Mapping adult literacy performance — support document

Shelley Gillis[1]
Mark Dulhunty[2]
Margaret Wu[2]
Leanne Calvitto[2]
Geri Pancini[1]

[1]Victoria University
[2]Educational Measurement Solutions

This document was produced by the authors based on their research for the report <*Does 1 = 1? Mapping measures of adult literacy and numeracy*>, and is an added resource for further information. The report is available on NCVER's website: <http://www.ncver.edu.au>.

This document should be attributed as Gillis, S, Dulhunty, M, Wu, M, Calvitto, L & Pancini, G 2013, *Mapping adult literacy performance: support document*, NCVER.

Mapping of Adult Literacy Performance – Stage 2

Draft Technical Report

Shelley Gillis, Victoria University

Mark Dulhunty, Educational Measurement Solutions

Margaret Wu, Educational Measurement Solutions

Leanne Calvitto, Educational Measurement Solutions

Geri Pancini, Victoria University

25th June, 2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

OVERVIEW

BACKGROUND

In 2010, the National Centre for Vocational Education Research (NCVER) conducted a preliminary study to determine the feasibility of mapping the performance levels of the international Adult Literacy and Life Skill Survey (ALLS) to those of the Australian Core Skills Framework (ACSF) using a Delphi technique (Circelli, Curtis, Perkins, 2011[1]). In that study, a small number of adult literacy and numeracy experts used their professional judgement to qualitatively align a sample of ALLS items to the ACSF levels. At the completion of the study, there was general consensus among the participants that:

- the mapping process was feasible for the:
    - Reading domain of the ACSF to the ALLS prose and document literacy domains; as well as the
    - Numeracy domains of the two frameworks.
- a larger-scale research study should be undertaken to empirically align the two frameworks onto a single scale for each of the two domains (i.e., Reading and Numeracy).

The National Centre for Vocational Education Research (NCVER) commissioned Victoria University (Shelley Gillis) in conjunction with Educational Measurement Solutions (Margaret Wu and Mark Dulhunty) to undertake the larger-scale research study.

AIMS AND OBJECTIVES

The main aim of the study was to empirically map the ALLS and ACSF frameworks onto a single scale of measurement for the two domains of 'Reading' and 'Numeracy'. The two frameworks were to be mapped by level, as opposed to individual indicators. The study was also limited to mapping the levels 1 to 5 within each framework to determine their relationship across frameworks, with the ACSF Pre-Level 1 Supplement outside the scope of the current investigation.

The study was designed to address the following research question:

- What is the relationship between the five levels on the ACSF and the five levels on the ALLS for the Numeracy and Reading domains of adult literacy?

To answer this research question, it was also necessary to empirically validate the structure of the ACSF to ensure that it was hierarchical, developmental and cumulative in its design.

1 Circelli, M., Curtis, D., & Perkins, K. (2011). Mapping Adult Literacy Performance, NCVER: Adelaide.

METHODOLOGY

The main purpose of the study was to empirically align the reading and numeracy skill areas within the ACSF and the ALLS so that the relationship between the five levels on each framework could be compared. To achieve this aim, a survey design method was employed in which individuals familiar with adult literacy and numeracy concepts completed an on-line survey. Within this research design, each participant (referred to hereon as 'rater') rated one or more anonymous learners, whose reading and/or numeracy level was familiar to the rater, against content drawn from both frameworks.

To select the framework content to be included in the on-line survey, each of the components within the ACSF and the ALLS were initially examined for potential use as stand-alone items[2]. In relation to the ALLS, each of the Level Descriptors within each skill area, as well as a sample of publicly available retired scaled items and a random sample of Numeracy Complexity Statements were included in the survey. In relation to the ACSF, a random sample of Performance Features and the total pool of Level Indictors were selected. The survey item pool comprised a total of 79 items for Reading (i.e., 34 items representing the ALLS and 45 representing the ACSF) and 86 items for Numeracy (i.e., 50 items representing the ACSF and 36 representing the ALLS) (see Appendix 10).

