Maximally efficient two-stage screening: Determining intellectual disability in Taiwanese military conscripts

Chia-Chang Chien[1]
Shu-Fen Huang[1,2,3,4]
For-Wey Lung[1,2,3,4]

[1]Department of Psychiatry, Kaohsiung Armed Forces General Hospital, Kaohsiung, Taiwan; [2]Graduate Institute of Behavioral Sciences, Kaohsiung Medical University, Kaohsiung, Taiwan; [3]Department of Psychiatry, National Defense Medical Center, Taipei, Taiwan; [4]Calo Psychiatric Center, Pingtung County, Taiwan

Correspondence: For-Wey Lung
Department of Psychiatry, Kaohsiung Armed Forces General Hospital,
No 2 Chung Cheng 1st Rd.,
Kaohsiung City, Taiwan
Tel +886 7 749 0056
Fax +886 7 749 3767
Email forwey@seed.net.tw

Objective: The purpose of this study was to apply a two-stage screening method for the large-scale intelligence screening of military conscripts.

Methods: We collected 99 conscripted soldiers whose educational levels were senior high school level or lower to be the participants. Every participant was required to take the Wisconsin Card Sorting Test (WCST) and the Wechsler Adult Intelligence Scale-Revised (WAIS-R) assessments.

Results: Logistic regression analysis showed the conceptual level responses (CLR) index of the WCST was the most significant index for determining intellectual disability (ID; FIQ $\leq$ 84). We used the receiver operating characteristic curve to determine the optimum cut-off point of CLR. The optimum one cut-off point of CLR was 66; the two cut-off points were 49 and 66. Comparing the two-stage window screening with the two-stage positive screening, the area under the curve and the positive predictive value increased. Moreover, the cost of the two-stage window screening decreased by 59%.

Conclusion: The two-stage window screening is more accurate and economical than the two-stage positive screening. Our results provide an example for the use of two-stage screening and the possibility of the WCST to replace WAIS-R in large-scale screenings for ID in the future.

Keywords: intellectual disability, intelligence screening, two-stage positive screening, Wisconsin Card Sorting Test, Wechsler Adult Intelligence Scale-Revised

Introduction

The prevalence of intellectual disability (ID) is increasing globally, especially in less developed countries.[1] Most of the maladaptive behaviors of ID appear before the person reaches 18 years of age.[2] The comorbidities of ID are associated with many kinds of diseases[3–5] and a varying number of people with ID present at police stations.[6,7] Military conscription is mandated in Taiwan, thus an economical and accurate means of screening men who are not competent for the military is vital. In fact, among various factors related to learning and adjustment, the weighting of intelligence is most considerable. People with ID may have poor communication skills, and also lack or be deficient in self-care, social skills, home living, community use, self-direction, health and safety, functional academics, leisure, and work.[8] Intelligence as a concept has many different definitions and measuring tools.[9] According to Wechsler's definition,[10] intelligence is interpreted as "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment". The Wechsler Adult Intelligence Scale-Revised (WAIS-R) is a widely used intelligence test used to measure the intelligence of those between 16 and 74 years of age. Although the reliability and validity of the WAIS-R are quite satisfactory,[9] the test is not economical for mass screening in military settings to identify intellectually impaired recruits.

On the other hand, the Wisconsin Card Sorting Test (WCST) is a neuropsychological test that is used generally as a measure of abstract reasoning ability and of the ability to

utilize environmental feedback to shift cognitive strategies.[11] The computerized WCST has been developed and is administered automatically, including stimuli-presenting and result-scoring, and the examiner only needs to monitor the assessment process. Generally, the computerized WCST is less expensive and requires less time than the WAIS-R and is much easier to administer and interpret.

