involved a deductive, quasi-experimental, methodology and statistical testing of this hypothesis. This was followed by a qualitative phase that used a systematic, inductive approach to understanding the real-life experiences of student with SpLD who used exam access arrangements with the aim of providing a deeper understanding and explanation of why and how the phenomenon of the research project occurs and thus act as a step towards the identification of solutions. This discussion chapter aims to synthesize the results of the quantitative phases of the research project with the qualitative phase of the research project and triangulate with extant literature to develop a model that may explain the exam underperformance of students with SpLD and identify potential interventions.

13.2 Length of exam scripts

Based on the argument that extra exam time potentially advantages candidates with SpLD as additional time results in longer answers and longer answers (as they include a greater wealth of detail and more fully address the question) result in higher marks (Bachan, 2017; Sokal & Wilson, 2017; Lovett & Lewandowski, 2015; Spenceley & Wheeler, 2016), this research project commenced with a comparison of length of exam scripts. This aimed to explore the question of whether students with SpLD who are granted the use of a word processor and/or 25% extra time produce significantly longer exam scripts (and hence receive an advantage in terms of length of answers) than their TD peers who sat the same exams under standard conditions. This phase of the research project focused on students who sat exams in Humanities subjects (and so produced essay style questions), rather than students who sat exams in STEM subjects, given that the exam scripts of the participants in the STEM subjects included formulae, calculations and mathematical or chemical notations, rather than continuous prose.

This current research project found that students with SpLD who handwrote their Humanities exam scripts with 25% extra time produced significantly shorter exam scripts than their TD peers, while participants with SpLD who were granted the use of the word processor in addition 25% extra time produced exam scripts of a similar length to their TD peers who sat the same exams under standard conditions. Thus, the findings of the between group comparisons of length of exam scripts provided sufficient evidence to reject the null hypothesis[144] in relation to the SpLD participants who handwrote their scripts with 25% extra time (as the SpLD group produced significantly fewer words). However, as no differences were identified in word count between the SpLD group who word processed their scripts with 25% extra time and their TD peers, the data failed to reject the null hypothesis in that instance.

The rationale for granting extra time in exams for students with SpLD pivots on the theory that individuals with SpLD process information more slowly (Snowling et al, 2020; Moll et al, 2020) and so require additional time in order to compensate for any 'substantial disadvantage' that would otherwise 'arise as a result of their disability' in a time-constrained exam situation (Equality Act 2010). To establish if the statistical data suggests that the underpinning deficits in processing speed that present core features of SpLD translate into the slower production of information in exams, this research project compared the number of words that the SpLD participants produced in each minute of the exam with the number of words per minute produced by the TD participants. This comparison removed any mitigating effect that additional time may have on the numbers of words produced by the SpLD group in the script as a whole and focused, instead, on the speed with which the exam responses of each group were produced. The findings demonstrated that both the participants with SpLD who handwrote their scripts and the participants with SpLD who word processed their exam scripts produced significantly fewer words in each minute of the exams than was the

[144] The null hypothesis is that there will be no significant differences between the length of exam answers of the SpLD group who were granted exam adjustments and their TD peers who sat the same exams under standard conditions.

case for their TD peers who sat the same exams. Thus, the null hypothesis (that no differences would exist between the SpLD and TD groups) was rejected and the findings support the rationale for granting extra time to candidates with SpLD in time-constrained exams in order to normalize their exam performance, in terms of *quantity* of output, to that of their TD peers.

Overall, the findings from this research project in relation to the length of exam scripts in Humanities exams fails to support the premise that granting additional time or additional time and the use of a word processor to candidates with SpLD results in longer answers (Lovett & Lewandowski, 2015). The data suggests that the use of the word processor in addition to extra time creates equity (but does not confer an advantage) in terms of quantity of output, whereas extra time alone fails to do so. The participants in the qualitative phase of the research project provided a number of explanations of this finding. They commented that the word processor can compensate for slow handwriting and lack of spelling automaticity[145] that students with SpLD experience in exams; characteristics which may result in an additional cognitive burden and higher time cost for students with SpLD who handwrite during exams. This observation is further supported by extant research that found that the cognitive burden caused by poor handwriting fluency had a significant impact on the higher order processes of essay writing in exams, resulting in slower production of work and reduced output as it leaves 'fewer cognitive resources available to the higher order processes needed for composition' (Alamargot et al, 2020; Barnet & Brunty, 200; Connolly, Dockrell & Barnett, 2005, p99; Suarex-Coalla et al, 2020).

13.3 Relationship between word count and mark

The theory that additional time results in longer answers and longer answers (as they include a greater wealth of detail and more fully address the question) result in higher marks (Lovett & Lewandowski, 2015; Zuriff, 2000), pivots on an assumption that a linear relationship exists between word count and mark (i.e. high word count correlates with high marks and low word count correlates with low marks). However, when the data from this current research project was statistically analysed to identify any

[145] The spell-check facility remained enabled during the exams

correlation between word count and mark, a non-linear relationship emerged, with the majority of marks located in a high 2:1 or above being associated with word counts between 2,500 and 4,500 words, while word counts above 4,500 or below 2,500 were more likely to be associated with marks in the low 2:1 range or below. An explanation of this non-linear relationship between word count and mark is shown diagrammatically in Figure 13.2.

TYPE OF RESPONSE	CATEGORY DESCRIPTOR	MEAN WORD COUNT	MARK
Insufficient argument & excessive detail	High word count	>5000	Low Mark (2:2 or below)
Cogent argument & optimal detail	Optimal word count	2,500 to 4,500	High Mark (2:1 or above)
Insufficient argument & insufficient detail	Low word count	<2000	Low Mark (2:2 or below)

Figure 13.2: Model of non-linear relationship between word count and mark

This outcome also reflects the findings of Benton (2017) who compared the length of essay responses in the scripts of 5,000 GCSE English Literature exams to the mark they received. Benton similarly found that very short scripts resulted in low marks, as the candidates had provided insufficient material to evidence their skills to the examiner, yet, at the same time, once essay length increased over a certain point, there was a 'flattening off' effect followed by a reduction in mark as the length continues to increase. Thus, Benton concludes, 'quantity certainly does not trump quality' (p 40).

A significant moderate negative relationship was found in the relationship between word count and marks for TD participants (that is to say, as word count increased marks decreased), while no significant negative relationship was identified for participants with SpLD. This indicates that the candidates with SpLD did not achieve the 'ceiling effect' even with accommodated exam conditions (so still had the potential to increase marks) and suggests that the TD participants were able to show the full extent of their knowledge and ability in the time available and, presumably, extra time (leading to higher word count) is unlikely to significantly improve the performance of students with no identified disability. These findings support the studies that suggest that non-disabled students do not experience barriers accessing the exam under standard conditions, are able to complete the test in the designated time to their full potential, and therefore do not show significant gains in their exam scores when permitted additional time (Crisp, Johnson & Nivakovic 2012; Goegan, 2014; Portolese, Kreuse & Bonner, 2016)

13.4 Differences in exam marks between SpLD and TD participants

A number of studies comparing the academic performance of university students with SpLD with those of their TD peers have found that students with SpLD consistently achieve lower degree classifications than their TD peers (Duncan & Purcell, 2017; Richardson, 2015). Richardson (2015), for example, compared degree outcomes for 4961 students with SpLD with those of their TD peers and found that less than 40% of students with SpLD were awarded a 2:1 classification or above in comparison to over 52% of their TD peers, leading to a conclusion that 'specific learning difficulties have deleterious consequences for the likelihood of academic progression and attainment' (p.335). Similarly, in this current research project statistically significant differences between the exam marks of the SpLD participants and those of their TD peers were apparent in the quantitative data across all three Humanities subjects, as well as in the Maths exams - with the participants with SpLD being awarded lower marks and lower classifications. Although no significant between-group differences were apparent in the overall marks in the Medicine exams, there were significant differences in the degree classifications awarded, with the SpLD group achieving, on average, lower classifications. Thus, the slight overall mark differences, albeit not statistically significant in themselves, were sufficient in magnitude for the TD group to cross the boundary into the 1st degree classification while the SpLD group remained in the 2:1 classification. Overall, therefore, the analysis of the quantitative data from this research project rejects the null hypothesis that there will be no significant between-group differences in exam scores. Between-group differences in exam mark and degree classification were identified, with the SpLD group achieving lower results than their TD counterparts. Therefore, rather than suggesting that the exam adjustments confer an advantage to students with SpLD, this data reveals that the exam adjustments are insufficient to level the playing field and that students with SpLD are disadvantaged in exams, resulting in an awarding gap despite the implementation of exam adjustments.

The participants in the qualitative phase of this research project posited that the difficulties they experience on account of their SpLD with accurate and automatic literacy skills, working memory, memorisation, recall, and processing speed, represent significant factors driving the awarding gap due to the impact of these access skills in exams. This perception is supported by Masterman (2018) who argues that in order to complete the higher-level processes required for essay writing in exams (i.e. critical thinking and application of knowledge), the individual requires automaticity in the lower-level processes involved in text production:

> Writers must (a) be fluent in generating ideas that can be written down and (b) write these ideas down quickly before they are forgotten. If writers are efficient in executing (a) and (b), they will be able to use the metacognitive processes … and other cognitive resources (e.g., genre and content knowledge…) to create reader-based prose. (Masterman, 2018, p.211)

Masterman (2018) suggests that essay writing requires a range of processes, including planning and organising ideas, retrieving facts and knowledge, as well as revising and editing the text and also that, in order to be represented in writing, ideas need first to be put into language and stored temporarily in working memory. Those ‘temporary mental representations’ then need to be translated ‘into more permanent external representations using the symbols of the writing system’ (Masterman 2018, p 205), which also involves ‘retrieving letter forms and familiar word spellings from long-term memory, strategically spelling novel words, and motor planning to produce the letters’ (Masterman 2018, p206). These processes of essay writing therefore depend on automatic literacy skills, effective working memory, efficient recall from long term memory and adequate processing speed – in other words, those very construct irrelevant skills that the participants with SpLD in the qualitative phase of this research project identified as areas of deficit. A model of the processes of essay writing is shown in Figure 13.3.

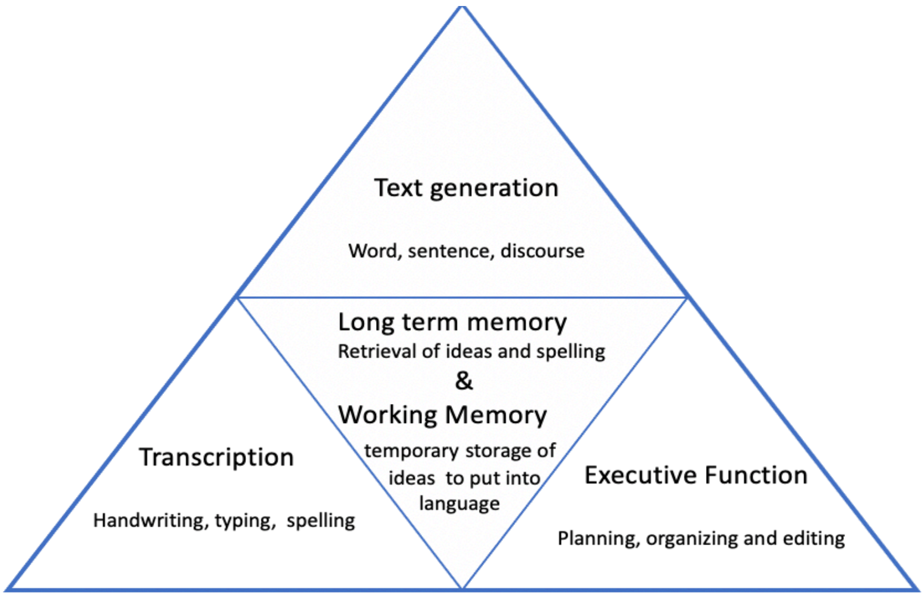


Figure 13.3: Simple view of writing adapted from Barnett (2014), Warren (2017), and Masterman (2018).

This current research project found that the impact on the exam of deficits in these construct irrelevant skills differed, to some extent, according to the requirements of the exams. For example, in Maths exams, the main problems described by participants related to the reliance on rote learning, memorization and recall of formulae. In Humanities exams, the participants ascribed their underperformance to the burden the exam placed on literacy skills, working memory, recall and sequencing/organization of ideas, text generation and transcription. Participants in all discipline areas concurred that these issues were not offset by the extra time and use of a word processor.

Out of all the subjects scrutinized in this research project, the Medicine exams showed the least difference in attainment between the SpLD and TD participants (no significant difference in mean exam mark was identified between the two groups, albeit the SpLD group achieved lower degree classifications than was the case for their TD peers). This outcome reflects similar findings by McKendree & Snowling (2011) and also Gibson & Leinster (2011) who also identified no significant effects of dyslexia on exam results of Medical undergraduate students who received extra time. This may be explained by the more diverse format of exams in Medicine (which includes a combination of multiple choice questions (MCQ), short answers and long answers, rather than entirely essay style responses) while the other discipline areas scrutinized in this research project were entirely essay style questions (Humanities) or long responses to a mathematical question (Maths). It may also be the case that the extremely exacting entry requirements for the course (minimum UCAS points[146] required for entry to the Medicine course = 168, by comparison to the minimum UCAS points required for entry to Humanities course, which is 141[147]) may filter out those individuals with SpLD who are more negatively affected by the construct irrelevant barriers in pre-University exams and attained lower UCAS points as a result, so these individuals do not feature in the sample group[148].

In addition, a number of reasons for the more equitable outcomes in Medicine were suggested by the participants in the qualitative phase of the research project to explain why Medicine exams appear to be less disadvantaging to candidates with SpLD. Participants commented that the questions in Medicine exams are presented as 'real-life' scenarios (the Medicine questions take the form of a clinical setting) and so privilege applied knowledge and understanding over rote memorization and abstract recall of information and also can be visualised in a more concrete way. In addition, the multiple choice questions require a tick box response and the questions in the short answer papers are well scaffolded with the organisation of the content, and the structure of the response, being built into the question (questions may be given in parts, with the candidate providing a single word or short sentence response to each part, or the candidate is asked to complete a table calculating a drug action over time, for example). Thus, the format of these exams places less emphasis on prose construction and the sequencing of ideas (reducing the burden on literacy skills and working memory) than

---

146 Due to limited places and the high competition for entry to the University, applicants commonly offer higher UCAS points than the minimum requirement (data supplied by Business Information Team, AFPA, Finance Division, University of Cambridge, 23.7.2021).

147 Mean UCAS points on entry in 2020/21 for students accepted into Humanities subjects (History, English and Law combined) at the site University was 189 for SpLD applicants and 217 for TD applicants. In the STEM subjects (Maths and Medicine combined) the mean UCAS points on entry in 2020/21 for SpLD applicants was 228 and for TD applicants was 244 (data supplied by Business Information Team, AFPA, Finance Division, University of Cambridge, 23.7.2021).

148 Although the SpLD students accepted on the Medicine degree course in 2020/21 entered with lower UCAS points (M=235) than the TD applicants (M=256), these differences were not of a magnitude to be statistically significant (p=0.202). However, the difference in mean UCAS points achieved by SpLD applicants accepted into English, History and Law degree courses (M=189) in comparison to those achieved by their TD peers (M= 217) were statistically significant ($p<0.01$), with the SpLD applicants achieving significantly lower UCAS points (data supplied by Business Information Team, AFPA, Finance Division, University of Cambridge, 23.7.2021).

is the case with the other subjects included in this research project. Interestingly, the data from the long answer paper in Medicine, which requires the candidate to structure an answer in a format closer to a short essay and so places a heavier reliance on literacy skills and working memory, revealed subtle between group differences, with the data from SpLD group showing a significant difference between their marks in the short and long answer questions (with lower marks in the long answer questions), while their TD peers showed no statistical differences between their marks on the short and long answer questions.

The finding of this current research project that the exam marks of Medical students with SpLD who are granted extra time in MCQ are similar to those of their TD peers reflects the findings in the field. Ricketts, Brice & Coombs (2010) investigated the impact of multiple choice questions on the performance of medical students with dyslexia. Their retrospective statistical analysis (ANOVA) of assessment data from 5 year groups (5 cohorts) in a UK medical school showed that, when students with dyslexia were granted 25% extra time in the multiple choice exams (N=900), their exam performance matched the performance of their TD peers who sat the same exams under standard conditions. They also found that the majority of students completed all the questions in the time permitted. Ricketts, Brice & Coombs (2010) concluded that 'properly-designed multiple-choice tests of medical knowledge do not systematically discriminate against medical students with specific learning disabilities' (p265).

13.5 Relationship between exam mark and dissertation mark

The statistical analysis of the data from this current research project suggests that exam marks and dissertation marks are closely correlated. That is to say, the students who achieved high scores in the exam (relative to the cohort), similarly achieved high scores in the dissertation (relative to the cohort) and vice versa. This relationship between exam marks and dissertation marks implies that the exam and the dissertation are measuring similar skills and further suggests that the dissertation is as reliable an assessment of a student's breadth of subject knowledge, critical thinking and competence in applying their knowledge as is the case for the exam. This close relationship between exam performance and dissertation performance is equally the case for both SpLD and TD participants, albeit the gap between the dissertation mark and the exam mark is greater for the students with SpLD than is the case for their TD peers. This suggests that timed, closed book, written exams discriminate against students with SpLD, whereas coursework does not. This outcome was mirrored in the qualitative phase of this research project, with all ten participants unanimously agreeing that they performed better in coursework and dissertations than was the case in exams. Participants' comments included, "exams are a lot worse than coursework. Coursework is different and is a lot better. Its more enabling" (Juliet); "definitely there is a difference for me between exams and coursework. My exams can go really badly at times, but my coursework has never tanked." (Anna); and "my exam scores were all lower than my essays and coursework, even though I had the extra time and computer" (Bella) - to cite just three examples.

13.6 Phase 2: students diagnosed part-way through the course

Ricketts, Brice & Coombs (2010) identified that, in 2005, 1.4% of all medical school applicants in the UK disclosed an SpLD. However, by graduation, around 3% of all medical students had disclosed an SpLD, representing a significant increase in SpLD being diagnosed after admission. This increase in students being diagnosed with SpLD post admission to degree courses is reflected in the research site institution where, in 2019, around 2.8% of all applicants disclosed a diagnosis of SpLD prior to admission and a further 3.4% were diagnosed part-way through their course. Shaw et al (2017) attribute the high proportions of students seeking a diagnosis of SpLD during their course to the increase in demands of study at degree level, which may render inadequate the compensatory strategies that these students had found to be effective when studying at school. Given that a large proportion of students with SpLD receive their diagnosis part way through their degree course, this provides the opportunity to compare their exam marks when they sat exams without exam arrangements (i.e. before their diagnosis) with the marks they achieved when they sat exams with accommodations (i.e. post diagnosis), and thereby quantify the impact that post-diagnosis support, including exam arrangements, have on academic performance.

13.7 Comparison of marks before and after receiving exam arrangements

The data from the second quantitative phase of this research project, comparing the exam performance of students with SpLD before and after they received exam arrangements, showed that their marks increased by 4% in the year that they were granted exam arrangements by comparison to the previous year (when they sat exams under standard conditions). Prior to the awarding of exam arrangements, the marks of the participants with SpLD increased by a mean of 1% each year, which mirrored the yearly increase of the marks of the TD participants (whose marks increased by a mean of 1% year-on-year). Thus, based on the 'normal maturation effect' demonstrated by the TD cohort, the marks of the students with SpLD were predicted to increase from a mean of 61% in year 1, to a mean of 62% in year 2 and a mean of 63% in year 3. This was, indeed, the case when the participants with SpLD did not receive exam arrangements, but, when exam arrangements were implemented, the marks in the year the adjustments were awarded increased by 4% and then returned to the standard upward trend of 1% for all subsequent years. This additional boost in marks following the implementation of exam adjustments clearly demonstrated that the exam adjustments had a significant positive impact on the exam achievement of the SpLD participants, narrowing (albeit not fully closing) the awarding gap. This was also reflected in the comments of the participants of the qualitative element of the research project:

> "Having extra time and being able to type meant that, compared to my first-year exams, I was able to answer every question and I got much better marks. In my first year I wasn't even able to even see the second half of the paper. I just wasn't able to read that much text, 'cos there's a lot of source material to read and I read slowly." (Daniel)

> "The extra time does level the playing field to some extent, but not completely. I still have to rush, and so I don't do myself justice, but the extra time really does help. I get worse marks without it." (Anna)

Interestingly, if the 'normal' upward trend of year-on-year increases in marks demonstrated by the TD participants can be attributable to the quality of the teaching and learning experience in tandem with student engagement during the course of study (Mourshed, Krawitz and Dorn, 2017), this significant 'boost', over and above the standard upward trend, of the marks of the SpLD participants following the granting of exam arrangements suggests that the awarding of exam access arrangements has a significantly greater impact on exam performance than any other aspect of teaching and learning during the course of study.

It should be pointed out, however, that it may not be possible to fully ascribe this identified improvement in marks of the participants with SpLD to the granting of exam arrangements alone, as there may be other contributing factors and confounding variables influencing this outcome. For example, the observed impact may be, in part, attributable to the suite of interventions that students identified with SpLD are likely to receive following diagnosis (including the awarding of assistive technology, access to study skills sessions, and adjustments to teaching practices recommended by the University's Disability Service), rather than solely to the implementation of exam adjustments. Nonetheless, the dramatic rise in exam results coincident with the awarding the exam adjustments (followed by the smaller, 'standard', increase in marks apparent in subsequent years), together with the finding that the data failed to identify any additional boost to exam marks as a result of study skills sessions, strongly suggests that the exam arrangements in themselves make a significant contribution towards their aim of redressing inequality inherent in the assessment practices that potentially discriminate against students with SpLD, narrowing the awarding gap between students with SpLD and their TD peers in timed, closed book, written exams, albeit they do not fully fashion a level playing field.

13.8 Differences in exam performance between diagnostic sub-groups of SpLD

Drilling down into the data of the 80 participants in the second quantitative phase of the research project (i.e. those who were diagnosed part way through their course), when the sub-groups of dyslexia, dyspraxia and both dyslexia & dyspraxia were separated out, no significant differences in performance were identified between the groups. That is to say, participants with dyslexia performed similarly to those with dyspraxia (and also similarly to those with a dual diagnosis) both when taking the exams under standard conditions as well as when taking the exams with adjustments. The outcome in relation to participants with a co-occurring diagnosis of dyslexia and dyspraxia fails to support extant studies in the field that suggest that having more than one diagnosis has a greater negative impact on learning outcomes than having one diagnosis, as the simultaneous presence of different specific learning difficulties result in multiple difficulties resulting in an additive effect (Kirby 2012; Cappa, Muzio & Giulivi 2012).