A 15 minute on-line survey was designed featuring multiple forms (i.e., three forms per skill area) with link items (i.e., common items across forms to enable the forms to be equated onto a single scale) to minimise rater workload, and at the same time, collect sufficient data on all 79 items for Reading and 86 items for Numeracy based on the expected sample size of raters. Each form contained approximately 50 items containing content drawn from both frameworks across three adjacent levels on each framework. Items were presented in random order so the raters were not able to obtain external cues about the level of an item (other than the wording of the item itself), and also to avoid any item positioning effect. As such, the complexity of an item could be determined solely upon the language contained within each item, as opposed to making an a priori assumption about the relative complexity of the item content according to its original positioning within the framework.

At the commencement of the survey, each rater was required to supply background information as well as to select an anonymous learner, whose literacy/numeracy levels were familiar to the rater, to form the basis of the ratings. The rater was first required to make an holistic judgement of the learner's ACSF level in either the Reading and/or Numeracy skill area. The holistic judgement was used to assign an appropriate form for the rater to complete. Each rater was then required to use a

2 An item is a collective term used within this report to describe the content drawn from the key components of the ACSF (e.g., Performance Features, Indicators) and/or the ALLS (eg. Level descriptors, scaled items) that was converted into a survey question.

three point rating scale (i.e., 'not very likely', 'somewhat likely' or 'very likely') to rate the likelihood that the learner would be able to independently perform the task described by each item.

For the Numeracy and Reading domains separately, the data was analysed and placed onto the same scale of measurement using Item Response Theory[3] which enabled the relative complexities of the five levels on the ACSF and the five levels of the ALLS to be determined. Explicitly, an IRT rating scale model was applied to analyse the data using the ConQuest computer software program[4] for item response theory applications. As the survey design contained items common to multiple survey forms, all items across all three forms per domain could be mapped onto the same scale using common item equating. The outcome of the analyses was that the complexity of each item and ability of each learner was estimated on a scale measured in logit units. A transformation was then applied to these estimates to make them more interpretable for end users of this report (analogous to converting inches into centimetres). Items were assigned a complexity estimate ranging from 100 (low complexity) to 200 (high complexity).

RESULTS

Four hundred and eleven surveys were completed, with the majority of ratings made against the Reading domain (72%). Although all states and territories were represented by the raters, the majority were located in Victoria (40%) and New South Wales (23%), with very few raters located in the Australian Capital Territory (1%). As expected, the raters tended to be employed within an Education and/or Training Institute (80%) and had more than 5 years' experience in Adult, Language, Literacy and/or Numeracy (65%). Only 7% of the raters had less than one year experience, indicating that majority of raters who participated in this study were very experienced within this field.

When rating the Reading domain, raters tended to select learners who were taking (or had previously undertaken) an ESL program (26%) or an LLN Program (22%); whereas with Numeracy, raters tended to select learners from an adult literacy and numeracy course (32%) or an LLN Program (22%). In both instances, raters tended to select learners who were at Levels 1 to 3 on the ACSF (as determined by the rater's initial holistic judgement), with less than 10% of raters selecting learners thought to be at Levels 4 & 5 in either Reading or Numeracy.

Subsequently, the majority of ratings were made against items that were aligned to Levels 2 and 3 for both Reading and Numeracy (ranging from 102 to 294 ratings made), with very few ratings made against the items that were at Level 5 on both frameworks. Given the large standard errors of measurement that were found for the Level 5 items, it was recommended that caution be

[3] Rasch, G., 1960. Probabilistic Models for some Intelligence and Attainment Tests. University of Chicago Press, Chicago, IL USA.

4Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. (2007). ConQuest (Version 2.0) [Computer Software]. Camberwell, Australia: ACER.

exercised when comparing the highest complexity levels on both frameworks, particularly for the Numeracy domain.