A two-stage screening process entails using a more inexpensive, less invasive test at the first stage (eg, computerized WCST), and a more expensive or more invasive, but more sensitive and specific test at the second stage (eg, WAIS-R).[12] The two-stage screening can be subtyped into two-stage positive and two-stage negative screening. Two-stage positive screening implies that only those who screened positive on the first test will be scheduled for a later assessment, which is more expensive or more invasive. Participants who screen positive in both stages will be referred for further intervention programs (Figure 1a). Two-stage negative screening picks out more participants and is used for more life-threatening disorders (eg, acquired immunodeficiency syndrome). Participants who screen positive in the first stage will be directly referred for treatment. Additionally, participants who screen negative in the first stage need to be screened in the second stage. Those who screen positive in the second stage will also be referred for further treatment or exempted from military service in this study (Figure 1b).[13]

We combined the concepts of two-stage positive and two-stage negative screening in this study to create a new screening approach called the "two-stage window screening" (Figure 2). In this approach, we selected two cutoff points from a receiver operating characteristic (ROC) curve with maximized area under the curve (AUC), one in the higher and one in the lower area. We assumed those scoring higher than the higher cutoff point as having normal intelligence, and those scoring lower than the lower cutoff point as intellectually disabled; the only group that needs to be further tested is the group between the two cutoff points. This group has a higher probability of being misclassified. Using this method of screening would decrease the number of people who need to test in the second stage. Previous investigations have shown that when the prevalence of a disease was enhanced, the positive predictive value (PPV; the proportion of true positives among individuals who tested positive) would be increased. When a disease has a higher prevalence, the screening instrument will show a better accuracy in predicting the disease.[12] Other than increasing the accuracy of prediction, decreasing the amount of people needed to test and will further decrease the cost of testing.

In Taiwan, approximately 150,000 young men are conscripted into the military every year, for a two-year enlistment. In the past, the Symptom Checklist (SCL-90-R)[14,15] was used as a screening instrument for all military conscripts, but it was costly and inefficient. At present, only those who have shown maladjusted behavior are brought to the psychiatrist, and further assessed by psychologists using the WAIS-R. The purpose of this study was to investigate whether the application of two-stage window screening in large-scale screening of military conscripts for ID, using the WCST in the first stage and the WAIS-R in the second stage, would be more accurate and cost-efficient than the two-stage positive screening method.

Designs and methods
Participants

The protocol of this study was approved by the committee of research ethics (the Institutional Review Board) at Kaohsiung Armed Forces General Hospital in southern Taiwan. We collected 99 conscripted soldiers from the largest branch of the military in southern Taiwan, whose educational levels were senior high school level or lower, to be the participants. In Taiwan, only males were conscripted, thus all participants were male between the ages of 20 and 24 (M = 20.24, SD = 0.624). After detailed explanation regarding the purpose of the study, the participants gave their consent. Subsequently, all participants completed
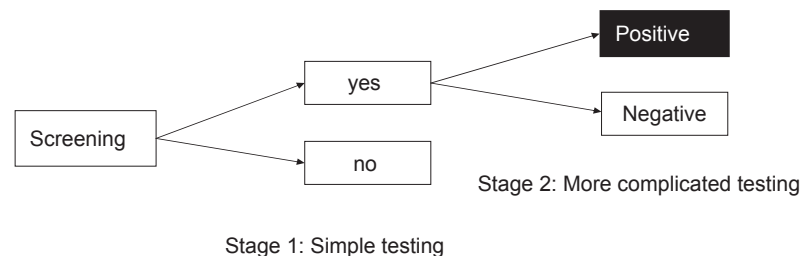

Figure 1A Two-stage positive screening.
Note: Black box indicates the group of participants referred for further intervention.

Stage 1: Simple testing
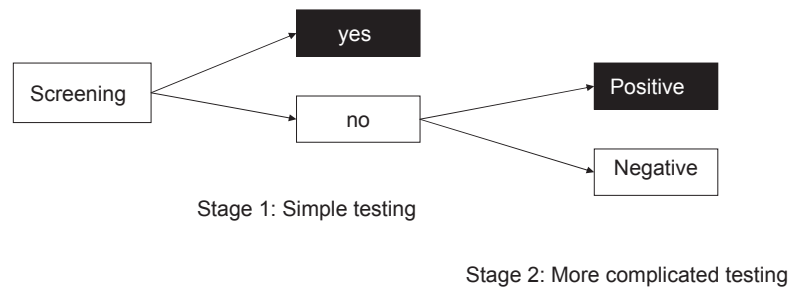
Stage 2: More complicated testing

Figure 1B Two-stage negative screening.
Note: Black boxes indicate groups of participants referred for further intervention.

the WAIS-R and the computerized WCST assessment. No participant had any clinical evidence of psychiatric disorders according to the Chinese versions of the Structured Clinical Interview for DSM-III-R (SCID).[16]

