

This is the peer reviewed version of the following article:

McKenzie, K., Murray, G.C., & Rae, H. (2007) Changes made to the Wechsler Intelligence Scale for Children – Third Edition by Psychologists working in Child Services. *Clinical Psychology Forum*, 178, 25-28. The final published version is available via : <http://www.bps.org.uk/networks-and-communities/member-networks/division-clinical-psychology>

Changes made to the Wechsler Intelligence Scale for Children – Third Edition by Psychologists working in Child Services

Abstract

Psychologists were found to amend the way they used the WISC-III, by either regularly missing out sub-tests or not following the standardised manual instructions. The majority felt these changes would impact on the test scores of the individual child.

Word count= 3414 (including references)

Keywords: WISC III, Administration, Psychologists, Child services

Introduction

The Wechsler Intelligence Scale for Children-Third Edition (WISC-III) (Wechsler, 1991) is a well-validated, standardised and objective intellectual assessment that is commonly used in clinical and educational services (Watkins et al., 1995, Stinnett et al., 1994). A strong relationship has been found between an individual's global IQ and academic achievement (Neisser et al, 1996). Intellectual assessments have also been commonly used to help establish an individual's cognitive strengths and weaknesses, learning support needs and if they have a profile that is suggestive of specific learning difficulties e.g. dyslexia, or learning disability (Canivez & Watkins, 1998; Evers & Hill, 1999). As well as looking at overall profiles, scores on individual sub-tests of the Wechsler scales have also been used to indicate specific learning difficulties (Kaufman, 1979) and neuropsychological deficits (Rourke et al., 1983).

Some researchers have, however, argued that subtest profiles are not predictive of academic performance (Watkins & Glutting, 2000). Intellectual assessments do, however, have important implications for the child in terms of accessing learning support and other educational services. The identification of a child as having special educational needs also triggers a legislative process whereby the child receives follow-up reviews within set periods. These processes can have a crucial impact on the educational career and attainment of the child and it is, therefore, crucial that the assessments are valid, reliable and used only by appropriately trained and qualified professionals. This is formalised in the UK by professional guidance that ensures that any assessment of an individual's intellectual functioning is carried out by, or under the supervision of an

applied psychologist, using an individually administered, standardised psychometric assessment which is reliable and valid (British Psychological Society, 2001).

To increase the probability of the assessment outcome being reliable and valid, the WISC III manual recommends that the clinician follow the instructions for administration and scoring as closely as possible. Psychologists have, however, been found to make a number of errors when administering the tests, including incorrect scoring of answers, not recording responses and asking inappropriate questions (Moon et al. 1991, Slate et al, 1992). A more recent study by Alfonso et al. (1998) found that psychologists incorrectly reported the child's Full Scale and Verbal IQs. Such mistakes can lead to an over-estimate of the child's IQ (Slate et al., 1991).

As well as examining the effects of mistakes in administration, research has also examined the impact of factors which are deliberately introduced by the assessor. An early study by Witmer et al (1971) found that being given verbal approval during testing led to a significantly higher score on the Wechsler Intelligence Scales for Children (WISC, Wechsler, 1949). This result was replicated by Saigh (1981) using the Wechsler Intelligence Scales for Children-Revised (WISC-R, Wechsler, 1974). More recent research, which examined the extent to which clinical psychologists amended the use of the WAIS III with adults with a learning disability (McKenzie et al. 2004), found that over 83% deliberately amended the administration. These changes included regularly missing out particular sub-tests, simplifying the language and providing additional instructions, encouragement or praise.

The present paper aims to extend the work of McKenzie et al. (2004) to examine the extent to which deliberate changes in the administration of intellectual assessments by psychologists extend beyond learning disability services to child services. The aim of the present paper was, therefore, to explore the way the WISC III is currently administered by psychologists working in child and educational services in Scotland.

Method

This was a self-report study, consisting of a postal questionnaire which asked the following:

- How long have you worked in child services?
- Do you use, or have you previously used, the WISC III?
- If so, how useful do you find it? This was measured on a scale of 1 to 5, with 1 equalling not useful at all and 5 equalling very useful.
- Do you follow the instructions exactly as written in the manual when assessing a child?
- If yes, do you find any difficulties with this?
- If no, please note the types of change you make for each subtest.
- Do you miss out any of the sub-tests on a regular basis?
- If yes, please note which ones and the reasons why
- Do you think that any of the changes you make impact on the outcome of the assessment

- If yes, please note the reasons why.
- Any other comments.

Participants were provided with a table listing all of the subtests, to be completed in relation to each relevant question. All responses were anonymous.

