SUPPORT DOCUMENT

Skill matches to job requirements: Support document

CHRIS RYAN AND MATHIAS SINNING

AUSTRALIAN NATIONAL UNIVERSITY

This document was produced by the authors based on their research for the report *Skill matches to job requirements*, and is an added resource for further information. The report is available on NCVER's website: <http://www.ncver.edu.au>

Contents

Tables and figures

Tables

Figures

Development of scales

Measures of job tasks

Measures of job tasks reflect reports by individuals of the frequency with which they undertook literacy and numeracy tasks at work. Respondents in both surveys were asked a partially overlapping set of questions about the literacy and numeracy tasks they undertook at work. These included, for example, how often they wrote "reports or articles", or "letters or memos", or how often they filled in forms such as "bills, invoices or budgets", or how often they calculated "prices, costs or budgets". Appendix 2 includes a description of the job task measures and lists the underlying items that were used to construct them.

Empirical Approach

Since the job task measures may not be observed directly, they have to be measured indirectly through their effects on items that are observed in the data. In order to model the relationship between the item responses and the unobserved measures, statistical models of the item response theory (IRT) may be applied.

The most common IRT model is the Rasch model (Rasch 1960, 1961), which may be used for items with dichotomous responses (e.g. right or wrong answer). In the Rasch model, the probability of a positive response is modelled as a function of an item parameter and a person parameter. While the item parameter represents the difficulty of an item, the person parameter reflects the person's magnitude of the unobserved measure. The probability of a positive response is typically modelled as a logistic function of the difference between person and item parameter. The Rasch model places persons and items on a common scale. While the probability of a correct response increases with the ability of a person, it decreases with the item difficulty and the probability of a correct response is 0.5 when a person's ability is equal to the difficulty of the item.

Since the available items that may be used to create the job task measures include ordered response categories rather than dichotomous responses, it was necessary to consider an extension of the Rasch model that may be applied to ordered response categories. The partial-credit model (Masters 1982) represents such a model. It specifies the response probability as a function of the person ability and a step parameter associated with each category of a certain item.

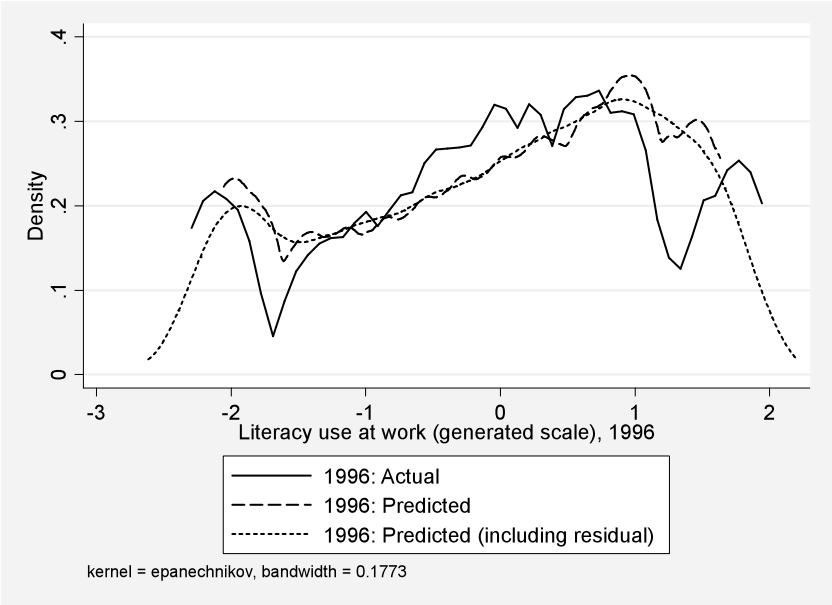
The partial-credit model for ordinal items can be placed within the generalized linear latent and mixed modelling (GLLAMM) framework and fitted by using Stata® (see Rabe-Hesketh, Skrondal & Pickles 2004; Zheng & Rabe-Hesketh 2007). After estimating the parameters of the partial-credit model, the expected a posteriori scores may be derived for each individual. These scores are used to obtain the job task measures, which are rescaled to take on values between 0 and 500.

