ORIGINAL RESEARCH

An investigation of the use of standardised leaving certificate performance as a method of estimating pre-morbid intelligence

E. Costello[1,2,*], T. Burke[1], K. Lonergan[2], T. Burke[2,3], N. Pender[2,3] and M. Mulrooney[2]

[1] *School of Psychology, Dublin City University, Dublin 9, Ireland*
[2] *Department of Psychology, Beaumont Hospital, Dublin 9, Ireland*
[3] *Academic Unit of Neurology, Trinity College Dublin, Dublin, Ireland*

Background: In cases of brain pathology, current levels of cognition can only be interpreted reliably relative to accurate estimations of pre-morbid functioning. Estimating levels of pre-morbid intelligence is, therefore, a crucial part of neuropsychological evaluation. However, current methods of estimation have proven problematic.

Objective: To evaluate if standardised leaving certificate (LC) performance can predict intellectual functioning in a healthy cohort. The LC is the senior school examination in the Republic of Ireland, taken by almost 50 000 students annually, with total performance distilled into Central Applications Office points.

Methods: A convenience sample of university students was recruited ($n = 51$), to provide their LC results and basic demographic information. Participants completed two cognitive tasks assessing current functioning (Vocabulary and Matrix Reasoning (MR) subtests – Wechsler Abbreviated Scale of Intelligence, Second Edition) and a test of pre-morbid intelligence (Spot-the-Word test from the Speed and Capacity of Language Processing). Separately, LC results were standardised relative to the population of test-takers, using a computer application designed specifically for this project.

Results: Hierarchical regression analysis revealed that standardised LC performance [$F_{(2,48)} = 3.90$, $p = 0.03$] and Spot-the-Word [$F_{(2,47)} = 5.88$, $p = 0.005$] significantly predicted current intellect. Crawford & Allen's demographic-based regression formula did not. Furthermore, after controlling for gender, English [$F_{(1,49)} = 11.27$, $p = 0.002$] and Irish [$F_{(1,46)} = 4.06$, $p = 0.049$] results significantly predicted Vocabulary performance, while Mathematics results significantly predicted MR [$F_{(1,49)} = 8.80$, $p = 0.005$].

Conclusions: These results suggest that standardised LC performance may represent a useful resource for clinicians when estimating pre-morbid intelligence.

Introduction

Estimating pre-morbid intelligence is a crucial process in neuropsychological evaluation (Franzen *et al.* 1997; Teng & Manly, 2005). An understanding of cognitive deficits, relative to a baseline, facilitates clinicians to contextualise their findings allowing for the most suitable diagnosis and/or treatment to be considered. Knowledge of pre-morbid intelligence is also of crucial importance in cases of litigation related to impairment (Reynolds, 1997). Unless an individual has previously completed neuropsychological testing prior to brain injury, estimating pre-morbid functioning and, as a result, the degree of change is difficult and prone to error.

In practice, clinicians typically use one or more of five established methods to estimate pre-morbid

intelligence: 'hold–don't hold' tests (Wechsler, 1958), best performance method (Lezak, 1983), reading ability tests such as the National Adult Reading Test (NART; Nelson, 1982) or the Test of Pre-morbid Functioning (TOPF; Wechsler, 2011), demographic-based regression formulae (e.g. Crawford & Allan, 1997) and combined demographic and current performance formulae [e.g. Oklahoma pre-morbid intelligence estimate (OPIE); Krull *et al.* 1995]. Research has shown that each of these methods has a number of limitations that must be taken into consideration when estimating pre-morbid functioning, not least of which is the degree of statistical error inherent in the procedures.

Hold–don't hold tests, best performance method and reading ability tests are dependent on the assumption that specific cognitive abilities, that is, verbal ability, are more resistant to brain injury not directly impacting a language area. However, numerous studies have shown that 'hold' test performance and reading ability performance are negatively affected by a number of

*Address for correspondence: E Costello, Department of Psychology, Beaumont Hospital, PO Box 1297, Dublin, Ireland.
(Email: emmet.costello23@mail.dcu.ie)

brain pathologies (Russell, 1972; Kaufman & Lichtenberger, 1990; Patterson *et al.* 1994; Johnstone & Wilhelm, 1996; Reynolds, 1997). The best performance technique has also received significant criticism, with studies showing that it produces systematic overestimation of pre-morbid functioning (e.g. Mortensen *et al.* 1991). Demographic-based regression formulae, while objective and easily quantifiable, are limited by their tendency to regress towards the mean. More critically, these formulae rarely perform beyond chance levels in predicting IQ range (Klesges & Troster, 1987). These formulae are also limited by the occupation classification systems they use, which often exclude students, members of the armed forces and homemakers. Combined demographic-performance approaches are argued to be better predictors of IQ than purely demographic formulae (Krull *et al.* 1995; Bright & van der Linde, 2018). However, Axelrod *et al.* (1999) found no significant improvement in prediction accuracy. Worryingly, almost all methods fail to acknowledge the significant variability that individuals have across different cognitive abilities (see, e.g. Franzen *et al.* 1997; Binder *et al.* 2009). See Griffin *et al.* (2002) for a comprehensive review and critique of methods of pre-morbid estimation.

