# 4. Results and Analysis

## 4.1 Overview

### I - Objectives

The objectives of the analysis were to enable the research hypotheses (sub-section 1.4) to be addressed: firstly, by comparing ABC data from the groups of dyslexic and non-dyslexic students; secondly, by comparing ABC data from the dyslexic students of the Control subgroup with the non-dyslexic students in the Base subgroup; finally, ABC data for the quasi-dyslexic students of the Test subgroup were compared with those in the Control subgroup. Students were filtered into subgroups according to their levels of dyslexia-ness, which was determined by the output of the Dyslexia Index Profiler section of the self-report questionnaire (sub-section 4.3(III)). Through use of unique, continuous range input sliders that displaced more conventional fixed anchor-point Likert-style scales, extent of agreement with the questionnaire's scale-item dimension statements were converted into quantitative data. A text-input section of the questionnaire collected additional qualitative data, optionally provided by participants.

### II - Analysing quantitative data - rationales:

#### 1. Internal consistency (reliability) - Cronbach's 'alpha' (α)

Both the metrics in this study gauged constructs that used linear scales. The existing, ABC Scale operationalized academic confidence, and the Dyslexia Index Profiler was developed especially for this study to assess participants' levels of dyslexia-ness. Each scale comprised multiple dimensions, collectively designed to assess their respective underlying construct. In order to have confidence that the data generated was meaningful, it was important to assess the consistency (reliability) and validity of the scales (where validity is the precision of the scale). Cronbach's ɑ coefficient of internal reliability is typically used in social science research, notably in psychology (Lund & Lund 2018). An ɑ value within the range 0.3 < α < 0.7 is considered as acceptable with preferred values being closest to the upper limit (Kline, 1986). On this basis, precedents have shown acceptable levels of internal reliability for the ABC Scale, determined by Cronbach's ɑ > 0.7, (Putwain & Sanders, 2016; Shaukat & Bashir, 2015; Nicholson et.al., 2013; Aguila Ochoa & Sander, 2012; Sander & Sanders, 2009; Sanders & Sander, 2007). Clearly, as the Dx Profiler has been developed for this current study, no prior measures of the scale's internal reliability or validity are available.

However, several features of the ɑ coefficient assessment imply that outputs derived from using it to gauge a scale's internal reliability should be considered tentatively. In the first instance, excessively high levels of ɑ (i.e. > 0.9) may indicate scale-item redundancy, that is, where some items (dimensions) are measuring very similar traits (Streiner, 2003; Panayides, 2013). There is a lack of agreement, however, about which level of ɑ should be chosen as the critical value for this interpretation, with ɑ > 0.7 frequently considered as the popular 'rule of thumb' (e.g.: Morera & Stokes, 2016). This is despite (computational) evidence that a scale with more items, supposedly gauging the same underlying dimension, will naturally increase the value of ɑ (Cortina, 1993; Nunnally & Bernstein, 1994; Tavakol & Dennick, 2011). Secondly, it is important to note that Cronbach's ɑ tests the consistency of responses within a datapool as opposed to the reliability of the scale per se, and therefore is attributable to a specific use of the scale (Streiner, 2003; Boyle, 2015; Louangrath, 2018). Thirdly, and especially when used in conjunction with dimension reduction techniques, it is reasonable to suppose that the factors which emerge from such a reduction, that is, the sub-scales, should also be evaluated for their reliability, and these outcomes cited together with the ɑ value for the complete scale.

However, frameworks have been suggested for improved reporting and interpretation of internal consistency estimates that may present a more comprehensive picture of the reliability of data collection procedures, particularly data elicited through self-report questionnaires. In particular, and consistent with the approach adopted in this current study for reporting effect size differences (see (2), below), reporting an estimate for a confidence interval for ɑ in addition to the single-point value, noting particularly the upper-tail limit is considered to be one improvement (Onwuegbuzie and Daniel, 2002). The idea of providing a confidence interval for Cronbach's α is attractive because the value of the coefficient is only a point estimate of the likely internal consistency of the scale (and hence the construct of interest). Interval estimates are stronger, not least as the point estimate value, α, is claimed by Cronbach in his original (1951) paper to be most likely a lower-bound estimate of score consistency. This implies that the traditionally calculated and reported single value of α is likely to be an under-estimate of the true internal consistency of the scale, were it possible to apply the process to the background population. Hence the upper-limit confidence interval can be reported in addition to the point-value of Cronbach's α because this is likely to be a more generalizable report about the internal consistency of the scale.

This principle is adopted in this current study, with confidence intervals calculated using Fisher's (1915) transformation which maps the Pearson Product-Moment Correlation Coefficient, r, (upon which Cronbach's α) is derived) on to a value, Z', which was shown to be approximately normally distributed and hence, confidence interval estimates could be constructed. Therefore it follows that Fisher's Z' can be used to transform Cronbach's α and subsequently create confidence interval estimates for α. This process enabled a more complete reporting of the internal consistency of the ABC Scale, the Dx Profiler, and their respective sub-scales (identified through dimension reduction) for the datapool, and for each of the research groups, ND, DI (Table 12-14(check), sub-section 4.3(III, IV) below).

#### 2. Effect sizes

Effect size measures were used as the principal statistical evidence in this study. Effect size challenges the traditional convention that the *p*-value, an arbitrarily determined threshold derived from Null Hypothesis Significance Testing (NHST) (Vila et.al., 2016), is the most important data analysis outcome response to determine whether an observed effect is real or should be attributed to chance events (Maher et al., 2013). Effect size values are a measure of either the magnitude of associations or the magnitude of differences, depending on the nature of the data sets being analysed. Effect size is an absolute value measure (as opposed to the significance) of an observed effect (Cumming 2012), and provides a generally interpretable, quantitative statement about the magnitude of a difference in [or association between] observations (Fritz, et.al., 2012). When clearly defined in a study's methodology, and reported together with their respective confidence intervals, effect sizes provide an improved way to interpret data (Ferguson, 2016). Effect size is easy to calculate, and when used to gauge the between-groups difference between means, effect size is generally reported as Cohen's *d* (Cohen, 1988).If the groups being compared have dissimlar sample sizes (as is the case in this current study), the unbiased estimate of *d* is used, alternatively referred to as Hedges' *g,*(Hedges, 1981), calculated using the weighted, pooled standard deviations of the datasets. Effect size is increasingly prevalent in quantitative analysis (Gliner, et.al., 2001; Sullivan & Feinn, 2012; Carson, 2012; Maher, et.al., 2013), and is particularly useful when observed measurements have no intrinsic meaning, such as with data formulated from Likert-style scales (Sullivan & Feinn, 2012). Guidance for researchers on the use and reporting of effect size is becoming more widely available (Ferguson, 2016; Lorah, 2018; Funder & Ozer, 2019), possibly due to some leading social science journals requiring effect size to be part of the data analysis in studies submitted for publication (Fritz, et.al., 2012; Funder & Ozer, 2019). The use of effect size as a method for reporting statistically important analysis outcomes is especially gaining traction in education, social science and psychology research (Kelley & Preacher, 2012; Rollins, et al., 2019), not least in studies about dyslexia, where it is claimed to be a vital statistic for quantifying intervention outcomes designed to assist struggling readers (ibid).

#### 3. Null-hypothesis significance testing (NHST); ANOVA

Notwithstanding (2) above, the effect size data analyses were supported by measures of the statistical significance of the difference between independent sample means, determined through Student's t-test outcomes, to acknowledge the continued value of null-hypothesis significance tests in social science research. Thus, when taken together with effect sizes and their confidence intervals, comprehensive and pragmatic interpretation of the experimental outcomes could be discussed. One-tail t-tests were conducted in accordance with the alternative hypotheses stated (sub-section 1.4). Homogeneity of variances was established using Levene's Test, and according to the output, the appropriate *p*-value was taken, with the conventional 5% level being adopted as the significance boundary value. It is recognized that the application of ANOVA to this data may have been appropriate had dyslexia-ness been categorized into 'high', 'moderate', 'low', or other sub-gradations, that is, that the independent variable was categorical in nature (Moore & McCabe, 1999; Lund & Lund, 2016). However, the Student's t-test was regarded as a better choice because it is easier to interpret, commonly used, and appropriate when the independent variable (in this case, Dyslexia Index), is continuous in nature (ibid).

#### 4. Dimension reduction

Although these statistical processes outlined so far proved sufficient to address the research hypotheses, dimension reduction by principal components analysis (PCA) was applied later as a secondary process to determine whether meaningful factor structures could be established for both the ABC Scale and the Dyslexia Index metric. The original objective was to explore the influences of groups of similar dimensions of dyslexia-ness (Dx factors) on academic confidence to search for more nuanced explanations for differences in ABC, although in the light of PCA outcomes for the Dx Profiler, this was later modified (see below).

The PCA process is said to be useful to explore whether a multi-item scale that is attempting to evaluate a construct can be reduced into a simpler structure with fewer components (Kline, 1994, Kanyongo, 2005), although there remains considerable debate about how to best identify the most appropriate number of factors to retain from those which emerge from dimension reduction (e.g.: Velicer, et.al, 2000). As a precedent, Sander and Sanders (2003) recognized that dimension reduction may be appropriate for their original, 24-item ABC Scale. Their procedure generated a 6-factor structure, the components of which were designated as Grades, Studying, Verbalizing, Attendance, Understanding, and Requesting. By combining datasets from their earlier studies, a subsequent analysis found that the ABC Scale could be reduced to 17 items with 4 factors, designated as Grades, Verbalizing, Studying and Attendance (Sander & Sanders, 2009). The remaining dimensions of the reduced, 17-item ABC Scale were unamended. Hence, retaining the the full, 24-item scale in this current study enabled dimension reduction to be applied to consider whether a meaningful, local sub-scale structure was likely. It was also possible to calculate alternative 17-item overall mean ABC values simultaneously so that both sets of results were available to consider against the research hypotheses.

Now just as Cronbach's ɑ can offer a measure of internal consistency to a local construct scale (and identify scale item redundancy), factor analysis is ascribable to the dataset onto which it is applied. It was considered therefore that the Sander and Sanders factor structures may not be the most appropriate for the data in this current study, despite being widely used by other researchers in one form (ABC24-6) or the other (ABC17-4) (e.g.: de la Fuente et al., 2013; de la Fuente et al., 2014; Hilale & Alexander, 2009; Ochoa et al., 2012; Willis, 2010; Keinhuis et al., 2011; Lynch & Webber, 2011; Shaukat & Bashir, 2016). Indeed when reviewing the ABC Scale, Stankov et.al., (in Boyle et.al., 2015) implied that more work should be done to consolidate some aspects of the ABC Scale, not so much by levelling criticism at its construction or theoretical underpinnings, but more so to suggest that as a relatively new measure (> 2003) it would benefit from wider applications in the field, and subsequent scrutiny about how it is built and what it is attempting to measure. In the event, only one study was found (Corkery et.al., 2011) which appeared to share this cautious approach for adopting the ABC Scale per se, choosing instead to conduct a local factor analysis to determine the structure of the Scale according to their data, setting a single precedent for taking the same course of action in this current study.

However, it also remained unclear from the Sander and Sanders original, and subsequent studies, whether the components analyses adopted for both the individual and the later, combined datasets, were compared with a factor structure that may have been just as likely to have occurred by chance. Indeed, from the body of literature examined where the ABC Scale has been used either as the principal metric or as an additional aspect of the analysis processes, no studies' data analysis appear to suggest that any comparisons with a factor structure which may have occurred randomly were conducted. Common practice to determine the number of factors to retain in these, and in numerous other studies where component analysis has been applied, use either a visual inspection of the scree plot of eigenvalues against components (Cattell, 1996; Horn & Engstrom, 1979) looking for the point where the slope changes markedly as a means to determine the number of components to declare; or otherwise choose components which present initial eigenvalues > 1 in the table of total variance explained, as those to be included in the final factor structure (Kaiser, 1960). Both processes are not without their difficulties: In the first instance, determining the the number of components to include from visual inspection of the scree plot relies on subjective judgement (e.g.: Zwick & Velicer, 1982), despite common convention; and when relying on eigenvalues > 1 in the table of total variance explained, when no clear distinction exists between two (or more) components that are very close to this critical value, it becomes difficult to decide which components to include and which to omit.

In this current study, early iterations of the process suggested that solutions of four, five, or six factors for both ABC and for Dx could be reasonably supported, determined from both the eigenvalues > 1, and visual interpretations of the scree plots criteria. In the event, five-factor solutions for both variables were initially adopted, based on realistically determining outcomes that could lead to a meaningful interpretation of the data. However, a parallel analysis of multiple randomized versions of the raw data (Eigenvalue Monte Carlo Simulations) was subsequently conducted to examine a factor structure that could have emerged by chance, to consider against the initial iterations of the PCA applied to the ABC and Dx Scales. This was conducted in SPSS according to the guidance provided by O'Connor (2000), and also served to take account of the likelihood of assumption violations unduly influencing solutions for retaining factors (Hutchinson & Bandalos, 1997; Kanyongo, 2005). This later re-analysis of the data suggested that a three-factor solution may be a better model (sub-section 4.5). Whilst outcomes for the ABC Scale were robust, dimension reduction results for the Dx Profiler Scales were inconclusive and thus, speculative. It is possible, if not likely, that this could be because the metric was developed especially for this current study, and hence, only the 166 datasets collected from participants were available. Precedents for dimension reduction processes (i.e. for the ABC Scale), suggest that combining similar-source datasets from several studies is likely to increase confidence in the robustness of sub-scales that emerge, eventually leading to a more standardized scale and sub-scales which can be applied confidently to individual studies. Thus, application of the process to determine a possible sub-scale structure for the Dx Profiler would benefit from additional data from other studies before outcomes can be meaningful. Hence, it was considered that a more nuanced, factorial analysis of the ABC data collected in this study could be confidently conducted. However, to apply unstable Dx Profiler factors to sub-divide outcomes further was considered unwise, and may lead to conclusions of dubious worth, not least due to the small sample sizes of the data subgroups. Development of this aspect of the enquiry will be a topic for a subsequent study.