Estimates of learners' ability were highly reliable in both Reading and Numeracy, with reliability estimates of ≥0.971 produced for both domains. Furthermore, there was also strong evidence of the construct validity of the measures as indicated by the replication of the levels structure within each framework. That is, even though the 50 or so items presented to raters were presented randomly, and drawn from two frameworks spanning three levels each, the calibration of item complexities (using Item Response Theory) replicated the expected sequencing of the set of items within each level. In fact, only 7 of the 165 items were found to produce unexpected complexity estimates, and these items were subsequently excluded from further analysis in which the complexity levels across frameworks were compared.

When comparing the levels across frameworks, a similar pattern was evident for both Reading and Numeracy. For example, Level 1 on both frameworks appeared to be similar in their complexity, whereas Levels 2 & 3 on the ALLS were found to be more complex than ACSF Level 2 and ACSF Level 3, respectively on both domains; with the difference more pronounced for Reading. That is, ACSF Reading Level 3 was closely aligned to ALLS Reading Level 2; and ACSF Reading Level 4 was closely aligned to ALLS Reading Level 3. For Numeracy, the difference between frameworks at Levels 2 and 3 was still evident, but not as pronounced as seen in the Reading domain. The small number of ratings made against Level 5 Numeracy items meant the relationship between the two frameworks at this level could not be determined reliably. A table summarising the indicative empirical relationship between the levels based on the results of this study is contained in Table 1.

Table 1: Empirical alignment of ACSF to ALLS by ACSF Level

Reading		Numeracy	
ACSF Level	**ALLS Level**	**ACSF Level**	**ALLS Level**
1	1	1	1
2	1-2	2	1-2
3	2	3	2-3
4	3	4	3-4
5	4-5	5	Uncertain

The findings of the current study had direct implications for future use and refinement of the ACSF, as well as future mapping to other similar programs (e.g., AMEP) and frameworks (e.g., the Core Skills for Employment Framework).

ORGANISATION OF THE TECHNICAL REPORT

The main aim of this technical report was to

- Document the process to provide a transparent account of the study and to enable the study to be replicated for other frameworks that may require mapping (e.g., Adult Migrant English Program)
- Provide a report on the outcomes of the analyses, including an overview of the aims of the study, methodology, results, conclusions and recommendations

The process employed and the outcomes of this study have been presented in five chapters.

Chapter 1 describes the research team's understanding of the similarities and differences between the ACSF and ALLS which informed the research design.

Chapter 2 documents the methodology employed, including the rationale for the research design, a description of the target population, a brief outline of the survey design features as well as a description of the data analysis techniques employed.

Chapter 3 details the process used to design and develop the on-line survey, including consideration to content selection, survey functionality and format, as well as alternative forms and link items.

Chapter 4 presents the findings from the study in which the relationship between the five levels within each skill domain was determined.

Chapter 5 outlines the major findings from the study and the implications for further research and development activities.

CHAPTER 1: INTRODUCTION

To inform the research design of the current study, existing documentation in relation to both the ACSF and ALLS were reviewed to ascertain the similarities and d ifferences between the two frameworks in terms of structure, purpose and assessment requirements. A summary of the findings has been presented below.

THE STRUCTURE

ACSF

- The ACSF has been designed to illustrate a developmental learning pathway for five core skills: Learning, Reading, Writing, Oral Communication and Numeracy.
- For each core skill area, there are five levels on the developmental continuum.
- Within each core skill of the ACSF, there are a number of statements that describe each level. Such statements
    - can be found within the Indicators, Performance Features and Sample Activities;
    - describe typical performance at a particular level;
    - should be hierarchical and cumulative across levels;
    - should be clear and explicit in their descriptions of increasing proficiency within a core skill area; and
    - should be internally coherent in that the set of statements should describe a single underlying construct.
- The ACSF is based on a broad range of theoretical understandings, and whilst it has been subjected to extensive consultation throughout its development, it has yet to be empirically validated for its validity and reliability.

