Ethnic Group Differences in Cognitive Ability Test Scores within a New Zealand Applicant Sample

Nigel Guenole
Selector Group
Paul Englert
OPRA Consulting Group
Paul J. Taylor
Chinese University of Hong Kong

Given the widespread use of cognitive ability tests for employment selection in New Zealand, and overseas evidence of substantial ethnic group differences in cognitive ability test scores, a study was conducted to examine the extent to which cognitive test score distributions differ as a function of ethnicity within a New Zealand sample. An examination of 157 Maori and 82 European verbal and numeric ability test scores from within a New Zealand government organisation revealed sizeable and statistically significant mean differences between the two ethnic groups on two of three cognitive tests evaluated. Specifically, Maori scored, on average, 0.55 standard deviations lower than European applicants on a measure of verbal reasoning, and 1.79 standard deviations lower on a measure of numerical business analysis. No mean difference was observed between ethnic groups for a test of general numeric reasoning. In light of these substantial differences on two of the three tests, we discuss strategies that organisations using cognitive tests can employ to minimize adverse impact on Maori applicants, as well as further research that is needed.

Many organisations in New Zealand strive to achieve multiple objectives with their personnel selection procedures, including maximizing both predictive validity and selection utility (i.e., cost effectiveness), as well as achieving and maintaining an ethnically diverse workforce. These goals, however, can conflict, such that selection methods that achieve one goal (e.g., predictive validity) work against another goal (e.g., diversity).

Overseas research suggests that a cognitive ability test is such an example of a selection method that supports one goal at the expense of another (Huffcutt & Roth, 1998; Roth, Bevier, Bobko, Switzer & Tyler, 2001). Cognitive ability tests have been found to be one of the most valid forms of predicting future job performance for a wide range of jobs (Schmidt & Hunter, 1998). For this reason, some authors have suggested that abandoning their use in employment decisions would result in a substantial sacrifice in workforce productivity (Gottfredson, 1994; Hunter & Hunter, 1984). Indeed, Schmidt and Hunter (1998) have argued that, since cognitive ability tests have such high predictive validities, other selection methods should simply be considered as adding incremental validity to the selection decision once the cognitive ability of the candidate has been assessed, implying that cognitive ability testing should be a major component of a thorough selection practice for many jobs. While more recent reviews of the validity of alternative selection methods suggest that, when estimates of range restriction and criterion reliability are standardized across studies, structured employment interviews are at least as valid (Robertson & Smith, 2001), if not slightly more valid (Hermelin & Robertson, 2001) than cognitive ability tests, cognitive ability tests clearly remain one of the more valid predictors of job performance.

Cognitive ability tests play a prominent role in the personnel selection systems of many organisations both overseas and in New Zealand. In a recent survey of selection practices within 100 randomly selected New Zealand organisations and 30 recruitment firms, Taylor, Keelty and McDonnell (2002) found that almost one-half of the organisations sampled use cognitive ability tests for selecting managerial personnel – over twice the proportion used a decade ago (Taylor, Mills & O'Driscoll, 1993) - and that almost two-thirds of recruitment firms use cognitive tests in selection. In fact, the use of cognitive tests in personnel selection is now greater in New Zealand than in many other countries, according to a recent cross-national survey of staff selection practices in 18 countries. This survey found that the prevalence of cognitive ability test use in New Zealand was greater than all but three other countries (Ryan, McFarland, Baron & Page, 1999).

The value of cognitive ability testing for employee selection does not, however, come without costs and some controversy. In the United States, for example, the use of cognitive ability testing has been found to adversely impact African American and Hispanic applicants as a result of substantial differences in mean test scores (Sackett, Schmitt, Ellingson and Kabin, 2001). Large-scale meta-analyses have confirmed that African Americans score approximately one standard deviation lower than Whites on measures of quantitative ability, verbal ability, and comprehension and that similar, and slightly smaller differences (.7 - .8 standard deviations) have been found between Hispanic and White applicants (see, for example, Roth, Bevier, Boko, Switzer, & Tyler, 2001). Consequently, where staff selection is based largely on cognitive ability test scores, members of affected minority groups, such as African Americans and Hispanics in the USA, receive fewer employment opportunities than Whites and some other minority groups (Scientific Affairs Committee, 1994). This situation has led to a dilemma in the USA, in which relying on cognitive ability tests has been seen by many organisations and researchers as sensible from the perspective of maximizing predictive validity, but doing so threatens the achievement of social objectives, such as overcoming past social inequities, pluralism, and creating an ethnically diverse workforce (Sackett & Wilk, 1994; Sackett et al., 2001).