While not without caveats, relatively few studies have recognised standardised exam performance as a possible method of estimating pre-morbid intelligence. Numerous studies have, however, shown that a strong relationship exists between American college test scores (e.g. ACT, SAT, etc.) and measures of intelligence (e.g. Wikoff, 1979; Follman, 1984; Wechsler, 1991). Based on their research, Baade & Schoenberg (2004) proposed that the predicted-difference method (Shepard, 1980) could be used to estimate IQ scores. Their approach utilises a regression equation based on an individual's standardised exam performance and known correlations between college board tests and measures of current intellectual ability to predict pre-morbid intelligence. An exam performance-based approach to predicting IQ in adults has many advantages over other methods: records of academic performance are easily attainable and require no additional testing, academic tests have, generally, been completed prior to the brain injury and are not reliant on current performance and examination of different subjects acknowledges an individual's cognitive variability. Despite the advantages of this approach, no research, to date, has examined if a similar exam performance-based approach to estimating pre-morbid functioning might be useful outside of the United States.

The Irish education system is particularly suited to an exam performance-based approach to predicting IQ as all state examinations are standardised and results are, therefore, normally distributed. Leaving certificate (LC) examinations take place under incredibly well-controlled conditions. Students are examined in a standardised setting, under strict supervision, with restricted time allowed. On the other hand, they do not control for the numerous unknown factors that inevitably affect exam performance, such as mood, effort or the wider context of the individual taking the exam. Students typically sit 7–9 exam subjects, with varying levels of difficulty (Foundation, Ordinary or Higher level). Core exam topics include English, Maths and Irish, while other subjects such as History, Biology, French, etc., are chosen by each individual student. Performance on each subject is graded into standard categories, traditionally A1 being the best and F being the worst. In 2018, this categorisation was replaced by a new system, H1 being the highest grade for higher level and H8 being the lowest, and O1 being the highest grade for ordinary level and O8 being the lowest. Overall performance is also measured by a points system [referred to as Central Applications Office (CAO) points]. The highest grades earn the most points, with each declining grade earning fewer points (e.g. A1 = 100 points, A2 = 90 points, B1 = 80 points). Overall CAO points for each student are the sum of their best six exam results. In order to gain admission into one's preferred college course, students must meet the minimum points total required for that course.

Over 50 000 students complete the LC examinations each year and the Irish Central Statistics Office provides a full statistical breakdown of examination results. Given the ready availability of standardised results (overall performance based on CAO points and subject-specific scores), this approach would permit us to examine overall performance in addition to the variability between an individual's grades on different subjects. This could then, at least potentially, be used in the formation of a unique profile, indicating a person's specific strengths and weaknesses. Statistics released by the State Examinations Commission have shown that there are significant gender differences in the subjects chosen by males and females and in their overall performance (State Examinations Commission, 2015). Subjects such as art, music and home economics are mostly undertaken by females, compared to subjects such as physics, engineering and construction studies, which are more male-dominated. For most subjects, females outperform males, with mathematics and physics the exception. Academic performance is driven by a range of factors, such as motivation, personality and socio-economic status (O'Connor & Paunonen, 2007; Turner *et al.* 2009; Farooq *et al.* 2011). However, the largest predictor of academic performance is intellectual functioning (Heaven & Ciarrochi, 2012).

The current study

The aim of this study was to determine whether and to what extent standardised LC performance can predict current intellectual functioning in a group of university-based healthy controls, specifically university students for whom current regression-based formula approaches are problematic.

In this study, the predictive ability of standardised LC performance (both overall and subject specific) was compared to the predictive ability of two other pre-morbid estimation methods, specifically Crawford & Allan's (1997) demographic-based regression equation and the Spot-the-Word task (Baddeley *et al.* 1992). Additionally, analyses were carried out on standardised subject-specific scores, to examine which exam subjects relate more closely with performance on tests of specific cognitive domains [i.e. Vocabulary and Matrix Reasoning (MR) subtests from the Wechsler Abbreviated Scale of Intelligence, Second Edition (WASI-II); Wechsler, 2011]. Given the exploratory nature of this study, no *a priori* predictions were made as it was considered that to make predictions as to the nature of possible relationships would be unfounded and might limit the scope of the study.