One hundred and thirty six questionnaires were sent out to qualified clinical and educational psychologists working in child and educational services in Scotland. Of these eight respondents replied to say that they no longer worked in child services. Fifty-eight completed questionnaires were returned, giving a response rate of 45%. Forty were educational psychologists and the remaining 18 were clinical psychologists. The amount of experience respondents had working in child/education services ranged from 1 to 37 years, with a mean of 19.47 (SD=9.73).

Results

Of the 58 respondents, 46 said that they currently, or had previously used the Wechsler Intelligence Scale for Children– 3rd Edition, (WISC-III). There was no significant difference found between clinical and educational psychologists in terms of their likelihood to use the WISC III. Of those 46 who used the WISC III, 21 used it at least every three months while 14 used it every 6-12 months. The remaining respondents used it less than once a year. In terms of usefulness, the minimum score was 1, the maximum was 5 and the mean was 3.52 (SD = 1.13). Thirty-seven respondents rated the usefulness of the WISC III at 3 or above, indicating that it was considered to be a useful to very

useful assessment. The clinical psychologists were significantly more likely to rate the WISC III as more useful than the educational psychologists ($z=-2.276$, $p<0.05$).

Omitting Subtests

No significant differences were found between the clinical and educational psychologists in terms of omitting subtests. Twenty seven respondents who used the WISC III reported that they regularly missed out subtests. Only two sub-tests were never omitted by any of the respondents who used the WISC III. These were Block Design and Verbal Comprehension. Four subtests were regularly omitted by seven or more respondents. Mazes was omitted by 26 respondents, Symbol Search by 18, Digit Span by 11 and Object Assembly by 7. The most common reason for omitting a sub-test was due to time constraints (30), while additional reasons were that the information provided by the subtest was not relevant (12) or that the information was obtained by other means (11).

A Cochran's Q test found a significant difference between the sub-tests in terms of the likelihood of being omitted $Q=168.49$, $df =12$, $p<0.0001$. The Binominal test found that the Symbol Search sub-test was significantly more likely to be omitted than any of the other sub-tests, ($n=38$, $p<0.0001$), with the exception of Picture Arrangement and Object Assembly. The Mazes subtest was significantly more likely to be omitted than any other subtest ($n=38$, $p<0.0001$).

Administering/Presenting Sub-tests

When asked about administering the WISC III, 30 of respondents who used the WISC III, said that they did not follow the standardised instructions in the WISC III manual

exactly. No significant differences were found between the clinical and educational psychologists in this area. Table 1 illustrates the number of respondents making each type of change in relation to each of the WISC III subtests.

Table 1: The number of respondents making each type of change in relation to each of the WISC III subtests.

Test	Change language (e.g. simplify, make more concrete)	Do not use strict timings	Change order of presentation of items	Don't always follow rule as specified	Give additional instruction, clarification or praise	Total
	No.	No.	No.	No.	No.	No.
Picture Completion	12	7	0	8	19	46
Information	8	2	0	9	14	33
Coding	11	0	0	2	16	29
Similarities	17	1	1	8	18	45
Picture Arrangement	11	1	1	6	17	36
Arithmetic	8	5	1	6	16	36
Block Design	10	3	2	6	17	38
Vocabulary	9	1	1	8	16	35
Object Assembly	10	2	3	4	17	36
Comprehension	11	1	1	6	19	38
Symbol search	7	0	0	1	9	17
Digit Span	7	0	1	2	12	22
Mazes	5	0	0	1	7	13

A Cochran's Q test was used to examine if any one sub-test was significantly more likely to be used in a non-standardised way than the others. When an adjustment was made to allow for multiple comparisons, the only significant differences were found in relation to giving additional instructions, clarification or praise ($Q=70.023$, $df=12$, $p<0.0001$). Binominal tests showed that respondents were less likely to provide additional guidance or praise in relation to the mazes subtest as compared with Picture Completion and Comprehension ($n=40$, $p<0.0001$).

Experience and Responses

The relationship between respondent's experience of working in child services and their rating of the usefulness of the WISC III was examined. No significant relationship was found. Similarly, no relationship was found between experience and total of number of changes that respondent's made to the WISC III administration.

The impact of changes to WISC III administration

Thirty-five participants felt that changes in the administration would have an impact on the results. Some gave more than one reason, but in total, eight felt it would reduce the reliability of the assessment, while 11 felt it would provide a more accurate picture of the child's abilities. Twenty felt it would increase the child's understanding of the task and 12 felt it would increase rapport or child confidence.