#### 5. Multiple Regression Analysis

Finally, a tentative multiple regression analysis was conducted to add an additional perspective to the statistical evidence generated thus far to address the research hypotheses. Precedents suggest that multi-variable regression analysis can be valuable in dyslexia research (see sub-section 5.#) to add substance to the rationales which underpin the multi-factorial approaches to understanding dyslexia (sub-section 2.1(II/6)). Hence regression analysis was considered to have value in this current study where the objective was to examine differences between observed and expected ABC outcomes according to Dx inputs, rather than to suggest predictive models for indicating levels of ABC based on Dyslexia Index. The purpose was to use the generated regression equations to determine whether quasi-dyslexic students return higher than expected levels of ABC than their dyslexia-identified peers.

### III - Analysing qualitative data - rationales

Qualitative data was not formally analysed, instead, elements of these data were used to elaborate the discussion element of the thesis where apposite (see Section 5). However, the principles for applying an Interpretative Phenomenological Analysis (IPA) to these data were considered, as IPA is typically used to explore, interpret and understand a phenomenon in people - dyslexia in students in this current study - from the perspectives of the lived-experiences of the individuals of interest (Reid et al., 2005). But an IPA approach was deferred for three reasons: firstly, understanding how students with dyslexia make sense of their learning and study experiences at university and how they attach meaning to the life events that occur in this context (e.g.: Smith et al., 2009), was not the main focus of the research. Instead, the research aim was quite specific, that is, to use the dyslexia-ness continuum approach to examine how dyslexia-ness impacts on academic confidence. Secondly, these (qualitative) data were only acquired from students in the dyslexic group. This was not by design, merely that no participants in the non-dyslexic group provided any data in this form. Hence it was considered that formal, qualitative analysis would have been skewed, and not generalizable across the datapool. Lastly, although IPA attempts to uncover themes in qualitative data, it is conventionally conducted with small, purposive samples of typically fewer than ten participants (Hefferon & Gil-Rodriguez, 2011), with analysis being overly descriptive rather than more deeply interpretative (ibid). In this study, the qualitative data was drawn from a moderately large dataset (n=68) rather than by selecting a small, representative sample. Hence, although some elements of IPA are utilized, for example in identifying thematic narratives, these are used to support the quantitative outcomes of the data analysis, the formal process was not adopted. That said, the data provided an extensive representation of the challenges and difficulties faced by dyslexic students at university, and hence may be used in a more focused study later.

## 4.2 Terminology

Refer to the List of Abbreviations for the meanings of labels, terms, acronyms and designations used in the reporting and discussion of the data, results and analysis.

## 4.3 Results

### I - Demographics

A total of n=183 questionnaire replies were received. Seventeen were discarded due to Dx Profiler data less than 50% complete and so to determine these individuals' Dyslexia Index was considered unrealistic.

The demographic distribution of the datapool according to dyslexia status, gender; home residency, and study level is shown in Table 4. The equivalent distributions for the Test and the Base subgroups, which were both subsets of the non-dyslexic students’ group; and for the Control subgroup, which was a subset of the dyslexic students’ group, are presented in Table 5.

##### Distribution by gender

Overall, female participants (n=113, 67%) outnumbered male participants (n=55, 33%) by a factor of approximately 2 to 1. Amongst the dyslexic participants, females (n=53, 78%) outnumbered males (n=15, 22%) by more than 3 to 1. Of students recruited through the open invitation to all students, and who subsequently formed research group ND (n=98), the distribution by gender showed females (n=60) substantially outnumbered males (n=38) (39%). It is not known whether this is representative of the gender distribution of students more widely at this university as those data were not available

##### Distribution by domicile

Participants were asked to declare whether they were a 'home' or an 'international/overseas' student. Non-UK EU students were not identified as a distinct subgroup. National data for 2016/17 (HESA, 2018) demonstrated a broadly similar distribution although those data were for student enrolment for that academic year rather than a measure of the domicile distribution of all students studying at UK institutions at that time. It is reasonable to assume that the ratio of 'home' students to non-UK students would not be substantially different were an aggregated figure used (which was unavailable).

##### Distribution by study level

Data about level of study was collected to determine whether the datapool represented a reasonable cross-sectional match to student communities attending UK universities more generally. Although a wider selection was available in the questionnaire for participants to choose the level of study which most closely matched their own, these data were grouped as either study at up to, and including level 6 (equivalent to final-year undergraduate), or higher than level 6. Those participants who indicated study for professional or vocational qualifications were grouped with post-graduates, and that to be consistent with national levels, those studying at Foundation/Access level also included those studying at pre-level 4 (pre-1st year undergraduate). National data for 2016/17 (HESA, 2018) showed that 54% of the UK student population were undergraduates, 12% were attending Foundation or Access courses, 31% were studying on post-graduate taught programmes and 3% were post-graduate researchers. Hence, where study at level 6 or lower accounted for 66% of the student population nationally, undergraduate respondents in this study (n=124, 75%) are slightly over-represented, and that the proportion studying at post-graduate level is under-represented (n=42, 25%).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Datapool | Dyslexia status | | Home domicile | | Gender | | Study level‡ | |
|  | Dyslexic | Non-Dyslexic | UK | Non-UK | M | F | ≤ L6 | > L6 |
| 166 | 68 | - | 65 | - | 15 | - | 9 | 5\* |
| - | 50 | 42 | 7\* |
| - | 3 | 0 | - | 0 | 0 |
| - | 3 | 0 | 3 |
|  | | | | | | | |
| - | 98 | 72 | - | 29 | - | 23 | 6 |
|  | 43 | 34 | 8✟ |
| - | 26 | 9 | - | 8 | 1 |
| - | 17 | 8 | 9 |
|  |  |  |  |  |  |  |  |  |
| *subtotals* | 68 | 98 | 137 | 29 | 53 | 113 | 124 | 39 (42, see notes) |
| *totals* | 166 | | 166 | | 166 | | 166 | |

Table 4: Demographic distribution of the datapool by dyslexia status, home domicile, gender and study level

**‡** Study level according to the Regulated Qualifications Framework for England and Wales (Ofqual, 2015) \* +1 respondent study level not disclosed; ✟ +1 studying for Professional or Vocational qualification

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Research Group | Research Subgroup | | Home domicile | | Gender | | Study level‡ | |
| ND | Test  (Dx ≥ 592.5) | Base  (Dx ≤ 400) | UK | Non-UK | M | F | ≤ L6 | > L6 |
| 98 | 18 | - | 16 | - | 9 | - | 6 | 3 |
| - | 7 | 4 | 3 |
| - | 2 | 0 | - | 0 | 0 |
| - | 2 | 1 | 1 |
| *check totals:* | | *18* | | *18* | | *18* | |
| - | 44 | 27 | - | 12 | - | 10 | 2 |
| - | 15 | 10 | 5 |
| - | 17 | 4 | - | 3 | 1 |
| - | 13 | 7 | 6 |
| *check totals* | | *44* | | *44* | | *44* | |
|  | | *36* | | *36* | | *36* | |
| DI | Control  (Dx ≥ 592.5) | - |  |  |  |  |  |  |
| 68 | 47 | - | 44 | - | 10 | - | 7 | 3 |
| - | 34 | 26 | 8 |
| - | 3 | 0 | - | 0 | 0 |
| - | 3 | 0 | 3 |
| *check totals* | | *47* | | *47* | | *47* | |
| *residue (400 < Dx < 592.5)* | | *21* | | *21* | | *21* | |
| *datapool* |  | | *166* | | *166* | | *166* | |

Table 5: Demographic distribution of Test, Base and Control research subgroups by home domicile, gender and study level

**‡** Study level according to the Regulated Qualifications Framework for England and Wales (Ofqual, 2015)

### II - How students with dyslexia learned of their dyslexia

#### the impact of a diagnosis of dyslexia on Academic Behavioural Confidence

This study's hypotheses were grounded on the premise that the dyslexia label may be one of the contributing factors to reduced ABC in students with dyslexia, and which may be especially likely when this label emerged from diagnosing dyslexia as a disability (see Section 2.1(IV)). Thus, one aspect of the enquiry explored how dyslexic students were told about their dyslexia. A sub-hypothesis was constructed to test whether students whose dyslexia was diagnosed to them as a disability had substantially lower levels of academic confidence when compared with students who were told about their dyslexia, otherwise. Hence a null sub-hypothesis was constructed to test against alternatives:

H0: the terminology used to inform dyslexic students of their dyslexia has no impact on their academic confidence;

AH1: students whose dyslexia is diagnosed to them as a disability (or as a difficulty (=AH2); or as a disability or a difficulty (=AH3)) show lower levels of academic confidence in comparison to those who are told about their dyslexia in other ways.

Participants in this current study who declared their dyslexia were invited to report how they were informed about their dyslexia by selecting options to complete a simple statement (Figure 13).

Figure 6: Dyslexic students completed a verb-noun option sentence to indicate how they learned of their dyslexia

|  |  |  |  |
| --- | --- | --- | --- |
| ‘My dyslexia was | choose one …▼ | to me as a learning | choose one …▼ |
|  | disclosed |  | disability |
|  | described |  | difference |
|  | identified |  | weakness |
|  | diagnosed |  | strength |
|  |  |  | deficit |
|  |  |  | difficulty |

It was reasonable to assume that the 68 students who declared their dyslexia had participated in a formal dyslexia screening and/or assessment at university, or during their earlier years in education, and 64/68 (94%) provided data (Table 6). 22/64 (34%) respondents said that their dyslexia was diagnosed to them as a disability; 40/64 (64%) respondents said that their dyslexia was diagnosed to them as a disability or a difficulty. 15/64 (23%) students learned of their dyslexia by one of the other alternatives offered, with 3/15 (< 5% of the total) had their dyslexia described or identified as a difference. Of the 4 students with dyslexia who did not respond, it is not known whether this was due to a reluctance to disclose, or that an option that matched their recollection about how they learned of their dyslexia was not available.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | [disability] | [difference] | [weakness] | [strength] | [deficit] | [difficulty] | ∑ |
| [disclosed] | 1 | 0 | 0 | 0 | 0 | 2 | 3 |
| [described] | 2 | 1 | 0 | 0 | 0 | 5 | 8 |
| identified] | 5 | 2 | 0 | 0 | 1 | 4 | 12 |
| [diagnosed] | 22 | 0 | 0 | 0 | 1 | 18 | 41 |

Table 6: Summary of dyslexia self-report sentence: ‘My dyslexia was [ … ] to me as a learning [ … ]’

The 64 datasets were sorted into subgroups comprising: those whose dyslexia was diagnosed to them as a disability (subgroup DS); those whose dyslexia was diagnosed to them as a difficulty (subgroup DF); leaving the remainder to be aggregated into a third subgroup E.

The full, 24-item ABC Scale was used, and mean average values were calculated both overall, and for each of the three ABC24 Factors (determined through PCA (see below, sub-section 4.5(II)) was calculated for each subgroup and also for subgroups DS and DF combined. Unbiased effect size differences (Hedges '*g*') were calculated, supported by *t*-test outcomes. In accordance with the hypotheses, one-tail tests were applied at the 5% significance level. Levene's Test for homogeneity of variances was applied and where violated, the outcome for unequal populations variances is quoted. (Table 7).

Moderate to large effect size differences in mean ABC24-overall values are indicated between subgroup E, and subgroups DF, DS, and DF+DS combined (*g*=0.704, 0.627, 0.639 respectively), supported by *t*-test outcomes indicating significant differences between mean values in all cases. Hence, students whose dyslexia was diagnosed as a disability or as a difficulty (or either), returned significantly lower overall ABC24 mean values when compared with students who were told of their dyslexia in any of the alternative ways. Thus the null hypothesis is rejected in favour of each of the alternatives, respectively.