Instruments

Wechsler Adult Intelligence Scale-Revised
WAIS-R has been widely used in Taiwan in recent years to measure the intelligence of those from 16 to 74 years of age. It comprises 11 subtests: six subtests on verbal intelligence and five subtests on performance intelligence.[17] This instrument was used as the gold standard in this study. The WAIS-R generally takes 60 to 90 minutes to administer.[17]

Wisconsin Card Sorting Test
The computerized WCST was adopted for this study, this computerized version was developed based on the standardized criteria of Heaton and colleagues[11] by Tien and colleagues in 1996 using Turbo Basic 1.1.[18] The assessment illustrations comprise trends in speech sound and written Chinese characters. The computerized system can be adjusted by the test administrator based on the conditions of the testee. The participants were given a pack of 128 response cards and told to match each response card to one of the stimulus cards. After matching 10 cards using the first sorting category (color), the participants shifted to the second category (form), and then to the third (number). The procedure was repeated twice, until the participants completed six categories, or until all 128 response cards were used. The WCST index consists of the following: percent preservative errors (PPE), trials to complete the first category, failure to maintain the set (FMS), learning to learn, categories achieved, and conceptual level responses (CLR). Higher scores in the first three indices indicate a worse performance. In contrast, higher scores in
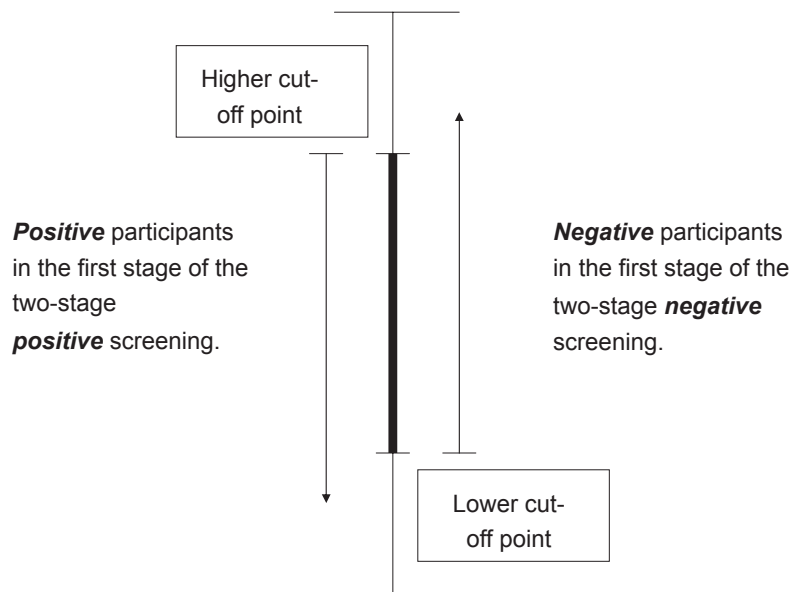

Figure 2 Two-stage window screening.
Note: Participants whose scores were in the thickened area have to be tested in the second stage.

last three indices indicate a better performance. Generally, the manual WCST takes 20 to 30 minutes to administer,[11] administered by those with psychological assessment training, the computerized version provides a more reliable and administration, data acquisition accuracy and scoring than the manual version.[18]

Statistical analysis

Data were analyzed with SPSS for Windows (version 15.0; SPSS Inc., Chicago, IL) and a computer program for accuracy analysis.[19] According to the DSM-IV definition, if the FIQ of the individual is equal to or lower than 70, the individual is defined as having ID; if the FIQ of the individual is between 71 and 84, then the individual is defined as having borderline intelligence.[20] Hence, this study used the results of the WAIS-R as the "gold standard" to estimate whether the testee had ID. The individual would be placed in the ID group if the FIQ was equal to or lower than 84, and in the normal group if the FIQ was greater than 84.

After using multiple logistic regression analysis to verify the indexes of the WCST that could be used to predict ID, we used ROC analysis to determine the optimum cutoff point of the index. ROC analysis is a nonparametric approach that was first developed in psychophysics.[21] This method of analysis described a function that summarizes all possible performances of a case faced with the task of detecting a signal from noise under a curve.[22] The principle of ROC analysis has been applied to screening psychiatric illness.[23–25] The area under the ROC curve (AUC) could be regarded as the probability of correct prediction.[26,27] We first selected the largest AUC value to be the optimum cutoff point of the predictive index, and the point was set as the basic (or fixed) point. We tried to compute the AUC of the basic and other points of the predictive index. The largest AUC value would be the optimum cutoff point of the predictive index.