Method

Design

This study was carried out using a non-experimental, quantitative research design. Participants were a convenience sample assessed individually on a range of standardised cognitive tests. Given the lack of previous research in the area, an exploratory approach was taken to data analysis.

Participants

Individuals over the age of 18 years were recruited on a voluntary basis from the Dublin City University (DCU) student population. In total, 51 students took part in the study (26 males and 25 females). Participants had a mean age of 20.80 years (S.D. ± 1.04), ranging from 19 to 22 years. Individuals were recruited from a range of college courses such as psychology, engineering, multimedia, science and communication studies. On average, participants had completed 15.96 years of education (S.D. ± 1.20).

Materials

In order to obtain a measure of IQ, all participants completed the Vocabulary (V) and MR subtests of the WASI-II (Wechsler, 2011). The Spot-the-Word (version 1) subtest from the Speed and Capacity of Language Processing test (Baddeley *et al.* 1992) was administered to obtain a measure of word recognition and to serve as a proxy for pre-morbid IQ.

In addition, participants provided demographic information relative to age, gender, years in education completed, course of study, occupation and details of their LC results. Further detail on the cognitive tests is provided below.

Vocabulary and MR tests measure an individual's current intellectual functioning. The Vocabulary subtest requires participants to define 28 words. This task assesses a person's word knowledge and verbal concept formation. The MR subtest requires participants to view 30 incomplete series of pictures. Participants must complete each series by selecting one of five possible response options to complete the series. This measure is considered to tap into an individual's fluid intelligence, broad performance-based intelligence and classification and spatial ability (Wechsler, 2011).

The Spot-the-Word test is a measure of verbal intelligence specifically designed as an alternative to the NART (Nelson & Willison, 1991), often used to estimate pre-morbid intelligence. The test requires participants to make a lexical decision by identifying real words from pseudo-words. Performance on this test correlates strongly with verbal intelligence and Vocabulary (Baddeley *et al.* 1992).

Procedure

For all participants, testing took place in a small, quiet room in the DCU School of Nursing and Human Sciences. Participants submitted their LC results to the study supervisor and received an ID number. They were then given a self-administered test booklet to record their responses. This booklet included a cover sheet, demographic questionnaire, test instructions and a response sheet. Participants were instructed to complete the demographic questions sheet before moving on to the Vocabulary, MR and Spot-the-Word tests. This procedure ensured that the researcher was unaware of participant's LC results and performance. LC performance was standardised using a Microsoft Excel application specially designed for the project. Participant LC performance was compared to the national cohort (stratified by year, gender and subject) and their percentile scores were calculated. Percentile scores were then converted to standard scores in order to carry out further data analysis.

Data handling and analysis

All cognitive tests were scored according to the criteria specified in the test manuals and converted to standard scores. WASI-II Full Scale IQ (FSIQ)-2 scores were calculated by combining Vocabulary and MR scores for each participant. Crawford &

Table 1. *Means, standard deviations and ranges for CAO, Spot-the-Word, demographic-based regression formula FSIQ and WASI-II FSIQ-2 standard scores*

Test score (standard score)	M	S.D.	Range	Paired *t*-test with WASI-II FSIQ-2 *t*(df)	*p*
CAO score	119.33	6.69	105–135	4.23 (50)	<0.001
IQ					
Spot-the-Word	105.94	12.02	81–135	−4.88 (49)	<0.001
Regression formula FSIQ	93.43	3.36	86–102	−13.82 (50)	<0.001
WASI-II FSIQ-2	113.57	10.23	88–137	–	–

Allan's (1997) regression-based formula was applied to participant's demographic data to generate FSIQ estimates. This formula is presented below:

$$\text{Predicted FSIQ} = 87.14 - (5.21 \times \text{occupation})$$
$$+ (1.78 \times \text{education}) + (0.18 \times \text{age})$$

Occupation was coded based on the Office of Population Censuses and Surveys Classification of Occupation (Boston, 1980), in line with Crawford & Allan's (1997) procedure. However, as this system fails to account for students, individuals in full-time education were assigned to the semi-skilled category. Crawford & Allen's (1997) formula was originally designed to predict Wechsler Adult Intelligence Scale – Revised (WAIS-R) scores. Therefore, in order to compare predicted scores to our measure of current intellectual functioning, these scores were converted to WASI-II scores using the Wechsler interpretation manuals (Wechsler, 1997, 2008, 2011).