At a more granular level, examining the outcomes for differences in ABC24 at a factorial level reveals a slightly more complex picture. Moderate, or moderate to large effect sizes were indicated between mean ABC24 factor values for each of the three subgroup comparisons, and although these were not universally supported by significant differences in means, most t-test outcomes were less than, (i.e. significant), or in the region of the 5% critical value (i.e., marginal). See Section 5 for an interpretation of these results. Examining outcomes generated using the alternative ABC Scales accommodated in this current study (see sub-section 4.3(IV) below), will be a topic for further study later.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| (sample size) | ‘My dyslexia was […] to me as a learning […]’ | |  | | | |
| ABC24 Overall | ABC24 Factor 1:  Study Efficacy | ABC24 Factor 2:  Engagement | ABC24 Factor 3:  Organization & Planning |
| subgroup E(15) | [disclosed / described / identified] | [difference / difficulty] | 66.59 | 68.78 | 58.49 | 74.52 |
| subgroup DF(19) | [diagnosed] | [difficulty] | 55.44 | 57.72 | 48.85 | 61.30 |
| subgroup DS(22) | [diagnosed] | [disability] | 58.26 | 61.06 | 51.37 | 63.91 |
| subgroup DS+DF(41) | [diagnosed] | [disability / difficulty] | 56.95 | 59.52 | 50.20 | 62.70 |
| comparing: | subgroup E against subgroup DF | effect size: Hedges g | 0.7040 | 0.4667 | 0.4906 | 0.6708 |
|  | t-test | t(31) = 2.139  p = 0.0202 | t(32) = 1.703  p = 0.0490 | t(32) = 1.506  p = 0.0710 | t(35) = 2.120  p = 0.0206 |
|  | subgroup E against subgroup DS | g | 0.6268 | 0.3828 | 0.3969 | 0.6180 |
|  | t-test | t(33) = 1.930  p = 0.0311 | t(32) = 1.339  p = 0.0950 | t(30) = 1.249  p = 0.1107 | t(35) = 2.120  p = 0.0206 |
|  | subgroup E against subgroup DF+DS | g | 0.6394 | 0.4924 | 0.4598 | 0.6231 |
|  | t-test | t(33) = 2.419  p = 0.0106 | t(29) = 1.769  p = 0.0437 | t(27) = 1.576  p = 0.0634 | t(44) = 2.646  p = 0.0056 |
| † homogeneity of variances violated (Levene’s Test) | | | | | | |

Table 7: Comparing ABC mean values of dyslexic students according to how they learned of their dyslexia

### III The Dyslexia Index Profiler

#### 

#### I Internal reliability of the Dx Profiler - the Dx20 and Dx16 scales

The Dx Profiler was at first, a 20-item scale, later found to have a possible, 3-factor, sub-scale structure (see sub-section 4.5, below). The levels of internal reliability of the scale and of the sub-scales were assessed using the Cronbachs's ɑ criterion. According to the conventional interpretation of ɑ values (see 4.1(II/1) above), the Dx Profiler overall, together with each of the sub-scales, presented acceptable levels of internal reliability for examining the datasets in this datapool, although there was some concern about the low levels of reliability of the Factor 3 sub-scale in comparison to both the other factors and to the scale overall. Later evidence from dimension reduction analysis confirmed the unstable nature of the 3-factor sub-scale structure for the Dx Profiler, as based on only the data in this current study (see sub-section 4.5(III) below).

Furthermore, the reliability analysis also suggested that some dimensions in the 20-item scale may be redundant by contributing minimally to the overall Dyslexia Index value for each respondent - considered as possible, additional evidence of uncertainty about a sub-scale structure for the metric. Interpretation of the matrix of correlation coefficients (not shown) to identify pairs of dimensions that showed a correlation of r > 0.7, enabled each of the potentially redundant dimensions to be eliminated in turn and in permutations, to permit corresponding re-runs of the reliability analysis. Several iterations of this process subsequently enabled similar, acceptable levels of reliability to be established by reducing the 20-item scale to 16 items. The ɑ coefficients for both scales were calculated for the datapool and also for the two primary research groups. The 95% upper boundary of confidence intervals for ɑ are also provided [~, upper boundary] (Table 8).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scale | Sample | ɑ, complete scale | ɑ, Factor 1 | ɑ, Factor 2 | ɑ, Factor 3 |
| Dx20 | datapool | 0.849 [~,0.887] | 0.875 [~,0.907] | 0.617 [~,0.703] | 0.481 [~,0.590] |
|  | ND | 0.850 [~,0.897] | 0.879 [~,0.917] | 0.498 [~,0.634] | 0.569 [~,0.690] |
|  | DI | 0.723 [~,0.820] | 0.740 [~,0.832] | 0.370 [~,0.559] | 0.666 [~,0.780] |
| Dx16 | datapool | 0.887 [~,0.916] | 0.841 [~,0.881] | 0.802 [~,0.850] | 0.441 [~,0.556] |
|  | ND | 0.869 [~,0.882] | 0.829 [~,0.882] | 0.795 [~,0.858] | 0.401 [~,0.555] |
|  | DI | 0.762 [~,0.847] | 0.637 [~,0.760] | 0.659 [~,0.776] | 0.442 [~,0.616] |

Table 8. Cronbach's ɑ reliability coefficients for the Dx20 and Dx16 scales.

The α value for the 16-item scale exceeded that for the 20-item scale for the datapool and also for the dyslexic and non-dyslexic groups, although the ɑ values for both versions of the scale were within 0.04 of each other for the datapool and for both groups respectively. Hence it was reasonable to assume that either scale, or indeed, both, were likely to be providing reliable indicators of dyslexia-ness amongst the respondents in this datapool. Note that the dataset composition of the three, comparison subgroups (Base, Test, and Control), showed slight variations depending on whether the Dx20 or Dx16 scales were used to calculate respondents' Dx values. This differences impacted slightly on the corresponding ABC outcomes (see sub-section 4.4, below).

The four scale items that were identified as redundant from the 20-item Dx scale were:

* Dx 03: ‘I find it very challenging to manage my time efficiently;
* Dx 05: ‘I think I am a highly organized learner’;
* Dx 07: ‘I generally remember appointments and arrive on time’;
* Dx 13: ‘I find following directions to get to places quite straightforward’.

These dimensions had been identified at an earlier stage of the data collation process as potentially troublesome, demonstrated by a very wide disparity in Dx dimension values across the datapool which appeared to be independent of students' dyslexia status. A cursory scale-reliability analysis of these four dimensions taken together indicated them to be unlikely to comprise a unique factor scale - further supported by the dimension reduction analysis of the Dyslexia Index metric later (see 4.5 below). It was possible to identify dimension Dx13 as the most disruptive of these four dimensions by examining the impact of removing this dimension on both scale reliability, and also on the dimension reduction factor identification process. A more confused, rather than clearer picture emerged, suggesting that despite its wayward nature, a more stable scale could be established when Dx13 was removed together with the other three redundant dimensions.

Examining scale reliability was an important part of the development process for the Dyslexia Index Profiler although the emergence of two scales, Dx20 and Dx16, led to a more complex analysis of ABC outcomes later (see sub-section 4.5 below). In the absence of more data being available to verify which version of the metric is likely to be the more precise gauge of dyslexia-ness, both were retained for the reporting of results. This slight ambivalence indicates that although the Dx Profiler was adequate for the data analysis in this current study, it would benefit from more development work, either based on a more extensive datapool, or through a process of meta-analysis of other studies' results following publication of the metric later. This is to be expected for a metric that had been specifically designed for use in this current study in the absence of standardized, readily available alternatives.

#### II Dx Profiler distributions and basic statistics

Visual inspections of both distributions indicated them to be approximately normal by broadly exhibiting the characteristic bell-shaped outline (Figure 14), although the distribution for the non-dyslexic group presented elements of bimodality. This was an anticipated outcome, confirming the likelihood of the quasi-dyslexic subgroup. Nevertheless, the Shapiro-Wilks test (p>0.05) indicated normality in both distributions according to conventional interpretations, which was further supported by examination of Q-Q plots (Figure 15 (Dx20 plots shown)) where the datapoints for each group are generally positioned approximately along the diagonal. There were no outliers in either distribution, determined by examination of the respective box-plots and application of the +/- three standard deviations criteria (Lund & Lund, 2018).

There were marked differences between Dx values for the two groups where both the sample mean Dx and median Dx are much lower for the non-dyslexic students using either scale (Table 9). For the Dx20 scale, a very large effect size of *g* = 1.34 [95% CI: 1.00, 1.68] (Sullivan & Feinn, 2012) between the Dx sample means, was supported by an NHST outcome indicating a significantly lower mean Dx for students with dyslexia ( t(161) = 8.81, p<0.001), assuming unequal population variances as indicated by violation of Levene’s test for homogeneity of variances (F(164) = 7.65, p=0.006). Outcomes from the reduced item, Dx16 scale were similar (Table 9), although a wider Dx range for the non-dyslexic group using this version of the metric, together with greater differences in the measures of central tendency between the two groups, may be indicating that better discriminative granularity was demonstrated with this version of the scale. Interpretations of outcomes from both scales suggest that the Dx Profiler is returning the expected, high Dx values for the majority of students who declared their dyslexia, and a much lower value for the substantial proportion of those who declared no dyslexic learning challenges, with these marked differences being clearly visible when the distributions were plotted on the Dyslexia-ness Continuum (Figure 16). Thus, it was reasonable to conclude that the Dx Profiler discriminated well between the two groups, correctly detected dyslexic students in the dyslexic group, and hence exhibited good sensitivity.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Scale | Research Group | n | Dx range | sample mean Dx | 95% CI for μ | median Dx |
| Dx20 | ND | 98 | 32 < Dx < 888 | 430 | 393 < Dx < 466 | 398 |
|  | DI | 68 | 326 < Dx < 960 | 670 | 635 < Dx < 706 | 700 |
|  |  |  | Hedges *g*: | 1.44 |  |  |
|  |  |  | t-test outcome: | t(160) = 9.29;  p < 0.001 |  |  |
| Dx16 | ND | 98 | 81 < Dx < 831 | 446 | 414 < Dx < 479 | 425 |
|  | DI | 68 | 349 < Dx < 933 | 646 | 615 < Dx < 677 | 661 |
|  |  |  | Hedges *g*: | 1.34 |  |  |
|  |  |  | t-test outcome: | t(161) = 8.81;  p < 0.001 |  |  |

Table 9. Dyslexia Index summary according to research group and Dx scale.

A screenshot of a cell phone

Description automatically generated

Figure 7: Distributions of Dyslexia Index for the Non-Dyslexic, (RG:ND), and the Dyslexic (RG:DI), groups.

A screenshot of a social media post

Description automatically generated

Figure 8. Research groups located on the Dyslexia-ness Continuum using the Dx20, and Dx16 scales

A screenshot of a social media post

Description automatically generated

Figure 9: Normal Q-Q plots for Dyslexia Index.

#### III - Setting boundary values for Dx

##### 1. Dx boundary value for the Test, and Control subgroups

Some studies suggest that the proportion of known dyslexics studying at university is likely to be much lower than the true number of students with dyslexia or dyslexia-like study characteristics (e.g.: Richardson & Wydell, 2003; MacCullagh et al., 2016; Henderson, 2017). This current study was grounded on this (amongst other) research outcomes, and the core of the research design was to devise a robust mechanism to detect such quasi-dyslexic students so that their academic confidence could be compared to the other groups and subgroups which emerged from the datapool. Hence, to establish this Test subgroup of quasi-dyslexic students, it was necessary to define a boundary Dx value, or at least a boundary region, in the group of non-dyslexic students above which datasets would be filtered into the Test subgroup, with the same protocol being applied to datasets in the dyslexic group to establish the Control subgroup. At the design stage, setting a value of Dx = 600 as the filter was considered intuitively reasonable because this corresponded to an average 60% agreement with the 20 dyslexia-ness dimensions of the original Dx Profiler. The scale was set so that higher percentage dimension-statement agreement was the marker for higher levels of dyslexia-ness.

For datasets derived from the 20-item scale, applying the Dx > 600 boundary value to the non-dyslexic group initially generated a Test subgroup of n=20 quasi-dyslexic students - that is, individuals with no previously reported dyslexia but who appeared to be presenting similar levels of dyslexia-ness to students in the dyslexic group. Applying the same Dx filter value to datasets in the dyslexic group established the Control subgroup of students presenting similarly high levels of dyslexia-ness, which numbered 47 out of the 68 students with declared dyslexia.

However, in order for the academic confidence of the Test and Control subgroups to be justifiably compared later (through ABC Scale outcomes), it was important to establish that the defining, Dx parameters for each of these two subgroups were similar, that is, statistically *not* significantly different from each other. At the Dx = 600 filter boundary level, the mean Dx20 for the Test and Control subgroups were Dx = 676, 716, respectively. These were shown to be significantly different ( t(43) = 2.374; p = 0.011) and thus, a more appropriate boundary was required. By selecting different values close to Dx = 600, (with the added consequence of some datasets then included or omitted into the respective subgroups accordingly), it became clear that to set a fixed boundary Dx value was not a realistic objective. This was due to the subgroup means being unduly affected by extreme Dx values, predominantly from amongst the datasets in the dyslexic group where the highest Dx20 value recorded was Dx=933, compared to Dx=831 in the non-dyslexic group (Table 9). Although these values were not identified as notable outliers from inspection of the distributions' box-plots, it was necessary to consider them as such so that the mean Dx values of datasets in the Test and Control groups respectively would not not significantly different.

Consequently, some datasets from the upper end of the dyslexic group's range were omitted from the Control subgroup, leading to the lower boundary Dx values for the Test and Control subgroups to emerge at slightly different points on the dyslexia-ness continuum - although both remained close to the intuitively determined value of Dx=600. For the 20-item scale this process subsequently determined the Dx20 mean values for the Test and the Control groups at Dx=683, Dx=705 respectively, outcomes which emerged as not significantly different ( t(31) = 1.352; p = 0.093). The complete process was repeated for the 16-item scale (Table 10).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Scale | Subgroup | sample size, n | mean Dx | Dx range |
| Dx20 | Test | 18 | 683 | 623 < Dx < 832 |
|  | Control | 40 | 705 | 614 < Dx < 812 |
|  |  | t-test outcome: | t(160) = 1.32; p < 0.093  no significant difference |  |
| Dx16 | Test | 19 | 696 | 611 < Dx < 889 |
|  | Control | 43 | 724 | 604 < Dx < 830 |
|  |  | t-test outcome: | t(160) = 1.48; p < 0.073  no significant difference |  |

##### 2. Dx boundary value for the Base subgroup

A lower boundary value was required to filter the additional comparator subgroup of students from the non-dyslexic group who presented low levels of dyslexia-ness - the Base subgroup. It was considered also intuitively reasonable to set this value at Dx = 400, thus representing a mean average agreement of 40% with the dyslexia-ness dimensions in the Profiler. Using the Dx20 scale, this generated a Base subgroup of n=44, representing 45% of the non-dyslexic students, or 55% of the remaining non-dyslexic students after the Test subgroup had been filtered out. Using the Dx16 scale, the Base subgroup comprised n=50 students (51%, 63% respectively).