Other than that, we also compared the sensitivity (the ability of a test to correctly identify those who have the characteristic of interest), specificity (the ability of a test to correctly identify those who do not have the characteristic of interest), PPV (the proportion of true positives among those who test positive), negative predictive value (NPV; proportion of true negatives among those who test negative),[28] AUC (probability of correct prediction),[26,27] and assessment expense of the two-stage positive screening and two-stage window screening (each WAIS-R assessment costs US$21.50).

Results

The distribution of the WAIS-R full IQ and WCST scores and the correlation among each sub-test is presented in Table 1. As assessed with the WAIS-R, 38 (38.4%) out of 99 participants had a FIQ lower than or equal to 84, which means they were intellectually disabled. Further, multiple logistic regression analysis results showed that the CLR index in the WCST is a valid predictor for ID ($\beta = -0.047$, $p = 0.000$) (Table 2), indicating that a lower CLR score would result in a higher possibility of ID.

The optimum cutoff point of the CLR index obtained from the ROC curve was 66. When a CLR index of 66 as the cutoff point was used in two-stage positive screening, 61 were determined to be in the intellectual disability group and 38 were in the normal intelligence group (Table 3). This group of 61 persons with positive result for ID will be screened in the second stage, and further tested with the WAIS-R. When this method was used, the results indicated that the AUC was 0.71, sensitivity was 0.87, specificity was 0.54, PPV was 0.54, and NPV was 0.87.

The two optimum cutoff points of the CLR index obtained from the ROC curve were 66 and 49 (AUC = 0.74). When the two-stage window screening method was used, with the CLR index and two cutoff points of 66 and 49, it was found

Table 1 The distribution of the WAIS-R and WCST scores, and the correlation among each subtest

Correlation	Range	M (SD)	FIQ	CA	PPE	TCFC	CLR	FMS	LL
FIQ	47–122	88.57 (13.67)		0.454**	−0.469**	−0.209**	0.551**	−0.079**	0.273**
CA	0–6	4.48 (1.88)			−0.696**	−0.547**	0.904**	−0.084	0.218**
PPE	5.71–57.03	15.48 (8.29)				0.439**	−0.737**	−0.070	−0.325**
TCFC	10–102	24 (20.88)					−0.584**	−0.021	0.435**
CLR	3.13–91.89	56.74 (20.75)						0.088	0.162
FMS	0–5	1.19 (1.18)							−0.046
LL	−28.28–16.30	0.62 (6.90)							

Notes: *$p < 0.05$; **$p < 0.001$.
Abbreviations: CA, categories achieved; CLR, conceptual-level responses; FIQ, WAIS-R Full IQ; FMS, failure to maintain the set; LL, learning to learn; PPE, percent preservative errors; SD, standard deviation; TCFC, trials to complete the first category; WAIS-R, Wechsler Adult Intelligence Scale-Revised; WCST, Wisconsin Card Sorting Test.

Table 2 Results of multiple logistic regression analysis of the indexes of the WCST and ID

The WCST index	β	Standard error	p value
CLR	−0.047	0.012	0.000**
Constant	2.123	0.709	0.003**

Notes: *p < 0.05; **p < 0.001
Abbreviations: CLR, conceptual level responses; ID, intellectual disability; WCST, Wisconsin Card Sorting Test.

Table 3 Results of using one cut-off and two cut-off points in each group showing intellectual status (Unit: person)

		Total (%)
One cut-off point	CLR > 66	38 (38.4%)
	CLR ≤ 66	61 (61.6%)
Two cut-off points	CLR > 66	38 (38.4%)
	CLR = 49–66	25 (25.3%)
	CLR < 49	36 (36.4%)

Abbreviation: CLR, conceptual level responses.

that 25 participants had CLR scores between the two cutoff points (Table 3). Eleven of the 25 participants, whose FIQ with the WAIS-R was equal to or lower than 84, could have been true-positives in the two-stage window screening, and the others, whose FIQ WAIS-R results were higher than 84, could have been true-negatives. After re-setting the classification, the AUC was 0.82, sensitivity was 0.87, specificity was 0.77, PPV was 0.70, and NPV was 0.90.