Separately, Microsoft Excel files were designed to generate overall and gender-based standardised LC performance scores. These files were based on data that are freely available online (State Examinations Commission, 2015). For the purpose of this study, the files contained a full breakdown of LC national statistics for each year (2012–2015), with both a gender-specific and a non-gender-specific breakdown of results. This allowed for the calculation of percentile rankings of overall and subject-specific LC results for each participant.

Pearson's correlations were carried out to examine the correlation between WASI-II FSIQ-2 scores and both CAO and Spot-the-Word score. Three hierarchical linear regressions were carried out to determine whether and to what extent standardised CAO points (overall and subject specific), Spot-the-Word score and demographic-based regression formula score predict WASI-II FSIQ-2 scores. To control for its potentially confounding effect, age was entered into each model as the first predictor variable. Each estimation method was then entered as the second factor in their respective models. Additionally, two-way mixed, absolute agreement intraclass correlation coefficients (ICCs) were carried out to determine the agreement between each estimation method and WASI-II FSIQ-2 scores.

Results

Descriptives

Initial descriptive analysis was carried out to compare the accuracy of CAO, Spot-the-Word task and demographic-based regression formula scores in predicting WASI-II FSIQ-2. A summary of performance on these tests is provided above (see Table 1).

As can be seen from this table, IQ estimates vary depending on method used with the lowest scores observed with regression formula FSIQ (M = 93.43, S.D. ±3.36) and the highest when objective assessment is used (WASI-II FSIQ-2: M = 113.57, S.D. ±10.23). Regression-based estimates were almost two standard deviations below WASI-II FSIQ-2 scores. Paired *t*-tests found that all methods differed significantly from WASI-II FSIQ-2 scores. CAO performance was significantly higher than WASI-II FSIQ-2 [*t*(50) = 4.23, *p* < 0.001] score, while Spot-the-Word [*t*(49) = −4.88, *p* < 0.001) and demographic-based regression formula [*t*(50) = −13.82, *p* < 0.001] scores were significantly lower than WASI-II FSIQ-2 score.

Prediction of IQ

Pearson's coefficients were calculated to determine the correlation between WASI-II FSIQ-2 and both CAO and Spot-the-Word scores (see Fig. 1). Significant positive correlations were observed between WASI-II FSIQ-2 score and Spot-the-Word score (*r* = 0.44, *p* = 0.002) and between WASI-II FSIQ-2 score and CAO score (*r* = 0.34, *p* = 0.015). Three hierarchical

Table 2. *Hierarchical linear regression model summaries of CAO, Spot-the-Word and regression formulas, respectively, in predicting WASI-II FSIQ-2*

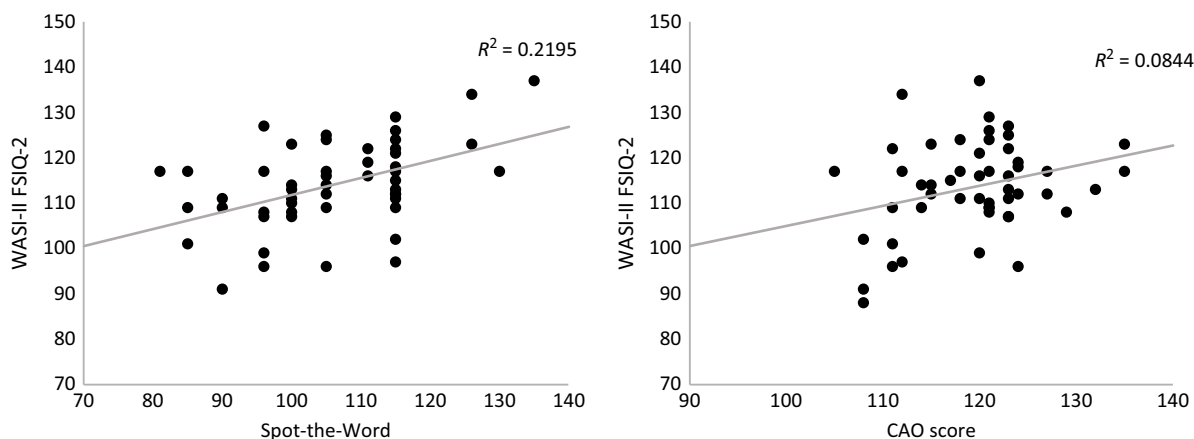|	F(df)	t(df)	β	R^2/R^2 change
Model 1	5.88(2,47)			0.20**
Age		0.79(48)	0.61	0.02
Spot-the-Word		3.26(48)	0.33	0.20**
Model 2	0.46(2,48)			0.02
Age		0.37(49)	0.42	0.01
Regression formula FSIQ		0.47(49)	0.25	0.01
Model 3	3.90(2,48)			0.14*
Age		1.17(49)	1.00	0.01
CAO		2.65(49)	0.53	0.14*

*$p<0.05$, **$p<0.005$.