By contrast, only five (Dx20) or three(Dx16) students with declared dyslexia presented values of Dx < 400. Given that these datasets were not identified as outliers to be excluded from further analysis, these remained anomalous results for other reasons, although no additional information about these students was available to enable any conclusions to be stated.

It is of note that sizeable minorities of non-dyslexic students presented Dx levels between the upper boundary value of the Base subgroup (Dx20, Dx16 = 400) and the lower boundary value of the Test subgroup (Dx20 = 623, Dx16 = 611). Using the 20-item scale, 36 students fell into this category whereas the 16-item scale identified slightly fewer (n=29). In either case these data suggest approximately one-third of the students in this datapool presented levels of dyslexia-ness that placed them in the central area of the Dyslexia-ness Continuum. This is discussed below (Section 5).

### IV The ABC Scale

#### 

#### I - Internal reliability of the ABC scales

There are currently two versions of the ABC Scale widely available to researchers: a 24-item scale which emerged out of the earlier, Academic Confidence Scale (Sander & Sanders, 2003) together with a later, 17-item scale developed through a meta-analysis of several studies, item redundancy analysis conducted through scale reliability interpretations, and dimension reduction processes (Sander & Sanders, 2009).

In a relatively early study using the 24-item scale, Sander and colleagues reported it to possess an internal reliability of ɑ = 0.88 (2007), based on data acquired from a sample of 284 participants drawn from two UK universities. All other studies using the ABC Scale found to date, appear to have either relied on this ɑ-value, or only report the internal reliability of the ABC Scale's sub-scales, as derived by prior dimension reduction (op cit). With one exception, no other studies were found that indicated item redundancy analysis nor dimension reduction of the ABC 24-item scale as a mechanism for a more nuanced analysis of local data. The exception was a short conference paper detailing a statistical evaluation of the factor structure of the preceding, Academic Confidence Scale, that used data collected from a local university (Corkery, et.al., 2011), and although no overall measure for scale reliability was indicated, coefficients for the three subscales were presented, with values ranging from 0.711 < ɑ < 0.880.

In this current study, data were collected using the original, 24-item scale because this permitted 17-item scale outputs to be generated simultaneously, as this version of the scale had merely discarded the redundant, seven items, leaving the remainder unamended. Hence, either scales could be used to address the research hypotheses. Reliability analysis was conducted on both versions, and this process also permitted scale item redundancy to be considered for the 24-item scale based on local data. Items were identified as redundant using the same protocols as for the Dx Profiler (see 4.3(III/I above), that is, by inspection of the matrices of item correlation coefficients (not shown) and adopting the *r* > 0.7 criterion. Results suggested two possible alternatives to the existing ABC Scales, one comprising 17 items (co-incidentally), the other a 21-item scale. The local 17-item scale emerged as similar but not identical to the Sander and Sanders version (for the differences, see sub-section 4.5, below). Scale and sub-scale reliability coefficients all exceeded ɑ > 0.7, widely considered as an appropriate critical value for indicating a reasonable balance between strong levels of internal reliability, and possible scale item redundancy (Table 11).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ABC Scale | ɑ, Scale | ɑ, Factor 1 | ɑ, Factor 2 | ɑ, Factor 3 |
| ABC24 | 0.914 | 0.877 | 0.839 | 0.820 |
| ABC17 | 0.874 | 0.889 | 0.770 | 0.727 |
| ABC21-L (local) | 0.899 | 0.855 | 0.807 | 0.794 |
| ABC17-L (local) | 0.877 | 0.844 | 0.786 | 0.707 |

Table 11: Reliability coefficients for the ABC Scales and sub-scales

Hence a variety of alternatives were available, both at scale and sub-scale level, for relating the data collected in this study to the focus of the enquiry, the research questions, and hypotheses being explored. It was considered that the results that emerged from all versions of the ABC Scale with the three, comparison subgroups (Test, Control and Base) defined according to both Dx Profiler Scales, was a strength of the study because interpretation of the differences in outcomes that emerged contributed positively to the discussion element of this thesis (Section 5).

#### II - Differences in mean ABC values

The principal focus of this study was to explore differences in levels of academic confidence between students with dyslexia and their non-dyslexic peers, and secondly to try to determine whether the quasi-dyslexic students sifted into the Test subgroup through the Dyslexia Index profiling process, demonstrated higher levels of academic confidence than their dyslexia-identified peers. Were this to emerge as substantial or significant, it may imply that at least part of the reason for lower levels of academic confidence amongst students with dyslexia may be accounted for by the identification of the dyslexia itself.

In the event, the primary data analysis outcomes showed notable differences in mean ABC values (Table 12) and effect sizes (Tables 13, 14) between the non-dyslexic and dyslexic groups overall, and also between the Test and Control subgroups, and between the Base and Control subgroups, when these subgroups were determined by either the Dx20 or the Dx16 Profiler Scales.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| group / subgroup | Dx Scale | n | mean ABC24 | mean ABC17 | mean ABC21-L | mean ABC17-L |
| ND (all non-dyslexic students) | - | 98 | 67.21 | 67.51 | 68.08 | 68.30 |
| DI (all dyslexic students) | - | 68 | 58.45 | 58.40 | 59.48 | 60.48 |
| ND / Base | Dx20 | 44 | 72.44 | 72.61 | 73.44 | 73.28 |
| " | Dx16 | 50 | 72.54 | 72.79 | 73.45 | 73.52 |
| ND / Test | Dx20 | 18 | 61.72 | 62.08 | 62.60 | 63.21 |
| " | Dx16 | 19 | 63.69 | 63.68 | 64.46 | 65.95 |
| DI / Control | Dx20 | 40 | 58.05 | 57.92 | 59.20 | 60.83 |
| " | Dx16 | 43 | 57.94 | 57.83 | 59.17 | 60.53 |

Table 12: Summary of ABC mean values by research group and subgroup according to ABC, and Dx scales.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ABC Scale | Dx20 Scale: group / subgroup | ABC24 Group DI | ABC 24 Control | ABC17 Group DI | ABC17 Control | ABC21-L Group DI | ABC21-L Control | ABC17-L Group DI | ABC17-L Control |
| ABC24 | Group ND | 0.604 |  |  |  |  |  |  |  |
|  | Base |  | 1.027 |  |  |  |  |  |  |
|  | Test |  | 0.242 |  |  |  |  |  |  |
| ABC17 | Group ND |  |  | 0.621 |  |  |  |  |  |
|  | Base |  |  |  | 1.029 |  |  |  |  |
|  | Test |  |  |  | 0.268 |  |  |  |  |
| ABC21-L | Group ND |  |  |  |  | 0.598 |  |  |  |
|  | Base |  |  |  |  |  | 1.034 |  |  |
|  | Test |  |  |  |  |  | 0.226 |  |  |
| ABC17-L | Group ND |  |  |  |  |  |  | 0.532 |  |
|  | Base |  |  |  |  |  |  |  | 0.876 |
|  | Test |  |  |  |  |  |  |  | 0.184 |

Table 13: ABC Scales' effect size differences (Hedges' *g*) between the non-dyslexic (ND) group, Test, and Base subgroups, and the dyslexic group (DI), and Control subgroup, when the subgroups were defined according to the Dx20 Profiler.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ABC Scale | Dx16 Scale: group / subgroup | ABC24 Group DI | ABC 24 Control | ABC17 Group DI | ABC17 Control | ABC21-L Group DI | ABC21-L Control | ABC17-L Group DI | ABC17-L Control |
| ABC24 | Group ND | 0.604 |  |  |  |  |  |  |  |
|  | Base |  | 1.086 |  |  |  |  |  |  |
|  | Test |  | 0.406 |  |  |  |  |  |  |
| ABC17 | Group ND |  |  | 0.621 |  |  |  |  |  |
|  | Base |  |  |  | 1.079 |  |  |  |  |
|  | Test |  |  |  | 0.400 |  |  |  |  |
| ABC21-L | Group ND |  |  |  |  | 0.598 |  |  |  |
|  | Base |  |  |  |  |  | 1.086 |  |  |
|  | Test |  |  |  |  |  | 0.378 |  |  |
| ABC17-L | Group ND |  |  |  |  |  |  | 0.532 |  |
|  | Base |  |  |  |  |  |  |  | 0.953 |
|  | Test |  |  |  |  |  |  |  | 0.392 |

Table 14: ABC Scales' effect size differences (Hedges' *g*) between the non-dyslexic (ND) group, Test, and Base subgroups, and the dyslexic group (DI), and Control subgroup, when the subgroups were defined according to the Dx16 Profiler.

## 4.4 Relating results to hypotheses

With four possible ABC Scales available to gauge the academic confidence of non-dyslexic and dyslexic students overall, and between the Test, Control and Base subgroups of quasi-dyslexic, strongly dyslexic, and strongly non-dyslexic students respectively when these subgroups were identified from either of two Dyslexia Index scales, it was possible to relate the outcomes to the research questions and hypotheses (sub-section 1.4). The picture that emerged was not straightforward:

#### I - Differences in ABC between the non-dyslexic and the dyslexic groups and subgroups:

The greatest, absolute difference in mean ABC values between the non-dyslexic and the dyslexic groups was 9.11 percentage points generated from the ABC17 Scale (Table 12), which suggests that when ABC is gauged according the criteria on that scale, non-dyslexic students in this datapool are expressing, on average, 16% higher levels of academic confidence relative to their non-dyslexic peers. The corresponding, least absolute difference of 7.82 percentage points (ABC17-L Scale, Table 12) still suggests a 13% relative difference. Taking into account distribution variances and sample sizes, effect size differences emerged of *g*= 0.621 and 0.532 respectively (Table 13) with corresponding 95% confidence interval upper boundaries of *g*= [~,0.945], [~,0.938], respectively. These were considered as 'moderate' using the conventional criteria for describing effect sizes (Cohen, 1985), although such a label must be considered as tentative in the absence of other, similar measures being available more widely in this research domain that would provide comparisons (Schafer & Schwarz, 2019).

Hence, it was concluded that:

1. In comparison with their non-dyslexic peers, (RG:ND), students with a declared dyslexic learning difference (RG:DI) presented a significantly lower mean ABC (67.21-68.30, 58.40-60.48 respectively), indicated by a moderate effect size (0.532 < *g*< 0.621; [~, 0.938-0.945]), supported by a significant difference in sample means (ABC17: t(134) = 3.86, *p*< 0.001; ABC17-L: t(137)=3.33, *p*< 0.001). Thus, Null Hypothesis (1), that there is no difference in mean ABC between the two groups, is rejected in favour of Alternative Hypothesis (1), that non-dyslexic students present a higher, mean ABC than their non-dyslexic peers.

#### II - Differences in ABC between the non-dyslexic (Base), and the dyslexic (Control) subgroups:

When students presenting particularly high levels of dyslexia-ness in the Control subgroup were compared to non-dyslexic students with low levels of dyslexia-ness in the Base subgroup, established using either of the Dx Profiler Scales, the differences are more marked. These criteria established the greatest absolute difference in mean ABC at 14.96 percentage points (ABC17, Dx16 Scale, Table 12). The least difference between students in these subgroups of 12.45 percentage points (ABC17-L, Dx16) still represented a substantial difference. In the most extreme case, those values represented a 26% relative difference between the academic confidence of strongly dyslexic students and their strongly non-dyslexic peers when levels of dyslexia-ness are taken as the gauge. When sample size and distribution variances were taken into account, the greatest effect size of *g*= 1.0864, [~,1.526] emerged when the 24-item ABC Scale was used to gauge datasets sifted into the Control and Base subgroups using the Dx16 Scale. Within the caveats stated above, this effect size was considered as large to very large.

Hence, it was further concluded that:

2. In comparison with their strongly non-dyslexic peers in the Base subgroup, students in the Control subgroup of identified, dyslexic students present a significantly lower mean ABC (72.44-73.44, 57.83-60.83, respectively), indicated by a large effect size (0.876 < *g* < 1.086 [~, 1.329-1.385]), supported by corresponding NHST outcomes (ABC17-L, Dx20: t(77)=3.98, *p* < 0.001; ABC24, Dx16: t(83)=5.16, *p*< 0.001).

#### III - Differences in ABC between the quasi-dyslexic (Test), and the dyslexic (Control) subgroups:

With attention focused on differences between students in the quasi-dyslexic, Test, subgroup and the dyslexic students in the Control subgroup, the outcomes were less marked but still of interest. Overall, when any of the ABC Scales were applied to datasets in these subgroups, whether sifted according to the Dx20 Profiler Scale or the Dx16 alternative, mean ABC values were higher for the quasi-dyslexic students when compared with their identified-dyslexic peers.