While much research has been conducted in the United States on ethnic differences on cognitive tests used in employment selection, we know of no prior published research on the topic in New Zealand. Such research is important, given both the prevalence of cognitive testing for employment in New Zealand, and government policy programs aimed to decrease social and economic disparity between Maori and non-Maori. If differences in the distributions of occupational cognitive ability test scores between Maori and Europeans are near-zero (i.e., so small as to have no practical significance), employers can use such tests with the confidence that doing so will result in no adverse impact on achieving an ethnically diverse workforce. If, on the other hand, substantial differences in cognitive test scores are found, as they have been overseas, then organisations wishing to employ an ethnically diverse workforce must carefully consider whether and how to use such tests. The purpose of the present study was to investigate whether ethnic group differences exist among a sample of New Zealand job applicants, using historical recruitment data from an organisation that had administered cognitive tests as part of their staff selection procedure.

Method

Participants

Archival test score data were available from a large New Zealand government organisation on applicants who had completed one or more of three cognitive ability tests while applying for professional level positions, such as analysts, senior analysts, or finance positions.

Participants in this research were applicants for analyst, senior analyst, and finance positions. The data used were historical, and were collected as part of the recruitment process. The testing for the candidates in this analysis occurred over the period 1997-2001, including test scores for 239 candidates. These included 76 Maori male test scores, 74 Maori female test scores, 40 European male test scores, 39 European female test scores, and 10 scores (7 Maori and 3 European) of unknown gender. The breakdown for each test is presented in Table 1.

For the purposes of this research, job candidates were classified as Maori if they had identified themselves when applying as either Maori or Maori and any other origin, including European. Only candidates who indicated solely New Zealand European heritage were classified as European for the purposes of this research. Excluded from this analysis were all test scores for candidates who indicated an ethnic origin other than NZ European, or Maori and any other ethnic origin (that is, any ethnic origin other than that identified for inclusion in this analysis), as sample sizes were inadequate for meaningful analysis.

Given that the focus of this paper is on mean differences in test scores, we have not examined employment offer data. Consequently, inferences about whether adverse impact resulted in this particular organisation are impossible to identify. Any adverse impact is likely to have been minimized by the existing recruitment approach that provided equal weighting to structured employment interviews, cognitive testing, and work sample tests (provided ethnic differences were not as prominent on these other recruitment methods).

As the data were from a government organisation, all positions were advertised in external newspapers, jobs websites, or both. However, information with regard to whether candidates were internal or external was unavailable. Ideally, in future research, this variable should be controlled. This would minimise any chance that results observed are due to different ethnic composition of the internal and external samples, if, for example, internal candidates have an advantage over external candidates.

Table 1. Breakdown of Sample for Each Test by Ethnicity and Gender

	Maori			European		
Test	Male	Female	Gender Unknown	Male	Female	Gender Unknown
Verbal reasoning test	54	49	6	25	27	3
Numerical business analysis test	6	13	0	6	4	0
General numerical reasoning test	16	12	1	9	8	0

Note: Some applicants sat more than one of the three tests, and so samples for each test were not independent.

Even if the ethnic composition of internal applicants was different from the composition of the external applicant pool, which is doubtful, such differences could be expected to have limited effect on cognitive ability test scores.

Materials

The cognitive tests investigated in the present study consisted of three, standard (i.e., not customized), commonly used occupational tests provided by Saville & Holdsworth New Zealand Ltd. All three tests were paper and pencil based. Brief descriptions of each of the tests follow:

VMT3 Verbal Reasoning Test. This test measures the ability to understand and interpret complex written information, such as written reports and policy documents in the subject organisation. It consists of a series of passages, followed by questions asked to address a broad range of verbal analysis skills such as summarising, drawing appropriate inferences and logical reasoning. The verbal reasoning test was administered to applicants for policy analyst and senior policy analyst positions (i.e all candidates within this test condition were applying for similar level positions).

NMT4 Numerical Business Analysis Test. This test measures the ability to interpret and use complex, business-related, numerical information, such as identifying trends across a wide range of data and combining statistical information from different departments to establish new information. The test presents a series of charts, tables and graphs relating to a commercial organisation, and candidates are required to interpret the data and combine the information from different sources in order to answer particular questions.