However, the outcome of this current research project may be explained by the small number of participants in the second phase of this research project who received a dual diagnosis (N=11), which limits the ability of the research project to reliably detect an effect. Interestingly, this low number of participants disclosing dual diagnosis of both dyslexia and dyspraxia also fails to support research studies that suggest that co-occurrence is more common than single presentation (Kirby 2012; Kirby & Sugden 2007; Kaplan 1998). One explanation of the low numbers of participants with dual diagnosis in this phase of the research project could be that, if co-occurrence is the norm, coupled with the fact that having more than one SpLD has a greater negative impact than having one SpLD alone, students with co-occurring SpLDs are arguably more likely to have been identified and have received a diagnosis earlier in their academic career. The participants in this research project had received late diagnoses, and so it could be argued that these late diagnoses potentially are the result of more subtle 'endophenotype' of SpLD, where the characteristics have not been sufficiently problematic prior to studying at degree level to cause the student to seek a diagnostic assessment at an earlier stage (which is, arguably, less likely to be the case for individuals who have co-occurring SpLDs). This potential explanation is supported by Callens & Brysbaert (2019) who suggest that 'as HE is not compulsory, individuals with dyslexia entering this sector of education are likely to be a specific, highly motivated subgroup with possibly less severe symptoms' (p194). Nonetheless, despite these contradictions, the lack of any statistically significant differences in the exam marks between the three SpLD sub-groups (dyslexia, dyspraxia and dyslexia + dyspraxia) in this current phase of the research project suggests that the exam performance of the participants with dyslexia, both with and without exam arrangements, are sufficiently similar to that of the participants with dyspraxia (and also to those with co-occurring diagnoses) to justify amalgamating these sub-groups into a single SpLD group for the purposes of this research project's data analysis.

13.9 The awarding gap: Identifying the barriers

The findings from the first two phases of this research project are consistent in suggesting that the exam arrangements of 25% extra time, or 25% extra time and use of word processor, fail to level the playing field in terms of word count, mark or degree classification. The second phase of the research project suggests that the exam arrangements narrow the awarding gap between SpLD candidates and TD peers in exams, but do not fully close it. In order to explain this outcome and suggest interventions, it is necessary to pinpoint the specific barriers to learning that the exam arrangements are failing to overcome. That is to say, identifying which systems, practices and challenges continue to act as barriers preventing students with SpLD from being able to fully demonstrate their potential in their exams, despite the awarding of exam adjustments, represents the first step towards the identification of solutions.

A number of studies suggest that the neurobiological and cognitive differences associated with SpLD lead to deficits in construct irrelevant skills needed in exams (Gooch, et al, 2016; Jones et al, 2016; Moll et al, 2016; Perry et al, 2019; Snowling, et al, 2020). These construct irrelevant deficits are experienced in the following key areas:

- Working memory
- Efficient recall of factual information
- Spelling accuracy/fluency/speed / automaticity
- Reading accuracy/fluency/ speed / automaticity
- Writing fluency/speed/ legibility / automaticity
- Planning and organization of ideas
- Sentence structure, grammar and punctuation
- Word retrieval
- Expressive language/ concise written expression
- Processing speed
- Proofreading
- Time management

These deficits associated with SpLD were explicitly identified by the participants in the qualitative phase of the research project as having a negative impact on their performance in timed, closed book, written exams (despite not comprising the constructs that the exam was testing) and were ameliorated, but not fully overcome, by the exam arrangements of a word processor and/or 25% extra time. Thus, as identified in the second quantitative phase of the research, the performance of the participants with SpLD is significantly improved by the implementation of the exam adjustments, but not fully normalized to the performance of their TD peers.

The theory of the underpinning construct irrelevant deficits associated with SpLD, in conjunction with the descriptions of the real-life experiences of the participants with SpLD in the qualitative phase of the research project, suggests that the use of a word processor in the exams should result in higher marks for those candidates with SpLD who used the word processor by comparison to the candidates with SpLD who handwrote their scripts. However, an unexpected finding of this current research project was that, although the participants in the qualitative phase of the research project attested to the benefits of using a word processor in the exams (in terms of reducing the cognitive burden imposed by the difficulties they experience with handwriting fluency and spelling as well as with structuring, organising, and expressing their ideas), and the quantitative phase also showed that the use of the word processor levels the playing field in terms of quantity of text production, the statistical data showed that there were no significant differences between the exam marks achieved by the SpLD group who used a word processor with extra time and the exam marks achieved by the SpLD group who handwrote their scripts with extra time. That is to say, the use of the word processor in addition to extra time made no statistical difference to marks by comparison to extra time alone, albeit it improved the quantity of output and the participants' perceived quality of the exam response. If the perception of the participants was that the word processor improved the structure and fluency of their writing in exams, why did this not make a statistical difference to their marks?

One explanation could be that, by virtue of their need for a word processor in addition to extra time, the participants with SpLD who used a word processor may inherently have a greater severity of deficits in handwriting fluency and the ability to structure ideas in writing than is the case for their SpLD counterparts who required extra time only. If this is indeed the case (and the use of the word processor in addition to extra time normalized their performance to that of their SpLD peers who had extra time only) it would be expected that the participants who used a word processor would potentially achieve lower marks than their SpLD counterparts who only used extra time when neither group were allowed access arrangements. However, the statistical analysis of the data in the second quantitative phase of this research project revealed no significant differences in the exam marks either before or following the implementation of exam arrangements for the SpLD group who handwrote their exam script with 25% extra time by comparison with the SpLD group who word processed their exam scripts with 25% extra time (p = .165). Thus, the lack of any 'differential boost' resulting from the use of a word processor in addition to the extra time, is not explained by any differences in sub-group characteristics. An alternative explanation of the lack of differential improvement in performance occasioned by the use of a word processor, despite the perceptions of the participants to the contrary, may relate to the unconscious effect that typed scripts can have on examiners' decision making when marking. Mogey & Purcell (2016), Mogey & Sarab (2006) and Russell & Tao (2004), for example, all found that examiners marked word-processed scripts more harshly than handwritten ones, due to an unconscious expectation that a word-processed script should be more 'polished', with Mogey & Purcell (2016) identifying that disabled students who use word processors are unduly marked down by around 2%.

The outcome of this current research project in relation to the lack of exam mark improvement attributable to the use of a word processor also conflicts with the findings of a similar study by Duncan & Purcell (2017) which showed that exam candidates with SpLD who word processed their scripts achieved higher marks in comparison to the SpLD participants who handwrote their exam scripts. Whilst the study by Duncan & Purcell (2017) used a similar methodology to this current research project, the sample size in the 2017 study was smaller (36 SpLD participants using a word processor in the 2017 study's sample group, by comparison to 121 SpLD participants using a word processor in this current research project), giving the earlier study a lower power[149] (80% power by comparison to 95% power in this current research project). Also, it should be pointed out that, although a significant difference was identified between the marks of the SpLD group who handwrote their scripts (mean of 64%) in comparison to

[149] A lower power 'reduces the likelihood that a statistically significant result reflects a true effect' (Button et al, 2013, p365).

the SpLD group who used a word processor (mean of 65%), this difference constituted only a small effect size (d = .21), suggesting little material difference in the marks between the two groups despite the use of the word processor. Interestingly, a review by Masterman (2018) of 46 studies comparing the marks achieved by candidates in handwritten exams with those achieved in typed exams concluded that the studies in this field are inconsistent and contradictory and that no consensus exists on whether or not candidates who use a word processor achieve higher marks than those who handwrite their scripts. Perhaps much depends on the keyboard skills of the individual candidate, as any lack of automaticity may impose a high cognitive burden. Kohler (2015), for example, argues:

> '…..lack of fluency in lower order cognitive processes such as keyboarding or handwriting constrain higher order cognitive processes….To this end, it might make sense that less fluent typists would be forced to spend more time on lower order processes as opposed to higher order processes that have to do with the content and organization of their ideas in essays' (Kohler, 2015, pp140-141).

While the unconscious bias associated with marking a typed script may have the effect of disproportionately penalizing candidates with SpLD who use a word processor in exams, it may also be the case that candidates with SpLD who handwrote their scripts are penalized for poor presentation and legibility. This marking bias may be one factor in the difference in exam marks between the SpLD and TD groups. For example, research by Greifeneder, Alt, Bottenberg, et al (2010) found a marking bias when handwritten scripts of university students are assessed, with examiners tending to mark more legible essays higher than those that were harder to read, and similarly penalized scripts where the handwriting was less fluent. Although this research did not focus on candidates with SpLD in particular, given that copious crossed out words, arrows pointing to the order of ideas, lower legibility and less fluent handwriting are characteristic of the written work of candidates with SpLD when under time pressure in exams (Suarex-Coalla et al, 2020), it may be surmised that these surface features are likely to be negatively influencing examiners' marking. Thus, it could be argued, the participants in this current research project with SpLD who had 25% extra time and handwrote their scripts are likely to be affected by marking bias due to the errors in presentation, while those who word processed their scripts with 25% extra time are also likely to have been affected by marking bias due to the expectation that their scripts should be more polished than the handwritten ones of their TD peers. As a result, both groups are being marked down for the presentation of their scripts (albeit for different reasons) – a situation that potentially occurs less commonly for their TD counterparts, whose scripts are handwritten, rather than typed, and also likely to contain fewer presentational errors than is the case for candidates with SpLD.

13.10 The awarding gap: Factors outside the exam

As the focus of this research project is on the difference between the marks achieved by students with SpLD and those of their TD counterparts in time-bounded, closed book, written exams, the tendency is to focus on barriers to performance arising in the actual exam itself. However, it should also be borne in mind, that the awarding gap between the exam mark of the SpLD candidates and those of the TD peers may not be wholly attributable to the exam itself. That is to say, while the awarding gap may be driven by deficits in exam construct irrelevant skills that are not being fully compensated for by existing exam access arrangements, problems that students with SpLD experience in accessing the course during the year on a level playing field with peers may also represent a contributory factor. For example, there is a general consensus that students with SpLD in HE tend to experience problems listening and taking notes in lectures simultaneously (and so have poorer lecture notes or learning from lectures), read slowly (and so fail to get through the same quantity of material/information as peers), have revision difficulties (due to problems with memorization), take longer to get through the revision material, and have poorly organized notes from their learning during the year from which to revise (Biotteau et al., 2017; Blank et al, 2019; Byrne, 2018; Cacola & Lage, 2019; Kendall, 2018; MacCullagh, Bosanquet & Badcock, 2017; O'Byrne, Jagoe & Lawler, 2019; Shaw & Anderson, 2018). All of these difficulties may mean that students with SpLD have not been exposed to the same quantity or quality of learning and revision as their TD peers on the same course when sitting the same exams. A diagrammatical depiction of the range of barriers to learning that exist for students with SpLD, which may negatively affect their performance in exams, is shown in Figure 13.4.

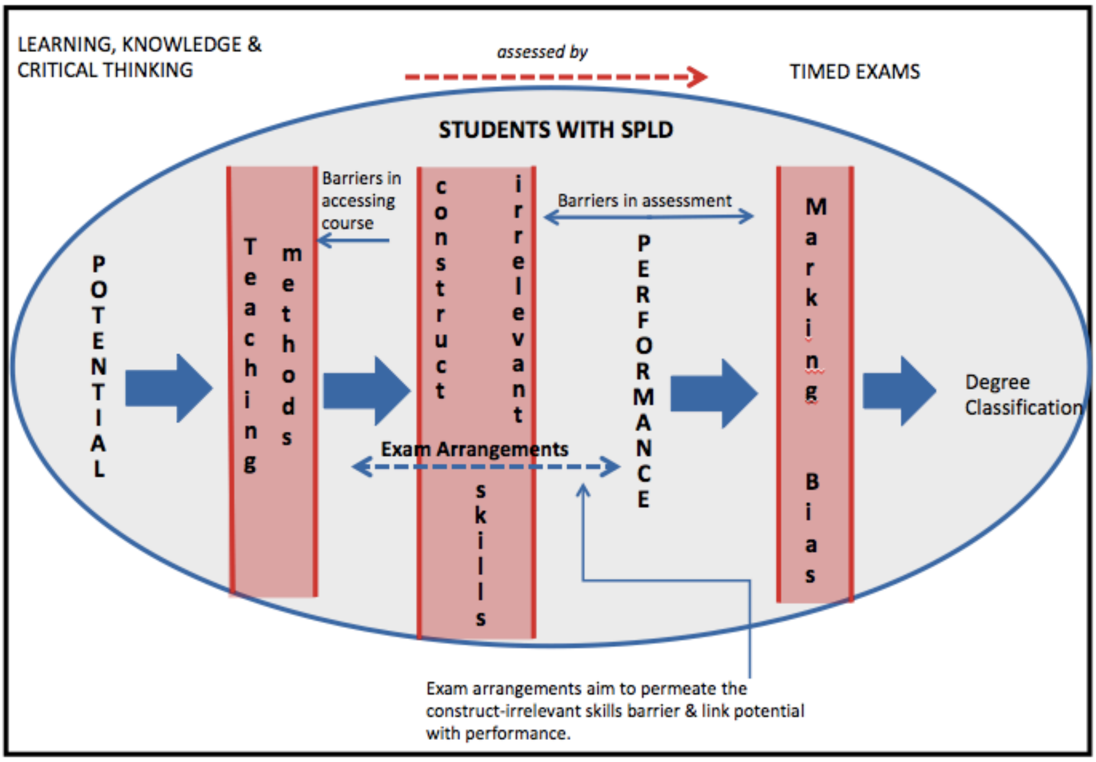


Figure 13.4: The barriers to learning that exist for students with SpLD

13.11 The awarding gap and academic self-concept

One area that has not yet been considered, but may potentially be a factor contributing to the awarding gap, relates to the impact of academic self-concept on academic achievement. Academic self-concept can be defined as an individual's self-perceptions of their academic ability and potential and is 'formed through experience with and interpretations of his or her environment' (Marsh & Hattie 1996, p.58) together with 'either estimates of how good one is at a given activity or expectations for one's future performance' (Wigfield et al. 1997, p.451). Thus, academic self-concept does not equate entirely with self-esteem or confidence but relates to the individual's perception of their own academic efficacy and ability as defined in comparison with their peers and other external frames of reference. Nandika (2020) and Cvencek et al (2018) posit that there is a substantial, bi-directional, relationship between academic self-concept and academic achievement and conclude that higher achievement is underpinned by a secure belief in one's capability to regulate one's own learning. This suggests that robust and resilient academic self-concept is vital to high academic achievement.

There is a general consensus that individuals with SpLD present with lower self-concept than their TD peers and that this may contribute to academic under-achievement (Elgendi et al, 2021; Livingston, Siegal & Ribary 2018; Stagg et al, 2018). Lower academic self-esteem and self-efficacy was also described by the participants in the qualitative phase of this current research project:

> "It [dyspraxia] makes me feel different to others and not as able, and so I have a real problem with imposter syndrome." (Alex)

> "You don't feel so worthy. How can I be part of this cohort when I have dyslexia and am constantly behind?" (James)

Participants explained how they found that the need to have different exam conditions to their peers had a negative impact on their academic self-concept and self-esteem, as this 'special treatment' acted as a visible reminder of their past experiences of being othered:

> "The exams adjustments make you feel that you are seen as 'special' as in 'special educational needs', not in a good way. All those SEN, special educational needs, feelings I had all throughout my schooling were right back present in my University studies. I felt exposed and it made me feel less able." (Joanna)

This experience was similarly identified in a study into students' self-perceptions of SpLD by Majer (2018). Majer found that the impact on self-esteem of the stigma associated with SpLD persisted throughout the individual's lifetime, rather than being time-limited and societal perceptions of a direct link between literacy skills and intelligence, emphasised by SEN policies, is internalised by the individual with SpLD from an early age, with the result that a susceptibility to negative academic self-esteem and self-concept remains lifelong and is readily triggered by subsequent experiences.

It may, therefore, be argued that the weakened academic self-concept and vulnerable academic self-esteem associated with SpLD may have played a role in the exam performance disparities identified in this current research project. This premise is supported by research by Gibson et al (2015) into the awarding gap between female and male Physics undergraduate students, which found that female students present with lower academic self-concept than their male counterparts and perform less well in exams as a result (achieving lower marks and lower degree classifications than their male peers in the exams, albeit no between group differences were identified in coursework). Gibson et al (2015) argue that the male students have inherently greater confidence in both their innate ability and their skill in using examination techniques to good effect, with the result that they prepare for exams by combining revision with developing exam techniques with the aim of producing an excellent exam performance. By contrast, the female students prepared for exams by revising hard, as their aim was to show the extent of their understanding. Gibson et al (2015) suggest that students who emphasize understanding and seek a creative exploration of multiple texts can stockpile large quantities of complex material, leaving them with little time to distil their knowledge and order their thoughts in the time-bounded exam and so may find themselves overwhelmed by a wealth of details from which it is difficult to extricate a clear argument. Therefore, given that these differential approaches to learning and exams arise from differences in academic self-concept between the two groups, and contribute to a gender awarding gap, it could be argued that this finding that low academic self-concept has a deleterious effect on exam performance may be applied to students with SpLD who, like the female students in the study by Gibson et al (2015), have lower academic self-concept than their TD peers.

13.12 Conclusion

The purpose of this research was to explore whether or not the practice of granting exam access arrangements to students with SpLD achieves equity of opportunity or confers an advantage in time-constrained, closed book, written exams. This research is located in the debate between the statutory duty to relieve students with SpLD of 'any substantial disadvantage that might arise as a result of their disability' in the exam on the one hand (Equality Act 2010), and the need to maintain academic standards and safeguard the integrity of the exam for the whole student body on the other. That is to say, exam arrangements should ensure fair access to all, but should not give one group of students an advantage over others.

This research project tested the hypothesis that, as a result of being granted extra time or additional time and the use of a word processor in exams there will be significant differences between the length of exam answers and marks of students with SpLD when compared to their typically developing peers. The statistical analysis rejected the null hypothesis, as, rather than identifying no difference in the marks or length of exam script between the two groups, the SpLD sample group achieved statistically significantly lower marks, lower classifications and, in the case of those who handwrote their scripts, lower overall word counts on their exam scripts. The research project failed to reject the null hypothesis in relation to the word count on scripts of the sub-group of the SpLD participants who used a word processor (rather than handwrote their scripts) and the marks of the SpLD participants in Medicine exams overall (which were not statistically different to those of their TD peers), albeit statistically different degree classifications were awarded to the SpLD and TD groups in Medicine exams (with the SpLD group receiving lower degree classifications than their TD peers).

Overall, the data from this research project does not suggest that students with SpLD are advantaged by the exam adjustments. When students with SpLD were granted the use of a word processor and/or 25% extra time they did not produce longer exam scripts or achieve higher marks than their TD peers who sat the same exams under standard conditions. In fact, quite the reverse, with the data revealing that participants with SpLD who were granted exam arrangements achieved significantly lower marks and lower classification than their TD peers and, in the case of those who handwrote their answers with extra time, produced significantly fewer words on the exam scripts. Thus, concerns that the granting of exam adjustments (use of word processor and/or 25% extra time) to students with SpLD over-inflates their marks and undermines academic standards (Bachan, 2017; Sokal & Wilson, 2017; Lovett & Lewandowski, 2015; Spenceley & Wheeler, 2016) are not supported by the findings of this research project.

The purpose of exam access arrangements is to ensure that all students have an equal opportunity to demonstrate their potential in exams, regardless of disability, by removing barriers caused by a lack of access skills. However, the findings of this research project suggest that the exam arrangements received by students with SpLD in University exams (the use of a word processor and/or 25% extra time), while reducing inequality and improving accessibility, fail to fully achieve their purpose of leveling the playing field. Thus, in addressing the question ‘Equity or Advantage?’ the research project has found that the exam access arrangements conferred **neither** equity **nor** advantage to students with SpLD in university exams, and, worryingly, the exam adjustments proved insufficient to offset the disadvantage that candidates with SpLD experience in the exams, resulting in a persistent awarding gap.

14: CONCLUSION

This chapter presents the thesis' conclusions. The chapter commences by outlining the circumstances that led to the conception of this research project and then goes on to broaden out the implication of research project's findings to the national debate around exam and assessment processes in HEI in general. The chapter also considers the wider, societal, ramifications of the identified awarding gap experienced by students with SpLD and explains how this research project has meaningfully contributed to knowledge. This chapter concludes by making recommendations for developments in assessment policy and practice, and also identifies where further study would be helpful.

14.1 From 'unfair advantage' to 'mind the gap': The journey of this research

The path that led me to carry out this research project arose from my day-to-day experiences as a disability practitioner working with students with SpLD in Higher Education. My professional role focuses on identifying and reducing the barriers which students with SpLD experience when engaged with their learning, and ensuring that these students enjoy fair and equitable access to their education. Reasonable adjustments[150], including exam access arrangements, play a vital role in mitigating

---

[150]Section 20 (1) of the Equality Act 2010 imposes the duty on HEIs to make reasonable adjustments where 'a criterion or practice … puts a disabled person at a substantial disadvantage…in comparison with persons who are not disabled… A provision, criterion or practice does not include the application of a competence standard' (Equality Act 2010). This enshrines the requirement to apply reasonable adjustments to assessment practices and how exams are administered, provided those adjustments do not conflict with the competence standards being tested. Reasonable adjustments are defined as 'positive steps to ensure that disabled students can fully participate in the education … provided for students…to avoid as far as possible by reasonable means the disadvantage which a disabled student experiences

barriers to learning and allowing students with SpLD to demonstrate their academic potential on a level playing field with their TD peers, and these adjustments can have a significant impact on the academic experiences and outcomes of students with SpLD (Equality Challenge Unit, 2020; Equality & Human Rights Commission, 2019).

In my professional role, I regularly experienced intellectual challenges from university academic and support staff in relation to recommendations for the awarding of extra time (or extra time with the use of a word processor) to students with SpLD in timed, unseen, closed book, written exams. This opposition was located in concerns related to the upholding of academic rigour, fairness or parity. It was posited that the awarding of exam adjustments potentially conferred an unfair advantage as students with SpLD were then able to produce longer scripts with a greater amount of detail and enjoy additional time to consider and refine their responses; which would translate into over-inflated marks. In my professional view, this argument appeared rooted in a misconstruction of the nature of the impact of SpLD in exam settings (and of the effect of the exam adjustments that aim to reduce that impact), on an unquestioning acceptance of unsubstantiated reports that exam adjustments result in an unfair advantage, as well as on an uncorroborated assumption that additional time necessarily equates to higher (over-inflated) grades.