Whichever ABC Scale was used, differences between ABC means were greater than 5 percentage points when the Test and Control subgroups were generated from the Dx16 Profiler, with the greatest, absolute difference in mean ABC of 5.85 percentage points when the ABC17 Scale was applied (Table 12). Taking into account distribution variances and sample sizes, effect sizes were in the range 0.378 < *g*< 0.406, with confidence interval upper boundaries falling in the range [~,0.924] to [~,0.954]. These represent moderate-to-low effect sizes although the true effect sizes may be substantially larger, as indicated by the upper boundaries of the confidence intervals. Given the small sample size of the Test subgroup, (n=19) in comparison to the Control subgroup (n=43), this degree of imprecision is not unexpected. Effects were smaller when datasets were generated and sifted according to the Dx20 Profiler, with absolute differences ranging between 2.38 and 4.16 percentage points, corresponding to an effect size range of 0.184 < *g*< 0.268, ([~,0.744] to [~,0.828]), with uncertainty likely to be related to sample sizes (Test: n=18, Control: n=40). However, in all cases, the mean ABC for the quasi-dyslexic subgroup exceeded levels for the dyslexic subgroup.

Thus, it was not possible for conclusions to be quite so robust in this case:

3. In comparison with students in the Control subgroup of identified, dyslexic students, quasi-dyslexic students in the Test subgroup presented a higher mean ABC (57.83-60.83, 61.72-65.95 respectively), indicated by a low-to-moderate effect size ( 0.184 < *g <* 0.406; [~, -0.744-0.954]). None of the NHST outcomes indicated differences to be significant, although the outcome generated from the ABC17 Scale and the Dx16 Profiler was marginal (t(38)=1.504, *p*=0.0703). Hence evidence to reject the Null Hypothesis (2) was also marginal when based on students in this datapool. However, the differences that did emerge presented a clear pattern, with quasi-dyslexic students consistently presenting higher levels of academic confidence on average, than their dyslexia-identified peers.

See Section 5 for a discussion about the implications of these results.

## 4.5 Further analysis: Dimension reduction

### Applying dimension reduction to the ABC Scales and the Dx Profilers

The ABC Scales and the Dx Profilers are multi-dimensional, continuous variable, linear scales. Of the many dimension reduction techniques available to explore possible sub-scale structures, Principal Component Analysis (PCA) was chosen as the most appropriate firstly, because all precedents for dimension reduction applied to the ABC Scale had used this process, and hence guidance was available; secondly, a factor structure that emerged from PCA on the data in this current study could then be considered alongside existing factor structures for the ABC Scale determined from similar processes for comment.

To maintain consistency of dimension reduction approach, and also to minimize computational complexity, PCA was also the preferred choice for the Dx Profiler Scales. However, because this metric was developed uniquely for this current study, only the local data collected from n=166 participants was available. Hence, it was expected that any sub-scale structure that might emerge would be speculative in the absence of data from other sources. Consequently, the factor analysis of the Dx Profiler was considered as unlikely to be sufficiently reliable to contribute to a deeper interpretation of the data in this study. However, the process was completed to assess whether any early indications of a possible sub-scale structure emerged. Reserving data until they can be supplemented from subsequent studies was considered the most prudent course of action.

Furthermore, it was anticipated that once the dimensions reduction processes were completed, determining the number of factors to retain was likely to be far from straightforward, not least in the light of controversy in research communities about the best criteria to adopt. Hence, parallel analyses using randomized raw score data simulations were conducted to aid this process (Eigenvalue Monte Carlo Simulations).

#### Assumptions and preliminary work

Although complete-scale outcomes have enabled the research hypotheses to be addressed and conclusions drawn, precedents set for the ABC Scale indicated that applying dimension reduction to explore any factor structure which may emerge could reveal more nuanced outcomes, subsequently permitting a deeper interpretation of the data collected. In this current study, 4 possible ABC Scales emerged as contenders for the most appropriate for analysing data, together with two versions of the Dx Profiler. Assumptions and preliminary work was carried out for all of these, but as exemplars, details are reported for the ABC24 Scale and for the Dx16 Profiler, although identical processes were conducted for all scales which produced similar results.

For a PCA to be valid, it is considered that a scale-item variable that presents a correlation of *r* ≥ 0.3 with at least one other scale-item variable is worthy of inclusion in the analysis (Hinton et al., 2004). An analysis of the inter-variable correlation matrix for both metrics showed that for the ABC24 Scale, 138 out of the 300 possible correlations returned a coefficient of *r* ≥ 0.3 with all variables returning at least one correlation of r ≥ 0.3. For the Dx16 Profiler, of the 120 possible correlation outcomes, 80 returned a Pearson correlation coefficient of *r* ≥ 0.3, also with all variables returning at least one correlation of *r* ≥ 0.3 with any other variable.

Furthermore, sufficient sampling adequacy is fundamental to PCA, but this adequacy is a function of the total number of observations rather than to the sample sizes(s) per se. Statistical conventions indicate that at least 150 observations would be a sufficient condition (Guadagnoli & Velicer, 1988) although a later study suggests that aspects of the variables and the study design have an impact on determining an appropriate level of sampling adequacy, recommending that this is improved with a higher number of observations (McCallum et al., 1999). In this current study, 4,032 observations for the ABC24 Scale, and 2,656 for the Dx16 Profiler were recorded. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy produced a value of 0.866 for the ABC24 Scale, and KMO = 0.889 for the Dx16 Profiler. Measures of sampling adequacy for individual variables were examined to ensure that these also confirm the appropriateness for factor analysis. For ABC24, the individual variables measures returned values of 0.753 ≤ KMO ≤ 0.929, and for Dx16, the corresponding outcome was 0.563 ≤ KMO ≤ 0.924. According to Kaiser's (1974) own classification, KMO values can range from 0 to 1, with a value of KMO ≥ 0.5 considered to be desirable (Hinton el.al, 2004). Finally, the null hypothesis that there are no correlations between any of the variables was tested using Barlett’s Test of Sphericity where a rejection is sought as determined by a *p*-value of p < 0.05. When applied to both the ABC Scale and to Dyslexia Index, the test returned values of p < 0.001.

Thus, for both metrics, the null hypothesis that there are no correlations between the metrics' variables is rejected, hence suggesting that there *are* correlations between the variables; and that sufficient levels of sample adequacy were also achieved. Therefore, justification for running the PCA on both metrics is met.

### Eigenvalue Monte Carlo Simulations

No studies found to date which have used the ABC Scale as their principal metric have indicated that sub-scale structures that emerged through PCA, based on the data within them, were tested against structures that might have occurred by chance. Indeed, of the numerous studies that that were examined, with one exception, all have relied on the factor structures developed by the originators of the Scale, where 6 factors were determined for the 24-item scale, reduced to 4 factors in the ABC17 Scale. Only Corkery et.al. (2011) reported a 3-factor solution for their local data. Early analysis of the data in this current study for the ABC24 Scale indicated that possible 6-factor, 5-factor or 4-factor solutions may provide an appropriate PCA outcome to permit a more in depth examination of the data (sub-section 4.1/II(4)). This was based on both the eigenvalues > 1 principal, and the visual inspection of the scree plots.

With all of these solutions showing equal merit, a parallel (simulation) analysis (Horn, 1965) was conducted to determine how any (or all) of these solutions compared with a sub-scale structure likely to have occurred by chance. Also known as the Eigenvalue Monte Carlo Simulation (O'Connor, 2000), this statistical process determines the number of factors which are likely to have occurred through use of random data (or, more usefully, a randomized version of the experimentally acquired data in a study). By running multiple, simulated reductions, not only are mean value eigenvalues produced in the table of total variance explained, but also critical values are generated where the boundary is conventionally set at the 95th percentile, in much the same fashion as for NHST. Comparison is then made between the eigenvalues generated in the PCA of the raw data with the 95th percentile boundary values for eigenvalues generated from the parallel analysis. Data-generated values that are greater than their parallel analysis, 95th percentile critical value, are statistically significant, and therefore are unlikely to have occurred by chance. Hence only those factors are retained in the final solution.

#### I - Eigenvalue Monte Carlo Simulation for ABC24 and ABC17 Scales:

For both the ABC24 and ABC17 Scales, parallel analysis simulation of 1000 random permutations of the raw data collected in this study identified three components to retain based on their eigenvalues exceeding the equivalent random data 95th percentile critical value (Tables 15, 16). This was also illustrated by the points of intersections of the comparative scree plots for both scales (Figure 17; scree plot for ABC17 Scale not shown), which occurred at eigenvalues between those for the third and fourth components. Hence this suggested that a three-factor solution was the most appropriate extraction to use for the ABC data in this current study, thus displacing the earlier solutions derived from either the eigenvalues > 1, or the visual inspection of the scree plot criteria (or both).

|  |  |  |  |
| --- | --- | --- | --- |
| ABC 24 Component | Eigenvalue (raw data) | Eigenvalue (mean value, simulations) | 95th %ile Eigenvalue (simulations) |
| 1 | 8.3899 | 1.7682 | 1.8949 |
| 2 | 2.3844 | 1.6377 | 1.7309 |
| 3 | 1.9199 | 1.5400 | 1.6158 |
| 4 | 1.2106 | 1.4553 | 1.5209 |
| 5 | 1.1295 | 1.3797 | 1.4367 |
| 6 | 0.9625 | 1.3132 | 1.3678 |
| 7 | 0.8943 | 1.2503 | 1.3049 |
| 8 | 0.8353 | 1.1909 | 1.2389 |
| 9-24 | < 0.761 | < 1.136 | < 1.181 |

Table 15: Parallel Analysis: Principal components and raw data permutations for the ABC24 Scale

|  |  |  |  |
| --- | --- | --- | --- |
| ABC 24 Component | Eigenvalue (raw data) | Eigenvalue (mean value, simulations) | 95th %ile Eigenvalue (simulations) |
| 1 | 6.009 | 1.6011 | 1.7274 |
| 2 | 1.9909 | 1.4724 | 1.5661 |
| 3 | 1.7972 | 1.3769 | 1.4481 |
| 4 | 1.0532 | 1.2937 | 1.3589 |
| 5 | 0.9552 | 1.2212 | 1.2783 |
| 6 | 0.8058 | 1.1517 | 1.2509 |
| 7 | 0.6917 | 1.0881 | 1.1418 |
| 8 | 0.6329 | 1.0286 | 1.0786 |
| 9-24 | < 0.5592 | < 0.9703 | < 1.10187 |

Table 16: Parallel Analysis: Principal components and raw data permutations for the ABC17 Scale

A screenshot of a cell phone

Description automatically generated

Figure 10: Scree plot of raw data and Eigenvalue Monte Carlo Simulations for the ABC24 Scale

#### II - Eigenvalue Monte Carlo Simulation for the Dx Profiler Dx20 and Dx16 Scales:

For the Dx20 version of the Profiler a similar, parallel analysis simulation of again, 1000 random permutations of the raw data clearly identified two significant eigenvalues with a third falling on the 95th percentile critical boundary (within three significant figures) (Table 17, Fig.18)). The same simulation applied to the scale variables of the 16-item Dx Profiler conversely indicated that no sub-structure scale could be reasonably determined (Table 18) and that the Dx16 Profiler was best considered as a single-factor gauge of dyslexia-ness for the datasets in this current study.

|  |  |  |  |
| --- | --- | --- | --- |
| Dx20 Component | Eigenvalue (raw data) | Eigenvalue (mean value, simulations) | 95th %ile Eigenvalue (simulations) |
| 1 | 6.3364 | 1.6774 | 1.8068 |
| 2 | 1.9701 | 1.5470 | 1.6289 |
| 3 | 1.5215 | 1.4505 | 1.5238 |
| 4 | 1.2009 | 1.3650 | 1.4304 |
| 5 | 1.0600 | 1.2909 | 1.3485 |
| 6 | 0.9878 | 1.2237 | 1.2779 |
| 7 | 0.8007 | 1.1614 | 1.2102 |
| 8-20 | < 0.7598 | < 1.1024 | < 1.1481 |

Table 17: Parallel Analysis: Principal components and raw data permutations for the Dx20 Profiler.

|  |  |  |  |
| --- | --- | --- | --- |
| Dx16 Component | Eigenvalue (raw data) | Eigenvalue (mean value, simulations) | 95th %ile Eigenvalue (simulations) |
| 1 | 6.2341 | 1.5782 | 1.7075 |
| 2 | 1.2992 | 1.4408 | 1.5304 |
| 3 | 1.1998 | 1.3503 | 1.4223 |
| 4 | 0.9675 | 1.2667 | 1.3283 |
| 5 | 0.8751 | 1.1937 | 1.2469 |
| 6 | 0.8202 | 1.1259 | 1.1793 |
| 7 | 0.7607 | 1.0632 | 1.1125 |
| 8-16 | < 0.6525 | < 1.0017 | < 1.0485 |

Table 18: Parallel Analysis: Principal components and raw data permutations for the Dx16 Profiler.

A screenshot of a cell phone

Description automatically generated

Figure 11: Scree plot of raw data and Eigenvalue Monte Carlo Simulations for the Dx20 Profiler Scale.