NMG2 General Numerical Reasoning Test. This test measures the ability to make correct decisions or inferences from numerical data. The tasks and data presented are relevant to a range of management jobs.

Both numerical tests were administered to candidates applying for positions in finance and monitoring/evaluation sections of the organisation (i.e all candidates within this test condition were applying for similar level positions). . Candidates were permitted to use calculators for both tests.

Procedure

The data used in this study were drawn from archival job vacancy files at the subject organisation, in which raw test scores, ethnicity and gender for job applicants were entered in a spreadsheet for analysis. Test administration was standardised within positions, with all candidates for similar jobs sitting the same test(s) and between candidates, where all candidates received standardised instructions and testing conditions.

Analysis

Mean ethnic group differences were computed using the effect size statistic, *d*. The *d* statistic is commonly used and considered the most appropriate means of comparing group differences in test scores (Roth et al., 2001). It represents the standardized difference between group means, and is calculated as the mean ethnic group difference divided by the pooled standard deviation. The primary interpretation of *d* is its magnitude, although statistical significance of group differences can also be interpreted. Effect sizes (*d*) and associated statistics were computed using *Comprehensive Meta-Analysis* software (1998). Computations of effect sizes and confidence intervals were also cross-checked using formula from Hedges and Olkin (1985), producing similar results.

Results and Discussion

Group means and standard deviations for the three cognitive ability tests are reported in Table 2, as well as effect size statistics for group differences. Sizeable and statistically significant ethnic group differences emerged for both the verbal reasoning test (d = .55) and numerical business analysis test (d = 1.79), with Maori mean scores lower than European mean scores for both tests. No difference in mean test scores was apparent for the general numerical reasoning test (d = .01).

In interpreting these differences, it is important to keep in mind that they represent *mean* differences across groups, and that they do not suggest differences for all individuals. In fact, the distributions have considerable overlap, and prior research suggests that differences in cognitive ability test scores between individuals *within* ethnic groups is far greater than differences found *between* ethnic groups (Roth et al., 2001). Hence individuals scoring very high on cognitive ability tests can be found in all ethnic groups.

These findings generally converge with overseas research, indicating that mean ethnic group differences in cognitive ability test scores can be quite large. The implication of such differences for New Zealand organisations is that, if such differences exist within an organisation's pool of applicants, the use of cognitive ability tests for selection decisions could result in employing smaller

Table 2. Test Means, Standard Deviations and Effect Size Statistics for Ethnic Group Differences

Test	Maori			European			Effect Size Statistics for Ethnic Group Differences			
	N	Mean	SD	N	Mean	SD	d	Standard Error	95% Confidence Interval	
Verbal reasoning test	109	19.28	5.35	55	22.2	5.09	.55*	.17	.22	- .89
Numerical business analysis test	19	8.21	3.34	10	14.90	4.43	1.79*	.46	.85	- 2.72
General numerical reasoning test	29	14.72	5.76	17	14.76	5.55	.01	.31	-.61	- .62

* indicates that difference between groups is significant at p < .01

proportions of Maori applicants than European applicants. Furthermore, findings from overseas research suggests that ethnic differences in cognitive ability test scores are associated with similar (though less pronounced) ethnic differences in other selection methods that have a large cognitive component, such as in-basket exercises within assessment centres (Goldstein, Yusko & Nicolopoulos, 2001).

Why a substantial mean difference was found for one numerical test (the numerical business analysis test) but not the other (the general numeric reasoning test) is not entirely clear. The difference in findings for these two tests may be due to differences in the amount of business knowledge and experience required by the tests. The principle difference between the two tests is that the test on which no difference was observed does not assume any business knowledge, while the test on which differences were observed assumes the person has had exposure to business terminology (e.g. net operating profit and gross profit). If Maori in the sample had less exposure to business terminology included in the test on which differences were observed this lack of experience may have accounted for the effect that was observed.[1]

While the present study identified ethnic differences in test scores, these results do not necessarily constitute test *bias*. Test bias is observed when systematic differences are found between ethnic groups not only in mean test scores, but also *in how tests predict job performance ratings*. For example, the verbal analysis test would only be considered *biased* if it was found to differentially predict job performance for Maori and European applicants (e.g., on average, Maori scored lower on the test but performed the job just as effectively as Europeans).