Discussion

The present study examined the extent to which qualified clinical and educational psychologists working in Scotland consciously altered the way in which the WISC III was administered. Forty-six of the psychologists who participated had experience of using the WISC III and 21 used it at least every three months. In addition, it was rated as a useful tool by the majority of participants. This finding was very similar to the profile found for clinical psychologists using the WAIS III in learning disability services (McKenzie et al., 2004).

Many psychologists were found to routinely miss out one or more subtests on the WISC III, with half routinely missing the Symbol Search and Mazes subtests. The latter test

does not contribute to the calculation for either the IQ scores or Index scores and this may be the reason that it is so commonly omitted. The main reasons given for tests being omitted were that the sub-tests were time-consuming and the information derived from the tests was not considered to be useful. Only two sub-tests were never missed out by any respondents. These were Block Design and Verbal Comprehension. This suggests that individual psychologists are developing their own short-forms of the WISC III, to meet the time constraints and clinical demands of their work. This result was again similar to that found by McKenzie et al. (2004) in relation to the use of the WAIS III in learning disability services. Here three quarters of participants regularly missed out at least one sub-test.

Research suggests that efficiency is one of the main reasons that practitioners adopt short-forms (Donders, 1997). The development of short-forms, however, requires the instrument to meet the necessary standards for reliability, validity and underlying factor structure of the original instrument. Donders (1997) proposes an 8 subtest short form which has been evaluated in respect of these criteria. The participants in the present study, however, do not appear to be using short-forms based on previously evaluated models and there is the danger that the results of such shortened assessments may be less reliable and valid as a result. Hishinuma (1995), for example, suggests that missing out tests in a battery should be categorised as a modification that is likely to have a significant psychometric impact.

Thirty of the psychologists also reported that they consciously changed the way in which they administered the WISC III. The psychologists were most likely to give additional instructions, clarification and praise or change the way that instructions were worded. Research has indicated that providing more explicit instructions about a timed subtest can lead to a significant increase in scores when compared with standard instructions (Joncas & Standing, 1998). The psychologists in the present study may be being influenced by the likelihood that, if the child does not understand the instructions for a test, then the assessment may simply reflect the child's verbal comprehension, rather than the other constructs that the sub-test was designed to measure. McKenzie et al. (2004) found that improving understanding of the assessment was one of the main reasons given by participants who changed the administration instructions when using the WAIS III with clients with a learning disability. In that study 20 of the 24 participants changed the instructions when administering the WAIS III.

Participants in the present study were equally as likely to change the presentation of any of the subtests in the WISC III with one exception. The respondents were less likely to provide additional guidance or praise in relation to the Mazes subtest as compared with Picture Completion and Comprehension. This may be an artefact of the finding that the Mazes sub-test was significantly more likely to be omitted than any of the other sub-tests. Experience of working in child services was not found to be a factor in terms of changing the way the WISC III was administered, with those with more experience being equally as likely to make changes as those with less experience.

In terms of the perceived impact that these changes would have, thirty-five participants felt that changes in the administration would have an impact on the results. The majority saw this in positive terms as a way of increasing rapport, the child's confidence and the usefulness of the results. Only eight respondents felt that it would reduce the reliability of the results. There is, however, strong evidence to suggest that altering the way the WISC III is administered is likely to impact on the reliability of the results. In particular, praise has been shown to significantly increase children's scores on both the WISC (Wechsler, 1949) and the WISC-R (Wechsler, 1974) (Witmer et al, 1974; Saigh, 1981).

Changes in administration may also be partly responsible for the finding that both sub-test (Carnivez & Watkins, 1998) and Index scores (Stavrou & Flanagan, 1996) on the WISC III can change significantly over time from first to second testing. Failure to adhere to the standardised administration procedure for the WISC III may result in the same child being tested on a second occasion by the same psychologist, but in a different way or a different psychologist in a different way. This may in turn impact on the level of support the child receives. A study by Canivez & Watkins (1998), for example, spanned nearly a three year period, and included the results of 667 children, the majority of whom were receiving special education. The authors note that only 298 of these children retained a diagnosis of learning disability across both test administrations. While no detailed reasons for this change are given by the authors, one possibility is that the scores of some children increased significantly on retesting due to changes in test administration, leading them to lose their 'learning disability' diagnosis.

It could be argued that these increased scores may, in fact, be a more accurate reflection of the child's true abilities. Unfortunately, however, if the test is not administered in a standardised way across all psychologists and for all children, then some children may be unfairly disadvantaged in terms of accessing educational support. One solution may be to amend the WISC III administration instructions to reflect the reality of the practice of psychologists and to systematically evaluate the impact this has on the reliability and validity of the tests. An initial step suggested by Hishinuma (1995) was that the professional administering the test rates the changes made in terms of the impact they have had on the psychometric robustness of the test. Unfortunately, until the research is conducted which identifies exactly what impact these modifications have on an individual's performance, the clinician will have little empirical guidance for making these classifications.