ALLS

- The ALLS has been designed to document the developmental progression of learning for four skills of adult literacy and numeracy, namely prose literacy, document literacy, numeracy and problem solving. Each skill has been scaled using Item Response Theory (IRT).
- As with the ACSF, each of the ALLS skills has also been underpinned by a theoretical framework which provides the basis for assessment task development.
- Each skill is also reported according to five levels of achievement.
- Each level of achievement has a set of descriptors which describe the type and complexity of tasks that typically operate at each level on the framework within each skill. Hence, the descriptors are task driven in their definition.

THE PURPOSE OF EACH FRAMEWORK

- The ACSF was designed to be used for a range of different purposes (e.g., formative and summative) in a low-stakes assessment environment in Australia in which reporting typically occurred at the student and/or course level.
- The primary purpose of conducting the ALLS was to identify and measure literacy of adult populations within and across a number of participating OECD countries.

ASSESSMENT AGAINST EACH FRAMEWORK

- The ACSF allows support to be provided during assessment whereas the ALLS had been designed to be administered under exam like conditions in which no prompting/support from the invigilator(s) was permitted[5].

- In addition to the descriptor statements for each level within a skill, the ALLS had a range of calibrated assessment tasks (referred to as scaled items) that had been empirically scaled (and therefore had a difficulty level measure that corresponded to one of the five levels within the framework) that could be used within the current study to measure literacy (prose and document) and numeracy.

- At the time of this study, there were no available sets of assessment tasks for the ACSF that had been subjected to the same psychometric testing procedures as those items used to measure literacy and numeracy within the ALLS[6].

- Whilst there were no standardised, calibrated test items available for the ACSF, Sample Activities that provided specific examples of what a person may be able to do at each level had been developed through extensive consultation as part of the ACSF endorsement process.

A summary of the similarities and differences of the two frameworks has been displayed in Table 2.

- [5] It was therefore assumed (as was the case with the preliminary, feasibility study undertaken by Circelli, Curtis and Perkins (2011)), that when judging performance against the ACSF descriptors, the survey raters should assume that the learner has attempted to perform the task independently.

- [6] Note that although assessment tasks for the ACSF were available at the time of this investigation, such tasks had only been qualitatively mapped to the framework and further item development would need to be undertaken to supplement those available to ensure full coverage across all ACSF levels and skill areas. Furthermore, such tasks would also need to be subjected to paneling, piloting and trialing procedures to examine their psychometric properties prior to use.

Table 2: ACSF versus ALLS: Summary of the similarities and differences in design and use.

Theme	Feature	ALLS	ACSF
The structure	Reading domain	Prose Literacy Document Literacy	Reading
	Numeracy domain	Numeracy (previously referred to as Quantitative Literacy)	Numeracy
	No. of Levels	Five	Five
	Nature of the descriptors	Task driven	Competency driven
	Validation process	Empirical	Qualitative
	Nature of the framework	Developmental, hierarchical and cumulative	Developmental, hierarchical and cumulative
	Level of specificity of descriptors	Summary statements only	Detailed and extensive
Assessment	Primary Purpose	Monitoring & profiling within & across countries	Multiple purposes
	Target Population	Adult populations within and across participating countries	Post compulsory education learner within Australia
	Reporting	Population and sub-population estimates	Learner and cohort level performance
	Conditions	Independent completion of tasks under examination conditions	Support and continuous, ongoing assessment
	Task availability	Empirically calibrated test items with psychometric data available	Sample activities that have been subjected to widespread consultation.

For the purposes of this investigation, the terms 'Reading' and 'Numeracy' domains will be used throughout this study to refer to the overarching constructs that underpin both frameworks. The relationship between the different constructs within this investigation has been presented in Table 3.

It should also be noted that although there were five levels on each framework, it was uncertain as to how each of the levels related to each other. Determining such a relationship was the ultimate aim of the current investigation.