In order for evidence of test bias to be established, additional information is needed on applicants' actual job performance once employed, so that differences in regression slopes and intercepts can be assessed when job performance ratings are regressed on test scores within ethnic groups. Job performance ratings were not available for the present study, and so we were unable to explore evidence of test bias in the present study. While overseas research both in the area of scholastic achievement and job performance indicates that tests of cognitive ability predict equally well across ethnic groups – despite the fact that sizeable mean group differences exist (Roth et al., 2001), we know of no data published to date on test bias in New Zealand. This is an important area in need of future research.

Practical Significance of Mean Test Score Differences

Regardless of whether mean test score differences found in the present study reflect test bias, such differences would still lead to adverse impact for Maori as long as personnel selection decisions are based, even in part, on scores on such tests.[2] If the present findings generalize to other organisations that place considerable weight on cognitive ability test scores in their staff selection processes, Maori applicants, as a group, are less likely than Europeans to be selected.

The practical implications of these findings can be substantial, even if cognitive ability tests are simply used as a screening device, where those applicants who fail to achieve a particular cut-off score are removed from the selection process. For example, consider the case in which the standardized mean difference between Maori and European cognitive test scores is $d = .5$, similar to the ethnic group difference we found for the verbal reasoning test in the present study ($d = .55$), which of the three cognitive tests assessed, resulted in a mean score difference in between the differences found for the other two tests (1.78 and .01). If, with a mean group difference in selection test scores of $d = .5$, a minimum cut-off score is established which allows 50% of the European applicants through, only 30% of the Maori applicants would pass through to the next hurdle of the selection process. Furthermore, as the minimum cut-off score is set to be more stringent, the effect becomes even more pronounced, e.g., if the test cut-off score is set at a point which allows only 10% of the European applicants through, the same mean group difference in test scores ($d = .5$) would result in a pass-rate of only 4% of Maori applicants.[3]

Strategies for Minimizing Adverse Impact

Many occupational psychologists and organisations in New Zealand are concerned that their selection practices should not inadvertently disadvantage Maori, and so these findings present a dilemma for those who use cognitive ability tests in making hiring decisions. Next we consider strategies that have been recommended or used to minimise the adverse impact of cognitive ability tests.

One set of procedures that has been used by organisations (primarily overseas) involve within-group score adjustments. These include (a) providing bonus points on tests based on ethnic group membership; (b) within-group norming (i.e., standardizing individuals' scores, or converting them to percentiles, based on the distribution of scores for the individual's ethnic group); (c) establishing different score cut-offs for different ethnic groups; or (d) filling quotas for positions established for each group using top-down selection (i.e., offering the position/s first to the highest scoring applicant/s) from separate lists of applicants. While none of these strategies maximize predictive validity as effectively as top-down selection from the entire group (i.e., regardless of ethnicity), they all take advantage to some degree of the test's predictive validity while at the same time they can achieve diversity objectives. Such methods, however, are controversial because decisions are based, at least in part, on ethnicity rather than merit, and while the legal implications of such procedures have not been tested in New Zealand, they have been the subject of intense legal debate in the United States, where such forms of score adjustments based on ethnicity have been outlawed in the Civil Rights Act of 1991 (Gottfredson, 1994; Sackett & Wilk, 1994).

A more recently developed strategy employed to arrive at a compromise between maximizing both validity and ethnic diversity when using cognitive ability tests for personnel selection is score banding. Banding involves defining a range of test scores that are treated as statistically equivalent (Scientific Affairs Committee, 1994). Bandwidth is usually either set arbitrarily or based on the standard error of the difference between scores, e.g., scores within 1.65

standard errors of the top score in a band are considered to be equivalent (i.e., not statistically different at the .05 level of significance) from the top score. Hence those applicants falling within a band are considered equivalent on the cognitive ability test and are then chosen based on other characteristics, such as scores on other assessments (e.g., interviews, personality tests, assessment centres, reference checks, or even ethnicity). While banding sacrifices some degree of validity (depending on the distance between scores at the top of the range), it can support the goal achieving a more ethically diverse workplace provided that decisions within each band are based on job-related characteristics for which there are not ethnic group differences. Generally, score banding has become one of the more accepted means of meeting diversity goals while using cognitive ability tests for which mean ethnic group differences exist (Sackett & Wilk, 1994).