### II PCA on Academic Behavioural Confidence

As a result of the Eigenvalue Monte Carlo Simulations suggesting a three-factor structure to be the most likely to provide meaningful outcomes, dimension reduction through PCA was applied to the four, ABC Scales that have emerged as useful in this current study. Data varimax rotation was applied, being an orthogonal rotation method which assumes that the factors in the analysis are uncorrelated, where rotation of the factors is a mathematical process, usually employed to determine the simplest factor structure that is most likely (Kieffer, 1998). Other rotations were considered, but in the interests of expediency, only the two most popular were explored further: that is, to determine whether these data were best analysed using an orthogonal (eg: varimax) rather than an oblique (eg: direct oblimin) rotation. For these data, the factor correlation matrix (not shown) derived through an oblimin rotation showed only one correlation to be (marginally) > 0.32, considered as the critical factor for determining whether an oblique rather than an orthogonal rotation is the most appropriate (Tabachnik & Fiddel, 2007), suggesting that although either rotation would generate meaningful outcomes, an orthogonal process is said to produce less sampling error (op cit) and hence was chosen.

Hence with an extraction that fixed the number of components (factors) at three, the rotated component matrix for the reduction of the ABC24 Scale shows reasonably distinct factors (Table 19), although some dimensions loaded onto more than one factor (factor loadings < 0.3 were supressed in the output). The reduced, ABC17 Scale is also shown, both to indicate which scale items were removed as redundant through the original re-analysis of the scale (Sander & Sanders, 2009), and how factor loadings were distributed across the three sub-scale solution derived in this current study. The two alternative, locally derived scales, ABC21-L and ABC17-L were also complex, again with some dimensions loading onto more than one factor (Table 20). In these cases, dimensions were attributed to the highest-loading factor (Table 23, below). Where factor loadings for a dimension were only marginally different, a reasonable judgement was made about which factor to select (Table 20).

#### Proportion of variance explained

##### I - ABC24 Scale

The PCA process determines the percentage contributions of the total variance made by each of the variables if all of the components are retained. For the ABC24 Scale, the three factors (components) which were retained from this analysis cumulatively accounted for 52.9% of the total variance, with the most significant influence from Factor 1, which explained almost 35.0% of the total variance (Table 21). Despite the

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ABC Scale Item | ABC24 Factor 1 | ABC24 Factor 2 | ABC24 Factor 3 | ABC17 Factor 1 | ABC17 Factor 2 | ABC24 Factor 3 |
| 15: produce coursework at the required standard | 0.828 |  |  | 0.835 |  |  |
| 07: attain good grades | 0.792 |  |  | 0.813 |  |  |
| 16: write in an appropriate styles | 0.771 |  |  | 0.771 |  |  |
| 20: pass assessment at the first attempt | 0.678 |  |  | 0.680 |  |  |
| 23: produce best work in coursework assignments | 0.645 |  | 0.349 | 0.698 |  |  |
| 01: study effectively in independent study | 0.560 |  | 0.406 | 0.617 |  | 0.309 |
| 04: manage workload to meet deadlines | 0.524 |  | 0.506 | 0.591 |  | 0.433 |
| 22: remain adequately motivated throughout university | 0.500 |  | 0.440 | 0.559 |  | 0.389 |
| 10: ask lecturers questions during a lecture |  | 0.798 |  |  | 0.793 |  |
| 03: respond to lecturers questions in a full lecture theatre |  | 0.791 |  |  | 0.845 |  |
| 12: follow themes and debates in lectures |  | 0.716 |  | x | x | x |
| 05: present to a small group of peers |  | 0.646 |  |  | 0.700 |  |
| 08: debate academically with peers |  | 0.608 |  |  | 0.559 |  |
| 11: understand material discussed with lecturers |  | 0.592 |  | x | x | x |
| 09: ask lecturers questions in one-to-one settings |  | 0.516 |  | x | x | x |
| 17: ask for help if you need it |  | 0.480 |  | x | x | x |
| 02: produce best work in exams | 0.327 | 0.470 |  | 0.305 | 0.564 |  |
| 24: attend tutorials |  |  | 0.721 |  |  | 0.765 |
| 06: attend most taught sessions |  |  | 0.721 |  |  | 0.773 |
| 18: be on time for lectures |  |  | 0.677 |  |  | 0.707 |
| 14: read recommended background material | 0.345 |  | 0.579 | x | x | x |
| 21: plan appropriate revision schedules | 0.529 |  | 0.535 | 0.579 |  | 0.451 |
| 19: make the most of university study opportunities | 0.342 | 0.303 | 0.529 | x | x | x |
| 13: prepare thoroughly for tutorials | 0.481 |  | 0.528 | x | x | x |

Table 19: Rotated component matrix for ABC24 and ABC17 Scales (Sander & Sanders) showing factor loadings, and which items were removed as redundant from the ABC24 Scale (x).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ABC Scale Item | ABC21-L Factor 1 | ABC21-L Factor 2 | ABC21-L Factor 3 | ABC17-L Factor 1 | ABC17-L Factor 2 | ABC17-L Factor 3 |
| 15: produce coursework at the required standard | x | x | x | x | x | x |
| 07: attain good grades | 0.812 |  |  | 0.789 |  |  |
| 16: write in an appropriate styles | 0.733 |  |  | 0.718 |  |  |
| 20: pass assessment at the first attempt | 0.709 |  |  | x | x | x |
| 23: produce best work in coursework assignments | 0.682 |  | 0.311 | 0.676 |  |  |
| 01: study effectively in independent study | 0.630 |  | 0.311 | 0.709 |  |  |
| 04: manage workload to meet deadlines | 0.564 |  | 0.436 | x | x | x |
| 22: remain adequately motivated throughout university | 0.581 |  | 0.340 | x | x | x |
| 10: ask lecturers questions during a lecture | x | x | x | x | x | x |
| 03: respond to lecturers questions in a full lecture theatre |  | 0.763 |  | x | x | x |
| 12: follow themes and debates in lectures |  | 0.742 |  |  | 0.718 |  |
| 05: present to a small group of peers |  | 0.638 |  |  | 0.653 |  |
| 08: debate academically with peers |  | 0.638 |  |  | 0.726 |  |
| 11: understand material discussed with lecturers |  | 0.604 |  | x | x | x |
| 09: ask lecturers questions in one-to-one settings |  | 0.539 |  |  | 0.611 |  |
| 17: ask for help if you need it | 0.370 | 0.421 |  | 0.386 | 0.467 |  |
| 02: produce best work in exams | 0.337 | 0.484 |  | 0.432 | 0.357 |  |
| 24: attend tutorials |  |  | 0.712 |  |  | 0.742 |
| 06: attend most taught sessions |  |  | 0.770 |  |  | 0.783 |
| 18: be on time for lectures |  |  | 0.700 |  |  | 0.733 |
| 14: read recommended background material | 0.374 |  | 0.545 | x | x | x |
| 21: plan appropriate revision schedules | x | x | x | 0.680 |  |  |
| 19: make the most of university study opportunities | 0.447 |  | 0.498 | 0.491 | 0.318 | 0.408 |
| 13: prepare thoroughly for tutorials | 0.493 |  | 0.498 | 0.617 |  |  |

Table 20: Rotated component matrix for locally derived ABC21-L and ABC17-L Scales, showing factor loadings, and which items were removed as redundant from the ABC24 Scale (x).

extraction being directed by the three-factor solution indicated by the parallel analysis simulation (above), it is notable that eigenvalues for the fourth, and fifth components are significantly above the eigenvalue > 1 criterion, often applied for determining the number of factors to extract from a PCA. This may suggest that were a larger datapool have been available for the randomized raw data parallel analysis, a four-factor, or even five-factor solution, may have been the outcome, leading to a forced, four or five factor extraction in the PCA for this data. Visual inspection of the scree plot (Fig. 21) shows a marked change in gradient at the fourth component, also suggesting that were this criterion applied, a four-factor solution would have been the likely conclusion.

|  |  |  |  |
| --- | --- | --- | --- |
| ABC24 Component | Eigenvalue | % of variance | Cumulative % of variance |
| 1 | 8.390 | 34.958 | 34.598 |
| 2 | 2.384 | 9.935 | 44.893 |
| 3 | 1.920 | 8.000 | 52.893 |
| 4 | 1.211 | 5.044 | 57.937 |
| 5 | 1.130 | 4.706 | 62.643 |
| 6 | 0.963 | 4.010 | 66.654 |
| 7 | 0.894 | 3.726 | 70.380 |
| 8 | 0.835 | 3.480 | 73.840 |
| 9 | 0.760 | 3.168 | 77.028 |
| 10 | 0.663 | 2.764 | 79.792 |
| 11-24 | < 0.630 | < 2.700 | > 82.400 |

Table 21: Total variance explained for the PCA on the ABC24 Scale.

A screenshot of a cell phone

Description automatically generatedFigure 12: Scree plot of eigenvalues for components (factors) of PCA on the ABC24 Scale

##### II - ABC17, ABC21-L, and ABC17-L Scales

As expected, the distributions of proportions of variance for the three alternative versions of the ABC Scale used in this study are similar (Table 22). Scree plots of eigenvalues also presented similar characteristics to the ABC24 Scale scree plot, and hence are not shown.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ABC  Component | ABC17  Eigenvalue | ~ % variance | ~ cumulative % | ABC21-L  Eigenvalue | ~ % variance | ~ cumulative % | ABC17-L  Eigenvalue | ~ % variance | ~ cumulative % |
| 1 | 6.009 | 35.348 | 35.348 | 7.245 | 34.500 | 34.500 | 5.821 | 34.239 | 34.239 |
| 2 | 1.991 | 11.709 | 47.057 | 2.104 | 10.020 | 44.520 | 2.107 | 12.392 | 46.631 |
| 3 | 1.797 | 10.572 | 57.629 | 1.555 | 7.403 | 51.923 | 1.524 | 8.962 | 55.593 |
| 4 | 1.053 | 6.195 | 63.824 | 1.203 | 5.728 | 57.651 | 0.986 | 5.800 | 61.390 |
| 5 | 0.955 | 5.619 | 69.443 | 0.966 | 4.600 | 62.250 | 0.913 | 5.371 | 66.764 |
| 6 | 0.806 | 4.740 | 74.183 | 0.915 | 4.359 | 66.610 | 0.838 | 4.927 | 71.691 |
| 7 - ... | < 0.693 | < 4.070 | > 78.2 | < 0.845 | < 4.015 | > 70.6 | < 0.770 | < 5.530 | > 762 |

Table 22: Total variance explained for the PCA on the ABC17, ABC21-L, and ABC17-L Scales.

#### ABC Factors

Subsequent to the three-factor solutions of dimension reduction processes on the ABC Scales, factors were designated thematically according to their dimensional composition. Although the groupings of dimensions into factors varied slightly across the four versions

of the ABC Scales, common themes emerged which enabled consistent factor names to be assigned as: Factor 1, Study Efficacy; Factor 2, Engagement; Factor 3, Organization and Planning (Table 23). Thus, with dimensions assigned to factors, the dimension reduction process for the ABC Scales was completed, permitting comparisons in ABC Factor levels to be made across the research groups and subgroups (sub-section 4.6).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ABC Scale Item | ABC21-L Factor 1 | ABC21-L Factor 2 | ABC21-L Factor 3 | ABC17-L Factor 1 |
| 01: study effectively in independent study | F1 study efficacy | F1 study efficacy | F1 study efficacy | F1 study efficacy |
| 02: produce best work in exams | F2 engagement | F2 engagement | F2 engagement | F2 engagement |
| 03: respond to lecturers questions in a full lecture theatre | F2 engagement | F2 engagement | F2 engagement | x |
| 04: manage workload to meet deadlines | F1 study efficacy | F1 study efficacy | F1 study efficacy | x |
| 05: present to a small group of peers | F2 engagement | F2 engagement | F2 engagement | F2 engagement |
| 06: attend most taught sessions | F3 org & planning | F3 org & planning | F3 org & planning | F3 org & planning |
| 07: attain good grades | F1 study efficacy | F1 study efficacy | F1 study efficacy | F1 study efficacy |
| 08: debate academically with peers | F2 engagement | F2 engagement | F2 engagement | F2 engagement |
| 09: ask lecturers questions in one-to-one settings | F2 engagement | x | F2 engagement | F2 engagement |
| 10: ask lecturers questions during a lecture | F2 engagement | F2 engagement | x | F2 engagement |
| 11: understand material discussed with lecturers | F2 engagement | x | F2 engagement | x |
| 12: follow themes and debates in lectures | F2 engagement | x | F2 engagement | F2 engagement |
| 13: prepare thoroughly for tutorials | F3 org & planning | x | F3 org & planning | F1 study efficacy |
| 14: read recommended background material | F3 org & planning | x | F3 org & planning | x |
| 15: produce coursework at the required standard | F1 study efficacy | F1 study efficacy | x | x |
| 16: write in an appropriate styles | F1 study efficacy | F1 study efficacy | F1 study efficacy | F1 study efficacy |
| 17: ask for help if you need it | F2 engagement | x | F2 engagement | F2 engagement |
| 18: be on time for lectures | F3 org & planning | F3 org & planning | F3 org & planning | F3 org & planning |
| 19: make the most of university study opportunities | F3 org & planning | x | F3 org & planning | F1 study efficacy |
| 20: pass assessment at the first attempt | F1 study efficacy | F1 study efficacy | F1 study efficacy | x |
| 21: plan appropriate revision schedules | F3 org & planning | F1 study efficacy | x | F1 study efficacy |
| 22: remain adequately motivated throughout university | F1 study efficacy | F1 study efficacy | F1 study efficacy | x |
| 23: produce best work in coursework assignments | F1 study efficacy | F1 study efficacy | F1 study efficacy | F1 study efficacy |
| 24: attend tutorials | F3 org & planning | F3 org & planning | F3 org & planning | F3 org & planning |

Table 22: Total variance explained for the PCA on the ABC17, ABC21-L, and ABC17-L Scales.