Derived scales

The following figures present the actual and predicted scales of the generated job task measures. The functions were smoothed by regression-based multiple imputation (Rubin 1987). Specifically,

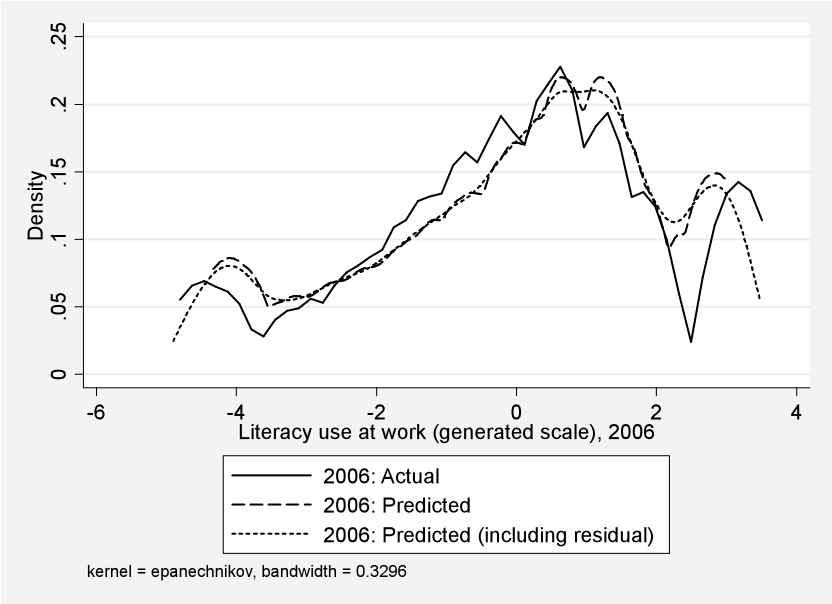
five randomly drawn residuals were used to obtain five different imputations, referred to as "implicates". Since implicates do not differ substantially from one another, each figure includes only one of them. To compare predicted and actual scales, the figures include the predicted scales before they were rescaled to take on values between 0 and 500.

Figure 1 Literacy use at work, actual and predicted scales, 1996



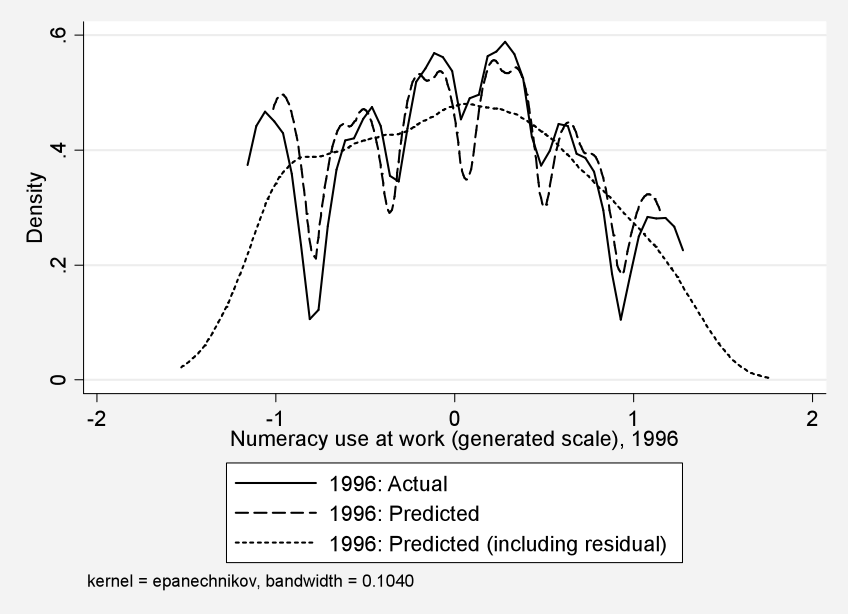
Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0.