Fig. 1. Scatterplot, Pearson's correlation coefficient and line of best fit between WASI-II FSIQ-2 and Spot-the-Word and CAO scores. *$p < 0.05$, **$p < 0.01$.

linear regression analyses were carried out to examine the amount of variance in observed IQ, as assessed by WASI-II FSIQ-2, explained by each of the two traditional methods of estimating IQ and by use of CAO standardised scores. Summary statistics for these analyses are presented above in Table 2. Data were screened to ensure assumptions of linearity, normality, homoscedasticity, lack of multicollinearity and independence of errors were not violated. One participant's Spot-the-Word score was treated as a missing value as they did not complete enough test items for their score to be valid.

Spot-the-Word performance significantly predicted WASI-II FSIQ-2 F(2,47) = 5.88, $p = 0.005$, accounting for 20% of the variance in WASI-II FSIQ score. Regression-based scores did not significantly predict WASI-II FSIQ. Of note, CAO performance significantly predicted WASI-II FSIQ-2, F(2,48) = 3.90,

$p = 0.03$, accounting for 14% of the variance in WASI-II FSIQ scores. Effect size was small for CAO score and small to moderate for Spot-the-Word score. ICCs were calculated to determine the reliability between each method and WASI-II FSIQ. ICC indicated fair reliability for Spot-the-Word (ICC = 0.5, $p = 0.001$) and CAO points (ICC = 0.41, $p = 0.01$), and poor reliability for demographic-based regression formula scores (ICC = 0.06, $p = 0.29$).

Predicted–obtained discrepancy scores

Given the significant association between WASI-II FSIQ-2 and Spot-the-Word and between WASI-II FSIQ-2 and CAO, linear curve estimation was applied to CAO and Spot-the-Word scores to generate predicted FSIQ scores. This computation utilised the significant regression equations below to generate predicted FSIQ:

Table 3. *Linear regression model summaries of English, Irish, Math and Spot-the-Word in predicting Vocabulary and MR standard scores, using non-gender-specific and gender-specific standard scores*

Criterion variable	F(df)	t(df)	β	R^2	F(df)	t(df)	β	R^2
		Non-gender-specific				Gender specific		
Vocabulary								
English	6.09(1,49)	2.47(49)	0.46	0.11*	11.27(1,49)	3.36(49)	0.6	0.19**
Irish	1.53(1,46)	1.24(46)	0.20	0.03	4.06(1,46)	2.01(46)	0.31	0.08*
Spot-the-Word	11.60(1,48)	3.41(48)	0.37	0.20**				
Matrix reasoning								
Mathematics	8.80(1,49)	2.97(49)	0.62	0.15**	5.32(1,49)	2.31(49)	0.50	0.10*
Spot-the-Word	6.44(1,48)	2.54(48)	0.33	0.12*				

*$p<0.05$, **$p<0.005$.

$$\text{CAO Predicted FSIQ} = 28.78 + \text{Age}$$
$$+ (0.526 \times \text{CAO T score})$$

$$\text{Spot-the-Word Predicted FSIQ}$$
$$= 4.986 + (0.61 \times \text{Age})$$
$$+ (0.44 \times \text{Spot-the-Word T score})$$

Using the predicted-difference method, the accuracy of each method's predicted score was then examined and compared. Paired sample *t*-tests were carried out to evaluate if CAO-predicted scores were more accurate than Spot-the-Word-predicted scores (i.e. whether predicted–obtained discrepancy scores were lower for CAO estimates). Based on this analysis, the mean absolute errors of CAO-predicted FSIQ (M = 4.73, s.d. ± 3.44) and Spot-the-Word-predicted FSIQ (M = 4.18, s.d. ± 3.59) did not differ significantly, $t(49) = 1.37$, $p = 0.18$.