In accordance with the principles of evidence-based practice, I turned to the extant literature to resolve the debate, by seeking empirically supported responses to the question of whether granting exam access arrangements to students with SpLD promotes equity or confers an advantage. However, far from settling the debate, my survey of the literature revealed that this was an international discourse with deeply entrenched and strongly held views on either side, but with no unified conclusions or overall consensus. A review of the literature exposed an intense and persistent dissensus around the question of fairness of exam adjustments; that is to say, whether or not the granting of extra time simply compensates for the student's disability and removes the barrier that they would otherwise experience in being able to perform to their potential,

---

because of their disability' and aim to redress disadvantage 'where a provision, criterion or practice puts disabled students at a substantial disadvantage'(Equality & Human Rights Commission, 2019).

or if it benefits all candidates and so undermines the validity of the exam by over-inflating scores (Gregg & Nelson, 2012; Foster, 2014; Payne-Tsoupros, 2020). At the same time, there was a notable paucity of empirical research conducted in the context of real-life university timed, closed book, unseen, written exams. Thus, the extant studies lacked sufficient ecological validity to generalize the findings to a university setting and so reconcile the University's debate on equity or advantage. This unresolved discourse around the potential effect of exam arrangements on students with SpLD among university staff, alongside the lack of non-contradictory, ecologically valid, empirical, research that could help reconcile the debate, in concert with my day-to-day experiences of the impact barriers in exams have on students with SpLD, formed the backdrop against which this research project was conceived and galvanized me into conducting primary research to address this gap.

14.2 Broadening out the research project: The national picture

Despite having its initial inception in a local university debate, this research project speaks to a much wider, national, discourse. Most Higher Education Institutions (HEIs) in the UK subscribe to the widening participation objectives of The Office for Students (OfS) and implement a range of widening participation initiatives in order to broaden the diversity of the student cohort. This, alongside increasing rates of diagnosis in the general population and a greater willingness among students to disclose a disability, may account for the increase in the proportion of students from previously under-represented groups, including students with SpLD, attending HEIs[151] (Advance HE, 2020). With this objective of redressing inequality of participation and attracting a cohort that is representative of the population (including students with SpLD) comes a duty to ensure that all enrolled students have equal opportunities to reach their potential in their course of study (Rodger et al, 2015; Equality Act, 2010).

Despite the changes in the demographics of matriculated students resulting from the widening participation policy (including increased numbers of students disclosing an SpLD), testing via, timed, closed book, unseen, written, exams remains the predominant element of the assessment strategy and continues to enjoy a largely

[151] In 2019/20, 119,600 students with SpLD were enrolled in Higher Education. This represents almost 5% of the entire HE student body in the UK (Advance HE, 2020).

unquestioned standing in HE (Bengtsson, 2019; Villerroel, et al, 2019). While the goal of this pedagogic model is to assess the knowledge, skill and ability of all of the candidates in the cohort in a uniform way, the process by which these exams are administered may act as construct-irrelevant barriers to students with SpLD and prevent them from demonstrating their knowledge, skills and ability in a way that reflects their potential (Asghar et al, 2018; Chong, 2018; MacFarlane, 2019; Payne-Tsoupros, 2020).

The Equality Act 2010 obliges HEIs to provide reasonable adjustments to disabled students (including students with SpLD) in order to offset the consequences of their impairment on their performance and enable them to equitably participate in their education (Equality Act 2010). Thus, in order to reduce the risk of being unfairly disadvantaged by the conditions under which the exam is conducted, and ensure that exam results are valid measurements of the knowledge, skills and ability of all students in the cohort, HEIs make adjustments to the administration conditions for students with SpLD in exams; the most common of which are allowing candidates with SpLD the use of a word processor and/or 25% extra time (Ofqual, 2020). The purpose of these exam access arrangements is to ensure that the exam process does not prevent students with SpLD from demonstrating their true knowledge, ability and skills where standard exam conditions may discriminate against a candidate with SpLD. That is to say, these adjustments are intended to remove construct-irrelevant barriers and promote parity of assessment, where time is not the construct being measured by the exam (Biotteau et al., 2017; Blank et al, 2019; MacCullagh, Bosanquet & Badcock, 2017).

However, granting exam access arrangements to candidates with SpLD has the effect of altering the standardised testing conditions, thereby changing the administrative conditions for certain candidates and thus violating the principle of assessing all candidates in a uniform way. This provokes questions of fairness, the maintenance and validity of assessment standards, and academic rigour, as it raises the concern that exam arrangements could, for those that receive them, potentially change what the test measures. As a result, the policy of granting exam adjustments to students with SpLD is contentious, with opponents of the practice arguing that, rather than creating parity of assessment, exam access arrangements may inflate the results of those students to whom they are granted and thereby lead to an unfair advantage, undermine the integrity of the exam for all candidates and compromise the comparability of exam results across the student cohort (Bachan, 2017; Sokal & Wilson, 2017; Spenceley & Wheeler, 2016). This tension between the requirement to ensure equal opportunities on the one hand and the need to uphold academic standards on the other, that operates across all HEIs, was at the heart of this research project's examination of the effect of granting exam access arrangements to students with SpLD.

14.3 Research objectives: Summary of findings and implications

The overall aim of this research project was to explore the debate of 'equity or advantage' and address the gap in extant research by investigating whether permitting students with SpLD to use a word processor and/or 25% extra time in exams creates parity or imparts an advantage. This was examined by testing the following hypothesis:

> H_1: As a result of being granted extra time or additional time and the use of a word processor there will be significant differences between the length of exam answers and marks of students with SpLD when compared to their typically developing peers

The hypothesis was tested through a comparison of the differences in exam performance between students with SpLD who were granted exam arrangements and TD students who took the same exam under standard exam conditions. The main areas of exam performance scrutinized were the length of exam paper (to ascertain whether the SpLD candidates granted the use of a word processor and/or extra time produced longer scripts than their TD peers) and mark (to ascertain whether the SpLD candidates granted the use of a word processor and/or extra time achieved higher marks than their TD peers). This was based on the premise that additional time results in longer answers and longer answers (as they include a greater wealth of detail and more fully address the question) result in higher marks (Lovett & Lewandowski, 2015; Zuriff, 2000).

The research outcomes demonstrated that, statistically speaking, the candidates with SpLD who were granted exam access arrangements (use of a word processor and/or 25% extra time) *did* perform differently to their TD peers who took the same exams under standard conditions. However, rather than over-inflating their performance, despite the extra time, the students with SpLD achieved significantly lower exam marks and degree classifications than their TD peers, and, in the case of those Humanities students who handwrote their scripts, lower total word count on their exam scripts. Only those Humanities students who word processed their scripts produced exam scripts of a statistically similar length to those of their TD peers. A comparison of the exam marks achieved by the SpLD and TD participants in this research project as a whole revealed that the TD participants received more marks in the upper 2:1 (above 65%) and 1st degree (above 70%) category than the SpLD participants, while the SpLD participants were awarded more marks than their TD peers in the lower 2:1 (60 – 64%), 2:2 (50 – 59%) and 3rd degree (below 50%) categories. This is shown diagrammatically in Figure 14.1.

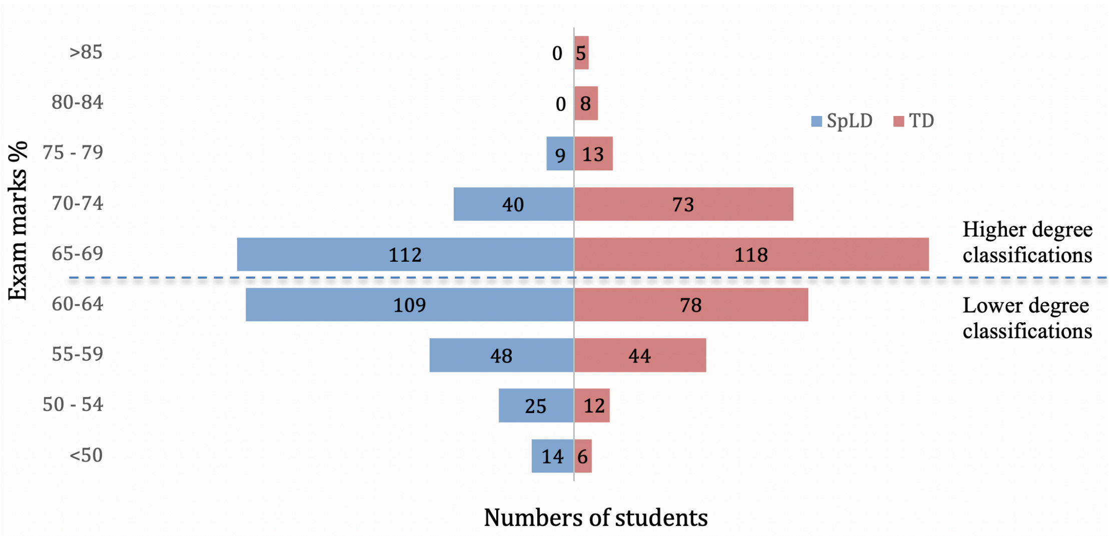


Figure 14.1: A comparison of all exam marks achieved by SpLD and TD participants (714 participants in total).

In addition, the data from this research project did not support the theory that longer answers equate with higher marks. Both very long scripts and very short scripts in the Humanities exams achieved lower marks, with a ‘sweet spot’ of between 2,500 and 4,500 words on the script (between 850 – 1500 words per essay) correlating to marks in the 2:1 or above category. This was the case both for the SpLD and the TD participants. However, a significant moderate negative relationship was found in the relationship between word count and marks for TD participants, whereas none was identified in the case of the SpLD participants. This suggests that TD participants reached the ceiling of their knowledge and ability in the standard time available and so, theoretically, extra time (leading to higher word count) was unlikely to significantly improve their performance, while the SpLD participants were unable to achieve this ‘ceiling effect’ and so further extra time, theoretically, leads to an improved performance. Thus, overall, these outcomes fail to reject this research project’s hypothesis and they refute the commonly held belief that the granting of exam access arrangements to students with SpLD confers an advantage (Bachan, 2017; Sokal & Wilson, 2017; Spenceley & Wheeler, 2016; Lovett & Lewandowski, 2015; Zuriff, 2000). Instead, the statistical outcomes of this research project support the findings of studies that suggest that, rather than conferring an advantage, the awarding of additional time is insufficient to enable students with SpLD to achieve at a similar level to their TD peers in timed, closed book, unseen, written exams (Camilleri et al, 2019; Cameron, 2016; Checuti et al, 2019).

Furthermore, the exam marks of the participants with SpLD, unlike those of their TD peers, fell significantly below that of their coursework marks, despite the coursework marks of both the SpLD and TD groups being statistically similar[152]. The between-group similarity in coursework marks suggests that there is no ability difference between the two groups, while the differences in *exam* marks between the two groups suggests that the candidates with SpLD are disadvantaged in timed, closed book, unseen, written exams, despite the exam adjustments. Given that the exam adjustments clearly fail to level the playing field, this raises the question of the extent to which the extra time offsets the disadvantage of the exam format for students with SpLD, if at all. To explore this question, the second phase of the project compared the exam marks that the participants with SpLD achieved when they sat their exams under standard conditions (i.e. before they were diagnosed with SpLD) with the exam marks that these same individuals achieved when they sat their exams with exam adjustments (i.e. after they were diagnosed with SpLD) and compared these outcomes to the marks achieved by their TD peers who sat the same exams under standard conditions over the same time period. The outcomes of this second phase of the project revealed that the implementation of exam access arrangements make a significant positive difference to the exam marks of candidates with SpLD and significantly reduces, albeit fails to fully eliminate, the awarding gap that exists between the SpLD and TD participants.

[152] Both the SpLD and TD groups produced their coursework under similar conditions as no extra time was granted to those with SpLD and both groups used a word processor.

Given the persistent disparity in exam performance between SpLD and TD candidates identified by this research project, despite the implementation of exam adjustments aimed to level the playing field, it would appear that the current system is failing to adequately remove the disadvantage which a student with SpLD experiences because of their disability. Thus, despite the extra time (or the use of a word processor with extra time) barriers that prevent students with SpLD from being able to demonstrate their full potential in timed, closed book, unseen, written, exams remain. This finding moved the central question of this research project away from a focus on whether or not the exam arrangements achieve parity or confer an unfair advantage to students with SpLD, to a scrutiny of the reasons why an awarding gap persists despite the implementation of exam adjustments.

As a first step towards identifying the factors driving the awarding gap that were not overcome by the exam adjustments, 10 students with SpLD who had been granted exam arrangements were interviewed about their exam experiences. The aim of these interviews was to glean insights into the barriers to exam performance that are both overcome by exam arrangements, and also the barriers that remain despite the granting of exam access arrangements. This, it was hoped, could also act as a first step towards identifying potential interventions aimed at resolving the awarding gap.

The participants in this third, qualitative phase of the research project, reported that the extra time was valuable in improving their exam performance, as it allowed them more time to go through the various stages of compiling an essay or exam response, enabled them to complete more of the paper, and reduced anxiety, which, in itself, helped them to think more coherently during the exam. However, they felt that 25% extra time was insufficient to compensate for the extent of their slow cognitive processing speed, although they also reported that more extra time would be counter-productive due to the cumulative impact of fatigue[153]. In addition, the participants described a number of barriers in exams that were not addressed by extra time or the use of a word processor. These included the emphasis closed book, unseen, written exams placed on fluent literacy skills and working memory, as well as on the need to be able to efficiently plan, structure, sequence and edit ideas, recall factual information, and be able to readily

[153] Exams were commonly held on consecutive days, over the course of a 2 week period.

identify the expectations of an exam essay. In addition, the participants were prohibited from using their assistive technology in exams (such as text-to-speech software, planning software, speech-to-text software, etc.) and so the strategies and techniques they had relied on, and honed, during the year for essay writing were not available to them in exams.

The participants also reported that a number of factors contributing to the awarding gap could be traced back to inaccessible teaching practices that pervaded their course of study and which created a 'backwash' effect on their exam performance. In addition, due to their slower speed of working they devoted more time to study and revision than was the case for their TD peers, but were still unable to cover as much material as their peers. Thus, they entered the exam room already at a disadvantage. Described in this way, it appears that the negative impact of having an SpLD in an educational context is iterative in nature, building up from being disadvantaged by a lack of fully inclusive teaching practices during the year, compounded by inefficient revision process, and further aggravated by construct-irrelevant barriers in the exam – culminating in an underperformance in the exam itself. In contrast, however, participants felt that the longer deadlines assigned to the production of coursework provided them with adequate time to write fluently, research ideas, incorporate feedback, revise and edit their work and so formulate and express ideas cogently. Thus, the participants with SpLD unanimously agreed, they are capable of demonstrating their potential in coursework, but not in the time-constrained environment of traditional closed book, unseen, written exams.

14.4 The wider ramifications of the research

Exams are not 'islands' existing in a bubble and the impact of the barriers that students with SpLD experience in demonstrating their potential in exams has far-reaching consequences. Students with SpLD are at higher risk than their TD peers of discontinuing study prior to completion[154], and have lower incidences of progression to postgraduate study as a result of poorer exam performance (AHEAD, 2017). At the same time, the class of qualification students obtain can affect employment

[154] The non-continuation rate for students with SpLD is reported to be around 9%, compared with 3% for their TD peers. This is a statistically significant difference (Hubble & Bolton, 2020)

opportunities, access to training scholarships[155] and enrolment on competitive postgraduate courses. Thus, exams play a substantial and influential part in the discourse of meritocracy in education and beyond. That is to say, students achieving the highest exam marks in school gain access to the most prestigious universities, and then those who achieve the highest degree classification in university are at an advantage in a decidedly competitive employment market and may more readily access highly rewarded careers (O'Byrne, et al 2018). Viewed in this way, the ramification of this research extends beyond the importance of equal opportunities in the context of education to equal opportunities in the workplace where academic achievement influences the choices and possibilities available to individuals subsequent to graduation.

14.5 Contribution to knowledge

It is clear from the earlier literature review chapters that there is a lack of consistent, ecologically valid, reliable, research into the effect of granting exam arrangements to HE students with SpLD, with the result that policymakers are required to make decisions relating to the appropriateness of exam access arrangements without sufficient empirical evidence to inform their decisions. This research project has aimed to address the lack of ecologically valid, empirical data by evaluating the actual exam performance of students with SpLD who were granted exam access arrangements and comparing this to the actual exam performance of their TD peers who took the same exam under standard conditions. This research project has thereby elicited unique empirical evidence that can now be used to inform decisions relating to the granting of exam arrangements to students with SpLD.

Already, the findings of this research are contributing to policy decisions relating to exam access arrangements at the research site HEI. Policymakers at the research site HEI convened an Exam Review Working Party, which conducted a formal review of the University's exam processes. The findings of this research have been presented to

[155] A degree classification of 2:1 or above is required in order to apply for a scholarship to study a PGCE or to be considered for the Civil Service Fast Stream, for example (https://getintoteaching.education.gov.uk/; https://www.prospects.ac.uk/job-profiles/civil-service-fast-streamer)

the Exam Review Working Party and are being used to inform their decisions and shape assessment design. For example, under consideration is this research's recommendation for Universal Design approaches with the result that the HEI is piloting the use of a word processor as the standard exam conditions for all candidates[156]. The Head of Education Policy and the Board of Exams are evaluating the outcomes of this pilot, together with the data from the current research project, to determine the viability of allowing all students the option to use a word processor in their exams. In addition, the Department of Medicine, has embedded the findings of this research project in a current root and branch review of its curriculum delivery and assessment design, and is developing pedagogic approaches that aim to reduce the burden on unnecessary rote memorization and improve the accessibility of the course and assessment.

In January 2018 the Department for Education founded the Office for Students (OfS) to regulate the Higher Education Sector in the UK and ensure fair access and participation of all students. As part of the compliance with the OfS regulatory framework, universities are required to develop Access and Participation Plans (APP), which set out the institutions' commitment to equality of opportunity for the entire student cohort (including disabled students). The APP contains targets on access and makes a commitment to closing awarding gaps. This current research has fed into the research site's APP by providing empirical data demonstrating the awarding gap that exists between students with SpLD and their TD peers alongside recommendations for redressing this inequality, as part of the process of meeting OfS regulatory requirements in relation to disadvantaged groups.

In addition to the contribution of this project in informing the reviews of assessment practices within the university where the research was conducted, this research is also contributing to knowledge more broadly in the field. This research was presented at the 2018 International Conference of the National Association of Disability Practitioners (NADP), the 2019 Inclusive Assessment: Innovations in Practice Conference (University of Bath), the 2021 National Conference of the SpLD Assessment Standards Committee (SASC), and the 2021 DCD International Conference, (Cardiff University),

[156] The use of the word processor, with spell-check enabled, for all candidates is currently being piloted by the Faculty of Law

which allowed the opportunity to share the findings with peers in the field, including practitioners and researchers across the international disability community.

14.6 Recommendations: Facilitating a successful university experience.

Education acts as a key driver of social justice, predicated on egalitarianism (Hou, 2020). That is to say, all students, including those with SpLD should be able to access their education equally and have equal opportunities to fulfil their potential. However, this research project suggests that timed, closed book, unseen, written exams discriminate against students with SpLD and that the adjustments provided to redress this inequality (namely, the awarding of the word processor and/or 25% extra time) fail to fully eliminate the awarding gap or place students with SpLD on a level playing field with their peers. At the same time, the data from the participants' dissertation results suggest that the disparity in exam marks between the SpLD and TD groups is not the result of differences in academic ability, but rather stems from barriers to performance in timed exams experienced by students with SpLD due to the core features of their diagnosis. A number of factors driving this underperformance have been posited in this research project, including inaccessible teaching practices during courses of study, and assessment practices that pivot on construct irrelevant access skills that represent key deficits associated with SpLD. Thus, it seems that academic practices that disadvantage or marginalise students with SpLD are woven into the very fabric of conventional teaching and learning approaches in tertiary education, which impacts not only on the educational performance and future opportunities of students with SpLD, but also on their self-esteem and self-concept. This runs counter to the social justice model of education.

Given the increasing numbers of students with SpLD attending University, coupled with the Office for Students' Access and Participation strategy that requires HEIs to identify and respond to awarding gaps in degree outcomes for disabled students, it behoves HEIs to ensure that teaching and learning are fully accessible and guarantee that all students have equal opportunities to reach their potential in their course of study. Separate exam adjustments that require the student to be removed from their cohort to sit exams in a different venue to their TD peers (segregation) highlight specific students as 'other'. These pedagogic practices have the tendency to engender negative feelings of self-worth and low academic self-esteem in students with SpLD, and a sense of being 'less able' than peers, with a resultant negative impact on academic performance and mental well-being (Sarid et al, 2020). Equity of opportunity and social justice, the participants with SpLD in this research project suggest, need to be a fundamental and inexorable principle of the education system predicated on full participation and practices that ensure the inclusion of all groups, rather than a system that separates out and 'others' disabled learners.

Given that the awarding of exam arrangements is contentious, fails to eliminate the awarding gap, and identifies a particular group as 'other', alternatives could be considered that may achieve the purpose of ensuring equity for all students without either raising concerns about conferring an advantage or negatively impacting students' self-efficacy. Assessment approaches founded upon the principles of Universal Design for learning (UDL), diversified assessments and a wider range of accommodations and assessment modalities could be adopted as a more equitable and accessible approach to assessing students' abilities. These approaches have the potential to more fully overcome the barriers faced by students with SpLD and thereby close the identified awarding gap (Ketterlin-Geller, 2005). Universal Design is a framework of inclusive practices aimed at developing environments that are flexible enough to accommodate individual differences. From the point of view of exams, Universal Design principles require the assessment process to be as inclusive as possible to 'allow participation of the widest possible range of students, and to result in valid inferences about performance for all students who participate in the assessment' (Thompson et al, 2002). Viewed from within this context, a policy of universal design would aim to ensure that assessment is designed in a way that is equitable to all candidates, minimizing time constraints, such that separate accommodations are rendered unnecessary.

UDL fosters a move away from a mainly logocentric method of teaching and learning, that overburdens working memory and pressurizes literacy skills, in a quest to develop practices that support the needs of all learners, including students with SpLD. As candidates with SpLD commonly experience working memory and processing speed difficulties, assessment approaches that reduce the pressure on working memory and processing speed[157] could be explored in order to compensate for the negative impact on the pedagogical aims of the exam that can result from the pressure of a timed exam. These approaches could include open book exams, non-time bounded exams (i.e. take-home exams), or exams that ensure the time allocated is sufficient for all students to complete the paper. In addition, to promote the accessibility of the exam, the readability and comprehensibility of the exam questions and carrier language should be evaluated, in order to lessen the impact of the language demands of the paper and reduce irrelevant score variance that arise from the differences in literacy skills between SpLD and TD candidates. Universal test design entails the development of assessments that are fair and valid for all students and allow all students equal access to the exam and equal opportunities to demonstrate their full skill levels. They require test designers to explicitly determine the competencies that are the foci of the assessment and ensure construct validity[158]. For example, if speed of performance is a true competence of the exam, then the amount of time limit imposed on the exam should be empirically established. If speed is not a construct that the assessment is, at least in part, designed to measure, then time constraints should be abated (Lewandowski, Cohen and Lovett, 2013; Ghosh et al, 2017; Villarroel et al, 2019). Villarroel et al (2019) advocate the use of a 'backward design methodology' for exam development (p 41) whereby course and test designers analyse the course learning outcomes and competency standards that graduates in the discipline are expected to acquire and then ensure that the assessments are designed to give candidates the opportunities to demonstrate that they have met these learning outcomes and competencies.