When the two-stage positive screening was compared with the novel screening method, AUC, sensitivity, specificity, PPV, and NPV were found to be increased in the novel screening method. Moreover, the cost of the second-stage screening using the WAIS-R assessment decreased from US$1311.50 to US$537.50 (Table 4).

Discussion

Generally, those conscripts who have moderate, severe, and profound ID can be determined before they enlist in the military. Mild and borderline intellectual disability is not easily perceived during the health examination; so those conscripts can usually enter the military without a hitch. Previous investigations have shown that when the prevalence of a disease is enhanced, the PPV value is increased. When a disease has a higher prevalence, the screening instrument will show a better accuracy in predicting the disease.[12] Hence, the present study selected those participants whose educational

levels were on par or lower than the senior high school level, to increase the prevalence of ID. In this study, we found that the AUC and PPV of two-stage positive screening were larger than those of two-stage window screening. As a result, if we could increase the cutoff points of the CLR, it would contribute to a better accuracy in classification or a higher probability of prediction. Moreover, two-stage window screening would be more accurate and economical than two-stage positive screening for predicting ID. Two-stage window screening can save approximately 59% money than the two-stage positive screening method (1311.5/537.5 US$ = 0.41). Since the participants were sampled from the largest branch of the military in southern Taiwan, the results of this study can be generalized to military sites throughout Taiwan.

One limitation of the study is that the screening procedure ignores the cultural, ethnic, socioeconomic, and adjustment factors of these conscripts,[8,29] which have been proposed by the American Association on Mental Retardation (AAMR) as important factors in the definition of ID. This proposal was made because ID can disappear as the person enters society, if the initial diagnosis did not adequately consider adaptive function. Another limitation is the small sample size of the study, since only 99 conscripts were recruited,

Table 4 Comparison of results using two-stage positive screening and two-stage window screening

	Two-stage positive screening using the one cutoff point of CLR = 66	Two-stage window screening using the two cutoff points of CLR = 66 and CLR = 49
ROC area (AUC)	0.71	0.82
Sensitivity	0.87	0.87
Specificity	0.54	0.77
PPV	0.54	0.70
NPV	0.87	0.90
Total cost of WAIS-R (USD)	1311.50 (100%)	537.50 (41%)

Abbreviations: AUC, area under the curve; NPV, negative predictive value; ROC, receiver operating characteristic; PPV, positive predictive value; WAIS-R, Wechsler Adult Intelligence Scale-Revised.

the ROC-curve analysis can only describe a tendency, thus a larger scale research is needed to verify whether WCST can be used to screen the intelligence of military conscripts using two-stage window screening in the future.

The algorithm that we propose for further studies is described as follows: using nomination techniques, cadres can determine problem individuals through initial observation (ie, copying skills and adjustment abilities), and these individuals can be sent to the psychiatrist afterward. During the interview process, they will be required to take the WCST. The prevalence of ID will be higher than the random sampling in the initial observation of the military cadres; moreover, the PPV value of the WCST to predict ID will also increase. If the WCST is verified to be accurate and valid to predict ID, it can be used as a first screening in the testing process for ID. The WCST is easier to operate and time saving in estimating ID, and can be used for large-scale screening for ID in the military in future.

Conclusion

Using the two-stage window screening, the AUC and the PPV increased. Moreover, two-stage window screening is more accurate and economical than two-stage positive screening. Our results provide a foundation for the use of two-stage screening and the WCST to replace the WAIS-R in large-scale screenings for intellectual disability.