Gender-specific LC subject scores

Further analysis was carried out to explore the relationship between specific LC subjects and performance on specific WASI-II subtests (See Table 3 below for a full summary of regression models when using gender-specific and non-gender-specific standard scores to predict Vocabulary and MR performance). Analysis focused largely on core LC subjects (i.e. English, Irish and Maths). Linear regression analyses revealed that standardised English performance (without reference to gender) significantly predicted Vocabulary scores [F(1,49) = 6.09, $p = 0.017$], accounting for 11% of the variance in Vocabulary scores. When taking gender into account, English significantly predicted Vocabulary performance [F(1,49) = 11.27, $p = 0.002$] and accounted for a higher degree of variance (19%). Vocabulary performance was also significantly predicted by the

Spot-the-Word test [F(1,48) = 11.60, $p = 0.001$], with a similar effect size ($R^2 = 0.20$). Gender-standardised Irish was also a significant predictor of Vocabulary [F(1,46) = 4.06, $p = 0.049$], accounting for 8% of variance.

Standardised Math score (without reference to gender) significantly predicted performance on MR, [F(1,49) = 8.80, $p = 0.005$], accounting for 15% of variance. However, taking gender into account did not improve the amount variance in MR explained. Spot-the-Word performance also significantly predicted MR [F(1,48) = 6.44, $p = 0.014$], but accounted for a lesser degree of variance (12%).

Discussion

The results of this study suggest that LC performance is a useful predictor of current intellectual functioning, comparable to the Spot-the-Word test. Both overall CAO points and Spot-the-Word scores were predictors of WASI-II FSIQ-2, although they accounted for a relatively small amount of variance (14% and 20%, respectively). Crawford & Allan's (1997) regression formula failed to significantly predict WASI-II FSIQ-2, grossly underestimating the majority of participant's scores. Spot-the-Word and CAO scores had moderate positive correlations with WASI-II FSIQ-2 and ICCs revealed fair reliability for both. Application of the predicted-difference method suggested by Baade & Schoenberg (2004) revealed that CAO and Spot-the-Word predicted scores were very similar in terms of prediction accuracy. These findings suggest that CAO points are as useful a predictor of current intellectual functioning as is Spot-the-Word score.

The advantage of using LC performance in estimating pre-morbid intelligence in adults is that it is already completed by most people and is likely to be completed before a brain pathology occurs and, therefore, cannot

be influenced by its effects. In contrast, the Spot-the-Word relies on current performance following brain injury. Research has shown that numerous forms of brain damage (e.g. right hemisphere stroke) can negatively affect performance on tests of verbal intelligence and verbal functions, such as the Spot-the-Word test (Patterson *et al*. 1994; Johnstone & Wilhelm, 1996).

Analysis also revealed that numerous LC subject scores predicted performance on Vocabulary and MR subtests. Gender-specific standardised English and Irish significantly predicted Vocabulary scores. It is probable that English and Irish LC tests tap into an individual's verbal intelligence similar to the Vocabulary test. Spot-the-Word was also a predictor of Vocabulary performance accounting for similar degree of variance as English (20% and 19%, respectively). Standardised Math performance was a significant predictor of MR, surprisingly accounting for a greater degree of variance without taking gender into account. Mathematics is one of few mandatory LC subjects that rely very little on verbal ability, perhaps tapping into more non-verbal abilities similar to those assessed by MR. Spot-the-Word was also a significant predictor of MR, but accounted for a lesser amount of variance (12% *v.* 15%).

The advantage of using individual subject scores over Spot-the-Word performance is that they allow a clinician to build a specific profile of a person's relative strengths and weaknesses. Examination of an individual's Spot-the-Word score does not allow a clinician to compare a person's verbal and non-verbal ability. Given the results of this study, there is a scope for further study into the relationship between LC subjects and cognitive abilities that might help provide a more holistic view of a person's pre-morbid profile.

This study provides further support that reading ability tests are a fair measure of current intellectual functioning. However, research has shown that performance on measures of reading ability (such as the NART and TOPF) is vulnerable to brain damage (Patterson *et al*. 1994; Johnstone & Wilhelm, 1996) and may, therefore, underestimate pre-morbid functioning. These tests can also be complicated by the presence of a specific learning disability, that is, Dyslexia, which is somewhat controlled for with allowances during the LC exam process, and not subject to a single test. Further research comparing Spot-the-Word and LC performance in controls and patients with brain pathology is needed to evaluate if LC performance might be a more accurate predictor of pre-morbid functioning. Future research should also evaluate a combined demographic formula – 'hold' test approach (such as OPIE) relative to CAO points. While recent research has supported the use of a combined approach, this study highlights the serious problem of using any demographic formula in student populations.