### III PCA on Dyslexia Index

Parallel analysis simulations provided some helpful insight into the possible sub-scale structure for the ABC Scale. However when the approach was applied to the Dx Profiler Scales, Dx20 and Dx16, outcomes were mixed, leading to poor levels of confidence that on the basis of data collected in this current study, any meaningful sub-scale structure could be determined. Certainly for the Dx16 Scale, it seemed likely that this was best considered as a single-factor scale. The outcomes for the Dx20 Scale indicated that there may be two distinct sub-scales although the rotated component matrix showed that one of these comprised the four dimensions which had been previously identified as troublesome (Dx303, 305, 307, 313) and likely to be redundant according to reliability analysis (sub-section 4.3(III/1) above), thus establishing the Dx16 version of the Profiler. As the Monte Carlo simulation had indicated a possible, borderline third eigenvalue, a three-factor solution for the Dx20 Scale was considered, but the factor loadings were also of dubious merit for identifying a meaningful structure (Table 24).

#### Proportion of variance explained

##### Dx20 Scale

The table of proportions of variance explained by the eigenvalues generated through the PCA on the Dx20 Scale indicated that in the absence of testing the dimension reduction outcome by applying a parallel analysis, a four, or possibly five-factor solution may have been adopted as a sub-scale structure (Table 25) according to the eigenvalues > 1 criterion for retaining factors. Conversely, the scree plot (Fig.22) shows no substantial change in gradient beyond the second component, conventionally taken as an indicator for the number of factors to retain when this additional criterion is used. Hence, evidence to support further examination of a possible sub-scale structure for the Dx Profiler (in either version) was sparse or at best indeterminate, suggesting that using the Dx20 and the Dx16 Profiler Scales as single-factor scales would be a more prudent approach for further analysis of the data in this current study.

However, it is possible to speculate that were more data available to contribute to the dimension reduction process, a more robust, three-factor solution may have been the outcome, at least for the Dx20 Scale. Hints of this were present in the distribution of factor loadings, notable when dimensions are grouped factorially (Table 26), with suggestions for possible theme-based, factor names. Were it possible to adopt it, such a structure would demonstrate a neat alignment with components drawn from the BDA definition of dyslexia, identified above (sub-section 3.3(III/2, Section 2 (Part 2/II)).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dx20 Scale Item | Factor 1  (.../2) | Factor 2  (.../2) | Factor 1  (.../3) | Factor 2  (.../3) | Factor 3  (.../3) |
| 06: In my writing, I frequently use the wrong word for my intended meaning | 0.804 |  | 0.745 | 0.303 |  |
| 09: I have difficulty putting my writing ideas into a sensible order | 0.742 |  | 0.611 | 0.325 | 0.402 |
| 08: When I'm reading, I sometimes read the same line again, or miss out a line altogether | 0.735 |  | 0.709 |  |  |
| 01: When I was learning to read at school, I often felt I was slower than others in my class | 0.696 |  | 0.678 |  |  |
| 19: I get in a muddle when I'm searching for learning resources or information | 0.689 |  | 0.620 |  | 0.371 |
| 10: In my writing at school, I often mixed up similar letters, like 'b' and 'd', or 'p' and 'q' | 0.684 |  | 0.700 |  |  |
| 18: My tutors often tell me that my essays or assignments are confusing to read | 0.675 |  | 0.682 |  |  |
| 20: I get really anxious when I'm asked to read 'out loud' | 0.651 |  | 0.725 |  |  |
| 02: My spelling is generally good\* | 0.619 |  | 0.702 |  |  |
| 16: I find it really challenging to follow a list of instructions | 0.586 | 0.382 | 0.522 |  | 0.488 |
| 11: When I'm planning my work, I use diagrams or mindmaps rather than bullet points or lists | 0.540 |  | 0.399 | 0.454 |  |
| 17: I get my lefts and rights easily mixed up | 0.538 |  | 0.605 |  |  |
| 04: I can explain things to people much more easily verbally, than in my writing | 0.529 |  | 0.345 | 0.550 |  |
| 12: I'm hopeless at remembering things like telephone numbers | 0.496 |  | 0.511 |  |  |
| 15: My friends say I often think in unusual or creative ways to solve problems | 0.447 |  | 0.337 | 0.319 |  |
| 05: I think I'm a highly organized learner\* |  | 0.762 |  |  | 0.768 |
| 07: I generally remember appointments and arrive on time\* |  | 0.740 |  | -0.405 | 0.633 |
| 03: I find it very challenging to manage my time efficiently |  | 0.652 |  |  | 0.790 |
| 13: I find following directions to get to places quite straightforward |  | -0.366 |  | 0.457 |  |
| 14: I prefer looking at the big picture rather than focusing on details | 0.279 | -0.300 |  | 0.705 |  |

Table 24: Factor leadings for 2-factor, and 3-factor solutions of PCA on the Dx20 Profiler (\* dimension with reverse-coded data).

|  |  |  |  |
| --- | --- | --- | --- |
| Dx20 Component | Eigenvalue | % of variance | Cumulative % of variance |
| 1 | 6.336 | 31.682 | 31.682 |
| 2 | 1.970 | 9.851 | 41.532 |
| 3 | 1.522 | 7.607 | 49.109 |
| 4 | 1.201 | 6.005 | 55.144 |
| 5 | 1.060 | 5.300 | 60.444 |
| 6 | 0.988 | 4.939 | 65.383 |
| 7 | 0.801 | 4.003 | 69.387 |
| 8 | 0.760 | 3.799 | 73.185 |
| 9 | 0.703 | 3.517 | 76.702 |
| 10 | 0.643 | 3.217 | 72.919 |
| 11-20 | < 0.621 | < 3.100 | > 83.0 |

Table 25: Total variance explained for components generated by PCA on the Dx20 Profiler Scale.

A screenshot of a cell phone

Description automatically generatedFigure 13: Scree plot of eigenvalues for components (factors) of PCA on the Dx20 Profiler Scale.

|  |  |  |  |
| --- | --- | --- | --- |
| Dx20 Scale Item | Factor 1  Language & Literacy | Factor 2  Thinking & Processing | Factor 3  Time & Organization |
| 06: In my writing, I frequently use the wrong word for my intended meaning | 0.745 | (0.303) |  |
| 20: I get really anxious when I'm asked to read 'out loud' | 0.725 |  |  |
| 08: When I'm reading, I sometimes read the same line again, or miss out a line altogether | 0.709 |  |  |
| 02: My spelling is generally good\* | 0.702 |  |  |
| 10: In my writing at school, I often mixed up similar letters, like 'b' and 'd', or 'p' and 'q' | 0.700 |  |  |
| 18: My tutors often tell me that my essays or assignments are confusing to read | 0.682 |  |  |
| 01: When I was learning to read at school, I often felt I was slower than others in my class | 0.678 |  |  |
| 19: I get in a muddle when I'm searching for learning resources or information | 0.678 |  |  |
| 09: I have difficulty putting my writing ideas into a sensible order | 0.611 | (0.402) | (0.325) |
| 17: I get my lefts and rights easily mixed up | 0.605 |  |  |
| 16: I find it really challenging to follow a list of instructions | 0.522 |  | (0.488) |
| 12: I'm hopeless at remembering things like telephone numbers | 0.511 |  |  |
| 14: I prefer looking at the big picture rather than focusing on details |  | 0.705 |  |
| 04: I can explain things to people much more easily verbally, than in my writing | (0.345) | 0.550 |  |
| 13: I find following directions to get to places quite straightforward |  | 0.457 |  |
| 11: When I'm planning my work, I use diagrams or mindmaps rather than bullet points or lists | (0.399) | 0.454 |  |
| 15: My friends say I often think in unusual or creative ways to solve problems | (0.337) | 0.319 |  |
| 03: I find it very challenging to manage my time efficiently |  |  | 0.790 |
| 05: I think I'm a highly organized learner\* |  |  | 0.768 |
| 07: I generally remember appointments and arrive on time\* |  | (-0.405) | 0.633 |

Table 26: Possible 3 factor sub-scale structure for the Dx20 Profiler (\* dimension with reverse-coded data).

### IV Comparing ABC Factor Means

The determination of a possible, 3-factor sub-scale structure for the ABC Scales through dimension reduction enabled mean ABC levels to be calculated for each of the three factors. Comparisons could then be made between the non-dyslexic and dyslexic groups, between the Test and Base, and between the Test and Control subgroups. The dataset composition of the sub-groups varied according to whether the Dx20 or Dx16 Profiler Scales were used to determine participants' levels of dyslexia-ness, reflected in variances in raw score differences and effect sizes (below). Outcomes derived from each of the four versions of the ABC Scales that were used, according to each of the Dx Profilers, revealed notable differences between within-factor means (Table 27).

The results of interest are the difference in factor mean ABC values between the Base and Control subgroups, and between the Test and Control subgroups (Table 27, and Figure 23). With participants' levels of dyslexia-ness gauged by the the Dx20 Profiler, and for all ABC Scales' Factor 1: Study Efficacy, and Factor 2: Engagement, differences were substantial between the Base subgroup with low levels of dyslexia-ness (Dx < 400) and the Control subgroup of dyslexic students defined by levels of dyslexia-ness greater than the Dx=614 critical value for the Dx20 Scale. Only for Factor 3: Organization and Planning, were differences less pronounced, with outcomes from the ABC17 Scale, and the local, ABC17-L Scale almost negligible. This pattern was repeated for mean ABC differences in Factors 1 and 2 between the quasi-dyslexic, Test, subgroup and the Control subgroup, with Factor 3 outcomes returning negative differences - that is, mean ABC values were stronger for the Control subgroup than for the Test subgroup. When the sub-groups were defined by outputs from the Dx16 Profiler, a similar pattern emerged (Fig. 24). (Differences were calculated by subtracting the Control subgroup mean value from the Test, or the Base subgroup mean value).

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Dx20 | ABC factor means according to ABC Scale - Dx20 Profiler Outcomes | | | | | | | | | | | |
| group /subgroup | ABC Factor 1: Study Efficacy | | | | ABC Factor 2: Engagement | | | | ABC Factor 3: Organization & Planning | | | |
| ABC Scale-> | ABC24 | ABC17 | ABC21-L | ABC17-L | ABC24 | ABC17 | ABC21-L | ABC17-L | ABC24 | ABC17 | ABC21-L | ABC17-L |
| ND | 70.28 | 68.98 | 69.87 | 65.48 | 62.61 | 56.88 | 63.98 | 66.21 | 69.62 | 80.85 | 71.46 | 80.85 |
| BASE | 77.06 | 75.43 | 76.85 | 71.68 | 67.90 | 61.12 | 69.46 | 71.27 | 73.01 | 83.31 | 74.77 | 83.31 |
| TEST | 65.51 | 63.57 | 64.62 | 59.52 | 56.69 | 51.87 | 57.92 | 62.43 | 63.87 | 74.59 | 66.50 | 74.59 |
| DI | 60.50 | 59.20 | 59.93 | 54.96 | 51.44 | 43.77 | 52.82 | 56.71 | 65.11 | 80.39 | 67.83 | 80.39 |
| Control | 58.93 | 57.74 | 58.74 | 54.19 | 51.24 | 43.49 | 52.45 | 57.13 | 65.11 | 82.50 | 68.74 | 82.50 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
| Dx16 | ABC factor means according to ABC Scale - Dx16 Profiler Outcomes | | | | | | | | | | | |
| group /subgroup | ABC Factor 1: Study Efficacy | | | | ABC Factor 2: Engagement | | | | ABC Factor 3: Organization & Planning | | | |
| ABC Scale-> | ABC24 | ABC17 | ABC21-L | ABC17-L | ABC24 | ABC17 | ABC21-L | ABC17-L | ABC24 | ABC17 | ABC21-L | ABC17-L |
| ND | 70.28 | 68.98 | 69.87 | 65.48 | 62.61 | 56.88 | 63.98 | 66.21 | 69.62 | 80.85 | 71.46 | 80.85 |
| BASE | 76.64 | 75.12 | 76.41 | 71.45 | 68.56 | 62.33 | 69.98 | 71.91 | 72.97 | 83.39 | 76.64 | 83.39 |
| TEST | 64.68 | 63.47 | 63.92 | 61.07 | 58.29 | 54.18 | 59.18 | 63.35 | 69.51 | 80.18 | 72.13 | 80.18 |
| DI | 60.50 | 59.20 | 59.93 | 54.96 | 51.44 | 43.77 | 52.82 | 56.71 | 65.11 | 80.39 | 67.83 | 80.39 |
| Control | 58.66 | 57.46 | 58.21 | 53.86 | 51.36 | 44.11 | 52.97 | 56.62 | 65.59 | 81.82 | 68.55 | 81.82 |

Table 27: Comparison of ABC Factor Means for all ABC Scales; subgroups established from Dx20, and Dx16 Profiler Scales.

A picture containing pencil

Description automatically generated

Figure 14: Raw score differences in ABC factor means as defined by the Dx20 Profiler.

A screenshot of a cell phone

Description automatically generated

Figure 15: Raw score differences in ABC factor means as defined by the Dx16 Profiler.