Figure 2 Literacy use at work, actual and predicted scales, 2006



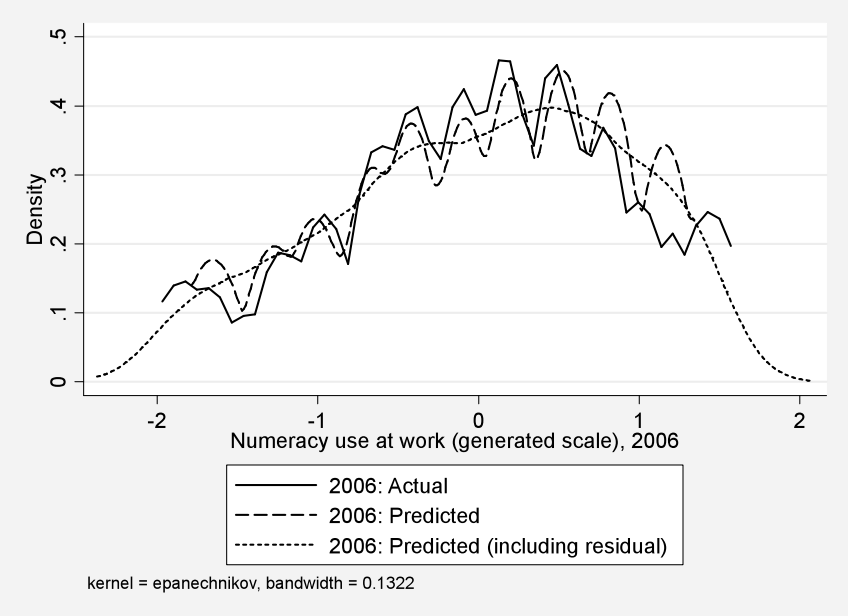
Source: ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

Figure 3 Numeracy use at work, actual and predicted scales, 1996



Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0.

Figure 4 Numeracy use at work, actual and predicted scales, 2006



Source: ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

Individual literacy measures

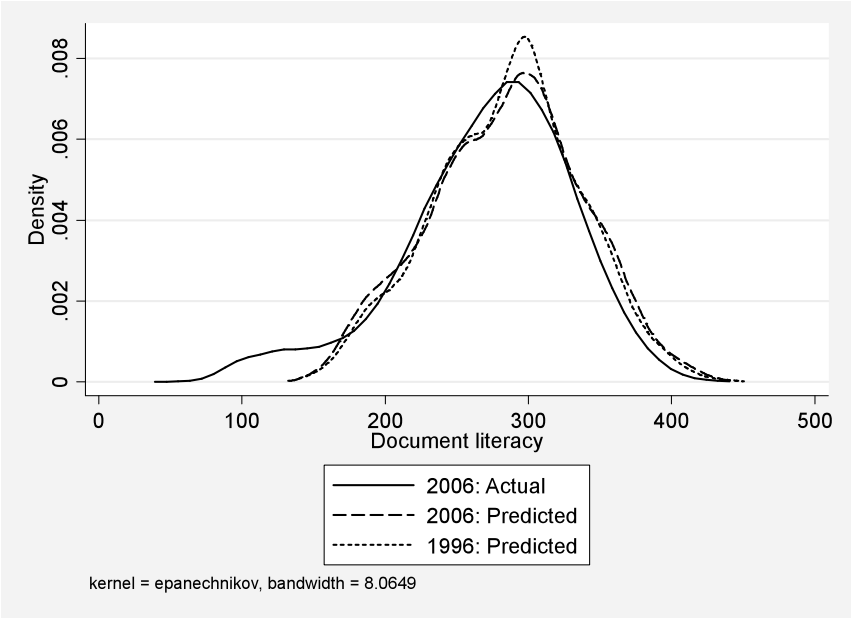
The measures of individual literacy we use are document literacy, prose literacy, numeracy (using scales contained in the data) and self-assessed skills (based on a scale we develop). We place all the scales we developed onto a 0-500 range, consistent with the literacy and numeracy scales provided in the ABS data.

While the measures of individual literacy in the 2006 data contain both an underlying, continuous score on a 0-500 range and a summary indicator in the form of a five point scale (with known thresholds from the underlying scale), the literacy skill levels of the 1996 survey were only published in Australia on the same summary five point scale used in 2006. To overcome this problem, we predict a continuous scale for 1996, given the observed five-point scale scores of individuals and a small set of other characteristics. The quantitative literacy domain, derived from the 1996 Survey of Aspects of Literacy (SAL), cannot be compared to the expanded measure of adult numeracy of the 2006 Adult Literacy and Life Skills Survey (ALLS) (ABS 2006). For that reason, the report only considers the numeracy measure of the 2006 survey.