The results of this study are promising but there are a number of limitations that must be acknowledged. Intellectual capacity is a necessary, but not sufficient, condition for academic achievement (Schinka & Vanderploeg, 2000). When examining academic performance, one must also consider factors that may result in a performance not reflecting intellectual capacity (e.g. peer pressure, lack of financial resources, personality factors and poor social support). One must also be cautious with CAO scores as an individual may only perform to the minimum requirements of their desired college course. If a person has no intention of further study or their course entry requirement is much lower than they are capable of achieving, they will likely be less motivated to perform to the best of their ability. Another limitation of this study is that Crawford's demographic-based regression formula was designed to predict WAIS-R scores and not WASI-II scores. While these scores were converted through WAIS interpretation manuals, this is not an ideal comparison. Lastly, this study was carried out on a very young, well-educated and highly intelligent cohort. As such, future study is needed to assess the predictive utility of LC performance in a range of age, education, occupation and socio-economic status groups.

As with any measure of pre-morbid intelligence, it is essential that clinicians use appropriate discretion and take into consideration numerous sources of information. As this study shows, even the best subjects are mild/moderate predictors of performance and should be interpreted with caution. Patient interviews, academic scores and current performance scores should be used in combination to build a cognitive profile of a person's pre-morbid functioning. Examination of LC performance may prove particularly useful in cases of litigation, where multiple sources of evidence may be required. It may also be a useful option to examine when a patient is unable to complete any tests but estimates of pre-morbid functioning are required (e.g. a patient in a minimally conscious state).

Exam performance and pre-morbid estimation are an unexplored area of neuropsychology. This study shows the predictive power of LC scores and although modest in terms of predictive power highlights the potential for further research. A large-scale study with a larger sample size could further elucidate the value of specific subjects using multiple regression. Unfortunately, due to a limited sample size, many subjects such as Art and Music were not completed by enough participants to meet the assumptions of analysis. A more extensive battery of current functioning would also allow for more detailed comparison of specific cognitive domains and exam subjects.

A potential extension of this study could evaluate efficacy of LC performance in other populations. Due

to the nature of recruitment, the majority of participants were in the Average to High Average IQ range. It would be interesting to examine if LC performance is a useful predictor in Low Average and non-student populations. Another promising area of investigation would be to look at Junior Certificate performance. Examination of an individual's Junior Certificate performance could provide concurrent evidence of academic achievement and helps a clinician develop a developmental profile of a person's educational performance. Further research could also explore the usefulness of mandatory primary school standardised tests (e.g. New Non-Reading Intelligence Test; Young & McCarty, 2012), which would extend a developmental profile. This may prove especially useful in establishing pre-morbid intelligence in children with brain injury.

Acknowledgements

Financial support

Conflict of interest

EC has no conflicts of interest to declare. TB has no conflicts of interest to declare. KL has no conflicts of interest to declare. TB has no conflicts of interest to declare. NP has no conflicts of interest to declare. MM has no conflicts of interest to declare.

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committee on human experimentation with the Helsinki Declaration of 1975, as revised in 2008. This study was approved by the DCU Research Ethics Committee (REC) under delegated authority to the School of Nursing and Human Sciences Undergraduate Psychology Ethics Committee (UPEC). Written informed consent was obtained from all participants in this study.

References

Axelrod BN, Vanderploeg RD, Schinka JA (1999). Comparing methods for estimating premorbid intellectual functioning. *Archives of Clinical Neuropsychology* **14**, 341–346.

Baade LE, Schoenberg MR (2004). A proposed method to estimate premorbid intelligence utilizing group achievement measures from school records. *Archives of Clinical Neuropsychology* **19**, 227–243.

Baddeley AD, Emslie H, Nimmo-Smith I (1992). *The Speed and Capacity of Language-Processing Test*. Thames Valley Test Company: London, UK.

Binder LM, Iverson GL, Brooks BL (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology* **24**, 31–46.

Boston G (1980). Classification of occupations. *Population Trends* **20**, 9–11.

Bright P, van der Linde I, (2018). Comparison of methods for estimating premorbid intelligence. *Neuropsychclical Rehabilitation* **2018**, 1–14.

Crawford JR, Allan KM (1997). Estimating premorbid WAIS-R IQ with demographic variables: regression equations derived from a UK sample. *The Clinical Neuropsychologist* **11**, 192–197.

Farooq MS, Chaudhry AH, Shafiq M, Berhanu G (2011). Factors affecting students' quality of academic performance: a case of secondary school level. *Journal of Quality and Technology Management* **7**, 1–14.

Follman J (1984). Cornucopia of correlations. *American Psychologist* **39**, 701.