Score banding, however, has several serious drawbacks, and has not been universally accepted. Critics argue that banding is logically flawed (Campion, et al., 2001; Schmidt, 1991), in that differences between scores within bands are considered meaningless and thus ignored, but similar or even smaller score differences between bands are treated as meaningful. Similarly, banding has been criticized as being inconsistent with the fundamental premise of occupational testing: namely the optimisation of performance prediction (Schmidt, 1991). The reduction in adverse impact achieved through banding is necessarily at the expense of using the test to most accurately predict performance, and herein lies the dilemma for those using cognitive tests for employee selection in New Zealand. With fairly large mean ethnic group differences, as found in the present study, quite wide bands must be formed in order to achieve diversity objectives, but because band width is associated with reductions in validity and hence utility, the cost-effectiveness of using cognitive ability tests with wide bands becomes questionable.

Another alternative strategy is to use, and place emphasis on, alternative selection methods that tap non-cognitive, job-related constructs. This approach has recently been advocated overseas (Sackett et al., 2001), particularly in light of the controversy surrounding test score banding (Campion et al., 2001). Non-cognitive performance constructs, such as interpersonal skills, organisational citizenship behaviours and team-related behaviours, have been viewed as increasingly important in today's workplace (see, for example, Ilgen & Pulakos, 1999), and these can largely be predicted by non-cognitive selection measures, such as personality tests, biodata instruments, employment interviews, team-based exercises, and reference checks. For example, structured employment interviews have been found to have both high predictive validity (McDaniel, Whetzel, Schmidt & Maurer, 1994; Taylor & Small, 2002; Wiesner & Cronshaw, 1988) and lower levels of ethnic group differences than either unstructured employment interviews or cognitive ability tests (Huffcutt & Roth, 1998). Similarly high levels of validity and low levels of ethnic group differences have been found for structured employment interviews within New Zealand (Gibb & Taylor, in press).

To the extent that non-cognitive aspects of performance are job related, the inclusion of such measures can both increase validity and reduce adverse impact. We note, however, that the present results suggest that, to the extent that measures of cognitive ability – either cognitive ability tests or other measures that have high levels of cognitive saturation – are included in an organisation's selection procedure, adverse impact will not be entirely removed.

Limitations and Conclusions

Limitations of the present study need to be considered when interpreting these findings, and these limitations suggest directions for future research on ethnic group differences on selection tests commonly used in New Zealand. The present study was conducted within a single New Zealand organisation, and so these results should be considered preliminary, and further research is needed in other settings, including other jobs and in private-sector organisations. Such research would create a larger database, providing more stable estimates of group differences, as well the opportunity to determine the extent to which differences vary as a function of job type and organisational setting. Overseas research, for example, suggests that ethnic group differences are even larger for jobs of lower complexity (Roth et al., 2001).

We were unable to obtain sufficient detail from the organisation on shortlisting procedures to preclude the possibility that a short-listing process was employed that resulted in substantial differences in the composition of Maori and European candidates sitting the cognitive tests than Maori and European applicants. For example, if the organisation short-listed a greater proportion of Maori applicants than European applicants using a measure which correlates highly with cognitive test scores (e.g., education level achieved), it is possible that the mean test score differences found in the present study could have resulted, in part, from differences in the short-listing practice, i.e., that such differences may not have existed had short-listing practices been the same for both groups. Future research should include measures of education level to explore this possibility.

The present study was based on three cognitive ability tests, and further research is needed on ethnic group differences in New Zealand using other cognitive tests. Perhaps more fluid measures of cognitive ability that do not involve vocabulary or knowledge of business concepts, such as the Ravens Progressive Matrices, would produce smaller ethnic group differences while also predicting job performance (Kline, 2000), and future research could investigate such measures.

As mentioned earlier, further research is needed which includes measures of job performance, in order to determine whether cognitive ability tests yield biased predictions of job performance. Future research might also include measures of prior education and experience, which could shed light on potential causes of observed differences. Finally, with the growing use of personality tests for personnel selection (Taylor et al., 2002), research is needed on personality test score differences across ethnic groups.

Unfortunately, research on ethnic group differences on occupational tests is difficult in New Zealand because of the relatively small numbers of Maori and European staff within particular job families in single organisations. Stable statistics require relatively large samples, and so meaningful interpretation of such data is virtually impossible within most New Zealand organisations. Consequently, the onus of responsibility for such research must fall on the very large organisations or on those firms that sell psychological tests in New Zealand and who are able to conduct analyses across organisations.