What could a more Universal Design approach to assessment look like? The participants with SpLD in the qualitative phase of this current research project tabled the value of working in partnership with course designers and test setters, using a co-design approach to inclusive assessment and inclusive curriculum design – a suggestion that mirrors the extant literature in the field of authentic assessment that similarly advocates for the co-creation of assessments, including marking criteria (López-Pastor

[157] Where memorisation or speed are not competence standards being tested by the exam

[158] Construct validity refers to the exam assessing what it purports to assess. That is to say, the results of exams that require highly proficient and automatic literacy skills, when those are not the foci of measurement, may reflect the weaker literacy skills of students with SpLD relative to their TD peers, rather than their ability and knowledge. This represents construct validity bias. (Villarroel et al, 2021).

& Sicilia-Camacho, 2017; Tai et al., 2017; Villerroel et al, 2019). The participants in this current research project also suggested designing exams that test knowledge from across topics within their discipline, rather than assessing in topic siloes, as is currently the case. They felt that this would more authentically reflect real life scenarios where multiple aspects of subject knowledge across the full range of their discipline come into play simultaneously (such as in Medicine, for example, when faced with a 'real-life' patient). Advice and training for teaching staff and exam markers was also seen by participants as an important aspect of improving the accessibility of assessment, as well as ensuring the fairness of marking. This view is supported by the literature which suggests, for example, that training for examiners potentially offers a solution to the unconscious marking bias of typed scripts. Mogey & Purcell (2016) found that when exam markers were made aware of typed script bias, the bias was reduced, while printing scripts in cursive fonts and offering the option to type exam scripts to all students (not just disabled students) reduced the typed script bias effect arising from unconscious ableism. Training exam setters in developing exam questions that are fully answerable in the standard time set may help ensure that all candidates have sufficient time to complete the exams, as completing the exam is not evidence that the questions are completable in the time. These are but two examples of areas where appropriate training for exam designers and exam markers may reduce the disadvantage experienced by candidates with SpLD that inadvertently arise from unconscious bias.

The participants with SpLD also commented that opportunities for skill development should be integrated into assessment, with exams located within a pedagogic framework of assessment *as* learning, rather than assessment *of* learning, with opportunities to practice higher order, critical thinking skills and knowledge application. This requires academic staff to be more explicit about the skills that are being explored in the exam and link these to the course learning outcomes (Bengtsson, 2019; Villarroel, et al, 2019). The participants commented that they would also benefit from an explicit link between the skills that they will use as a professional in their discipline and their responses to the exam questions (that is to say, how will they draw on the knowledge that they have written in response to the exam questions when they work as professionals in their field) as well as from explicit explanations of how the assessment is relevant to the application of the knowledge and skills that they have acquired. They also felt that receiving clearer guidelines on examiners' expectations from exam responses and better scaffolded questions would improve their performance in exams, which is consistent with the findings of Gibson et al (2015) who observed that scaffolded type questions significantly closed the gender awarding gap. The participants' perception of the potential benefits of scaffolding in exam questions is also consistent with the data from the quantitative phase of this current research project that revealed a smaller awarding gap in Medicine exam papers than was the case in the other subjects included in this research project, with the questions in Medicine exams being inherently more scaffolded than is commonly the case for other subjects.

Participants also expounded the benefits of diversified assessment, aligned to learning outcomes, to reduce the likelihood that the reliance on one type of assessment (such as end of year, time bounded, closed book, unseen, written exams) disadvantages a specific cohort of students (such as those with SpLD). This was also supported by the data from the quantitative phase of the research project that showed that students' dissertation marks act as a reliable predictor of exam mark (in terms of an individual's position on the distribution curve). Therefore, it could be argued, allowing candidates the option to submit coursework in place of exams is an equally robust measure of ability. The participants also suggested a range of alternative assessment methods that could viably replace some of the closed book, timed, unseen, written exams, thereby reducing the number of exams in each discipline and increasing student selection options. These alternatives included open book assessments, take-home exams, student presentations, teamwork and collaborative activities, oral exams, problem-solving activities, project work, and submitting a portfolio of work from across the year, for example. These recommendations for more diverse forms of assessment are supported by the literature that suggests that these approaches reduce test anxiety, promote higher-order thinking, allow time for reflection, are more equitable and enable assessment to act as a learning activity (Bengtsson, 2019; Pereira, Flores, & Barros, 2017).

The participants in this research project suggested reducing the dominance of timed, closed book, unseen, written exams in favour of diverse and multiple means of assessing the critical competences of their discipline. However, given the privileged position that timed, unseen, closed book exams continue to occupy in current educational assessment practices, it is unrealistic to suppose that this form of assessment will be entirely cast aside in favour of alternative assessment methods.

Nonetheless, the principles of Universal Design for learning (UDL) and authentic assessment practices can still be applied to the traditional exam format in order to substantially improve the fairness of this testing method. For example, the participants in this research project with SpLD felt that a shift away from the privileging of knowledge reproduction and the emphasis on rote memorization of factual knowledge[159] in favour of application of knowledge in exams would promote a more equitable assessment while still retaining competencies around core knowledge. These views are supported by Villarroel et al, (2019) and Bengtsson (2019) who argue that the memorization and reproduction of information represents superficial learning and exams that privilege this approach are testing:

> the lowest level of knowledge assessment and students quickly forget what they memorize… Instead, when students use higher-order cognitive skills to respond to an assessment, such as concluding, designing or evaluating, they gain a deeper understanding and show better stability in remembering what was learned. (Villarroel et al, 2019, pp38-39).

To encourage a shift away from a focus on memorization and knowledge reproduction, towards assessments of higher taxonomy levels[160], such as critical thinking skills, the application of knowledge, problem-solving techniques and innovation, the participants with SpLD suggested that the exams task should mirror more authentic 'real-world' activities. This would reflect, for example, the realism of, and cognitive challenges of, how one would practice in a professional setting (such as in performing the role of Medic or Lawyer, for example) which, they felt, would enable candidates to draw on knowledge and material in a deeper, more meaningful and autonomous way during the assessment and foster skills transferable to the workplace. This outcome is supported by Villerroel et al (2019), who recommend authentic assessment practices that simulate 'real-life' contexts and foreground analytical, problem solving and critical thinking skills alongside application of knowledge and innovation, rather than simply memorization skill and recall. Thus, authentic assessment can help promote the concept

---

[159] Flores, Veiga-Simao, Barros & Pereira (2015) & Villarroel et al, (2019) suggest that closed book, unseen, timed exams, inadvertently encourage students to adopt passive learning approaches that pivot on the memorisation rather than the understanding and application of the subject matter.

[160] Blooms Taxonomy sets out the phases of learning, from low-level taxonomy associated with memorisation and recall, through the higher taxonomies of 'understanding', 'applying', 'analysing', 'evaluating' and 'creating' (Prasad, 2021).

of 'assessment for learning' where the assessments themselves encourage students to define problems, predict, hypothesize, experiment, analyse, conclude, and think reflectively and innovatively. This has the added benefit of fostering the development of transferable problem-solving skills in preparation for the demands of the workplace or future scholarly endeavours (Bengtsson, 2019; Villerroeld et al, 2019).

Employing the principles of Universal design for learning (UDL) to ensure the full inclusion of students with SpLD may also have the potential to address the underperformance of students with SpLD arising from low or vulnerable academic self-concept. Yasar and Arslan (2017) in a quasi-experimental study into the impact of UDL on engagement, resilience and persistence found that UDL improved the self-efficacy of students with low academic self-concept, and normalised their self-concept and academic performance to the same level as the control group. Similarly, Kennedy et al (2014) found that students' academic self-concept was enhanced and reinforced through the positive learning experiences engendered by UDL and that differences in academic performance between disabled students and their TD peers diminished. A UDL approach to exams, that overcomes the necessity to assign separate exam conditions to specific groups, also has the potential to reduce the psychological impact of 'othering' that separate treatment can provoke and which undermines academic self-esteem and self-efficacy. Thus, course and assessment design predicated on principles of UDL has the potential to ensure that *all* students have equal opportunities to access their education and demonstrate their full academic potential in their course of study. Inclusive practices predicated on the principles of UDL, and the valuing of diversity throughout the institution, increasingly moves disability support away from the traditional SEN model (i.e. disabled students 'othered' and treated separately to other students) towards the prospect of a truly inclusive, empowering education for all. A model of the UDL practices and diversified assessment that address the range of barriers to learning that students with SpLD report is shown diagrammatically in Figure 14.2.

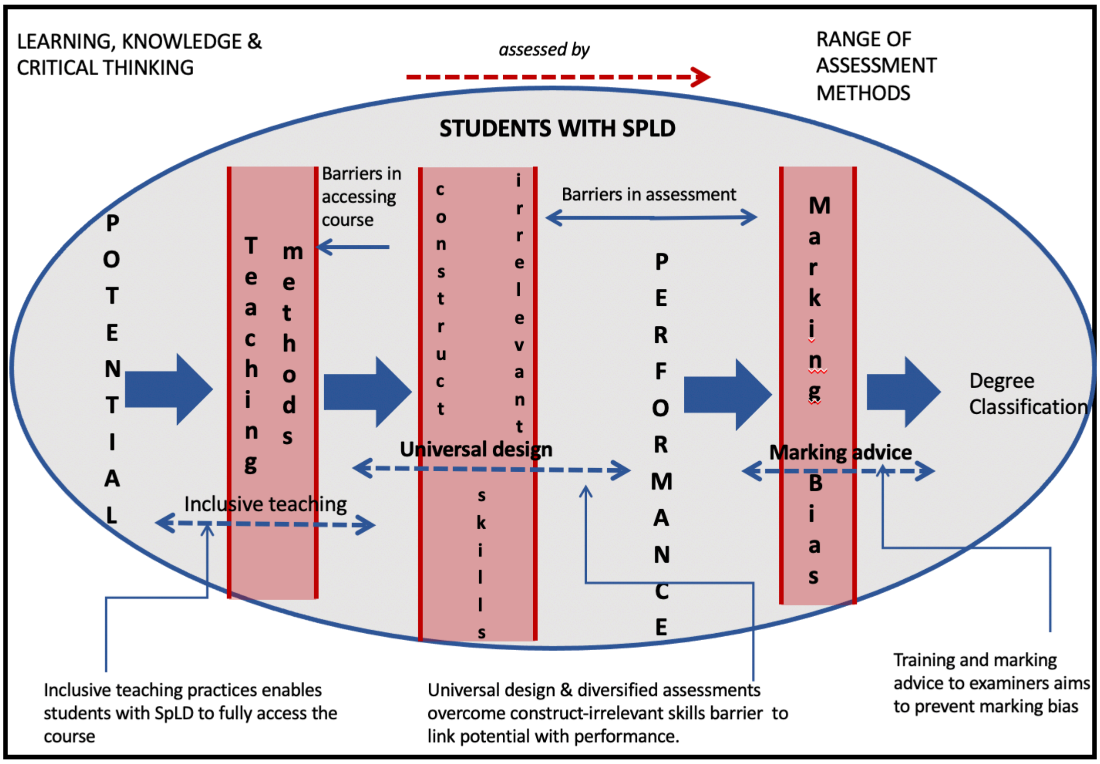


Figure 14.2: A model of barriers to the exam performance of students with SpLD addressed by principles of UDL, inclusive practices, staff training and diversified assessment.

14.7 Suggestions for future research

Overall, the review of the literature indicates that there is minimal research directly exploring the impact of access arrangements on HE students with SpLD conducted in the environment of formal, written, timed examinations. In addition, the research studies that currently exist yield contradictory findings, lack ecological validity or a homogenous sample group. This discernable gap, in combination with identified inconsistency in research outcomes, suggests opportunities for future investigation. While the findings from this current research project contribute to this field, in order to determine if the findings of this current research project are generalisable to HE students in other institutions or are confined to the cohort and pedagogic environment of the research site University[161], it would be useful for similar research to be conducted in randomly selected HEIs across the UK. The statistically lower UCAS points achieved by SpLD entrants to the University by comparison to the TD entrants suggest

[161] The research site institution is a Russell Group collegiate University

that the awarding gap may also be prevalent in exams at compulsory school level and research into this possibility is warranted[162]. In addition, no research appears to exist that provides empirical evidence for the determination of the amount of extra time that is needed to help normalize the performance of a student with SpLD to that of their TD peers in exams[163]. Research into developing a rubric for establishing the amount of extra time required by each individual, taking account of both their cognitive profile and the context of the exam (including the weighting of the construct irrelevant skills involved as well as the amount of standard time exam), is, therefore, much needed.

It would also be useful for future research to be conducted to explore some of the issues that have been raised by this research but lie outside its scope of investigation. This could include research into alternative exam arrangements, exam design or assessment models for ensuring parity in exam situations where a student may be disadvantaged solely on account of the core characteristics associated with SpLD. For example, changes to traditional examination practices, occasioned by the public health restrictions imposed during the COVID-19 pandemic[164], such as replacing timed, closed book exams with open book, 24 hour 'take-home' exams have been in operation for during the summer 2020 and the summer 2021 exam periods across HEIs. These examination environments were deemed to have adjustments for students with SpLD already 'baked in' (i.e. an inclusive model) and so no further extra time was granted to students with SpLD (all candidates used a word processor and had access to notes, texts and online resources). This model of examination is akin to the inclusive assessment practices suggested by the participants in this research and it would, therefore, be useful to scrutinize the data from these past two exam periods in order to compare the relative performances of the SpLD and TD cohort to ascertain if this assessment model has indeed had the positive effect on the awarding gap that was anticipated.

---

[162] See Appendix 6 for details of the UCAS points on entry for the participants

[163] The standard amount of extra time granted to students with SpLD varies from country to country, with the USA, for example, most commonly granting 50% or 100% extra time by comparison to 25% in the UK (Gosh et al, 2017; Periera et al, 2017; Villaroel, et al, 2020). The amount of extra time, therefore, seems to be influenced by the length of the exam (i.e. the longer the standard exam time, the shorter the amount of extra time allowed, to avoid the impact of excessively long exams) rather than established via any evidence based rubric.

[164] Public Health guidelines prohibited group exams taking place communally in examination halls.

The Office for Students (OfS) provides annual funding, through the disabled students' premium, to universities for the purpose of enabling HEIs to make their courses accessible to disabled students, support them to achieve, and thereby improve graduate outcomes for this disadvantaged group:

> Funding allocated through the disabled students' premium is earmarked to support the costs of activities that will promote inclusion and remove barriers to participation and success for disabled students (Office for Students, 2019).

In 2020/21, the OfS disabled students' premium allocated to HEIs stood at £40 million, while the cost to the UK taxpayer of Disabled Students Allowances[165] (DSA) for undergraduate students stood at over £105 million. The majority of this funding was allocated to support students with SpLD, who represent the largest single group of disabled students (OfS, 2021; Hubble & Bolton, 2020). However, there currently exists a lack of research evidence that measures the efficacy of specific interventions for which the OfS disabled students' premium funding or DSA is used. In general, HEIs collect data that shows overall trends on retention and degree classifications, but not an analysis of the efficacy of specific support activities, such as 1:1 study skills tuition:

> The majority of the knowledge about what works in terms of access, retention and success strategies is held at an institutional level by staff working directly with students and is not always systematically published, aggregated or discussed at national policy levels … In order to strengthen the evidence base, there is a need to understand what works, for whom and in what educational, institutional or regional context' (Bowes et al, 2015, p39).

Although this current research project touched on the impact of 1:1study skills sessions on the exam performance of student with SpLD, this was not its main objective. Only 32 participants in the research project accessed study skills sessions, which provided this part of the research project with a limited power to detect a true effect. The outcomes did not suggest that the study skills sessions improved the performance of the SpLD participants who used them to a greater extent than was the case for the SpLD participants who did not use study skills sessions. However, this non-significant finding, which contradicts much of the extant literature, may be confounded by the low

[165] DSA is a non-repayable, non-means tested, government grant that is available to cover the additional costs disabled students (including those with SpLD) incur during their education as a result of their disability. DSA provides funding for a range of support, including assistive technology and 1:1 study skills tuition, for example. https://www.gov.uk/disabled-students-allowance-dsa.

power of the research project and the limited take up of sessions by participants[166]. Given the requirement for accountability for public funds, the need to show that the funding makes a difference and the amount of funding that is currently allocated to support disabled students in university, further empirical research into the efficacy of study skills sessions provided to students with SpLD would be valuable. This future research could usefully interrogate the relationship between study skills sessions and student academic outcomes, evaluate different models of study skills support, and thereby provide a more rigorous evaluation of interventions in order to establish programme effectiveness.

On a final note, this current research project focused on academic outcomes for students with SpLD only. Therefore, future research using similar methodologies is warranted, investigating potential awarding gaps experienced by other, traditionally disadvantaged, groups of students[167]. By locating this future research within the context of the Office for Student's directive to identify and address awarding gaps prevalent in disadvantaged groups within the student body, the current debate may be steered away from questions about 'advantage' to a discourse on equality of educational opportunities for all students, regardless of identity, impairment or diagnostic profile: a paradigm shift to a model of social justice.

---

[166] Study skills strategies can take time to become effective and embedded (Saeidi et al, 2019) and the participants in this study only accessed study skills sessions in the second or third years their degree and only accessed, on average, 3 sessions each.

[167] Such as students from minority ethnic backgrounds, those experiencing socio-economic disadvantage, or those with other disabilities or impairments, including mental health conditions, for example.

15 REFERENCES

Aberson, C. L. (2010). Applied Power Analysis for the Behavioural Sciences. Routledge. New York

Adlof, S, M., & Hogan, T, P. (2018). Understanding Dyslexia in the Context of Developmental Language Disorders. Language, Speech and Hearing Services in Schools. 49, (4), 762-773.

Advance HE. (2020). Equality + Higher Education: Students statistical report 2020. https://www.advance-he.ac.uk/knowledge-hub/tags/equality-diversity-and-inclusion/disability

Agee, J. (2009). Developing qualitative research questions: a reflective process, International Journal of Qualitative Studies in Education, 22:4, 431-447, DOI: 10.1080/09518390902736512

Alamargot, D., Morin, M. & Simard-Dupis, E. (2020). Handwriting Delay in Dyslexia: Children at the End of Primary School Still Make Numerous Short Pauses When Producing Letters. Journal of Learning Disabilities. 53 (3), 163 – 175.

Alster, E. H. (1997). The effects of extended time on algebra test scores for college students with and without learning disabilities. Journal of Learning Disabilities. 30.

Alston J. (1994) Written Output and Writing Speeds. Dyslexia Review. 6 (2), 6-12

American College Testing Program. (1989). ASSET technical manual. Iowa City. IA

American Educational Research Association (AERA). (1999). Standards for educational and psychological testing. American Educational Research Association. Washington.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders, DSM-5, (5th ed.). Arlington, V.A: American Psychiatric Publishing

Amodeo, A., Marcus, L.A., Thornton, A.E., Pashley, P.J. (2009). Predictive Validity of Accommodated LSAT Scores for the 2002-2006 Entering Law School Classes. Law School Admission Council. LSAT Technical Report 09-01. LSAC Research Report Series.

Asghar, Z.B., Siriwardena, A.N., Elfes, C., Richardson, J., Larcombe, J., Neden, K.K., Salim, A., Smalley, D. & Blow, C. (2018). Performance of candidates disclosing dyslexia with other candidates in a UK medical licensing examination: cross-sectional study. Postgraduate Medical Journal, 94, 198-203.

Bachan, R. (2017). Grade inflation in UK higher education. Studies in Higher Education, 42(8), 1580-1600.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4, (11), 417-423.

Baddeley, A.D., Hitch, G.J. & Allen, R.J. (2019). From short-term store to multicomponent working memory: The role of the modal model. Memory & Cognition, 47, 575–588.

Barnes, C. (2012). in Watson, N & Vehmas, S (Eds). Understanding the social model of disability. Routledge Handbook of Disability Studies. Routledge. Oxon.

Barnett, A.L., Henderson, S.E., Scheib, B. & Schulz, J. (2009) Development and Standardization of a New Handwriting Speed Test: The Detailed Assessment of Speed of Handwriting. British Journal of Educational Psychology, 2 (6), 137–157

Barnett, A., Henderson, S., Scheib, B. & Schulz, J. (2010). Detailed Assessment of Speed of Handwriting 17+ (DASH). London, Harcourt.

Barnett A. L. & Prunty M. (2020). Handwriting difficulties in developmental coordination disorder (DCD). Current Developmental. Disorders Report. 8, 6–14.

Barrable, A., Papadatou-Pastou, M., & Tzotzoli, P. (2018). Supporting mental health, wellbeing and study skills in Higher Education: an online intervention system. International Journal of Mental Health Systems. 12 (54), https://doi.org/10.1186/s13033-018-0233-z

Bengtsson, L. (2019). Take-Home Exams in Higher Education: A Systematic Review. Education Sciences, 9 (4), 267. MDPI AG. Retrieved from http://dx.doi.org/10.3390/educsci9040267

Benton, T. (2017). How much do I need to write to get top marks? Research Matters: A Cambridge Assessment Publication, 24, 27-50.

Berger, C.L., & Lewandowski, L. (2013). The Effect of a Word Processor as an Accommodation for Students with Learning Disabilities. Journal of Writing Research, 4 (3), 299 – 318.

Bieder, A., Yoshihara, M., Katayama, S., Krjutskov, K., Falk, A., Kere, J., & Tapia-Paez, I. (2020). Dyslexia candidate gene and ciliary gene expression dynamics during human neuronal differentiation. Molecular Neurobiology. 57, 2944-2958.

Biotteau, M., Albaret, J.-M., Lelong, S., & Chaix, Y. (2017). Neuropsychological status of French children with developmental dyslexia and/or developmental coordination disorder: Are both necessarily worse than one? Child Neuropsychology, 23(4), 422–441.

Blank R., Barnett A.L., Cairney, J., Green, D., Kirby, A., Polatajko, H., Rosenblum, S., Smits-Engelsman, B., Sugden, D., Wilson, P., & Vinçon, S. (2019). International clinical practice recommendations on the definition, diagnosis, assessment, intervention, and psychosocial aspects of developmental coordination disorder. Developmental Medicine & Child Neurology. 61(3), 242-285.

Boedigheimer, R., Ghrist, M., Peterson, D., & Kallemyn, B. (2015). Individual oral exams in mathematics courses: 10 years of experience at the air force academy. Problems, Resources, and Issues in Mathematics Undergraduate Studies. 25(2), 99–120.

Bolt, S. E. (2004). Using DIF analyses to examine several commonly held beliefs about testing accommodations for students with disabilities. Paper presented at the annual conference of the National Council on Measurement in Education, San Diego, CA

Bolt, S. E., Decker, D. M., Lloyd, M., & Morlock, L. (2011). Students' perceptions of accommodations in high school and college. *Career Development for Exceptional Individuals*, *34*(3), 165-175. DOI: 10.1177/0885728811415098

Booth, A., Sutton, A., & Papaioannou, D. (2016). Systematic Approaches to Successful Literature Review (2nd Ed.). Sage Publications Ltd. London.

