×× Extremely poor on this standard. Should not be used until this is addressed
× Does not meet the standard.
● Appears to meet the standard but more information is neededl
✓ Meets the standard
✓✓ Exceeds the standard

Function	Standard	Findings	Rating
General	Background	The Wonderlic was first published in 1937, being developed by Charles Spearman as a measure of 'g' – a general intelligence factor that was thought to be heritable and unchangeable.	
Reliability	Reliability tells us how much error there is in measurement in the assessment. The more error, the less accurate it will be in measuring job performance. Standards are: • Does the internal consistency or split half reliability equal 0.75 or above (0.7 acceptable)? • Does the test re-test reliability equal .75 or above (0.7 acceptable).	Reliability is sound with Test-retest co-efficients of between 0.82 and 0.94 and alternate form reliabilities between 0.73 – 0.95.	✓✓
Validity	**Construct Validity** There are three distinct factors to construct validity: 1) Is there a good theoretical reason for why the construct should predict job performance? 2) Is the test delivering similar results to other tests also measuring that construct? and 3) Is the test really measuring everything it says it is measuring?	The underlying assumption of the assessment, that there is a general intelligence, has been widely disproven. More recent research has found eight independent intelligences, that can be reliably measured and predict specific aspects of people's performance.	×

Function	Standard	Findings	Rating
	• Has a sound theoretical framework, related to job performance • Correlates 0.6 with similar constructs in other measures • Is the structure verified through factor analysis	The Wonderlic appears to combine items that measure both fluid and crystallised intelligence. There is no independent analysis of these items to give an indication of what may be useful for different contexts – eg measure job performance or ability to learn. The test has sound correlations with other measures of general intelligence, ranging between 0.72 – 0.92.	
	Face Validity Face validity is a qualitative assessment, of whether the content looks relevant to the job. For example, are the types of questions being asked likely to be things the candidate has to deal with on the job, or are they of a personal, or non-occupational nature. The items in the test should comply with NZ legislation, especially the Human Rights Act and Privacy Act. These Acts stipulate that questions should not be asked which the candidate could see as discriminatory or irrelevant, even if they are not being used in that way. Even if not specifically illegal under the above Acts, the NZ Psychologists Ethics stipulate that unnecessarily intrusive questions should not be asked if there is a less intrusive way of equal validity which could be used.	The assessment has low face validity in a work context with a number of the items referring to proverbs and general knowledge which are related to a specific cultural context and are not work related. The assessment has items from a range of abilities, such as math, spatial and verbal areas. If these are not required for a role being applied for – for example spatial skills for an HR role – then applicants would likely see these as being irrelevant for the role.	✗
	Criterion Related Validity Criterion-related validity provides a measure of whether the assessment is measuring job relevant criteria. Tests that have no proven relationship to job performance or behaviour add no value to decision making processes. Legislation requires that employment decisions must be	Extensive validity evidence is available and shows sound validity coefficients, and range between 0.22 and 0.67. There are no published NZ studies.	✓✓

Function	Standard	Findings	Rating
	based on job relevant criteria – criterion related validity establishes this relevance. Standards are: • Two groups of at 50 people • Statistically significant results • New Zealand sample • Minimum of 0.3		
Norms	**Occupational** Occupationally relevant comparison groups enable the link to be made between assessment and job performance. Without a relevant comparison group, it is impossible to fairly or accurately gauge the level of a candidate's performance.	Occupational norms are available.	✓
	New Zealand Relevant NZ reference groups should also be available. Optimal sample sizes are dependent on the reliability of the assessment (the less reliable the greater sample size required). Standards are: • Sample of more than 100 • Occupational • Representative of the groups being used with	Published New Zealand occupational norm groups are not available.	✗✗
Diversity and inclusion	Information should be available to establish whether the test or questionnaire works for Maori, different age groups, genders and other groups as per the Human Rights Act. A test or measure should not produce large (greater than one standard error of measurement) systematic differences between different groups of people (on the basis of age, gender, or ethnicity). If these differences in test performance are also reflected in different job	As mentioned, the Wonderlic appears to combine items that measure both fluid and crystallised intelligence. Crystallised intelligence is tapping to knowledge that the applicant has gained, and to the extent that the knowledge is from a particular cultural context then this can exclude people who are not from the context. For example, the test asks people what RSVP means, what particular words mean (e.g. fall as in Autumn) and	✗✗

Function	Standard	Findings	Rating
	performance, then this can be justified. However, if the test differences are not reflected in job performance then variations greater than one standard error of measurement would result in real interpretation differences between the two groups which were not justified. Separate norm groups should be used where differences in test performance are not reflected in job performance. Where differences do occur, information should be provided as to their nature and magnitude. Standards are: • Information is available • Appropriateness for Maori established. • Differences less than 1 standard error of measurement	vacation - as in holiday), English sentence structure (what a verb is), the meaning of common English proverbs (Like father, like son), and what P.M. stands for. Perhaps related to the knowledge being measured in the assessment the Wonderlic manual report significant differences in test scores between 'White' (22.8), African (16.2) and Hispanic Americans (17.26). The popularity of the Wonderlic declined in the 1970's after several large law suites demonstrated adverse impact. Conversion tables are now available to correct for bias in these groups, however it is not apparent how these would be used in NZ. The test measures a range of verbal, math and spatial skills. Given that on average, males tend to perform better on spatial assessments, if spatial skills are not required for the job then this would be considered adverse impact. The test has 50 items to be completed in 12 minutes. This is considered highly speeded compared with other measures of ability where you expect between 20% and 80% of the candidates to complete the assessment (less than 1% of the WPT candidates attempt any of the last ten questions). Research indicates older applicants are not as fast as graduates, and respond better to slower more power orientated tests. Age adjustments are available for the test. The speed required would also discriminate against those whose first language is not English.	

Function	Standard	Findings	Rating
		The assessment has been criticised for its small print – 50 items on two sides of an A4, although the computerised version does not have this problem. People with sight impairment may struggle with the paper and pencil version of the assessment.	
		9th, 13th, and 14th Mental Measurements Yearbooks Eatwell, J.A.R and Wilson, I. The Effective Use of Psychometric Assessments in Decision Making. Published In W.W. Waitoki, J.S. Feather, N.R. Robertson, & J.J. Rucklidge (Eds.). (2016) *Professional Practice of Psychology in Aotearoa New Zealand*. 3rd edition. Wellington, NZ: New Zealand Psychology Society.	