In conclusion, psychological tests and other selection methods play a critical role as gate-keeper for desirable positions in organisations, and thus their use has important social consequences to both organisations and individuals. Understanding how such tests function in this gate-keeping role is necessary for making informed decisions about which methods are used, how they are scored, and how they should contribute to selection decisions. We hope that further research continues on this important topic.

References

Comprehensive meta-analysis (Version 1.023) [Computer software]. (1998). Englewood, NJ: Biostat.

Campion, , M.A., Outtz, J.L., Zedeck, S., Schmidt, F.L., Kehoe, J.F., Murphy, K.R., & Guion, R. (2001). The controversy over score banding in personnel selection: Answers to 10 key questions. *Personnel Psychology, 54,* 149-185.

Gibb, J.L., & Taylor, P.J. (in press). Past experience versus situational employment interview questions in a New Zealand social service agency. *Asia-Pacific Journal of Human Resources.*

Goldstein, H.W., Yusko, K.P., & Nicolopoulos, V. (2001). Exploring Black-White subgroup differences on managerial competencies. *Personnel Psychology, 54,* 783-807.

Gottfredson, L, S. (1994). The science and politics of race norming. *American Psychologist, 49,* 955-963.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, Fl: Academic Press.

Hermelin, E., & Robertson, I.T. (2001). A critique and standardization of meta-analytic validity coefficients in personnel selection. *Journal of Occupational and Organizational Psychology, 74,* 253-277.

Huffcutt, A. I., & Roth, P. L. (1998). Racial group differences in employment interview evaluations. *Journal of Applied Psychology, 83,* 179-189.

Hunter J.E., & Hunter R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Ilgen, D.R., & Pulakos, E.D. (Eds.). (1999). *The changing nature of performance: Implications for staffing, motivation, and development.* San Francisco: Jossey-Bass.

Kline, P. (2000). *Handbook of Psychological Testing* (2nd ed.). England: Routledge.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79,* 599-616.

Robertson, I. T., & Smith, M. (2001). Personnel selection. *Journal of Occupation and Organizational Psychology, 74,* 441-472.

Roth, P.L., Bevier, C.A., Bobko, P., Switzer, F.S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology, 54,* 297-330.

Ryan, A.M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology, 52,* 359-392.

Sackett, P, R. Schmitt, N. Ellingson, J, E., & Kabin, M, B. (2001). High-stakes testing in employment, credentialing, and higher education - Prospects in a post-affirmative action world. *American Psychologist, 56,* 302-318.

Sackett, P. R., & Wilk, S. l. (1994). Within-group norming and other forms of score adjustment in pre-employment testing. *American Psychologist, 49,* 929-954.

Scientific Affairs Committee (1994). An evaluation of banding methods in personnel selection. *The Industrial-Organisational Psychologist, 49,* 80-86.

Schmitt, F, l. & Hunter, J, E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124,* 262-274.

Schmidt, F.L. (1991). Why all banding procedures in personnel selection are logically flawed. *Human Performance, 4,* 265-277.

Taylor, P.J., Keelty, Y., & McDonnell, B. (2002). Evolving personnel selection practices in New Zealand organisations and recruitment firms. *New Zealand Journal of Psychology, 31,* 8-18.

Taylor, P., Mills, A., & O'Driscoll, M. (1993). Personnel selection methods used by New Zealand organisations and personnel consulting firms. *New Zealand journal of Psychology, 22,* 19-31.

Taylor, P. J., & Small, B. (2002). Asking applicants what they would do versus what they did do: A meta-analytic comparison of situational and past behaviour employment interview questions. *Journal of Occupation and Organizational Psychology, 75,* 277-294.

Wiesner, W., & Cronshaw, S. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology, 61,* 275-290.

Notes:

1. We were unable to determine the particular job analysis approach used by the organisation and test seller in order to select the most appropriate battery of tests for the family of jobs, i.e., whether the numeric test involving business knowledge is job-related. Given the sizeable ethnic group differences on this test, establishing the job-relevance of such a test would have become important.
2. We were unable to obtain details of how test scores were used in selection decisions within the particular organisation from which these data were collected.
3. These estimates were based on figures presented by Sackett & Wilk (1994).

Address for correspondence:
Nigel Guenole
PO Box 3218, Shortland Street
Auckland, New Zealand
Email: n.guenole@selectorgroup.com