Acknowledgements

References

1. World Health Organization. *Healthy aging: adults with intellectual disability*. Canberra: Australian Institute of Health and Welfare; 2000.
2. Wen X. *The definition and prevalence of intellectual disability.* Canberra: Australian Institute of Health and Welfar; 1997.
3. Hogg J, Northfield J, Turnbull J. *Cancer and people with learning disabilities: The evidence from published studies and experiences from cancer services.* Kidderminster: Bild Publications; 2001.
4. Splunder J, Stilma JS, Bernsen RMD, Evenhuis HM. Prevalence of visual impairment in adults with intellectual disabilities in the Netherlands: cross-sectional study. *Eye.* 2006;20:1004–1010.
5. Sullivan SG, Hussain R, Threlfall T, Bittles AH. The incidence of cancer in people with intellectual disabilities. *Cancer Causes Control.* 2004;15:1201–1205.
6. Gudjonsson GH, Clare IC, Rutter S, Pearse J. *Persons at risk during interviews in police custody: The identification of vulnerabilities. The Royal Commission on Criminal Justice. Research Study 12.* London: HMSO; 1993.
7. Murphy G, Harnett H, Holland AJ. A survey of intellectual disabilities among men on remand in prison. *Ment Handicap Res.* 1995;8:81–91.
8. Luckasson R, Coulter DL, Polloway EA, Reiss S, Schalock RL, Snell ME, Spitalnick DM, Stark JA. *Mental retardation: Definition, classification, and systems of supports (9th ed.)*. Washington, DC: American Association on Mental Retardation, 1992.
9. Gregory RJ. *Psychological Testing: History, Principles, and Applications (2nd ed.)*. Boston, MA: Allyn and Bacon; 1992.
10. Wechsler D. *The measurement of adult intelligence.* Baltimore, MD: William and Wikins; 1939.
11. Heaton RK, Chelune GJ, Talley JL, Kay GG, Curtiss G. Administration and scoring. In: Heaton RK, editors. *Wisconsin Card Sorting Test Manual: revised and expanded.* Odessa: Psychological Assessment Resources Inc; 1993. p. 5–20.
12. Gordis L. *Epidemiology*. Philadelphia, PA: W. B. Saunders; 1996.
13. Glascoe FP, Foster EM, Wolraich ML. An economic analysis of developmental detection methods. *Pediatrics.* 1997;99:830–837.
14. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale-preliminary report. *Psychopharmacol Bull.* 1973;9:13–28.
15. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. *Psychol Med.* 1983;13:595–605.
16. Chen YC, Ku CS. *Instruction Manual for the Structured Clinical Interview for DSM-III-R (SCID)*. Kaohsiung, Taiwan: Kaohsiung Medical College; 1993.
17. Wechsler D. *Wechsler Adult Intelligence Scale-Revised*. San Antonio, Taxes: The Psychological Corporation; 1981.
18. Tien AY, Spevack TV, Jones DW, Pearlson GD, Schlaepfer TE, Strauss ME. Computerized Wisconsin card sorting test: comparison with manual administration. *Kaohsiung J Med Sci.* 1996;12:479–485.
19. Swets JA, Pickett RM. *Evaluation of Diagnostic Systems: Method from Signal Detection Theory*. New York: Academic Press; 1982.
20. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorder, 4th ed*. Washington, DC: American Psychiatric Association; 1994.
21. Swets JA. *Signal Detection and Recognition by Human Observers*. New York, NY: John Wiley and Sons; 1964.
22. Egan JP. *Signal Detection Theory and ROC Analysis*. New York, NY: Academic Press; 1975.
23. Mari JJ, Williams P. A comparison of the validity of two psychiatric screening questionnaires (CHQ-12 and SRQ-12) in Brazil, using relative operating characteristics (ROC) analysis. *Psychol Med.* 1985;15:651–659.
24. Murphy JM, Berwick DM, Weinstein MC, Borus JF, Budman SH, Klerman G. Performance of screening and diagnostic tests. *Arch Gen Psychiatry.* 1987;44:550–555.
25. Lirio RB, Donderiz IC, Perez Abalo MC. Maximum likelihood estimation of signal detection model parameters for the assessment of two-stage diagnostic strategies. *Int J Biomed Comput.* 1992;31:117–126.
26. Grunkemeier GL, Jin R. Receiver operating characteristic curve analysis of clinical risk models. *Ann Thorac Surg.* 2001;72:323–326.
27. Swets JA. ROC analysis applied to the evaluation of medical imaging techniques. *Invest Radiol.* 1979;14:365–377.
28. Szklo M, Nieto FJ. *Epidemiology beyond the basics*. Sudbury, MA: Jones and Bartlett Publishers; 2004.
29. Mason J, Murphy G. People with an intellectual disability in the criminal justice system: Developing an assessment tool for measuring prevalence. *Br J Clin Psychol.* 2002;41:315–320.