Table 3: The relationship between the constructs under investigation

ACSF Construct	*ALLS Construct*	*Overarching Construct*
Reading	Prose Literacy Document Literacy	Reading domain
Numeracy	Numeracy (previously referred to as Quantitative Literacy)	Numeracy domain

CHAPTER 2: METHODOLOGY

This chapter is organised in five sections. Section 2.1 reports on the rationale for the research design. Section 2.2 describes the target population of the study and Section 2.3 describes the data collection procedures employed to recruit raters. Section 2.4 presents a brief outline of the design of the data collection instrument (i.e., the on-line survey) and Section 2.5 outlines the pilot study undertaken. Section 2.6 outlines the data analysis methods employed, and lastly, Section 2.7 outlines the approval processes required to undertake the research investigation.

2.1: RATIONALE FOR THE RESEARCH DESIGN

A common approach to mapping two different frameworks onto a single scale of measurement is to administer tests containing questions common to both frameworks to a large sample of learners. In this case, the ALLS has test questions that have been calibrated against the ALLS scale by Statistics Canada and OECD[7]. However, at the time of commencing this study, no equivalent, comprehensive standardised test was available that had been empirically aligned to the ACSF. Therefore, to empirically map the ALLS scale to the ACSF using standardised tests would require the development, paneling, piloting, trialling and finally administration of the new ACSF test items (as well as those available from Statistics Canada and Educational Testing Service for measuring literacy and numeracy within the ALLS) to a large group of learners with varying levels of proficiency in numeracy and reading. However, administering test items to learners would provide a measure of the learners' performance on the items. Such performance would then need to be generalised to reflect the skill descriptions as set out in the 'Reading' and 'Numeracy' frameworks. In this sense, administering tests provide an indirect measure of the 'Reading' and 'Numeracy' constructs under investigation, at least for the ACSF, as the levels of the ACSF are defined by a set of skill descriptors rather than by test items. Furthermore, such an approach would be very costly and time consuming to employ, especially when the key objective of the study was to map two frameworks rather than produce a new standardised test for the ACSF.

Given the costs associated with creating a new standardised test aligned to the ACSF, and that tests would only produce a sample of behaviours against the indicators of the ACSF core skills, the research team considered that the most direct and cost efficient way to empirically measure the ACSF was to obtain behavioural evidence on the ACSF Indicators and Performance Features themselves, rather than through test items which indirectly reflect the indicators.

---

[7] Statistics Canada and OCED (2005). Learning a Living: First results of the Adult Literacy and Life Skills Survey, OCED Publishing, Ottawa and Paris.

Statistics Canada (2006). International Adult Literacy Survey, Statistics Canada, Ottawa.

To address the research aim using the most direct approach to measuring each domain, the ALLS scale descriptors and scaled items as well as the ACSF Level Indicators and Performance Features needed to be r ank ordered within each framework, and matched between the frameworks (see Chapter 3). This is very similar to the situation where test items and learners are ranked, and learners' abilities are matched to the difficulties of the test items. An ideal approach to facilitate the empirical alignment of the frameworks was to use Item Response Theory[8].

Item Response Theory (IRT) is a par adigm for the design, analysis, and scoring of tests, questionnaires, and similar instruments for measuring abilities, attitudes, or in this case, scales that describe different levels of Adult Literacy and N umeracy. IRT is the preferred method for the development of tests such as the National Assessment Project – Literacy and Numeracy (NAPLAN), the OECD Programme for International Student Assessment (PISA) and the Trends in International Mathematics and Science Study (TIMSS). N ote that IRT underpinned the development of the ALLS scale to begin with and hence, is not a new paradigm in adult literacy and numeracy scale development.

Just as IRT is suited to calibrating test items and learners on the same scale, IRT is also a useful tool for calibrating items drawn from the ALLS and the ACSF onto the same scale. However, unlike learners who typically respond directly to test items, the research design adopted within this study required a rater to apply his/her professional judgement to rate a learner against skill descriptors drawn from the two frameworks. This is similar to the rating of learners in a performance task (e.g., writing or speaking in the language arts), where a teacher/assessor acts as a rater to generate item responses about a learner's level of performance against a task. The raters recruited in this study were acting as "raters" to provide their ratings of a post compulsory education literacy/numeracy learner against the descriptors.