Franzen MD, Burgess EJ, Smith-Seemiller L (1997). Methods of estimating premorbid functioning. *Archives of Clinical Neuropsychology* **12**, 711–738.

Griffin SL, Mindt MR, Rankin EJ, Ritchie AJ, Scott JG (2002). Estimating premorbid intelligence: comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology* **17**, 497–507.

Heaven PCL, Ciarrochi J (2012) When IQ is not everything: Intelligence, personality and academic performance at school. *Personality and Individual Differences* **53**, 518–522.

Johnstone B, Wilhelm KL (1996). The longitudinal stability of the WRAT-R reading subtest: is it an appropriate estimate of premorbid intelligence? *Journal of the International Neuropsychological Society* **2**, 282–285.

Kaufman AS, Lichtenberger EO (1990). *Assessing adult and adolescent intelligence*. Allyn 8c Bacon: Boston.

Klesges RC, Troster AI (1987). A review of premorbid indices of intellectual and neuropsychological functioning: what have we learned in the past five years? *International Journal of Clinical Neuropsychology* **9**, 1–11.

Krull KR, Scott JG, Sherer M (1995). Estimation of premorbid intelligence from combined performance and demographic variables. *The Clinical Neuropsychologist* **9**, 83–88.

Lezak M (1983). *Neuropsychological Assessment*, 2nd edn. Oxford University Press: New York.

Mortensen EL, Gade A, Reinisch JM (1991). A critical note on Lezak's 'best performance method' in clinical neuropsychology. *Journal of Clinical and Experimental Neuropsychology* **13**, 361–371.

Nelson HE (1982). *National Adult Reading Test (NART): For the Assessment of Premorbid Intelligence in Patients with Dementia: Test Manual*. NFER-Nelson: Windsor.

Nelson HE, Willison J (1991). *National Adult Reading Test (NART)*. NFER-Nelson: Windsor.

O'Connor MC, Paunonen SV (2007). Big five personality predictors of post-secondary academic performance. *Personality and Individual Differences* **43**, 971–990.

Patterson KE, Graham N, Hodges JR (1994). Reading in dementia of the Alzheimer type: a preserved ability? *Neuropsychology* **8**, 395.

Reynolds CR (1997). Postscripts on premorbid ability estimation: conceptual addenda and a few words on alternative and conditional approaches. *Archives of Clinical Neuropsychology* **12**, 769–778.

Russell EW (1972). WAIS factor analysis with brain-damaged subjects using criterion measures. *Journal of Consulting and Clinical Psychology* **39**, 133.

Schinka JA, Vanderploeg RD (2000). Estimating premorbid level of functioning. In *Clinician's Guide to Neuropsychological Assessment* (ed. RD Vanderploeg), pp. 39–67. Lawrence Erlbaum Associates Publishers: Mahwah, NJ.

Shepard L (1980). An evaluation of the regression discrepancy method for identifying children with learning disabilities. *The Journal of Special Education* **14**, 79–91.

State Examinations Commission (2015). *State examinations statistics* (https://www.examinations.ie/?l=en&mc=st&sc=r15). Accessed 10 October 2016.

Teng E, Manly J (2005). Neuropsychological testing: helpful or harmful? *Alzheimer Disease & Associated Disorders* **19**, 267–271.

Turner EA, Chandler MM, Heffer RW (2009). The influence of parenting styles, achievement motivation, and self-efficacy on academic performance in college students. *Journal of College Student Development* **50**, 337–346.

Wechsler D (1958). Mental deterioration and its appraisal. In *The Measurement and Appraisal of Adult Intelligence* (ed. J Matarazzo, 4th edn), pp. 199–213. Williams & Wilkins Co: Baltimore, MD.

Wechsler D (1991). *The Wechsler Intelligence Scale for Children: Manual*, 3rd edn. The Psychological Corporation: San Antonio, TX.

Wechsler D (1997). *WAIS-III: Wechsler Adult Intelligence Scale*, 3rd edn. The Psychological Corporation: San Antonio, TX.

Wechsler D (2008). *WAIS-IV: Wechsler Adult Intelligence Scale*, 4th edn. NCS Pearson: San Antonio, TX.

Wechsler D (2011). *WASI-II: Wechsler Abbreviated Scale of Intelligence*. Pearson: Bloomington, MN.

Wikoff RL (1979). The WISC-R as a predictor of achievement. *Psychology in the Schools* **16**, 364–366.

Young D, McCarty CT (2012). *New Non-reading Intelligence Tests 1-3 (NNRIT 1-3) Manual: Oral Verbal Group Tests of General Ability*. Hodder Education: London, UK.