Document and prose literacy were predicted for 1996 using the levels of document and prose literacy observed in 2006. Self-assessed skills and indicator variables of the five point scale were used as explanatory variables in the underlying regression models. Multiple imputation techniques were applied to consider the variations in the self-assessed skill measure and the five point scale. In order to assess the accuracy of scales that were predicted for 1996, they were compared to the scales predicted for 2006. The distributions presented in Figures 6 and 7 reveal that the predicted scales of document and prose literacy do not differ greatly between the two years, suggesting that the approach produced fairly accurate results.

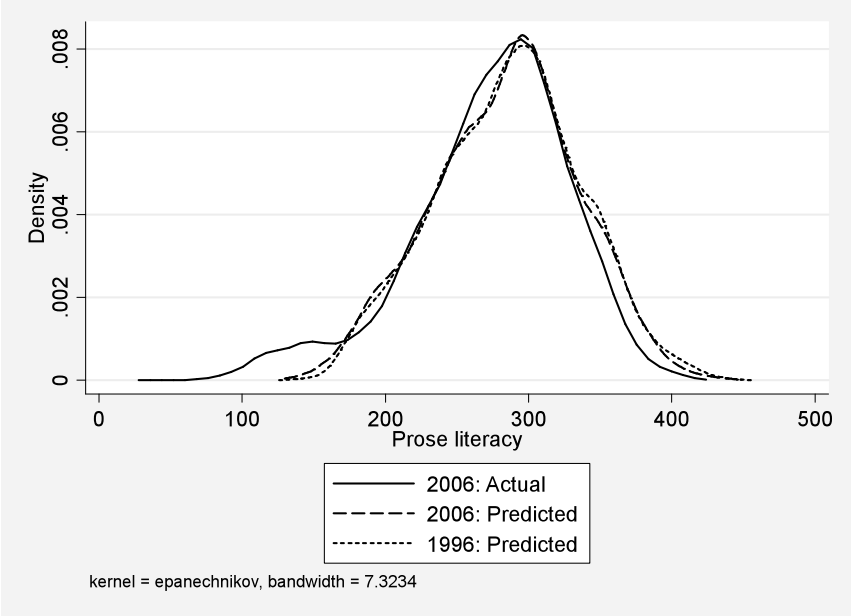
Finally, the measure of self-assessed skills was developed by using the partial-credit model described above. Appendix 2 includes a description of the scale and lists the underlying items that were used to construct it. Figures 8 and 9 include the actual and predicted scales of the generated self-assessed skill measure.

Figure 6 Actual and predicted document literacy, 1996 and 2006



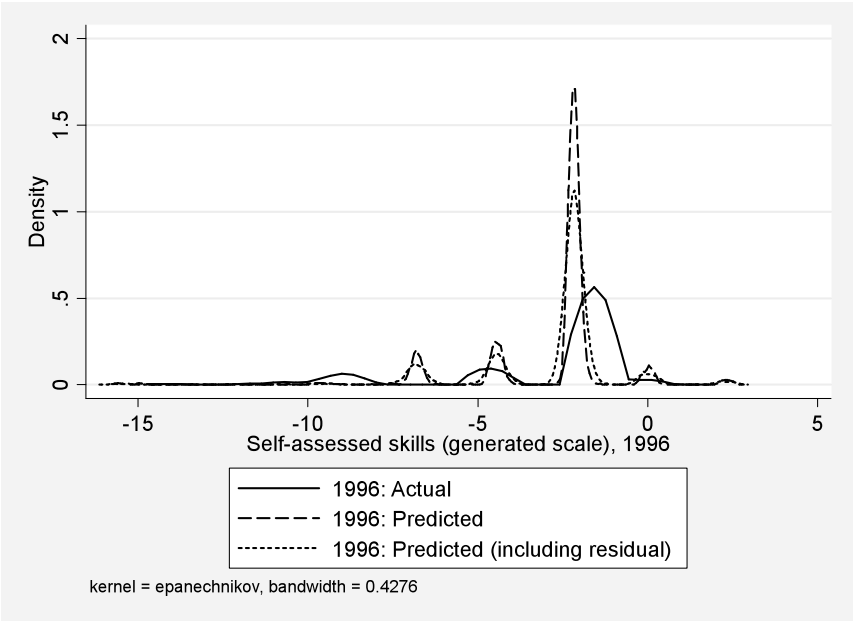
Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0; ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