Bowden, J., Hart, G., King, B., Trigwell, K.,& Watts, O. (2000). Generic Capabilities of ATN University Graduates. Canberra: Australian Government Department of Education, Training and Youth Affairs.

Boyes, M. E., Leitao, S., Claessen, M., Badcock, N. A., & Nayton, M. (2016). Why are reading difficulties associated with mental health problems? Dyslexia, 22(3), 263–266.

BPS Code of Human Research Ethics (2nd edition, 2014). [online] https://www.bps.org.uk/news-and-policy/bps-code-human-research-ethics-2nd-edition-2014 [Last accessed 2.6.2021]

Brandenburg, J., Klesczewski, J., Fischbach, A., Schuchardt, K., Büttner, G., & Hasselhorn, M. (2015). Working memory in children with learning disabilities in reading versus spelling: Searching for overlapping and specific cognitive factors. Journal of Learning Disabilities, 48(6), 622–634.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101

Braun, H, I., Kaplan, B., & Ragosta, M. (1986). The Predictive Validity of the Scholastic Aptitude Test for Disabled Students. Wiley On-line Library. ETS Research Report Series.1986 (2). DOI: 10.1002/j.2330-8516.1986.tb00193.x

Bridgeman, B., Cline, F.,& Hessinger, J. (2003). Effect of Extra Time on GRE Quantitative and Verbal Scores. GRE Research. Educational Testing Service. Princeton. NJ

Brimo, K.; Dinkler, L.; Gillberg, C.; Lichtenstein, P.; Lundström, S.; Johnels, J.Å. (2021). The co-occurrence of neurodevelopmental problems in dyslexia. Dyslexia, 1–17.

British Dyslexia Association. What is dyslexia? [online] https://www.bdadyslexia.org.uk/dyslexia/about-dyslexia/what-is-dyslexia [Last accessed 11.7.2021].

Broggi, M., Ready, R.E., &. Moore, D.L. (2019). Screening for reading disability in university students with phonological processing and working memory tasks, *Dyslexia*. 25 (3), 256-266.

Brown, J.A., Fishco, V.V., & Hanna, G. (1993). Nelson–Denny Reading Test: Manual for Scoring and Interpretation, Forms G & H. Rolling Meadows, IL: Riverside Publishing.

Brown, J. I., Bennett, J. M., & Hanna, G. (1981). Nelson–Denny Reading Test. Boston: Riverdale

Brown-Lum M., Izadi-Najafabadi S., Oberlander T. F., Rauscher A., Zwicker J. G. (2020). Differences in white matter microstructure among children with developmental coordination disorder. JAMA Network Open 3:e201184. 10.1001/jamanetworkopen.2020.1184

Brydges, C.R., Gignac, G.E & Ecker, U.K.H. (2018). Working memory capacity, short-term memory capacity, and the continued influence effect: A latent-variable analysis. Intelligence, 69, 117-122.

Bryman, A. (2015). Social research methods. Oxford University Press.

Bryman, A. , Teevan, J., Bell, E. (2009). Social Research Methods. (2nd Ed.). Oxford University Press. Oxford.

Byrne, C. (2018) Why do dyslexic students do worse at university? The Guardian. https://www. The guardian. com/education/2018/dec/06/whydo-dyslexic-students-do-worse-atuniversity

Button, K., Ioannidis, J., Mokrysz, C., Nosek, B., Flint, J., Robinson, E. & Munafo, M. (2-13). Power failure: why small sample size undermines the reliability of neuroscience. Nature Reviews Neuroscience. 14, 365–376. https://doi.org/10.1038/nrn3475

Byrne, B.M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B.A. Bracken (Ed.), Handbook of self-concept: Developmental, social and clinical considerations. New York: Wiley.

Cacola, P & Lage, G. (2019). Developmental Coordination Disorder (DCD): An overview of the condition and research evidence. Motriz: Journal of Physical Education. 25(2), 1-6.

Cahalan, C., Mandinach, E.B., & Camera, W. J. (2002). Predictive Validity of SAT I: reasoning test for test-takers with learning disabilities and extended time accommodations. New York, NY: College Entrance Examination Board.

Cahalan-Laitusis, C., King, T. C., Cline, F., & Bridgeman, B. (2006). Observational timing study on the SAT Reasoning test for test-takers with learning disabilities and/or ADHD. (Research Report 2006-4). New York, NY: College Board.

Cahalan-Laitusis, C., Morgan, D.L., Bridgemand, B., Zanna, J., & Stone, E. (2007). Fatigue Effects from Extended-Time Accommodations on the SAT Reasoning Test. New York, NY: College Board.

Callens, M., & Brysbaert, M. (2019).Cognitive Profile of Students with Dyslexia Entering Postsecondary Education, The Wiley Handbook of Adult Literacy. Wiley Online Library, pp193-213. doi.org/10.1002/9781119261407.ch9.

Callens, M., Tops, W., & Brysbaert, M. (2012). Cognitive profile of students who enter higher education with an indication of dyslexia. PloS one, 7(6), e38081

Camara, W., Copeland, T., & Rothchild, B. (1998). Effects of extended time on the SAT I: Reasoning Test: Score growth for students with learning disabilities. College Board Research Report, 98 (7). New York: The College Board.

Cameron, H. (2016). Beyond cognitive deficit: the everyday lived experience of dyslexic students at university. Disability & Society, 31, 223 - 239.

Camilleri, S., Chetcuti, D., & Falzon, R. (2020). 'They labelled me ignorant': the role of neuroscience to support students with a profile of dyslexia. In El-Baz, A., & Suri, J (Eds.). Neurological Disorders and Imaging Physics, Applications in dyslexia, epilepsy & Parkinson's. Volume 5, Chapter 9, pp1-22. IOP Publishing Ltd. Web.

Camilleri, S., Chetcuti, D., & Falzon, R. (2019). 'They labelled me ignorant': Narratives of Maltese Youth with dyslexia on National examinations. Sage Open, 9 (2), 1-18.

Cappa, C., Muzio , C., & Giulivi, S. (2012). Understanding Dyslexia. Lifelong Learning Progamme. http://dyscovery.research.southwales.ac.uk/media/files/documents/2014-01-16/Module_2.pdf [Last accessed 22.12.20]

Casini, L., Pech-Georgel, C., Ziegler, J.C. (2017). It's about time: revisiting temporarl processing deficits in dyslexia. Developmental Science. 21 (2), 1-14.

Centanni, T.M. (2020). Neural and Genetic Mechanisms of Dyslexia. In: Argyropoulos G. (Eds). Translational Neuroscience of Speech and Language Disorders. Contemporary Clinical Neuroscience. Springer, Cham.

Charmaz, K. (2014). *Constructing grounded theory: a practical guide through qualitative analysis*. Thousand Oaks, CA. SAGE.

Charmaz K & Thornberg, R. (2020). The pursuit of quality in grounded theory. *Qualitative Research in Psychology.(*online) DOI: 10.1080/14780887.2020.1780357. [Last accessed 23.7.21].

Chaytor, N. & Schmitter-Edgecombe, M. (2003). The Ecological Validity of Neuropschological Tests: A Review of the Literature on Everyday Cognitive Skills. Neuropsychology Review. 13 (4)

Cheng, H., Chen, J., Tsai, C., Shen, M., & Cherng, R. (2011). Reading and writing performances of children 7–8 years of age with developmental coordination disorder in Taiwan. Research in Developmental Disabilities, 32, (6), 2589-2594

Chetcuti, D., Falzon, R. & Camilleri S. (2019) 'The Pebble in my shoe': Dyslexic students and their views of examinations. In C. Bezzina & S. Caruana (Eds.) Teacher Education Matters: Transforming, 433-447. University of Malta

Chinn, S. (2019). Maths Learning Difficulties: Dyslexia and Dyscalculia. (2nd Ed.). Jessica Kingsley Publishers. London.

Chong, C.S.J. (2018). Battling biases: How can diverse students overcome test bias on the multistate bar examination. University of Maryland Law Journal on Race, Religion, Gender and Class, 18(1), 31-97.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd Ed.). Lawrence Earlbaum Associates. New York

Cohen, A., Gregg, N., & Deng, M. (2005). The role of extended time and time content on a high-stakes mathematics test. Learning Disabilities Research and Practice. 20

Cole, E. V., & Cawthon, S. W. (2015). Self-disclosure decisions of university students with learning disabilities. Journal of Postsecondary Education and Disability, 28(2), 163-179. Retrieved from https://www.ahead.org/publications/jped

Connelly, V., Dockrell, J.E., & Barnett, J. (2005). The Slow Handwriting of Undergraduate Students Constrains Overall Performance in Exam Essays. Educational Psychology. 25 (1), 99-107

Connelly, V., Campbell, S., MacLean, M. & Barnes, J. (2006) Contribution of Lower Order Skills to the Written Composition of College Students With and Without Dyslexia. Developmental Neuropsychology, 29 (1), 175–196.

Corbin, J. & Strauss, A. (2015) Basics of Qualitative Research. Techniques and Procedures for Developing Grounded Theory (4th Ed.). Sage. London.

Couzens, D., Poed, S., Kataoka, M., Brandon, A., Hartley, J., & Keen, D. (2015). Support for students with hidden disabilities in universities: A case study.

International Journal of Disability, Development and Education, 62(1), 24-41. doi:10.1080/1034912X.2014.984592

Creswell, J. W. & Creswell, J.D. (2018). Research design: Qualitative, Quantitative and Mixed Methods Approaches. 5th Ed. United Kingdom. Sage.

Creswell, J. W., & Clark, V. L. P. (2017). Designing and conducting mixed methods research. Sage publications.

Creswell, J. W., Klassen, A. C., Piano Clark, V. L.P., & Smith, K. C. (2011). Best practices for mixed methods research in the health sciences. Bethesda (Maryland): National Institutes of Health, 2094-2103.

Creswell, J. W., & Miller, D. L. (2000). Determining validity in qualitative inquiry. Theory into practice, 39(3), 124-130.

Creswell, J. W., & Poth, C. N. (2017). Qualitative inquiry and research design: Choosing among five approaches. Sage publications.

Crisp, v., Johnson, M., & Novakovic, N. (2011). The effects of features of examination questions on the performance of students with dyslexia. British Educational Research Journal, 38 (5), 813-839.

Cvencek, D., Fryberg, S.A., Covarrubias, R. & Meltzoff, A.N. (2018). Self-Concepts, Self-Esteem, and Academic Achievement of Minority and Majority North American Elementary School Children. Child Development, 89 (8), 1099-1109

Darki, F., Peyrard-Janvid, M., Matsson, H., Kere, J., & Klingberg, T. (2012). Three dyslexia susceptibility genes, DYX1C1, DCDC2, and KIAA0319, affect temporo-parietal white matter structure. Biological Psychiatry. 72

Dawson, K., Antonenko, P., Lane, H., & Ahu, J. (2019). Assistive Technologies to Support Students with Dyslexia. Teaching Exceptional Children, 51 (3), 226-239

Dawson, P., & Guare, R. (2004). Executive skills in children and adolescents: A practical guide for assessment and intervention. New York: Guilford

Deckoff-Jones, A., & Duell, M. N. (2018). Perceptions of appropriateness of accommodations for university students: Does disability type matter? *Rehabilitation Psychology, 63*(1), 68–76.

Denscombe, M. (2008). Communities of Practice: A Research Paradigm for the Mixed Methods Approach. Journal of Mixed Methods Research, 2 (3), 270 – 283.

Denscombe, M. (2009). Ground rules for social research: Guidelines for good practice. McGraw-Hill Education (UK).

Denzin, N.K., & Lincoln, Y.S. (2018). The Sage Handbook of Qualitative Research (5th Ed.). Los Angeles. Sage Publications.

Department for Education (DfE). (2017). Inclusive teaching and learning in higher education as a route to excellence. Retrieved January 16, 2018, from https://www.gov.uk/government/publications/inclusive-teaching-and-learning-in-higher-education

Dewey, D., & Bernier, F.P. (2016). The Concept of Atypical Brain Development in Developmental Coordination Disorder (DCD)—a New Look. Current Developmental Disorders. 3, 161–169.

Dobson, S. (2018). A Documentary Analysis of the Support Services Offered to Adult Learners with Dyslexia in Higher Education. Journal of Further and Higher Education . doi:10.1080/0309877X.2018.1463359.

Dobson Waters, S. & Torgerson, C., J. (2020). Dyslexia in higher education: a systematic review of interventions used to promote learning. *Journal of Further and Higher Education*. DOI: 10.1080/0309877X.2020.1744545

Drever, E. (2003). Using Semi-Structured Interviews in Small-Scale Research. A Teacher's Guide. Edinburgh: SCRE.

Duncan, H., & Purcell, C. (2017). Equity or Advantage? The effect of receiving exam access arrangements in University exams on Humanities students with SpLD. Widening Participation and Lifelong Learning. 19 (2).

Dyspraxia Foundation (2019). Dyspraxia at a glance… What is dyspraxia? Retrieved from: http://www.dyspraxiafoundation.org.uk/about-dyspraxia/ [Last accessed 4.2.2021].

Elgendi, M.M., Stewart, S.H., MacKay, E.J., & Deacon, H. (2021). Two aspects of psychological functioning in undergraduates with a history of reading difficulties: anxiety and self-efficacy. Annals of Dyslexia, 71, 84–102, https://doi.org/10.1007/s11881-021-00223-3

Elliott, S. N., Marquart, A.M. (2004). Extended time as a testing accommodation: Its effects and perceived consequences. Exceptional Children, 70 (3), 349 – 367.

Ellis, A.W. (2016). Reading, Writing and Dyslexia: A Cognitive Analysis. Routledge, Oxon

Equality Act. (2010). London. [Online]. London: The Stationery Office. Available from: https://www.legislation.gov.uk/ [Last accessed 5.4.2021]

Equality Challenge Unit (ECU), Equality in higher education: students statistical report 2018. [Online]. Available from https://www.advance-he.ac.uk/sites/default/files/2018_HE-stats-report-students.pdf [Last accessed 2.8.2019].

Equality & Human Rights Commission (2019). [Online] https://www.equalityhumanrights.com/en/advice-and-guidance/what-are-reasonable-adjustments [last accessed 19.7.21]

Falzon, R. (2020). Dyslexia and Academic Life. *Dyslexia*. IntechOpen, p1-33. [Online].https://www.um.edu.mt/library/oar/bitstream/123456789/65470/1/Dyslexia%20and%20Academic%20Life.pdf [Last accessed 4.4.2021].

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2019). Learning disabilities: From identification to intervention (2nd ed.). Guilford Press.

Fletcher, J. M., & Miciak, J. (2019). The identification of specific learning disabilities: A summary of research on best practices. Texas Center for Learning Disabilities.

Flores, M. A., Veiga-Simão, M. A., Barros, A., & Pereira, D. (2015). Perceptions of effectiveness, fairness and feedback of assessment methods: A study in higher education. Studies in Higher Education, 40, 1523–1534.

Florida Department of Education. (2003). Florida Comprehensive Assessment Test. Tallahassee. Florida.

Foster, A.M. (2014). Reasonable Accommodations on the Bar Exam: Levelling the Playing Field or Providing and Unfair Advantage? Valparaiso University Law Review, 48, 661-693.

Fox, N., Hunn, A., Mathers, N. (2009) Sampling and Sample Size. The NIHR Research Design Service for the East Midlands. National Institute for Health Research. NHS.

Fredman, S. (2016). Substantive equality revisited. International Journal of Constitutional Law, 14(3), 712-738.

Fullarton, S. & Duquette, C. (2016). Experiences of Students with Learning Disabilities in Ontario Universities: A Case Study. International Journal of Special Education, 21 (1), 55-66.

Gathercole, S. E. & Packium-Alloway, T. (2004). Working Memory and Reading Disabilities. Dyslexia Review. 15 (9)

Gaudet, M. (2015). Increasing engagement through oral exams. Teaching Theology and Religion, 18 (1), 87-98.

Gentle, J., Brady, E., Woodger, N., Croston, S & Leonard, H.C. (2021). Driving Skills of Individuals With and Without Developmental Coordination Disorder (DCD/Dyspraxia). Frontiers in Human Neuroscience. 15, 105

Gibson, V., Jardine-Wright, L., Bateman, E. (2015). An investigation into the impact of question structure on the performance of first year physics undergraduate students at the University of Cambridge. European Journal of Physics, 36, 1 – 17, Gibson, S. & Leinster, S. (2011). How do students with dyslexia perform in extended matching questions, short answer questions and observed structured clinical examinations? Advanced Health Science Education Theory & Practice, 16, 395-404. https://doi.org/10.1007/s10459-011-9273-8

Glaser, B.G., & Holton, J. (2004). Remodeling grounded theory. Forum Qualitative Sozialforschung/Forum. *Qualitative Social Research,* (5), 1–22.

Goegan, L.d. (2014). The Effects of Extended Time on the Writing of Students with Learning Disabilities: Maximum Potential or Differential Boost? MSc thesis, University of Calgary. http://dspace.library.uvic.ca/handle/1828/5794?show=full

Gomez, A., Piazza, M., Jobert, A., Dehaene-Lambertz, G., Dehaene, S., & Huron, C. (2015). Mathematical difficulties in developmental coordination disorder: Symbolic and nonsymbolic number processing. Research in Developmental Disabilities, 43 (4), 167–178.

Gooch, D., Snowling, M., Hulme, C. (2011). Time perception, chronological skills and executive function in children with dyslexia and/or ADHD symptoms. Journal of Child Psychology and Psychiatry, 52 (2), 195 – 203.

Gooch, D., Moll, K., Gobel, S., Landerl, K., Snowling, M. (2016). Cognitive risk factors for specific learning disorder: Processing speed, temporal processing and working memory. Journal of Learning Disabilities, 49, 272 – 281.

Gosse, C. & Van Reybroeck, M. (2020). Do children with dyslexia present a handwriting deficit? Impact of word orthographic and graphic complexity on handwriting and spelling performance. Research in Developmental Disabilities. 97, 1 – 14. https://doi.org/10.1016/j.ridd.2019.103553

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. Language and Linguistics. 13 (5), 3 - 10

Goswami, U. (2019). A neural oscillations perspective on phonological development and phonological processing in developmental dyslexia. Trends in Cognitive Sciences. 15, 1 - 21

Georgiou, G.K., Ghazyani, R. & Parrila, R. (2018). Are RAN deficits in university students with dyslexia due to defective lexical access, impaired anchoring, or slow articulation?. Annals of Dyslexia, 68, 85–103. https://doi.org/10.1007/s11881-018-0156-z

Ghosh, S., Bowles, M., Ranmuthugala, D., & Brooks, B. (2017). Authentic assessment in seafarer education: Using literature review to investigate its validity and reliability through rubrics. WMU Journal of Maritime Affairs, 15, 317–336.

Gostic, M., Martinelli, A., Tucker, C., Yang, Z., Gasparoli, F., Ewart, J., Dholakia, K., Sillar, K.T., Tello, J.A., & Paracchini, S. (2019). The dyslexia susceptibility KIAA0319 gene shows a specific expression pattern during zebrafish development supporting a role beyond neuronal migration. Journal of Comparative Neurology. 527 (16), 2634 -2643.

Goulardins, J. B., Rigoli, D., Licari, M., Piek, J. P., Hasue, R. H., Oosterlaan, J., Oliveria, J. A. (2015). Attention deficit hyperactivity disorder and developmental coordination disorder: Two separate disorders or do they share a common etiology? Behavioural Brain Research. 292, 484-492.

Grant, D. (2009). The Psychological Assessment of Neurodiversity. In: Neurodiversity in Higher Education. Pollack, D. (Ed). Wiley-Blackwell, pp.33–62.

Gregg, N., Coleman, C., Davis, M., & Chalk, J.C. (2007). Timed essay writing: implications for high-stakes tests. Journal of Learning Disabilities, 40, 306–318.

Gregg, N., & Nelson, J.M. (2012). Meta-analysis on the Effectiveness of Extra Time as a Test Accommodation for Transitioning Adolescents with Learning Disabilities: More Questions than Answers. Journal of Learning Disabilities, 45 (128).

Grigorenko, E. L., Compton, D., Fuchs, L., Wagner, R. K., Willcutt, E., & Fletcher, J. M. (2019). Understanding, educating, and supporting children with specific learning disabilities: 50 Years of science and practice. American Psychologist, 75, 37–51.

Guba, E., & Lincoln, Y. (1989). Fourth generation evaluation. Newbury Park, CA: Sage.

Guba, E., & Lincoln, Y. (1994). Competing paradigms in qualitative research. In Denzin, N., Lincoln, Y. (Eds.), Handbook of qualitative research, pp. 105–117. London: Sage.

Guidi, L.G., Velayos-Baeza, A., Martinez-Garay, I., Monaco,A.P., Paracchini, S., Bishop, D.V., & Molná, Z. (2018). The neuronal migration hypothesis of dyslexia: a critical evaluation thirty years on. The European Journal of Neuroscience. 48 (10), 3212-3233.

Hackett, G., Betz, N. E., Casas, J. M., & Rocha-Singh, I. A. (1992). Gender, ethnicity, and social cognitive factors predicting the academic achievement of students in engineering. Journal of Counselling Psychology, 39(4), 527-538.

Hackett, R.A., Steptoe, A., Lang, R.P., & Jackson, S.E. (2020) Disability discrimination and well-being in the United Kingdom: a prospective cohort study. BMJ Open;10:e035714. doi:10.1136/ bmjopen-2019-035714

Hall, J., McGregor, K. K., & Oleson, J. (2017). Weaknesses in lexical-semantic knowledge among college students with specific learning disabilities: Evidence from a semantic fluency task. Journal of Speech, Language, and Hearing Research, 60, 640–653.

Halla, J, W. (1988). A psychological study of psychometric differences in Graduate Record Examination General Test scores between learning disabled and non-learning disabled adults. Unpublished doctoral dissertation, Texans Tech University, Lubbock. Cited in Zuriff, G. E. (2000). Extra Examination Time for Students with Learning Disabilities: An Examination of the Maximum Potential Thesis. Applied Measurement in Education, 13 (1), 99-117.

Hanson, W. E., Creswell, J. W., Clark, V. L. P., Petska, K. S., & Creswell, J. D. (2005). Mixed methods research designs in counseling psychology. Journal of Counseling psychology, 52(2), 224.

Hanson, M.R., Hayes, J.R., Schriver, K., LeMahieu, P.G., & Brown, P.J. (1998). A plain language approach to the revision of test items. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA, April 16, 1998. Cited in Thompson, et al (2002) NCEO Synthesis Report 44

Harding, J., Hofmann, R., Alcott, B. (2021). From reasonable adjustment to inclusive practice. Exploring the influence of lecture capture on the academic outcomes and learning experiences of disabled students in higher education. Unpublished EdD thesis.