In conclusion, the present study included nearly half of the clinical and educational psychologists working in child services in Scotland and found that the majority used the WISC III and found it a useful clinical tool. Of those who used it, 30 out of 46 consciously changed the way they administered it. This is a lower percentage than those psychologists working in learning disability services, of whom 83% changed the administration of the WAIS III (McKenzie et al., 2004). The results of the latter study and the present study, do, however, suggest that the practice of consciously amending the way that intellectual assessments are administered is relatively common in both the child

and learning disability specialties. This suggests that further research is important to clarify exactly what impact changing test administration has on test results.

References

Alfonso, V.C., Johnson, A., Patinella, L. & Rader, D.E. (1998). Common WISC-III examiner errors: evidence from graduate students in training. *Psychology in the Schools*, 35(2), 119-125.

British Psychological Society (Professional Affairs Board) (2001). *Learning Disability: Definitions and Contexts*. Leicester: BPS.

Carnivez, G.L. & Watkins, M.W. (1998) Long-term stability of the Wechsler Intelligence Scale for Children – Third Edition. *Psychological Assessment*, 10(3), 285-291.

Donders, J. (1997). A short form of the WISC III for clinical use. *Psychological Assessment*, 9(1), 15-20.

Evers, C. & Hill, J. (1999). Identifying and classifying learning disabilities. *Clinical Psychology Forum*, 129, 9-13.

Hishinuma, E.S. (1995) WISC-III accommodations: The need for practitioner guidelines. *Journal of Learning Disabilities*, 28(3), 130-135.

Joncas, J. & Standing, L. (1998). How much do accurate instructions raise scores on a timed test? *Perceptual and Motor Skills*, 88, 1257-1258.

Kaufman, A.S. (1979). *Intelligent testing with the WISC-R*. New York: Wiley-Interscience.

McKenzie, K., Murray, G.C. & Wright, J. (2004). Adaptations and accommodations: The use of the WAIS III with people with a learning disability. *Clinical Psychology (In Press)*.

Moon, G.W., Blakey, W.A., Gorsuch, R.L., Fantuzzo, J.W. (1991). Frequent WAIS-R administration errors: An ignored source of inaccurate measurement. *Professional Psychology: Research and Practice*, 22(3), 256-258.

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W. et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.

Rourke, B.P., Bakker, D.J., Fisk, J.L. & Strang, J.D. (1983) *Child Neuropsychology: An introduction to theory, research and clinical practice*. New York: Guilford Press.

Saigh, P.A. (1981). The effects of positive examiner verbal comments on the total WISC-R performance of institutionalised EMR students. *Journal of School Psychology*, 19(1), 86-91.

Slate, J.R., Jones, C.H. & Murray, R.A. (1991). Teaching administration and scoring of the Wechsler Adult Intelligence Scale-Revised: An empirical evaluation of practice administrations. *Professional Psychology: Research and Practice*, 22 (5), 375-379.

Slate, J.R., Jones, C.H., Coulter, C. & Covert, T.L. (1992). Practitioners' administration and scoring of the WISC-R: Evidence that we do err. *Journal of School Psychology*, 30(1), 77-82.

Stavrou, E. & Flanagan, R. (1996). The stability of WISC III scores in learning disabled children. Cited in Carnivez, G.L. & Watkins, M.W. (1998) Long-term stability of the Wechsler Intelligence Scale for Children – Third Edition. *Psychological Assessment*, 10(3), 285-291.

Stinnett, T.A., Havey, J.M. & Kaspar, J.C. (1994). Current test useage by practicing school psychologists: A national survey. *Journal of Psychoeducational Assessment*, 12, 331-350.

Watkins, C.E., Campbell, V.L., Nieberding, R. & Hallmark, R. (1995). Contemporary practice of psychological assessment by clinical psychologists. *Professional Psychology: Research and Practice*, 26, 54-60.

Watkins, M.W. & Glutting, J.J. (2000). Incremental validity of WISC-III profile elevation, scatter and shape information for predicting better reading and math achievement. *Psychological Assessment*, 12(4), 402-408.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: Psychological Corporation.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children – Third Edition*. San Antonio, TX: Psychological Corporation.

Witmer, J.M., Bornstein, A.V. & Dunham, R.M. (1971). The effects of verbal approval and disapproval upon the performance of third and fourth grade children on four subtests of the Wechsler Intelligence Scale for Children. *Journal of School Psychology*, 9(3), 347-356.