Each rater who completed the survey rated his/her nominated learner against multiple items. This process meant that a particular item was compared against multiple learners. A subsequent analysis using IRT enabled the items to be calibrated with respect to each other and with respect to the learners being rated. This was because the mathematical models used in IRT describe relationships between the learners' proficiency level and the complexity of the items against which the learners were rated[9].

---

8 Wu, M. & Adams, R. (2007). Applying the Rasch model to psycho-social measurement: A practical approach. Educational Measurement Solutions, Melbourne. www.edmeasurement.com.au

9 On a technical note, IRT uses a probability function to describe the likelihood that a rater would give a particular rating to a learner against a particular statement. Through the use of the probability function, conclusions can be drawn about the relative standing of items and learners.

In addition, the use of alternative questionnaire forms requires a process to pool the data together. Again, IRT provided a helpful tool for "equating" the questionnaire forms as the raters were presented with different forms, provided that there were common items that linked the different questionnaire forms.

2.2: TARGET POPULATION

The target population of raters for the study was defined as personnel responsible for developing, managing, delivering and/or assessing literacy and numeracy content to post compulsory education learners within Australia. This included those involved in accredited courses such as Certificates of General Education, English as a Second Language Access courses and other bridging and further education courses. Language, literacy and numeracy programs offered through qualifications delivered in vocational education, VET in Schools and the community sector, as well as those offered in workplaces and higher education institutions were also included in the target population. The survey was designed so that although raters did not have to be familiar with the ACSF or the ALLS per se, it was desirable that the target population was familiar with:

- adult literacy and numeracy concepts; as well as
- the literacy or numeracy ability level of at least one post compulsory education learner (either past or present).

Examples of the type of personnel who were part of the target population for each of the educational sectors have been displayed in Table 4.

Table 4: Target Population according to educational sector

Sector	Example of Positions/Organisations
Senior Secondary	• Staff involved in delivering literacy and/or numeracy units within VET in Schools programs • English teachers within senior secondary certificates of education • Mathematics teachers within senior secondary certificates of education • The Australian Association of Mathematics Teachers • Others
Vocational Education	• RTO Teachers/Trainers delivering LLN courses • Industry Skills Council personnel with responsibility for adult literacy and/or numeracy • Services Skills Australia - WELL Network • Others
Adult and Continuing Education and Stakeholders	• Teachers/Trainers delivering LLN courses • AMES • Community Colleges Australia (CCA) • ACE Vic • Australian Literacy Educators' Association (ALEA) members • Australian Council for Adult Literacy (ACAL) • Australian Council of TESOL (ACTA) • Others
Higher Education	• University Learning Support Units, Academic Skills units, Language Support Centers

To ensure that there was adequate data across all levels within each skill area, a sample size of 600 respondents per domain was sought (i.e., around 100 per level). As a large number of individuals within the target population would teach and deliver both literacy and numeracy content, raters were encouraged to complete a separate survey for each domain and/or learner. As raters could complete more than one survey, then the number of individuals required for the study could be fewer than 600 for each domain.

To increase the response rate and maintain rater engagement, the survey was designed to take approximately 15-20 minutes to complete and contained no more than 50 statements per survey. Furthermore, each time a participant completed an entire survey, s/he was also eligible to enter the draw to win one of six iPads (3rd generation 32GP Wi Fi). See Appendix 4 for Terms and Conditions for the draw.

2.2.1 Desired number of ratings

It should be acknowledged that for the purposes of calibrating the two frameworks, it is not so much how many raters participate in the study which is of concern (i.e., the sample size), but more importantly, the number of ratings received per item.