Harris, R.B., Mack, M.R., Bryant, J., Theobald, E.J. & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could life underrepresented students into a 'hyperpersistent Zone'. Science Advances. 6 (24), 1-8. DOI: 10.1126/sciadv.aaz5687

Harrowell, I., Hollén L., Lingam R., Emond A. (2018). The impact of developmental coordination disorder on educational achievement in secondary school. Research in Developmental Disabilities. 72, 13–22.

Hassanbeigi, A., Askari, J., Nakhjavani, M., Shirkhoda, S., Barzegar, K., Mozayyan, M., & Fallahzedeh, H. (2011). The relationship between study skills and academic performance of university students. Procedia – Social and Behavioural Sciences, 30, 1416-1424.

Hatcher, J., Snowling, M.J., & Griffiths, Y.M. (2010). Cognitive assessment of dyslexic students in higher education. British Journal of Educational Psychology. doi.org/10.1348/000709902158801

Hazen. H. (2020): Use of oral examinations to assess student learning in the social sciences, Journal of Geography in Higher Education, 1 – 16. DOI: 10.1080/03098265.2020.1773418

He, M., Wei, S., & Lu, S. (2018). Research on Diversified Assessment Methods of Engineering Graphics Based on Ability Assessment. The Guide of Science and Education, 8, 23-34

Higher Education Funding Council for England (HEFCE). (2018). HEFCE annual report and accounts 2017-18. [Online] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/710103/CCS207_CCS0318280518-

1_HEFC_ARA_2017_18_WEB_ACCESSIBLE__1_.pdf [last accessed 17.7.21]

Higher Education Statistics Agency . 2019. Detailed Disabled and Ethnicity Breakdowns. https://www.hesa.ac.uk/data-and-analysis/students/whos-in-he [Last Accessed 22.8 2020].

Hill, G.A. (1984). Learning disabled college students: The assessment of academic aptitude. Unpublished doctoral dissertation, Texans Tech University, Lubbock. Cited in Zuriff, G. E. (2000). Extra Examination Time for Students with Learning Disabilities: An Examination of the Maximum Potential Thesis. Applied Measurement in Education, 13 (1), 99-117.

Holmes, A., & Silvestri, R. (2012). Assistive technology use by students with LD in Postsecondary education: A case of application before investigation? Canadian Journal of School Psychology, 27 (1), 81-97

Holmes, A., & Silvestri, R. (2019). Extra time or Unused Time? What Data from a College Testing Center Tells Us About 50% Extra Time as an Accommodation for Students with Learning Disabilities. Psychological Injury and Law, 12, 7-16

Horn, T.D., Huber, T. (2020). Assistive Technologies and Academic Success for Students with Dyslexia: A Literature Review. International Journal of Educational Technology and Learning, 9 (1), 52-59.

Horne, J., Ferrier, J., Singleton, C. & Read, C. (2011) Computerised Assessment of Handwriting and Typing Speed. Education and Child Psychology, 28 (2), 58–66.

Hou, A., Hill, C., Hu, Z. & Lin, L. (2020). What is driving Taiwan government for policy change in higher education after the year of 2016 – in search of egalitarianism or pursuit of academic excellence? S*tudies in Higher Education,* (online). Available at DOI: 10.1080/03075079.2020.1744126 [last accessed 22.7.21]

Hubble, S & Bolton, P. (2020). Support for Disabled Students in Higher Education in England. Briefing House of Commons Paper number 8716. 1-20. https://dera.ioe.ac.uk/37390/13/CBP-8716%20%28redacted%29.pdf

Huesman, R. L., & Frisbie, D. A. (2000). The validity of ITBS reading comprehension test scores for learning disabled and non-learning disabled students under extended time conditions. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA

Jascenoka, J., Korsch, F., Petermann, F., & Petermann, U. (2015). Cognitive achievement profiles of preschool children with developmental coordination disorder and ADHD. Practice of Child Psychology and Child Psychiatry. 64(2), 117–134.

Johnson, J., Adkins, D, & Chauvin, S. (2020). A Review of the Quality Indicators of Rigor in Qualitative Research. *American Journal of Pharmaceutical Education*, 84 (1), 138 – 146.

Johnston, R. J. (1994). Resources, Student: Staff Ratios and Teaching Quality in British Higher Education: Some Speculations Aroused by Jenkins and Smith. Transactions of the Institute of British Geographers (New Series), 19 (3), 359–365.

Joint Council for Qualifications (JCQ). (2020). Access Arrangements and Reasonable Adjustments 2014-2015. [electronic version]. http://www.jcq.org.uk/exams-office/access-arrangements-and-special-consideration/regulations-and-guidance [Last accessed 12/11/20].

Jones, M. W., Snowling, M. J., & Moll, K. (2016). What automaticity deficit? Activation of lexical information by readers with dyslexia in a rapid automatized naming Stroop-switch task. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42(3), 465-474.

Kafle, E. A. (2015). Nursing students with learning disabilities: Perceptions and attitudes regarding the role of disability support program services and access to accommodations. Dissertation Abstracts International: Section A. Humanities and Social Sciences. 76/08(E).

Kaplan, B.J., Wilson, B.N., Dewey, D.M. & Crawford, S.G. (1998). DCD may not be a discrete disorder. *Human Movement* Science, 17, 471—490.

Kaufman, A. S., & Lichtenberger, E.O. (2009). WAIS-V Assessment. Wiley & Sons. New Jersey.

Kearns, D. M., & Whaley, V.M. (2019). Helping Students with Dyslexia Read Long Words: Using Syllables and Morphemes. *TEACHING Exceptional Children*, 51(3), 212-225.

Kearns, D. M., Hancock, R., Hoeft, F., Pugh, K. R., & Frost, S.J. (2019). The Neurobiology of Dyslexia. *TEACHING Exceptional Children*, 51(3), 175-188.

Kendall, L. (2018). Supporting students with disabilities within a UK university: Lecturer perspectives. Innovations in Education and Teaching International, 55(6), 694-703. https:// doi/full/10.1080/14703297.2017.1299630

Keogh, E., Bond, F.W., French, C.C., Richards, A., Davis, R.E. (2007). Test anxiety, susceptibility to distraction and examination performance. *The Journal of Anxiety, Stress & Coping,* 17 (3), 241-252

Kettler, R.J., Rodriguez, M.C., Bolt, D.M., Elliott, S.N., Beddow, P.A. & Kurz, A. (2011) Modified Multiple-Choice Items for Alternate Assessments: Reliability, Difficulty, and Differential Boost. Applied Measurement in Education, 24(3), 210-234, DOI: 10.1080/08957347.2011.580620

Ketterlin-Geller, L.R. (2005). 'Knowing What All Students Know: Procedures for Developing Universal Design for Assessment', The Journal of Technology, Learning and Assessment, 4 (2)

Kimel, E., & Ahissar, M. (2020). Benefits from morphological regularities in dyslexia are task dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(1), 155–169. https://doi.org/10.1037/xlm0000717

Kirby, A. (2005). Overlapping conditions: overlapping management: services for indivduals with developmental coordination disorder. In: Sugden, D., Cambers, E (Eds.). Children with Developmenal Coordination Disorder. London: Whurr.

Kirby, A. (2012) Overlapping developmental disorders. British Dyslexia Association Handbook 2012.

Kirby, A., & Sugden, D. (2007). Children with developmental coordination disorders. Journal of the Royal Society of Medicine, 100 (4), 182-186

Kirwan, B., Leather, C. (2011). Students' voices: A report of the student view of dyslexia study skills tuition. Support for Learning: British Journal of Learning Support, 26, 33 – 41. DOI 10.1111/j.1467-9604.2010.01472.x

Kohler, B. (2015). Paper-based or computer-based essay writing: differences in performance and perception. Linguistic Portfolio 129–146

Kvale, S. (1994). Ten standard objections to qualitative research interviews. Journal of phenomenological psychology, 25(2), 147-173

Kvale, S. & Brinkmann, S. (2009). Interviews: An introduction to qualitative research interviewing. Thousand Oaks, Calif: Sage Publications.

Law School Admission Council. (2007). LSAT & LSDAS registration & information book: 2006-2007. Newtown, PA.

Lawson, A. & Beckett, A. (2020). The social and human rights model of disability: towards a complementarity thesis. The International Journal of Human Rights. 25 (2), 348-379.

Layes, S., Lalonde, R. & Rebai, M. (2020) Reading-related abilities underlying phonological awareness: a cross-sectional study in children with and without dyslexia, *Logopedics Phoniatrics Vocology*. DOI: 10.1080/14015439.2020.1768283

Leonard, H.C. & Hill, E,L. (2015). Executive Difficulties in Developmental Coordination Disorder: Methodological Issues and Future Directions. Current Developmental Disorders Reports. 2(2), 141-49

Lesaux, N.K., Pearson, M.R., & Siegel, L.S. (2006). The effects of timed and untimed testing conditions on the reading comprehension of adults with reading disabilities. Reading and Writing. 19.

Levitt, J.M. (2017). Exploring how the social model of disability can be re-invigorated: a response to Mike Oliver. Disability & Society. 32 (4), 589-594.

Lewandowski, L. J., Cohen, J., & Lovett, B, J. (2013). Effects of Extended Time Allotments on Reading Comprehension Performance of College Students With and Without Learning Disabilities. Journal of Psychoeducational Assessment, 31(3), 326-336. DOI: 10.1177/0734282912462693

Lewandowski, L. J., Lovett, B.J., Parolin, R., Gordon, M., & Codding, R. (2007). The effects of extended time on mathematics performance of students with and without attention deficit hyperactivity disorder. Journal of Psychoeducational Assessment, 25 (1), 17 - 28

Lewandowski, L. J., Lovett, B.J., & Rogers, C.L. (2008). Extended Time as a Testing Accommodation for Students with Reading Disabilities: Does a Rising Tide Lift All Ships? Journal of Psychoeducational Assessment. 26 (315)

Li, Y., Li, N., Su, L. 2017. Reform and practice of diversified assessment of nursing and comprehensive experiment course under the drive of professional accreditation. Journal of Baotou Medical College, 33 (5), 47 – 54

Lichtenberger, E.O. & Kaufman, A.S. (2009) Essentials of WAIS-IV Assessment. Chichester. Wiley.

Liebel, S. W., & Nelson, J.M. (2017). Auditory and Visual Working Memory Functioning in College Students with Attention-Deficit/Hyperactivity Disorder and/or Learning Disabilities. Archives of Clinical Neuropsychology. 1 (12). doi: 10.1093/arclin/acx014

Lightfoot, A., Janemi, R., & Rudman, D. L. (2018). Perspectives of North American postsecondary students with learning disabilities: A scoping review. Journal of Postsecondary Education and Disability, 31, 57–74.

Lindstrom, J.H. (2010). Mathematics Assessment Accommodations: Implications of Differential Boost for Students with Learning Difficulties. Intervention in School and Clinic, 46 (5), 5-12.

Lindstrom, J. H., & Gregg, N. (2007). The role of extended tie on the SAT for students with learning disabilities and/or attention-deficit/hyperactivity disorder. Learning Disabilities: Research & Practice, 22, 85-95.

Livingston, E.M., Siegel, L.S., & Ribary, U. (2018) Developmental dyslexia: emotional impact and consequences, Australian Journal of Learning Difficulties, 23 (2), 107-135, DOI: 10.1080/19404158.2018.1479975

Lockwood, A.B., Farmer, R.L., Winans, S. & Sealander, K. (2021). Specific Learning Disability Identification Practices in the USA: a Survey of Special Education Administrators. Contemporary School Psychology, https://doi.org/10.1007/s40688-021-00375-4

López-Pastor, V., & Sicilia- Camacho, A. (2017). Formative and shared assessment in higher education. Lessons learned and challenges for the future. Assessment & Evaluation in Higher Education, 42, 77–97.

Lovett, B., J. (2010). Extended Time Testing Accommodations for Students with Disabilities: Answers to Five Fundamental Questions. Review of Educational Research. 80 (611)

Lovett, B.J. (2011). Extended Time Accommodations: What does the Research say? Communique: The newspaper of the National Association of School Psychologists (NASP). 39 (1)

Lovett, B. J., & Leja, A. M. (2015). ADHD symptoms and benefit from extended time testing accommodations. Journal of Attention Disorders, 19(2), 167-172. doi:10.1177/1087054713510560

Lovett, B. J. & Lewandowski, L. J. (2015). Testing accommodations for students with disabilities: research-based practice. American Psychological Association.

Lu, Y., & Sireci, S. G. (2007). Validity issues in test speededness. Educational Measurement: issues and Practice, 6, 29-37.

Lyman, M., Beecher, M. E., Griner, D., Brooks, M., Call, J., & Jackson, A. (2016). What keeps students with disabilities from using accommodations in postsecondary education? A qualitative review. Journal of Postsecondary Education and Disability, 29(2), 123-140. Retrieved from https://www.ahead.org/publications/jped

MacCullagh, L., Bosanquet, A., Badcock, N.A. (2017). University Students with Dyslexia: A Qualitative Exploratory Study of Learning Practices, Challenges and Strategies. Dyslexia, 23(1), 3-23.

Macfarlane, K. (2018, January 28). Making peace with testing accommodations. Ms. JD. https://ms-jd.org/blog/article/making-peace-with-testing-accommodations

Macfarlane, K. (2019, January 10). Testing accommodations are not a gift of extra time. Ms. JD. https://ms-jd.org/blog/article/testing-accommodations-are-not-a-gift-ofextra-time

Majer, Vanessa (2018): Perceptions of dyslexia held by students with dyslexia and their teachers within a secondary school. University of Leicester. Thesis. https://hdl.handle.net/2381/42927

Maki, K. E., & Adams, S. R. (2018). A current landscape of specific learning disability identification: Training, practices, and implications. Psychology in the Schools, 56(1), 18–31.

Mandinach, E.B., Bridgeman, B., Cahalan-Laitusis, C., & Trapani, C. (2005). The impact of extended time on SAT test performance. College Board Research Report 2005-8. New York. NY: College Board

Margari, L., Buttiglione, M., Craig, F.,Cristella, A., De Giambasttista, Matera, E., Operto, F., & Simone, M. (2013). Neuropsychopathological comorbidities in learning disorders. MBC Neurology, 13, 198. [Published online] doi: 10.1186/1471-2377-13-198

Marinkovich, J., Velásquez, M., Córdova, A., & Cid, C. (2016). Academic Literacy and Genres in University Learning Communities. Ilha do Desterro, 69 (3), 95-113. https://doi.org/10.5007/2175-8026.2016v69n3p95

Marshall, O.N. (2021). Applying the Social Model of Disablity: a Phenomenology of Initiating Change in Higher Education. *Unpublished thesis*.

Martinez-Garay I., Guidi, L.G., Holloway, Z.G., Bailey, M.A., Lyngholm, D., Schneider, T., Donnison, T., Butt, S.J., Monaco, A.P., Molnár, Z., & Velayos-Baeza A. (2017). Normal radial migration and lamination are maintained in dyslexia-susceptibility candidate gene homolog Kiaa0319 knockout mice. Brain Structure & Function. 222(3):1367-1384

Masterman, E. (2018). Typed versus handwritten essay exams. Is there a need to recalibrate the gauges for digital assessment? Australasian Society for Computers in Learning in Tertiary Education. Conference Proceedings, pp204-213. Deakin University, Australia.

McKendree, J & Snowling, M.J. (2011). Examination results of medical students with dyslexia. Medical Education, 45, 176-182. https://doi.org/10.1111/j.1365-2923.2010.03802.x

McKimm, S. (2012). Question marks over exam access arrangements for candidates with disabilities. Independent Schools Council Bulletin (ISC). 29.

Meachon, E. (2017). An Investigation of Dyspraxia: What do we know and why the research is so far behind. St Anne's Academic Review. 7, 47 – 52.

Mertens, D. M. (2003). Mixed methods and the politics of human research: the transformative-emancipatory perspective. Handbook of mixed methods in social and behavioral research, 135-164.

Mertens, D.M. (2007). Transformative Paradigm: Mixed Methods and Social Justice. Journal of Mixed Methods Research, 1(3), 212-225.

Mertens, D.M., Bledsoe, K.L., Sullivan, M, & Wilson, A. (2010). Utilization of Mixed Methods for Transformative Purposes. In A. Tashakkori, & C. Teddlie. (Eds.). (2010). Sage handbook of mixed methods in social & behavioral research. Sage.

Mills, J., Bonner, A., & Francis, K. (2006). The Development of Constructivist Grounded Theory. International Journal of Qualitative Methods, 5 (1), 25 – 35.

Mogey, N., Sarab, G., Haywood, J., Van Heyningen, S., Dewhurst, D., Hounsell, D., & Neilson, R. (2007). Journal of Computer Assisted Learning, 24, (1).

Moll, K., Gobel, S.M., Gooch, D., Landerl, K., Snowling, M.J. (2016). Cognitive Risk Factors for Specific Learning Disorder: Processing Speed, Temporal Processing, and Working Memory. Journal of Learning Disabilities, 49 (3), 272-281 https://doi.org/10.1177/0022219414547221

Moll, K., Snowling, M J.., & Hulme, Charles. (2020). Introduction to the special issue "comorbidities between reading disorders and other developmental disorders". Scientific Studies of Reading, 24(1), 1–6

Monagle, K. (2015). Beyond access: An examination of factors that influence use of accommodations by college students with disabilities. Dissertation Abstracts International: Section A. Humanities and Social Sciences. 77/08(E)

Mor, S. (2019, May). Access to higher education for students with disabilities: International and national developments. Paper presented at the 15th Interdisciplinary Conference in Disability Research, Copenhagen, Denmark.

Morgan, D.L. (2020). Pragmatism as a Basis for Grounded Theory. The Qualitative Report. 25 (1), 64-73.

Mortimore, T., & Crozier, W. (2006). Dyslexia and difficulties with study skills in higher education. Studies in Higher Education, 31 (2), 235 – 251. https://doi. org/10.1080/03075070600572173

Mortimore, T. (2012). Dyslexia in Higher Education: developing a fully inclusive institution. Journal of Research in Special Educational Needs,13(1), 38-47.

Mortimore, T. (2013). Learning disability/dyslexia and higher education: creating an inclusive institution in higher education. In Smith, L.R (Ed.). Higher Education: recent trends, emerging issues and future outlook. Education in a competitive and globalizing world. Nova, New York.

Moshkani, M., Edris, M., Maryam, L. & Ghorban, E. (2017). Effect of phonological awareness on improving reading and self-esteem of students with dyslexia. Journal of Child Mental health. 4 (3), 107-118.

Mountford, H., Hill, A., Barnett, A. & Newbury, D. (2021). Genome-Wide Study of Motor Coordination. Frontiers in Human Neuroscience. 15, 1- 13

Mourshed, M, Krawitz, M., Dorn, E. (2017, October 4th). How to improve student educational outcomes: New insights from data analytics . McKinsey & Company. Retrieved from https://www.mckinsey.com/industries/social-sector/our-insights/how-to-improve-student-educational-outcomes-new-insights-from-data-analytics

Munhall, P. (2001). Ethical considerations in qualitative research. In P. Munhall (Ed.), Nursing research: A qualitative perspective (3rd ed.), pp. 537-549. Sudbury, MA: Jones and Bartlett.

National Association of Disability Practitioners (NADP). https://nadp-uk.org/ [accessed 27.3.2021]

Nandika, K (2020). Self-Directed Learning Model to Develop Academic Self-Concepts of Class XI Students in 2019/2020 Academic Year. International Journal of Educational Studies, 13 (1), 61 – 80.

Nelson, J. M., Gregg, N. (2012). Depression and anxiety among transitioning adolescents and college students with ADHD, dyslexia, or comorbid ADHD/dyslexia. Journal of Attention Disorders, 16, 244–254.

Nelson, J.M., Lindstrom, W., & Foels, P.A. (2015). Test Anxiety among College Students with Specific Reading Disability (Dyslexia): Non-verbal Ability and Working Memory as Predictors. Journal of Learning Disabilities, 48 (4), 422-432.

Nelson, L. M., & Reynolds, T. W. (2015). Speech recognition, disability, and college composition. Journal of Postsecondary Education and Disability, 28(2), 181-197. Retrieved from https:// www.ahead.org/publications/jped

Newlands, F., Shrewsbury, D., & Robson, J. (2015). Foundation doctors and dyslexia: A qualitative study of their experiences and coping strategies, Postgraduate Medical Journal, 91(1073), 121–126.

Newman. I. (2019). When saying 'go read it again' won't work: Multisensory ideas for more inclusive teaching and learning. Nurse Education in Practice, 34, 12-16.

Newman, L. A., & Madaus, J. W. (2015). Reported accommodations and supports provided to secondary and postsecondary students with disabilities: National perspective. Career Development and Transition for Exceptional Individuals, 38, 173–181.

Norris, D., Page, M.P.A., & Hall, J. (2018). Learning nonwords: the Hebb repetition effect as a model of word learning. Memory, 26 (6), 852-857, DOI: 10.1080/09658211.2017.1416639

Nieminen, J. H. (2020). Disrupting the power relations of grading in higher education through summative self-assessment. Teaching in Higher Education, 1–16.

NVivo qualitative data analysis software; QSR International Pty Ltd. Version 12, 2018.

O'Brien, B.A., Van Orden, G.C. & Pennington, B.F. (2013). Do dyslexics misread a ROWS for a ROSE?. Reading and Writing, 26, 381–402. https://doi.org/10.1007/s11145-012-9373-8

O'Byrne, C., Jagoe, C., & Lawler, M. (2019). Experiences of dyslexia and the transition to university: a case study of five students at different stages of study. Higher Education Research & Development, 38, 1031 - 1045.

Office for Students. (2021). Consultation on recurrent funding for 2021/22. https://www.officeforstudents.org.uk/media/8610a7a4-0ae3-47d3-9129-f234e086c43c/consultation-on-funding-for-ay2021-22-finalforweb.pdf [last accessed 17.7.21]

Ofiesh, N., Mather, N., & Russell, A. (2005). Using speeded cognitive, reading, and academic measures to determine the need for extended test time among university students with learning disabilities. Journal of Psychoeducational Assessment, 23, 35-52.

Ofiesh, N., Moniz, E., & Bisagno, J. (2015). Voices of university students with ADHD about test-taking: Behaviors, needs, and strategies. Journal of Postsecondary Education and Disability, 28(1), 109-114. Retrieved from https://www.ahead.org/publications/jped

OfQual (2020). 'Access Arrangements for GCSE AS and A Level: 2019 to 2020 Academic Year: Additional Report'. Office of Qualifications and Examinations Regulation. [online] https://www.gov.uk/government/statistics/access-arrangements-for-gcse-as-and-a-level-2019-to-2020-academic-year/access-arrangements-for-gcse-as-and-a-level-2019-to-2020-academic-year [last accessed 13.7.21]

OfQual (2019). 'Examination Reform: Impact of Leniar and Modulear Examinations at GCSE'. Office of Qualifications and Examinations Regulation. Ofqual/19/6506/1. www.gov.uk/ofqual

Oliver, M. & Barnes, C. (2012). The New Politics of Disablement. Palgrave Macmillan. Basingstoke.