#### ABC Factor Means effect size differences

Converting the ABC factor means raw score differences into effect sizes to take account of sample sizes and standard deviations presented a clearer comparison between the Base, Test, and Control subgroups, defined according to Dx20 Profiler outputs (Table 28), and from the Dx16 Profiler (Table 29). (Note firstly, that effect sizes between the non-dyslexic (ND) and dyslexic (DI) groups are identical as these are not dependent on which Dx Profiler was used; and secondly, that Factor 3, Organization and Planning, comprised the same ABC dimensions in both of the 17-item scales, ABC17, and ABC17-L, so for this factor, these effect sizes were the same). The greatest effect size (*g*= 0.6689) between the non-dyslexic and dyslexic groups emerged for Factor 2, Engagement, when their academic confidence was gauged from the ABC21-L, locally derived scale, indicating a moderate-to-large effect size for the group of ABC dimensions that comprised this factor.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Effect sizes between groups, and between subgroups (as defined by the Dx 20 Profiler) | | | | |
| ABC Factor | ABC Scale | ND <-> DI  effect size | Test <-> Base effect size | Test <-> Control effect size |
| F1: Study Efficacy | ABC24 | 0.5470 | 1.0703 | 0.3473 |
| ABC17 | 0.5443 | 1.0325 | 0.3020 |
| ABC21-L | 0.5437 | 1.0586 | 0.3091 |
| ABC17-L | 0.6401 | 1.0741 | 0.2942 |
| F2: Engagement | ABC24 | 0.6452 | 0.9785 | 0.2908 |
| ABC17 | 0.6427 | 0.8472 | 0.3848 |
| ABC21-L | 0.6689 | 1.0446 | 0.3013 |
| ABC17-L | 0.5379 | 0.7560 | 0.2653 |
| F3: Organization and Planning | ABC24 | 0.2576 | 0.4205 | -0.1092 |
| ABC17 | 0.0234 | 0.0435 | -0.4121 |
| ABC21-L | 0.2081 | 0.3564 | -0.1287 |
| ABC17-L | 0.0234 | 0.0435 | -0.4121 |

Table 28: Effect size difference in ABC factor means between non-dyslexic (ND) and dyslexic (DI) groups, and between subgroups as defined from the Dx20 Profiler.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Effect sizes between groups, and between subgroups (as defined by the Dx 16 Profiler) | | | | |
| ABC Factor | ABC Scale | ND <-> DI  effect size | Test <-> Base effect size | Test <-> Control effect size |
| F1: Study Efficacy | ABC24 | 0.5470 | 1/0997 | 0.3273 |
| ABC17 | 0.5443 | 1.0641 | 0.3231 |
| ABC21-L | 0.5437 | 1.1021 | 0.3099 |
| ABC17-L | 0.6041 | 1.1099 | 0.4184 |
| F2: Engagement | ABC24 | 0.6452 | 1.0409 | 0.3847 |
| ABC17 | 0.6427 | 0.8862 | 0.4724 |
| ABC21-L | 0.6689 | 1.0843 | 0.3561 |
| ABC17-L | 0.5379 | 0.8601 | 0.3586 |
| F3: Organization and Planning | ABC24 | 0.2576 | 0.4430 | 0.2375 |
| ABC17 | 0.0234 | 0.0852 | -0.0917 |
| ABC21-L | 0.2081 | 0.3674 | 0.2247 |
| ABC17-L | 0.0234 | 0.0852 | -0.0917 |

Table 29: Effect size difference in ABC factor means between non-dyslexic (ND) and dyslexic (DI) groups, and between subgroups as defined from the Dx16 Profiler.

For both Factor 1: Study Efficacy and Factor 2: Engagement, effect sizes between non-dyslexic and dyslexic students were moderate whichever ABC Scale was used, with values ranging from 0.54 < *g* < 0.67, suggesting that non-dyslexic students presented substantially higher academic confidence than their dyslexic peers in both their capacity or power to produce strong academic outputs, and also the degree to which they participated in active dialogues with their lecturers and collaborated academically with their peers. However it was notable that less pronounced or negligible differences were observed in areas of organization and planning (Factor 3). As would be expected, all of the differences in Study Efficacy, and Engagement, were accentuated when comparisons were made between the Base subgroup of students who presented low levels of dyslexia-ness, and their strongly dyslexic peers in the Control subgroup, with an effect size range of 0.76 < *g* < 1.07, and 0.86 < *g <* 1.10, for subgroups defined from the Dx20, Dx16 Profilers respectively. It is of note that the reduced item, Dx16 Profiler generated slightly higher outcomes. Effect sizes between the Base and Control subgroups for Factor 3: Organization and Planning were moderate when gauged with the ABC24 or the ABC21-L Scales, again with very slightly higher values recorded when the Dx16 Profiler was used.

Of greatest interest, however, were differences between the quasi-dyslexic, Test, subgroup and their dyslexic peers in the Control subgroup. Whilst differences in Study Efficacy and in Engagement were modest, a similar trend in differences was observed with the quasi-dyslexic students presenting higher levels of academic confidence for these two factors when compared with their dyslexic peers. Effect sizes ranged from a low-to-moderate, *g* = 0.27, when the ABC17-L Scale and the Dx20 Profiler were used to gauge Engagement, to a moderate, *g* = 0.47 for the same ABC Factor, gauged with the ABC17, and Dx16 Scales. Although the sample size of the Test subgroup was small (n=18, Dx20; n=19, Dx16), and hence, inferences from these outcomes must be treated tentatively, this result did appear to add to the evidence presented above (sub-section 4.4), that quasi-dyslexic students exhibit higher levels of academic confidence in many of the dimensions gauged by the ABC Scale(s) than their identified, dyslexic peers. It was notable that the greatest contribution to differences in mean levels of ABC overall, was from dimensions related to study efficacy and academic engagement. Confidence related to aspects of organization and planning in academic studies indicated few, or negligible differences between groups and subgroups of students in this datapool.

## 4.7 Applying multiple regression analysis

Whilst the rationale for conducting regression analysis was not to attempt to create a prediction model between academic confidence and dyslexia-ness per se, it was considered appropriate to use the process to generate expected outcomes for ABC based on Dx inputs, thus enabling comparison with the observed values acquired experimentally from participants in this current study. The rationale was to explore whether a regression analysis might add further, supporting evidence that quasi-dyslexic students appear to present higher levels of academic confidence than might be expected, based on their levels of dyslexia-ness.

In the first instance, a simple, linear regression analysis between the full, ABC24 Scale and the complete, Dx20 Profiler validated a moderate association between ABC and Dyslexia Index with an R2 value (effect size) of 0.1895 (unbiased R2 = 0.1853), derived from Pearson’s coefficient of correlation, *r* = 0.4353 (Figure 24). Hence this suggested that lower levels of academic confidence might be expected from individuals presenting higher levels of dyslexia-ness, an outcome that has already been demonstrated as likely, based on data from students in this datapool and the analysis above (sub-sections 4.3-4.6). To contextualize this outcome, seven further scatterplots with trendlines were produced for all other combinations of ABC Scales and Dx Profilers (not shown), which indicated broadly similar R2 values, placing this result towards the upper end of the range (Table 30).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dx Profiler | ABC24 | ABC17 | ABC21-L | ABC17-L |
| Dx20 | 0.1895 | 0.1946 | 0.1923 | 0.1492 |
| Dx16 | 0.1574 | 0.1609 | 0.1600 | 0.1217 |

Table 30: Values of R2 for simple, linear regressions for all permutations of ABC and Dx Profiler Scales.

But the Dx Profilers are multi-item scales, so it was reasonable to assume that a multiple regression analysis may generate a better model for the data, and hence provide a more accurate mechanism for comparing model-generated, expected mean ABC values with experimentally derived data. However, this study has accommodated four versions of the ABC Scale to analyse data derived from two versions of the Dx Profiler. Hence it was considered that running multiple regression analyses on the eight possible models resulting from permutations of these ABC and Dx Scales would be more appropriately conducted in a subsequent study, with a clear, research design to focus exclusively on this aspect of the possible relationships between academic confidence and dyslexia-ness.

Nevertheless, one multiple regression analysis was conducted using the ABC24 Scale and the Dx20 Profiler, principally as a pilot exercise to determine the feasibility of the processes, to examine whether outputs were meaningful, and hence to indicate whether such a study would be worthwhile later.​ In total, three multiple regression analyses were conducted to generate distinct regression equations from which four outcomes were of interest:

I - to generate expected ABC24 for all groups and subgroups based on the regression equation derived from Dx20 using data from the complete datapool;

II - to generate expected ABC24 for non-dyslexic students based on the regression equation derived from Dx20 data from that research group;

III - to generate expected ABC24 for dyslexic students based on the regression equation derived from Dx20 data from that research group;

IV - to generate expected ABC24 for students in the Test subgroup, based on the regression equation derived from Dx20 data for the dyslexic group.

In each of the four models the objective was to compare the expected mean ABC24 to the observed mean ABC24 so that the closeness of match could be examined. As this was a pilot for a later, more detailed multiple regression analysis, calculating differences between observed and expected mean ABC values was considered sufficient for this purpose. A more analytical examination could be developed later as part of an appropriate research design. The greatest interest was in the output for model IV, which compared the observed mean ABC for the quasi-dyslexic students in the Test subgroup to their dyslexia-identified peers as both cohorts presented on average, similar levels of dyslexia-ness.

The analysis was considered valid as a consequence of preliminary assumptions and tests thus: According to the study design it was considered highly unlikely that observations would be related, confirmed by the Durbin-Watson test for independence of errors

A screenshot of a cell phone

Description automatically generated

Figure 16: Scatterplot of ABC24 against Dx20 for the complete datapool

(residuals), which generated an output of 1.881. A value close to 2 is considered sufficient to demonstrate this (Lund & Lund, 2016-18). Tests for linearity were conducted by observing scatterplots of the studentized residuals against the unstandardized predicted values for each of the five regressions. The residuals formed an approximately horizontal band in all scatterplots, so it was assumed that the independent variables collectively are linearly related to the dependent variable, (see Appendix 8.5). Homoscedasticity was demonstrated through a visual inspection of the scatterplots of studentized residuals against unstandardized predicted values. Interpretation of correlation tables showed that none of the correlation coefficients were > 0.7 for any of the regression models indicating no evidence of multicollinearity. This was further confirmed by consulting the Table of Collinearity Tolerances where none were less than the recommended critical value of 0.1 (ibid).

Significant outliers were not detected on the basis of standardized residuals being greater than +/- 3 standard deviations (SDs). Consulting the studentized deleted residuals also confirmed the unlikelihood of significant outliers as none were greater than +/- 3 SDs. Checking for any datapoints having undue influence on the regressions showed that 93% of the datapoints presented leverage values of < 0.2, considered the boundary criteria between ‘safe’ and ‘risky’ (ibid), with all datapoints < 0.289 leverage. As a further test for influential datapoints, Cook’s Distance values were examined and none showed a value >1, considered to be the criteria for testing influence (ibid).

Visual inspection of Normal P-P plots of the regression standardized residuals indicated that the distributions were approximately normal (see Appendix 8.5, Figure 43 for an example Normal P-P plot). To test the ‘goodness of fit’ of the regression models to the data, the proportion of variance explained by each regression model (adjusted R-squared) was I:43.6%; II:42.7%; III:31.6%, suggesting that all models were adequate. To determine the statistical significance of the models, that is, whether they are significantly better at predicating ABC than the mean model, the ANOVA outputs showed that all models returned a statistically significant result (Appendix 8.5, Table 39 (to update)).

The mean ABC values calculated from observed data for the complete datapool, each of the research groups, and subgroups, was compared with the expected mean ABC values generated from the models (Table 31). Differences between observed and expected mean ABC values are generally small for models used to test their own cohort's data (not unsurprisingly), which confirmed the overall validity of the models. For example, the observed mean ABC=58.45 for the dyslexic group is a slim, 0.01 percentage points below the expected mean ABC=58.46 using the regression equation built from this research group’s observed data.

However, the results of particular interest showed that the quasi-dyslexic students presented higher than expected levels of academic confidence whichever model was used. The disparity was greatest with Model IV, which indicated a +6.07 percentage-points, higher-than-expected, average result for these students. Hence, initial evidence from this multiple regression analysis pilot indicated that quasi-dyslexic students appeared to present average levels of academic confidence that were substantially higher than might be expected, given their levels of dyslexia-ness. Thus a subsequent study to conduct a more detailed analysis would be warranted.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cohort | Observed mean ABC24 | | Expected mean ABC24:  Model I (datapool) | | Expected mean ABC24:  Model II (ND) | | Expected mean ABC24:  Model III (DI) | | Differences: ABC24 observed - expected |
| Datapool | 63.62 | | 63.56 | |  | |  | | +0.04 |
| Non-dyslexic group (ND) | 67.21 | | 66.54 | | (67.15) | |  | | +0.67, (+0.06) |
| Dyslexic group (DI) | 58.45 | | 57.47 | |  | | (58.46) | | +0.98, (+0.01) |
| Base subgroup | 72.44 | 71.91 | | (72.60) | |  | | +0.53, (-0.16) | |
| Test subgroup | 61.72 | 57.47 | | (58.47) | | [55.65] | | +4.25, (+3.35), [+6.07] | |
| Control subgroup | 58.05 | 57.84 | |  | | (57.72) | | +0.21, (+0.33) | |

Table 31: Comparisons of mean ABC24 between observed and expected values according to pilot multiple regression models.