Figure 7 Actual and predicted prose literacy, 1996 and 2006



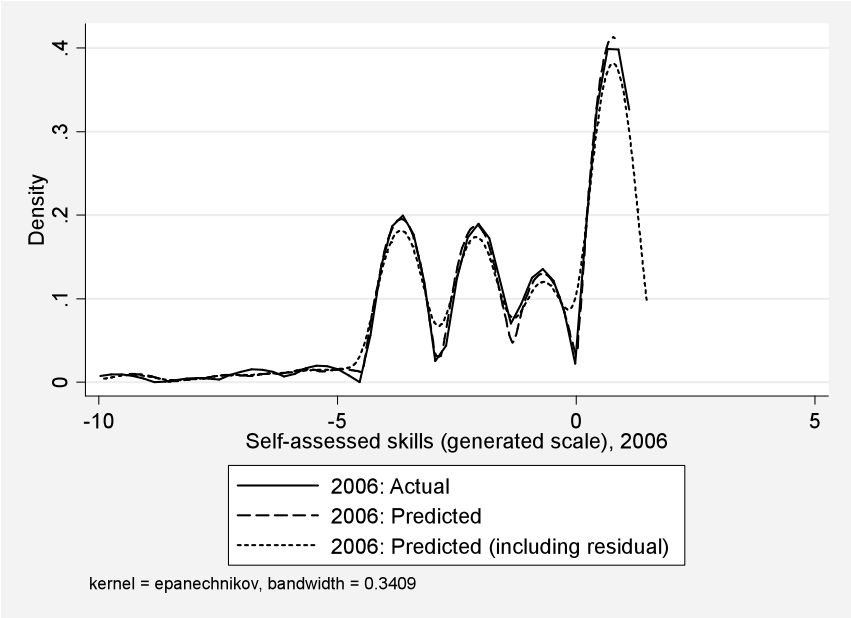
Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0; ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

Figure 8 Self-assessed skills, actual and predicted scales, 1996



Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0.

Figure 9 Self-assessed skills, actual and predicted scales, 2006



Source: ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

Definition of variables

The following table provides a description of the variables that were used in the empirical analysis of the report.

Table 1 Description of variables

Variable	
Person ID	Person ID
Year	Year indicator
Weight	Person weight
Literacy use at work (0-500)	Literacy use at work (generated scale), 1996: (i) How often reports or articles were written in main job, (ii) How often letters or memos were written in main job, (iii) How often directions or instructions for any products were used in main job, (iv) How often manuals or reference books were read or used in the main job, (v) How often reports, articles, magazines or journals were read or used in main job, (vi) How often letters or memos were read or used in main job; "At least once a week": 2, "Less than once a week": 1, "Never": 0; 2006: (i) How often reads letters, memos or emails, (ii) How often reads or uses reports, articles, magazines or journals, (iii) How often reads or uses manuals or reference books including catalogues, (iv) How often writes or fills in letters, memos or emails, (v) How often writes or fills in reports, articles, magazines or journals, (vi) How often writes or fills in manuals or reference books including catalogues; "At least once a week": 2, "Less than once a week": 1, "Never": 0; Scale takes on values from 0-500.
Numeracy use at work (0-500)	Numeracy use at work (generated scale), 1996: (i) How often arithmetic was used in main job to work out prices, costs or budgets, (ii) How often arithmetic was used in main job to measure or estimate the size or weight of objects, (iii) How often forms such as bills, invoices or budgets were filled out in main job (iv) How often bills, invoices, spreadsheets or budget tables were read or used in main job, (v) How often diagrams or plans were read or used in main job; "At least once a week": 2, "Less than once a week": 1, "Never": 0; 2006: (i) How often calculates prices, costs or budgets, (ii) How often measures or estimates the size or weight of objects, (iii) How often writes or fills in bills, invoices, spreadsheets or budget tables, (iv) How often reads or uses bills, invoices, spreadsheets or budget tables, (v) How often reads or uses diagrams or plans; "At least once a week": 2, "Less than once a week": 1, "Never": 0