Olofsson, A., Taube, K., & Ahl, A. (2015). Academic achievement of university students with dyslexia. Dyslexia, (21), 338– 349.

Olofsson, Å., Ahl, A., & Taube, K. (2012). Learning and study strategies in university students with dyslexia: Implications for teaching. Procedia Social and Behavioral Sciences, 47, 1184–1193.

Osborne, P. (2006). Pilot Study to Investigate the Performance of Dyslexic Students in Written Assessments. Innovations in Education and Training International, 36 (2), 155-160

Paolini, A. (2015). Enhanced Teaching Effectiveness and Student Learning Outcomes. The Journal of Effective Teaching, 15 (1), 20-33.

Pardy, B. (2016). Head Starts and Extra Time: Academic Accommodation on Post-Secondary Exams and Assignments for Students with Cognitive and Mental Disabilities. Education & Law Journal, 25, 191- 208

Payne-Tsoupros, C. (2020). A starting point for disability justice in legal education. Journal Committed to Social Change on Race and Ethnicity, 6 (1), 165 – 189.

Peer, L. and Reid, G. (2002). “Dyslexia and literacy: challenges in the secondary school”. In Dyslexia and literacy: theory and practice, Edited by: Reid, G. and Wearmouth, J. 241–250. Chicester: Wiley.

Pereira, D., Flores, M. A., & Barros, A. (2017). Perceptions of Portuguese undergraduate students about assessment: A study in five public universities. Educational Studies, 43, 442–463.

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. Psychological Science, 30(3), 386–395.

Peterson, M. C. (2016). Assistive technology management by disabilities services managers in higher education: A phenomenological study. Dissertation Abstracts International: Section B. Sciences and Engineering, 77/12(E).

Peterson, R., & Pennington, B.F. (2012). Developmental dyslexia. The Lancet. 379 (9830), 1997 – 2007.

Phillips, S.E. (1994). High-stakes testing accommodations: Validity versus disabled rights. Applied Measurements in Education, 7 (2).

Pino, M., & Mortari, L. (2014). The inclusion of students with dyslexia in higher education: a systematic review using narrative synthesis. Dyslexia, 20 (4), 346–369. https://doi.org/10.1002/dys.1484

Portolese, L., Krause, J. & Bonner, J. (2016). Timed Online Tests: Do Students Perform Better With More Time? *American Journal of Distance Education*, 30:4, 264-271, DOI: 10.1080/08923647.2016.1234301

Prasad, G.N.R. (2020). Evaluating student performance based on Bloom’s taxonomy levels. Journal pf Physics: Conference Series, 1797, 1-16.

Rabionet, S. E. (2011). How I learned to design and conduct semi-structured interviews: An ongoing and continuous journey. The Qualitative Report, 16(2), 563.

Ragosta, M., Braun, H., & Kaplan, B. (1991). Performance And Persistence: a Validity Study of The SAT For Students With Disabilities. Wiley Online Library. ETS Research Report Servies. 1991 (2). DOI: 10.1002/j.2333-8504.1991.tb01408.x

Rahim, N., & Meon, H. (2012). Relationship between Study Skills and Academic Performance. Conference Paper: The 20th national symposium on mathematical sciences: Research in Mathematical Sciences: A Catalyst for Creativity and Innovation. 1522. 10.1063/1.4801264.

Rauschenberger, M., Baeza-Yates,R., Rello,L. (2019). Technologies for Dyslexia. In Yesildaa, Y., & Harper, S. (Eds). Web Accessibility, pp 603 – 629. Springer. London.

Raven, J. (1981). Manual for Raven's Progressive Matrices and Vocabulary Scales. Research supplement no. 1: The 1979 British standardisation of the Standard Progressive Matrices and Mill Hill Vocabulary Scales, together with comparative data from earlier studies in the UK, US, Canada, Germany, and Ireland. Oxford, England: Oxford Psychologists Press/San Antonio, TX: The Psychological Corporation.

Reedy, A., Pfitzner, D., Rook, L. *et al.* (2021). Responding to the COVID-19 emergency: student and academic staff perceptions of academic integrity in the transition to online exams at three Australian universities. *Int J Educ Integr,* 17 (9). https://doi.org/10.1007/s40979-021-00075-9

Richardson, J.T.E. (2015). Coursework versus examinations in end-of-module assessment: a literature review. Assessment & evaluation in Higher Education, 40 (3), 439 – 455.

Richardson, J.T.E. (2015). Attainment in Students with Dyslexia in Distance Education. Dyslexia, 21, 323-337.

Ricketts, C., Brice, J., & Coombes, L. (2010). Are multiple choice tests fair to medical students with specific learning difficulties? Advances in Health Science Education, 15 (2), 265 – 275.

Riddell, S., & Weedon, E. (2006). What counts as a reasonable adjustment? Dyslexic students and the concept of fair assessment. International Studies in Sociology of Education, 16, 57– 73.

Rodger, J., Wilson, P., Roberts, H., Roulstone, A., Campbell, T. (2015). Support for Higher Education Students with Specific Learning Difficulties. Report to HEFCE by York Consulting & University of Leeds. [last accessed 24.12.19] Online https://dera.ioe.ac.uk/23654/1/HEFCE2015_spld.pdf

Rogers, C. M., Thurlow, M. L., Lazarus, S. S., & Liu, K. K. (2019). A summary of the research on effects of test accommodations: 2015-2016 (NCEO Report 412). Minneapolis, MN: University of Minnesota, National Center on Educational Outcomes.

Ruhkamp, R. (2015). Lived experiences of undergraduate and graduate students utilizing accommodations. Dissertation Abstracts International: Section A. Humanities and Social Sciences. 76/11(E).

Runyan, M. K. (1991). The Effect of Extra Time on Reading Comprehension Scores for University Students With and Without Learning Disabilities. Journal of Learning Disabilities. 24 (2), 327-340.

Runyan, M. K., & Smith, J. (1991). Identifying and accommodating learning disabled law school students. Journal of Legal Education. 41, 317–349.

Ryder, D., & Norwich, B. (2018). What's in a name? Perspectives of dyslexia assessors working with students in the UK higher education sector. Dyslexia, 24, 109-127.

Ryder, D., & Norwich, B. (2019). UK higher education lecturers' perspectives of dyslexia, dyslexic students and related disability provision. Journal of Research in Special Educational Needs, 19(3), 161-172. https://doi/ epdf/10.1111/1471-3802.12438

Saeidi, A., Hajipoour, R., Ghafouri, M., Gholinia, H., & Mousavi, M. (2019). Investigation of the Relationship Between Dental Students' Study Skills and Their Educational Status in Babol University of Medical Sciences. Biannual Journal of Medical Education Development Centre. Babol University of Medical Sciences, 7 (1), 47-5.

Sarid, M., Meltzer, Y. & Raveh, M. (2020). Academic Achievements of College Graduates With Learning Disabilities Vis-à-vis Admission Criteria and Academic Support. Journal of Learning Disabilities, 53(1), 60 – 74.

Sawilowsky, S (2009). "New effect size rules of thumb.". Journal of Modern Applied Statistical Methods. 8 (2), 467–474.

Sayeski, K.L. (2019). Dyslexia. Teaching Exceptional Children, 51(3), 170 – 171.

Shakespeare, T. (Ed.). (2015). Disability Research Today: International Perspectives. Routledge. Oxon.

Shaw, S. C. & Anderson, J. (2017). Doctors with dyslexia: a world of stigma, stonewalling and silence, still? MedEdPublish, 6 (1), 29. doi.org/10.15694/mep.2017.000029

Shaw, S. C. & Anderson, J. (2018). The experiences of medical students with dyslexia: An interpretive phenomenological study. Dyslexia, 24(3), 220 - 233.

Shaw, S. C. K., Anderson, J. L. & Grant, A. J. (2016). Studying medicine with dyslexia: a collaborative autoethnography. The Qualitative Report. 21 (11), 2036-54.

Shaw, S. C. K., Malik, M., & Anderson, J. L. (2017). The exam performance of medical students with dyslexia: a review of the literature. MedEdPublish. 6 (3:2) DOI: https://doi.org/10.15694/mep.2017.000116

Shetty, S. S.., Srinivasan, S.J. (2014). Effectiveness of study skills on academic performance of dental students. Journal of Education and Ethics in Dentistry. 4, 28 – 31. DOI: 10.4103/0974-7761.143175

Shrewsbury, D. (2016). Dyslexia in general practice education: considerations for recognition and support, Education for Primary Care, 27:4, 267-270, DOI: 10.1080/14739879.2016.1194170.

Sireci, S. G. & Parker, P. (2005). Validity on trial: Psychometric and legal conceptualizations of validity. Educational Measurement: Issues and Practice, 25, 27–34.

Sireci, S. G., Scaroati, S. E., & Li, S. (2005). Test accommodations for students with disabilities: An analysis of the interaction hypothesis. Review of Educational Research, 75 (4), 457-490.

Snowling, M.J., & Griffiths Y.M. (2004) Individual Differences in Dyslexia. In: Nunes T., Bryant P. (eds) Handbook of Children's Literacy. Springer, Dordrecht.

Snowling, M.J., & Hulme, C. (2012). Annual Research Review: The nature and classification of reading disorders. A commentary on proposals for DSM-5. Journal of Child Psychology and Psychiatry, 53, 593-607.

Snowling, M.J., & Hulme, C. (2016). Reading Disorders and dyslexia. Current Opinions in Paediatrics, 28 (6), 731-735.

Snowling, M.J., Hulme, C., & Nation, K, (2020). Defining and Understanding Dyslexia: Past, Present and Future. Oxford Review of Education, 46(4), 501- 513.

Sokal, L., & Wilson, A. (2017). In the nick of time: A pan-Canadian examination of extended testing time accommodation in post-secondary schools. Canadian Journal of Disability Studies, 6(1), 28- 62.

Spencer, M., Wagner, R. K., Schatschneider, C., Quinn, J. M., Lopez, D., & Petscher, Y. (2014). Incorporating RTI in a hybrid model of reading disability. *Learning Disability Quarterly*. 37 (3), 161-171.

SpLD Assessment Standards Committee (SASC). (2020). SpLD Working Group 2005/DfES guidelines and updates. https://sasc.org.uk/ [last accessed 11.7.21].

Stagg, S.D., Eaton, E., & Sjoblom, A.M. (2018). Self-Efficacy in undergraduate students with dyslexia: A mixed methods investigation. British Journal of Special Education, 45 (1), 26-42.

Stein, J. (2018). What is developmental dyslexia? Brain Sciences. 8 (2), 26 - 39.

Stein, J. (2019). The current status of the magnocellular theory of developmental dyslexia. Neuropsychologia. 130, 66-77.

Strauss, A.L., & Corbin, J.M. (1990). Basics of qualitative research: Grounded theory procedures and techniques. Newbury Park, CA: Sage.

Strauss, A.L., & Corbin, J.M. (1994). Grounded theory methodology: An overview. In N. Denzin & Y. Lincoln (Eds.), Handbook of qualitative research (pp. 273-285). Thousand Oaks, CA: Sage.

Strauss, A.L., & Corbin, J.M. (2015). *Basics of qualitative research: techniques and procedures for developing grounded theory*. 4th Ed. Thousand Oaks, CA. SAGE.

Suarex-Coalla, P., Afonso, O., Martinez-Garcia, C. & Cuetos, F. (2020). Dynamics of Sentence Handwriting in Dyslexia: The Impact of Frequency and Consistency. Frontiers in Psychology. 11, 319. doi: 10.3389/fpsyg.2020.00319

Sumner, E., Pratt, M.L. & Hill, E.L. (2016). Examining the cognitive profile of children with Developmental Coordination Disorder. Research in Developmental Disabilities. 56, 10-17.

Sumner, E., Crane, L., & Hill, E. (2021). Examining academic confidence and study support needs for university students with dyslexia and/or developmental coordination disorder. Dyslexia: An international journal of research and practice, 26 (3), 1-16

Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2017). Developing evaluative judgement: Enabling students to make decisions about the quality of work. Higher Education, 76, 467–481.

The Association of Dyslexia Specialists in Higher Education (ADSHE). https://adshe.org.uk/ [accessed 9.1.21]

Thornton, A.E., Reese, L.M., Pashley, P.J., & Delassandro, S.P. (2001). Predictive validity of accommodated LSAT scores. (Technical Report 01-01). Newtown, PA: Law School Admission Council. Newtown, PA. https://files.eric.ed.gov/fulltext/ED469183.pdf

Tie, Y.C., Birks, M., Francis, K. (2019). Grounded Theory Research. A design framework for novice researchers. Sage Open Med. (7). Published online doi: 10.1177/2050312118822927

Timmerman, L. C., & Mulvihill, T. M. (2015). Accommodations in the college setting: The perspectives of students living with disabilities. The Qualitative Report, 20(10), 1609-1625.

Tops, W., Callens, C., & Van Cauwenberghe, E. (2013). Beyond spelling: the writing skills of students with dyslexia in higher education. Reading & Writing, 26, 705–720. https://doi.org/10.1007/s11145-012-9387-2

Tseng, Y., Tsai, C., Chen, F. & Konczak, J. (2019) Position Sense Dysfunction Affects Proximal and Distal Arm Joints in Children with Developmental Coordination Disorder, Journal of Motor Behavior, 51 (1), 49-58.

Tyszko, J. A., & Sheets, R. G. (2019). Co-designing assessment and learning: Rethinking employer engagement in a changing world. Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Ukwueze, A. C. (2015). Effect of Dysgraphia on Students' Performance in Examinations: Implications for Counselling. Journal of the Nigerian Academy of Education, 11(2), 33-50.

Verleger, M. A. (2016), *Just Five More Minutes: The Relationship Between Timed and Untimed Performance on an Introductory Programming Exam.* Paper presented at 2016 ASEE Annual Conference & Exposition, New Orleans, Louisiana. 10.18260/p.25510

Villarroel, V., Boud, D., Bloxham, S., Bruna, D & Bruna, C. (2020) Using principles of authentic assessment to redesign written examinations and tests, Innovations in Education and Teaching International, 57(1), 38-49, DOI: 10.1080/14703297.2018.1564882

Visser, L., Röschinger, J., Barck, K., Büttner, G., & Hasselhorn, M. (2020) Learning Difficulties in Children with Symptoms of DCD And/or ADHD: Analyses from a Categorical and a Continuous Approach. International Journal of Disability, Development and Education. 1-17 DOI: 10.1080/1034912X.2020.1786023

Wade, M.G. & Kazeck, M. (2018) Developmental coordination disorder and its cause: The road less travelled. Human Movement Science. 57, 489-500.

Wagner, R.K., Zirps, F.A., Edwards, A.A., Wood, S.G., Joyner, R.E., Becker, B.J., Liu, G., & Beal, B. (2020). The Prevalence of Dyslexia: A New Approach to Its Estimation. Journal of Learning Disabilities. 53 (5), 354-365.

Walker, E., Shaw, S.C.K., Reed, M. & Anderson, J.L. (2021). The experiences of foundation doctors with dyspraxia: a phenomenological study. Advances in Health Science Education. 1-16

Warren, J, (2017). 25 words per minute – the expected undergraduate writing speed? Journal of Neurodiversity in Higher Education, 3, 5 – 17.

Weaver, S.M. (1993). The validity of the use of extended and untimed testing for postsecondary students with learning disabilities. Unpublished doctoral dissertation, University of Toronto, Canada

Wechsler, D. (1981). WAIS-R: Wechsler adult intelligence scale-revised. New York, N.Y. Psychological Corporation,

Wechsler, D. (2004). Wechsler Intelligence Scales for Children– Revision IV. Pearson. St Antionio. TX

Wechsler, D. (2008). Wechsler Adult Intelligence Scales – Revision IV. Pearson. St Antionio. TX

Weis, R., Dean, E.L. & Osbourne, K, J. (2016). Accommodation Decision Making for Postsecondary Students with Disabilities: Individually Tailored or One Size Fits All? Journal of Learning Disabilities, 5, 484 – 498.

Wibrowski, C., Matthews, W., & Kitsantas, A. (2017). The Role of a Skills Learning Support Program on First-Generation College Students' Self-Regulation, Motivation, and Academic Achievement: A Longitudinal Study. Journal of College Student Retention: Research, Theory & Practice, 19 (3), 317 – 332. DOI: 10.1177/1521025116629152

Wigfield, A., Harold, R. Freddman-Doan, C., Eccles, J., Yoon, K.S., Arbreton, A.J.A., Blumenfield, P.C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: a 3-year study. Journal of Educational Psychology, 89(3), 451-469.

Wiggins, G. (1990). The Case for Authentic Assessment. Practical Assessment, Research, and Evaluation. 2(2), p1–3.

Williams, A. D. (2015). Middle school students' experience of receiving test accommodations. Dissertation Abstracts International: Section A. Humanities and Social Sciences. 76/12(E).

Wilson, P.H., Smits-Engelsman, B., Caeyenberghs, K., Steenbergen, B., Sugden, D.A., Clark, J.E., Mumford, N. & Blank, R. (2017). Cognitive and neuroimaging findings in Developmental Coordination Disorder: New insights from a systematic review of recent research. Developmental Medicine & Child Neurology. 59(11), 1117-1129.

World Health Organisation (WHO). (2010). The International Classification of Impairments, Activities and Participation. ICIDH-2 https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health [last accessed 11.7.2021]

Yssel, N., Pak, N., & Beilke, J. (2016). A door must be opened: Perceptions of students with disabilities in higher education. International Journal of Disability, Development and Education, 63(3), 384-394. doi:10.1080/1034912X.2015.1123232

Yu, M., Novak, J. A., Lavery, M. R., Vostal, B. R., & Matuga, J. M. (2018). Predicting college completion among students with learning disabilities. Career Development and Transition for Exceptional Individuals, 41, 234–244.

Zambrano, A. (2016). The experience of student with disabilities in higher education. Dissertation Abstracts International: Section A. Humanities and Social Sciences. 77/11(E).

Zoccolotti, P., De Luca, M., Marinelli, C. V., & Spinelli, D. (2020). Testing the specificity of predictors of reading, spelling and maths: A new model of the association among learning skills based on competence, performance and acquisition. Frontiers in Human Neuroscience, 14, Article 573998. https://doi.org/10.3389/fnhum.2020.573998

Zurcher, R., & Bryant, D.P. (2001). The Validity and comparability of entrance examination scores after accommodations are made for students with LD. Journal of Learning Disabilities, 34 (5), 462-471.

Zuriff, G.E. (2000). Extra examination time for students with learning disabilities: An examination of the maximum potential thesis. Applied Measurement in Education, 13 (1), 99-117.

16: APPENDICES

16.1 Appendix 1: Literature review search strategy

16.1.1 Search strategy

In order to review the published studies exploring the impact of granting extra time in exams to students with SpLD, a search across a range of educational and multidisciplinary databases for conference and peer-reviewed journal papers was carried out (Booth, Sutton & Papaaioannou, 2016). The key terms used for this search included, 'exam', or 'test', or assessment' plus 'adjustments', or 'accommodations', 'arrangements', 'access', 'special arrangements', 'extra time', 'additional time, 'timed tests*'. This core search was combined with SpLD specific terms including, 'specific learning difficulties', learning disabilities', 'reading disabilities', 'dyslexia', 'dyspraxia', 'DCD'. The databases searched are shown in Table 16.1.

Table 16.1: Databases of Journals and conference papers searched

Database	Coverage
British Education Index (including conference abstracts)	Education (UK & Europe)
Australian Education Index	Education (Australasia)
Education Resources Information Centre (ERIC)	Education (USA)
PubMed	Multidisciplinary
Scopius	Multidisciplinary
Google Scholar	Multidisciplinary
Web of Science (including Social Sciences & Humanities conference data base)	Multidisciplinary
iDiscover (searches Cambridge University library collection of printed and electronic texts, journals, databases, e-journals & on-line resources)	Multidisciplinary
Education Research: Higher Education Academy	Education
PsychINFO, PsychARTICLES, PsychBOOKS	Psychology & related disciplines

16.1.2 Inclusion and exclusion criteria

The search included studies published between 1980 and 2018, and did not exclude any geographical locations. For the quantitative literature review, peer-reviewed papers that used an experimental or quasi-experimental design were included, as were articles that involved systematic reviews and meta-analyses. Due to the large number of articles identified by the key word search, studies with less than an 80% power to detect an effect were excluded. For the qualitative literature review, studies that used a qualitative approach and focused on tertiary education were included, but those using an experimental or quasi-experimental design were excluded. As the purpose of this paper is to explore the effect of exam arrangements in exams on students with SpLD, studies investigating the impact of test adjustments on students with other disabilities, including those with Attention Deficit Hyperactivity Disorder or Autism Spectrum

Disorder were excluded. The screening and inclusion/exclusion process is shown diagrammatically in Figure 16.2.