The aim is for each item to be rated many times. The more ratings received for an item the more confident one can be about the location of the item on the scale (as this will reduce the measurement error for the items). Given that the objective of the study is a broad mapping, approximately 200 ratings per item was sought to achieve the desired accuracies.

2.3 DATA COLLECTION PROCEDURES

Data collection involved collecting responses to an on-line survey administered to personnel responsible for developing, managing, delivering and/or assessing literacy and numeracy content to post compulsory education learners within Australia. The recruitment of raters was the responsibility of the National Centre for Vocational Education Research (NCVER). Through websites, email correspondence and newsletters maintained by the NCVER, raters were invited to voluntarily participate in the study. See Appendix 5 for an example of the flyer distributed to recruit raters. The NCVER also emailed peak bodies, association and stakeholders requesting them to disseminate the flyer through their networks, newsletters and website.

Data collection occurred over a 7-week period. Regular monitoring of participation rates per skill area and holistic ACSF Level was undertaken throughout the data collection period to identify skill areas and levels at risk of not meeting the desired quota. In such instances, the NCVER implemented targeted recruitment strategies with key personnel in the field.

2.4 DATA COLLECTION INSTRUMENTS

The main purpose of the study was to empirically align the ACSF and the ALLS onto a single common scale of measurement so that the relationship between each of the levels could be determined. To achieve the aims of the study it was proposed that the literacy and numeracy skills of the ALLS be mapped to the ACSF Reading and Numeracy core skills using a survey design method in which raters anonymously rated a learner, whose reading and/or numeracy levels were familiar to the rater, against statements and sample tasks drawn directly from both the ACSF and the ALLS. Chapter 3 provides a detailed description of how the on-line survey was designed, including selection of item content, item format, as well as the development of alternative, linked survey forms and survey functionality features.

In brief, a 15 minute on-line survey was designed to collect information about the relative complexity of a set of statements/items drawn from both frameworks across all levels. There were a total of 165 statements within the item pool, representing a random sample of Performance Features from the ACSF and all Level Indicators. In relation to the ALLS, each of the Level Descriptors within each domain, as well as a sample of publicly available retired scaled items and a random sample of Numeracy complexity statements were also included in the survey. The item pool comprised a total of 79 items for Reading (34 representing the ALLS and 45 representing the ACSF) and 86 items for Numeracy (50 representing the ACSF and 36 representing the ALLS) (see Appendix 10).

To minimise rater workload, and at the same time, collect sufficient data on 79 items for Reading and 86 items for Numeracy, three alternative forms per skill area were designed. In addition, the forms were designed to have a set of common items across each adjacent form to enable all three forms per domain to be calibrated onto a single scale. Each form had approximately 50 items per survey containing statements from both the ACSF and ALLS across three adjacent levels. Items were presented randomly within each form so the raters will not be able to get hints about the level of an item (other than the wording of the item itself), and also to avoid any item positioning effect. As a result, the ordering of the items (by item complexity) established by this study can be more validly attributable to the language of each item itself, independently assessed by the raters, and not assessed in reference with other items.

At the commencement of the survey, each rater was required to supply background information as well as select an anonymous learner, whose literacy/numeracy levels were familiar to the rater, to form the focus of the survey. The rater was then required to make an holistic judgement of the ACSF level of that learner in either reading and/or numeracy. The holistic judgement was used to assign an appropriate form for the rater to complete. Each rater was then required to rate the

likelihood that the learner would be able to independently perform the task described within each item, using a three point rating scale (i.e., 'not very likely', 'somewhat likely' or 'very likely')[10].

The pool of items were reviewed by a panel consisting of adult language and literacy specialists who checked, and where necessary, assisted with making modifications to the item content, format and instructions prior to the pilot study.

Refer to Chapter 3 for a more detailed description of the development of the on-line survey used to collect the data for this study.

2.5 PILOT STUDY

A small pilot study was undertaken to examine the usability and functionality of the on-line survey prior to data collection. The pilot study aimed to examine the:

- Appropriateness of the workload of each rater;
- Appropriateness of the background questions;
- Appropriateness and ease of completion of the item formats;
- Clarity of the instructions to complete each section of the survey;
- Ease of navigation throughout the on-line survey; and
- Ways in which the survey could be improved.