Table 1 continued

Variable	
Skills (0-500)	Self-assessed skills (generated scale), 1996: Self-perception of (i) English reading skills for the needs of main job, (ii) English writing skills for the needs of main job, (iii) English reading skills for the needs of daily life, (iv) English writing skills for the needs of daily life; "Poor": 0, "Good or moderate": 1, "Excellent": 2; 2006: (i) Has reading skills in English to do main job well, (ii) Has writing skills in English to do main job well, (iii) Self-perception of English reading skills for daily life, (iv) Self-perception of English writing skills for daily life; (i)-(ii): "Disagree / Strongly disagree": 0, "Agree": 1, “Strongly agree": 2, (iii)-(iv): "Poor": 0, "Good or moderate:" 1, "Excellent:" 2; Scale takes on values from 0-500.
Document literacy 1-5	Document literacy, level 1-5
Prose literacy 1-5	Prose literacy, level 1-5
Numeracy 1-5	Numeracy levels 1-5
Document literacy 1-5 (0-500)	Document literacy, continuous measure 1-5 (0-500)
Prose literacy 1-5 (0-500)	Prose literacy, continuous measure 1-5 (0-500)
Numeracy 1-5 (0-500)	Numeracy, continuous measure 1-5 (0-500)
Age	Five year age ranges
Birth cohort	Identifier for individuals from common birth cohorts in each survey
Sex	Male or female
State	State
Educational attainment	Highest level of completed schooling or post-school qualification; the following categories were used for 2006: [1] Year 8 or below, [2] Year 9, [3] Year 10, [4] Year 11, [5] Year 12, [6] Certificate I/II, [7] Certificate III/IV, [8] Advanced diploma/diploma, [9] Bachelor degree, [10] Postgraduate degree, graduate diploma/graduate certificate
Occupation (ANZSCO 2006)	Occupation of main job; the following categories were used in 2006: [1] Labourers, [2] Machinery Operators And Drivers, [3] Sales Workers, [4] Clerical and Administrative Workers, [5] Community and Personal Service Workers, [6] Technicians and Trades Workers, [7] Professionals, [8] Managers
Full-time employed	Indicator variable for full-time employment
Employer size	Employer size; number of persons employed at the location of the individuals’

Descriptive statistics

Table 2 Descriptive statistics, 1996

Variable	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
Job task measures				
Literacy use	284.1	108.4	45.6	450.6
Numeracy use	253.7	90.7	62.6	437.3
Individual literacy measures				
Document literacy	286.9	47.0	160.0	417.1
Prose literacy	287.7	47.1	163.8	416.8
Self-assessed skills				
Self-assessed skills	347.9	53.1	6.9	492.9
Age				
Age 15 - 19 years	0.049	0.217	0	1
Age 20 - 24 years	0.097	0.296	0	1
Age 25 - 29 years	0.127	0.333	0	1
Age 30 - 34 years	0.140	0.347	0	1
Age 35 - 39 years	0.152	0.359	0	1
Age 40 - 44 years	0.124	0.330	0	1
Age 45 - 49 years	0.118	0.323	0	1
Age 50 - 54 years	0.085	0.279	0	1
Age 55 - 59 years	0.055	0.228	0	1
Age 60 - 64 years	0.032	0.176	0	1
Age 65 - 69 years	0.011	0.108	0	1
Age 70 - 74 years	0.004	0.067	0	1
Gender				
Male	0.520	0.499	0	1
Female	0.479	0.499	0	1
State				
New South Wales	0.220	0.414	0	1
Victoria	0.200	0.400	0	1
Queensland	0.174	0.379	0	1
South Australia	0.113	0.317	0	1
Western Australia	0.148	0.356	0	1
Other	0.141	0.348	0	1

Notes: This table includes descriptive statistics (unweighted numbers) of the sample that was used in the empirical analysis of the report. Number of observations: 5459.

Source: ABS, Survey of Aspects of Literacy, Australia, Basic Confidentialised Unit Record File, 1996, 4228.0.