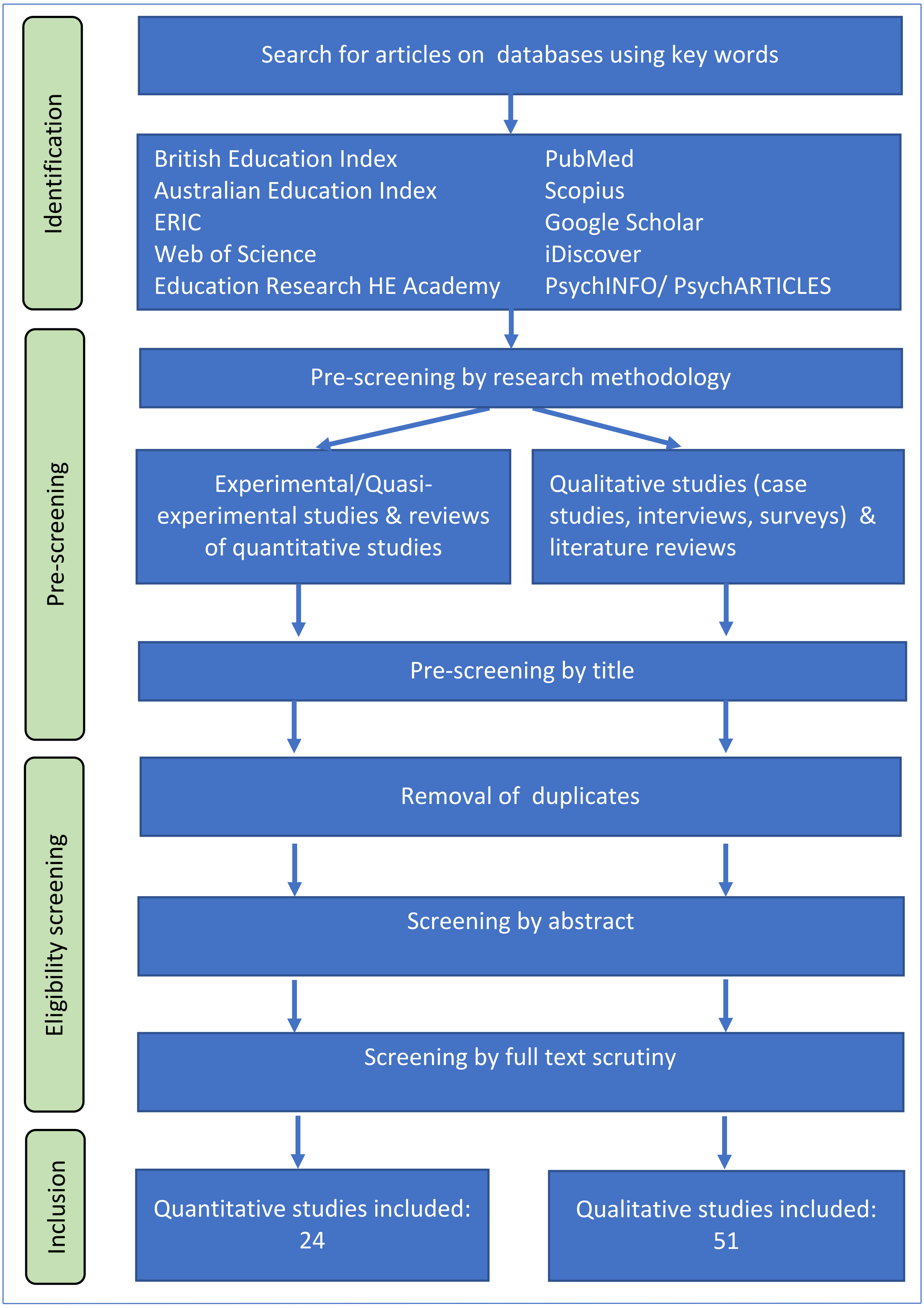


Table 16.2: Flow diagram of the literature search and screening process (adapted from Bengtsson 2019).

16.2 Appendix 2: Data collection and handling protocol

16.2.1 Quantitative Phase 1

- All quantitative data collected was obtained anonymously. The exam scripts was provided to the researcher by the University's Board of Exams and labelled by candidate number only, which ensured that the researcher was unable to identify the participants.

- The participant information sheets and consent forms were be sent to participants by the university's Board of Exams. The Board of Exams then collated the responses and forwarded anonymised exam papers to the researcher for those candidates who had given their consent (for opt-in consent). The Board of Exams also forwarded the exam papers of the index group participants who had not responded to withdraw from the project after 3 weeks of being informed of the research project. This process ensured that the researcher was not informed of the identity of any participant. The Board of Exams informed the researcher whether or not each exam paper was granted exam arrangements (and the nature of those arrangements).

- Unlike the researcher, the Board of Exams was able to identify those candidates who had disclosed a SpLD (and so required the opt-in form) separately from the TD candidates (who received the opt-out consent form). This ensured that anonymity was maintained while fully informed consent was obtained.

- Similarly, the Board of Exams was able to ensure that the students who were invited to participate did not include those who had disclosed a SpLD but who did not request exam access arrangements. Participants who were sent 'opt-out' consent were given 3 weeks to 'opt-out' before their scripts were sent to the researcher. All participants still had the right to withdraw their data from the research project for two years after this.

- The researcher replaced the candidate number with a participant number, so that the data obtained from each exam scripts could not be identified with a specific candidate through the candidate number.

16.3 Appendix 3: Ethical considerations

The guidelines of the British Psychological Society (British Psychological Society, 2014 have been followed throughout this research project. Specifically, care has been taken to protect the anonymity and confidentiality of the participants. This was achieved in this research project through the following procedures:

- All data collected is held anonymously.
- Electronic data is stored on a password protected computer, with the password known only by researcher.
- Only authorised persons such as researchers within the team, supervisors, and, for monitoring quality, regulatory authorities / R&D audit may have access to view both the candidate number and participant number.
- The data will be used for this research project only.
- This data will be retained for a minimum of 5 years and will be disposed of securely by the licensed confidential information disposal company used by the University for the purposes of disposing of confidential material.
- The handling, processing, storage and destruction of all data will be in accordance with the Data Protection Act 1998.

In phase 1 of the project, the following additional measures were also taken to protect the anonymity of the participants:

- All data collected was obtained anonymously. The exam scripts were provided to the researcher by the University's Board of Exams and labelled by candidate number only, which ensured that the researcher was unable to identify the participants.
- The participant consent forms were sent to participants by the University's Board of Exams. The Board of Exams then collated the responses and forwarded anonymised exam papers to the researcher for those candidates who had given their consent. This process ensured that the researcher was not informed of the identity of any participant.
- The researcher replaced the candidate number with a participant number, so that the data obtained from each exam script could not be identified with a specific candidate through the candidate number.

In phase 2 and 3 of the project, unlike in phase 1, the identities of the participants were known to the researcher. To protect the anonymity of the participants all data collected is held anonymously. The researcher replaced the participant's name with a participant number (phase 2) or pseudonym (phase 3), so that the data obtained cannot be identified with a specific participant through their name or any other identifying characteristics.

The final phase of the project (the qualitative phase) took the form of semi-structured interviews with the result that the identities of the participants were known to the researcher. To protect the anonymity of the participants all data collected is held anonymously with the identity of the participants known only to the researcher. The researcher replaced the participant's name with a code name, so that the responses obtained from each participant cannot be associated with any specific participant through their name or any other identifying characteristics. The interviews were carried out remotely (in accordance with public health guidance on physical distancing) via Microsoft Teams. All the students at the University from where the participants were recruited had been given access to MS Teams and were using it to access their online tutorials. It has good accessibility, has live captioning and a recording facility, and is familiar to the students. At the start of interview, the researcher reiterated the information about the focus of the research and why it was being conducted with the student. The researcher re-informed each participant that the interview would be audio recorded, but their information would be treated in confidence and the research would not name or identify them in any report of the findings. The researcher stressed that their participation was completely voluntary, and they could choose to stop the interview at any time and ask for their information to be removed from the research project at any point until the end of July 2021 (anticipated submission date for the thesis). Before commencing with the interview questions, the researcher checked that the student had read and understood the student participation sheet and had completed and signed the consent form. The researcher offered to answer any questions the participant had about the research project and their participation before starting and finally asked the student if they still wished to proceed.

Informed consent was obtained from all participants via accessible[168] information sheets (appendix 1) and consent forms (appendices 2). For those participants with a diagnosis of a SpLD, 'opt-in' consent was sought. For those participants who did not have a diagnosis of a SpLD 'opt-out' consent was sought. The participants were sent the participant information sheet in an accessible format (and in audio version via Sensus-access) in advance of the interview to allow them time to process the information and consider their involvement.

The four BPS ethical principles of respect, competence, responsibility and integrity were adhered to by ensuring that each candidate was provided with detailed information about the purpose of the research project, the reasons why they had been invited to participate, the voluntary nature of their participation, how their data and the results of the research would be used, how confidentiality would be maintained, the potential advantages and risks of participation, how to withdraw or make a complaint, and full contact details of the researcher (appendices 1 and 2). Full details of the data collection and data handling protocols for each phase of the project are provided in appendix 3.

One of the identified risks of this research project was the potential finding that exam access arrangements provide a significant advantage to students with SpLD. This could then provoke a change in policy in relation to the granting of the use of a word processor and /or extra time in exams, thus further marginalizing students with SpLD. However, this risk may be offset by citing national practices throughout the HE sector where such arrangements are commonly granted and regarded as 'reasonable adjustments' as well as the institution's legal obligations under the Equality Act 2010 to ensure that students with SpLD are not disadvantaged. In addition, proposing alternative models (such as Universal Design) that equally seek to relieve students with SpLD of significant disadvantage in a formal, timed exam situation aim to ensure that the participants are

[168] All participant forms were fully compatible with screen readers and were offered in alternative formats, including audio form.

not disempowered by the outcomes of the research project should a change in policy be provoked. It should also be stressed that the exam results of any participant would not be altered or reviewed should this research demonstrate any advantage has been conferred.

16.4 Appendix 4: Protocol for counting words on exam papers[169]

1. *Title:* Do not include words in the title

2. *Numbers:* Numbers of one digit or more are counted as one word.

3. *Times*: times are counted as one word. E.g. 9:04 or 12.30am

4. *Dates*: dates are counted as one word. E.g. 10/6/09

5. *Legible crossed-out words:* These are counted. This includes one letter words, e.g. 'I', 'A'

6. *Misspelt words*: these are counted

7. *Unfinished words*: these are counted if two or more letters have been written, but not if this is the final word

8. *Abbreviations/acronyms:* These are counted as one word, e.g. 'St.' 'etc.'

9. *Hyphenated words*: These are counted as two words, even if the hyphen is used to continue the word onto next line, e.g. further- more

10. *Deciding whether to score one or two words*: If there is a space between the words (even if they are normally one word, e.g. mid wife) score as two words. If in doubt about whether there is a space, score as two words. If words are close together but grammatically two separate words, count as two words, e.g. 'alot', 'aswell'.

11. *Totally illegible words:* Words that cannot be read (even given the context of the sentence) should be bracketed, counted up and total recorded. Crossed out illegible words are counted.

Record

- The number of legible words written on each paper
- The number of illegible words written on each paper
- The total number of words (legible plus illegible)

[169] Protocol taken from DASH-17+ (Barnett et al, 2010, pp40 – 41)

16.5 Appendix 5: Examples of exam paper questions

16.5.1 English

The paper selected from the English programme was the Practical Criticism paper. This is a compulsory paper that is taken by all undergraduate English students in each year of the degree programme and requires candidates to analyse the relationship between a critical quotation and previously unseen passages. Examples of questions are:

> Write a detailed analysis of B that draws on A in any way you see fit (where B is an excerpt from Aphra Benn's 'The Emperor of the Moon: A farce' and A is an excerpt from the Ecclogues of Virgil)

> Write an essay which responds to A through a detailed analysis of B. (where A is an excerpt from 'Poetry and the Fate of the Senses' by Susan Stewart and B is the medieval lyric 'Sumer is i-cumen in')

16.5.2 Law

The compulsory paper selected from the Law programme was Constitutional Law in Year 1, Contract Law in Year 2 and Equity paper in Year 3. Examples of questions are:

> To what extent, if any, does the uncodified nature of the UK constitution inhibit its capacity to embody and uphold the rule of law and separation of powers principle? [constitutional Law]

> 'The doctrine of penalties can be regarded as meeting the criticism levelled against unqualified freedom of contract, namely the possible inequality of bargaining power.' Consider critically the current state of the English law on penalties in the light of this statement.[Contract Law]

> The current law on trustee exemption clauses makes a mockery of the duties carefully developed by the courts over the centuries. Trustees should never be able to exclude liability for breach of trust.' Discuss. [Equity]

16.5.3 History

The compulsory paper selected from the History programme was History of Political Thought in Year 1, British Economic and Social History in Year 2, and Historical Argument and Practice in Year 3. Examples of questions are:

> How useful is the concept of 'embourgeoisement' to our understanding of consumer society in Britain before 1939?

> To what extent was government policy responsible for relative economic decline between 1880 and 1914?

> To what extent was the postwar Welfare State simply 'the application of

Beveridge's tidy mind to the jumbles of an obsolete poor Law'?

16.5.4 Maths

In the Mathematics exams, the subject papers selected for each participant in each year were:

Paper 1: Vectors and Matrices
Paper 2: Differential Equations
Paper 3: Groups and Vector Calculus
Paper 4: Numbers and Sets; Dynamics and Relativity

Each paper comprises 2 sections. Candidates are required to answer all four questions from Section I and 5 questions from section II out of a choice of 8. Questions in section I are marked out of 10 and questions in Section II are marked out of 20. In addition, alpha (α) and beta (β) awards are made according to the complexity of the question. A merit mark is then calculated by adding the raw score to the number of alphas awarded (multiplied by 30) plus the number of betas awarded (multiplied by 5) minus 120. The candidate's percentage mark is obtained through a linear scaling of merit marks.

Examples of questions in Section I are:

State and prove Lagrange's theorem. Let p be an add prime number, and let G be a finite group of order $2p$ which has a normal subgroup of order 2. Show that G is a cyclic group

With the help of definitions or equations of your choice, determine the dimensions, in terms of mass (M), Length (L), time (T) and charge (Q), of the following quantities:

(i) force;
(ii) moment of a force (i.e. torque)
(iii) energy;
(iv) Newton's gravitational constant G;
(v) Electric field E;
(vi) Magnetic field B;
(vii) The vacuum permittivity ε_0

Examples of questions in Section II are:

Define the sign, sgn (σ), of a permutation $\sigma \in$ Sn and prove that is it well defined. Show that the function sgn: Sn $\rightarrow \{1, -1\}$ is a homomorphism. Show that there is an injective homomorphism υ: GL2 ($\mathbb{Z} / 2\mathbb{Z}$) → S4 such that sgn o υ is non trivial. Show that there is an injective homomorphism ∅ : Sn→ GLn ($\mathbb{R}$) such that det (∅ (σ)) = sgn (σ)

For each of the following, either give an example or show that none exists:

(i) A non-abelian group in which every non-trivial element has order 2

(ii) A non-abelian group in which every non-trivial element has order 3
(iii) An element of S*9* of order 18
(iv) An element of S*9* of order 20
(v) A finite group which is not isomorphic to a subgroup of an alternating group

16.5.5 Medicine

The examination papers selected from Medicine consist of multiple choice (MCQ) papers, short answer papers and long answer papers.

a. The multiple choice (MCQ) paper requires candidates to answer 34 questions in 1 hour. All questions are compulsory and the paper carries 20% of the total marks for the FAB exam. Questions examine tissue anatomy, aspects of organogenesis, and the topographical, functional, and clinically applied anatomy of the human body. Examples of multiple-choice questions are:

   Which one statement is correct?
   A. Cytosine in DSA can be converted to uracil by deamination
   B. Uracil cannot base-pair with adenine
   C. Uracil is a purine
   D. Uracil is commonly found in eukaryotic DSA
   E. Uracil is not a component of eukaryotic RNA

b. The short answer paper comprises sub-divided questions that are centered on a particular theme. All candidates are required to answer all questions in 2 hours (no choice of question). The paper carries 20% of the total marks for the FAB exam. Examples of questions in the short answer paper:

   Microbiologists wonder whether the isolated plasmid contains a well-documented antibiotic resistance gene which codes for an acetyl transferase which adds an acetyl group to Antibiotic F rendering it inert. The acetyl transferase is a trimer with a monomer molecular weight of 25kDa. As the sequence of the Antibiotic F resistance gene is knows, the microbiologists next design PCR primers to this gene.

   a. Draw the first three cycles of a PCR reaction
   b. Design a pair of primers to amplify the coding region plus stop codon adding extra sequences to the primers so that a BamH1 restriction site is incorporated at the 5' end of the sense strand and an EcorR1 site is incorporated at the 3' end of the sends strand. The primers should be approx. 25 bp long.
   c. The microbiologists use these primers and carry out 30 cycles of the PCR reagent. How many molecules of PCR produce would be present at the end of 30 cycles?

c. In the longer answer paper, candidates are required to answer 3 questions in 2 hours (from a choice of 6 questions). This paper examines the ability to apply anatomical knowledge to a clinical situation or problem, to deduce basic clinical implications from anatomical principles, to show understanding of experimental physiology and histology and includes questions that require the analysis and interpretation of data from physiological experiments. The paper accounts for 60% of the total marks for the FAB exam. Examples of questions in the long answer paper:

> B6142 is a new antibiotic used to treat bacterial infection. In order to study the pharmacokinetic profiles of the drug, healthy volunteers were each given a single dose of 35mg B6142 by intravenous injection. The plasma concentrations of the drug were then determined at various times after injection. The results are shown in Table 1
>
> Table 1. Plasma concentrations of B6142 determined after intravenous administration

Time	*Plasma [B6142] ng/ml*
1	3360
2	2640
6	1640
12	1275
120	473
240	145
300	81

> Calculate the half-life, the plasma clearance and the volume of distribution of B6142.
>
> 240 minutes after the injection of the drug, the volunteers emptied their bladders and a urine sample was also collected after 360 mins. The average volume of urine in the samples was 250 ml and the concentration of the dug in the urine was 45 ng/ml. Calculate the renal clearance of the drug. Comment on the result.
>
> The therapeutic concentration of the drug is 1500 ng/ml. What infusion rate would be needed to achieve this concentration if the drug was given by constant intravenous infusion? What infusion rate would be needed to reach the therapeutic concentration in 45 mins?

d. The practical exams comprise a 90-minute ‘steeplechase’ paper with 30 stations. 3 questions are set per station, resulting in 90 questions in total. A total of 3 minutes per station is allowed (90 minutes in total). Stations include bones, prosected specimens, clinical photographs, and radiological images. Answers are given in verbal format to an examiner at each station.

15.6 Appendix 6: Demographic information for Participants in Phase 2 (quantitative)

16.6.1 Participant numbers

Category	Total Number of participants	No of participants receiving adjustments in year 1	No of participants receiving adjustments in year 2	No of participants receiving adjustments in year 3
SpLD	80	0	30	50
TD	80	0	0	0

16.6.2 Breakdown of courses studied by participants

Subject	Number of participants
Asian & Middle Eastern Studies	2
Classics	6
Computer science	4
Economics	6
Education	8
Engineering	8
English	20
Geography	4
Human social and political science	12
History	24
Law	18
Maths	8
Medicine	4
Music	6
Natural Sciences	22
Psychological and behavioural science	6
Theology	2

16.6.3 Gender

Category	No of Male participants	No of Female participants
SpLD	38 (47.5%)	42 (52.5%)
TD	38 (47.5%)	42 (52.5%)

16.6.4 Nationality and Ethnicity

Category	International	UK	BAME
SpLD	22 (27.5%)	58 (72.5%)	5 (6.25%)
TD	22 (27.5%)	58 (72.5%)	5 (6.25%)

16.6.5 Age at admission

The SpLD participants tended to be slightly older than TD participants on admission (63.75% of TD participants were aged 18 or below on admission, whereas 52.5% of SpLD participants were aged 18 or below. Yet 37.5% of SpLD students were aged 19 on admission by comparison with 28.75% of TD students). However, similar numbers in both groups were 21 or over on admission (2.5% SpLD students and 2.5% TD students):

Category	Age 17	Age 18	Age 19	Age 20	Age 21	Age 22	Age 23
SpLD	0	42 (52.5%)	30 (37.5%)	6 (7.5%)	0	0	2 (2.5%)
TD	2 (2.5%)	49 (61.25%)	23 (28.75%)	4 (5%)	1 (1.25%)	1 (1.25%)	0

16.6.6 Previous schooling

A higher percentage of students with SpLD attended independent schools than TD students (52.5% by comparison with 38.75%).

Participant Category	State grammar School	Independent School	Comprehensive school	Sixth form college	FE college
SpLD	8 (10 %)	42 (52.5%)	24 (30%)	6 (7.5%)	0 (0%)
TD	13 (16.25%)	31 (38.75%)	28 (35%)	7 (8.75%)	1 (1.25%)

16.6.6 Mean UCAS tariff points on admission

While both groups show high achievement at ‘A’ level, the between group difference in UCAS tariffs points suggests that the students with SpLD may be experiencing disadvantage in exams at school, not just in exams at University. The between group difference in UCAS tariff points is statistically significant (p=0.0455)

Participant Category	Mean UCAS tariff points on admission
SpLD	172
TD	201

16.7 Appendix 7: Semi -structured interview questions

Preamble:

The participants were sent the student information form in accessible format (and in audio version by sensus access) in advance of the interview. At start of interview, check that the student has read the student information form and signed the consent form and that the student clearly understands this information.

Read out the following information about the focus of the research and why it is being conducted with the student.

> *The aim of these interviews is to gather insights from students who have a diagnosis of SpLD and have had 25% extra time in exams (or 25% extra time and use of a word processor) on the impact that these adjustments have on their exam performance. There will also be some questions about your experience of using study skills sessions, if you had this support. This information will be added to a range of other data on the impact of exam arrangements and will be used to inform University policy.*
>
> *A voice recording will be made of this interview. However, your information will be treated in confidence and the research will not name or identify you in any report of the findings. Your participation is completely voluntary and you can choose to stop the interview at any time and ask for your information to be removed from the study.*
>
> *Do you have any questions about the study and your participation before we start?*
>
> *Do still wish to proceed with this interview?*

1. Tell me about the exam arrangements you received and how you used them

2. What do you consider is the purpose of the exam access arrangements that you received and do you feel that this purpose was achieved? If not, what else would have helped?

3. Do you feel that this purpose could be achieved in other ways, rather than through the awarding of extra time / use of a word processor? If so, please specify

4. How do you feel the exam adjustments affected your performance in the exams

5. How would you describe the differences for you between taking the exams with adjustments and when you took exams under standard conditions?

6. Do you feel that you are able to show your full potential in exams with the adjustments? If not, what do you feel are the main barriers in exams that prevent you from showing your potential?

7. Are there any other adjustments in addition to the extra time/ use of a word processor that you feel would have helped your exam performance?

8. Do you feel there are any disadvantages in having exam arrangements?

   Are there any other comments that you would like to make about your experience of using exam access arrangements?

9. Are there any differences between your performance & achievement in exams and your performance & achievement in your dissertation/assessed coursework? If so, why do you think this is?

10. Did you have any study skills sessions? If so, how many and when? If not, was there any reason for this?

11. If you accessed study skills sessions, did you use any to develop strategies for using exam access arrangements in exams? If so, what did you find most helpful and why?

16.8 Appendix 8: Timeline of impact

This research has fed into a number of projects driving policy change in assessment practices at the University of Cambridge. Specifically, it has led me to be a contributing member of the following working groups and projects:

2016 -2020: 5 year review of pan-university exam practices headed by the pro-vice chancellor for education

2017 to date: Working party on digital assessment. A pilot project granting all students the option to use a wordprocessor in exams)

2017 to date: Centre for Teaching and Learning 'Access and Participation plan. A project aimed at identifying and addressing awarding gaps among the university cohort

2017 to date: Centre for Teaching and Learning 'Diversifying assessment project'. A project aimed at developing inclusive, authentic and diversified assessment

2018 – to date: Skills advisory working party. Advising the university on embedding context appropriate study skills into academic courses.

2017 to date: Member of Exam and Assessment Committee. Advising university on issues relating to exams and assessment

2020 to date: Adviser to Foundation Year development Working group (widening participation project). Role: to advise on equitable assessment policy and practices in the proposed Foundation year under development

2021 to date: Member of the Assessment Working Group for the Medicine and Veterinary Medicine Curriculum Review. A project aimed at making the exams of the School of Medicine more inclusive and authentic