Table 3 Descriptive statistics, 2006

Variable	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
Job task measures				
Literacy use	297.6	116.6	27.1	473.8
Numeracy use	275.4	96.8	37.7	446.6
Individual literacy measures				
Document literacy	287.9	48.8	86.9	416.0
Prose literacy	286.4	47.0	89.4	405.8
Numeracy	281.9	50.8	73.2	418.6
Self-assessed skills				
Self-assessed skills	381.1	93.2	17.9	481.0
Age				
Age 15 - 19 years	0.042	0.202	0	1
Age 20 - 24 years	0.076	0.265	0	1
Age 25 - 29 years	0.100	0.300	0	1
Age 30 - 34 years	0.117	0.322	0	1
Age 35 - 39 years	0.132	0.338	0	1
Age 40 - 44 years	0.129	0.335	0	1
Age 45 - 49 years	0.123	0.329	0	1
Age 50 - 54 years	0.106	0.308	0	1
Age 55 - 59 years	0.091	0.287	0	1
Age 60 - 64 years	0.055	0.228	0	1
Age 65 - 69 years	0.016	0.128	0	1
Age 70 - 74 years	0.007	0.088	0	1
Gender				
Male	0.520	0.499	0	1
Female	0.479	0.499	0	1
State				
New South Wales	0.211	0.408	0	1
Victoria	0.195	0.396	0	1
Queensland	0.182	0.386	0	1
South Australia	0.118	0.323	0	1
Western Australia	0.143	0.350	0	1
Other	0.149	0.356	0	1
Educational attainment				
Postgraduate Degree, Graduate Diploma/Graduate Certificate	0.086	0.280	0	1
Bachelor Degree	0.181	0.385	0	1
Advanced Diploma/Diploma	0.102	0.303	0	1
Certificate III/IV	0.189	0.391	0	1
Certificate I/II	0.011	0.107	0	1
Year 12	0.164	0.371	0	1
Year 11	0.068	0.251	0	1
Year 10	0.130	0.336	0	1
Year 9	0.032	0.178	0	1
Year 8 or below	0.032	0.176	0	1

Notes: Table continued on next page.

Table 3 Continued

Variable	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*
Occupation				
Managers	0.163	0.369	0	1
Professionals	0.230	0.421	0	1
Technicians and Trades Workers	0.144	0.351	0	1
Community and Personal Service Workers	0.080	0.271	0	1
Clerical and Administrative Workers	0.157	0.363	0	1
Sales Workers	0.074	0.262	0	1
Machinery Operators And Drivers	0.059	0.236	0	1
Labourers	0.091	0.287	0	1
Full-time employment				
Full-time employment	0.708	0.454	0	1
Employer size				
Less than 20	0.128	0.334	0	1
20-99	0.258	0.438	0	1
100-499	0.143	0.351	0	1
500 and over	0.106	0.308	0	1

Notes: This table includes descriptive statistics (unweighted numbers) of the sample that was used in the empirical analysis of the report. Number of observations: 5286.

Source: ABS, Adult Literacy and Life Skills Survey, Australia, Basic Confidentialised Unit Record File, 2006, 4228.0.

References

ABS (Australian Bureau of Statistics) 2006, *Adult Literacy and Life Skills Survey: User Guide*, cat.no.4228.0.55.002, Australian Government Publishing Service, Canberra.

Masters, G 1982, 'A Rasch Model for Partial Credit Scoring', *Psychometrika*, vol.47, no.2, pp.149–174.

Rabe-Hesketh, S, Skrondal, A & Pickles, A 2004, 'GLLAMM Manual', University of California-Berkeley, Division of Biostatistics, Working Paper Series, Paper No. 160.

Rasch, G 1960, 'Probabilistic Models for Some Intelligence and Attainment Tests', Copenhage: Danmarks Paedagogiske Institut.

——1961, 'On general Laws and the Meaning of Measurement in Psychology', in Proceedings of the IV Berkeley Symposium on Mathematical Statistics and Probability, ed J Neyman, Berkeley, CA: University of California Press.

Rubin, DB 1987, 'Multiple Imputation for Nonresponse in Surveys', New York: John Wiley Sons.

Zheng, X and Rabe-Hesketh, S 2007, 'Estimating Parameters of Dichotomous and Ordinal Item Response Models with GLLAMM', *The Stata Journal*, vol.7, no.3, pp.313–333.