Ten individuals participated in the pilot study and were selected based upon their experience and expertise with adult literacy and numeracy concepts and/or the design and development of the Australian Core Skills Framework. Seven of the 10 pilot participants had previously been involved in the paneling exercise and were also participants of the preliminary mapping exercise undertaken by NCVER[11] (see Appendix 6).

In the first instance, each pilot participant completed the draft on-line survey as though it was part of the real data collection so that s/he could explore both the content and the functions of the survey from the perspective of a potential rater. Once familiar with the functionality, content and features of the survey, each pilot participant then completed a 5 minute on-line questionnaire designed to gather feedback on any problems encountered as well as recommendations for future improvements to the survey design and/or functioning. Overall, the findings were positive in terms of the usability and functionality of the on-line survey. The findings were used to make further refinements to the survey prior to data collection. A full report of the pilot study has been included in Appendix 9.

10 An item is a collective term used within this report to describe the statements drawn the ACSF (e.g., Performance Features, Indicators) and/or the ALLS (eg. Level descriptors, scaled items).

11 Circelli, M., Curtis, D., & Perkins, K. (2011). Mapping Adult Literacy Performance, NCVER: Adelaide.

2.6 DATA ANALYSIS

As described previously in Section 2.4, survey respondents were asked to rate whether a learner was 'not very likely', 'somewhat likely' or 'very likely' to succeed on the framework content contained within each item. For the purpose of analysis, respondents were considered 'raters' who were each responsible for rating the ability of a learner (nominated by the rater to form the focus of the rater's responses). For each 'rating' of a learner, a 'score' of 2 was assigned to ratings of 'very likely' indicating that the learner was very likely to succeed on the framework content contained in the item. Conversely, a 'score' of zero (0) was assigned to ratings of 'not very likely' which indicated that the learner was not very likely to succeed on the framework content contained in the item. A 'score' of 1 was assigned to ratings of 'somewhat likely'. That is, a rater taking a survey in effect was providing a rating of how likely it was that a learner would succeed on the framework content contained in each item. The 'score' indicated how able the learner was estimated to be. Higher scores indicated higher estimated ability (and conversely, low scores indicated low ability).

From a rater's point of view, the ratings provided an estimate of the perceived ability of the learner being rated. In addition, the ratings also provided information about the relative differences in complexity between the various pieces of framework content contained within each item.

Using Item Response Theory (IRT), the relative differences in the complexity between different items could be estimated regardless of the framework they originated from. Furthermore, the application of IRT to the data enabled all estimates of complexity to be mapped onto a single scale of measurement.

Under the umbrella of IRT, perhaps the simplest IRT model is the simple logistic (Rasch) model which is suitable for analysis of dichotomous items[12]. Dichotomous items are items with two response categories (e.g. correct / incorrect). This study utilised items with a common partial credit item format[13]. For each item, raters were required to select the most appropriate of three response categories: 'Not Very Likely'; 'Somewhat Likely'; and 'Very Likely'. Hence, analysis using the simple logistic (Rasch) model was not appropriate unless the response categories were collapsed / condensed together in some manner. Such an approach is not ideal since information about each item is lost during analysis. Instead, two different extensions of the simple logistic model were considered for use in the analysis: the Partial Credit model; and the Rating Scale model. Both the Partial Credit and Rating Scale models enable analysis of items containing more than two response

12 Rasch, G., 1960. Probabilistic Models for some Intelligence and Attainment Tests. University of Chicago Press, Chicago, IL USA.

13 Part A of the survey presented content to raters in a different format to Part B of the survey (statements from both frameworks vs ALLS calibrated items). To ensure ease-of-response for raters, item instructions were not verbatim between Part A and Part B of the survey but were still intended to provide the same substantive interpretation. Hence, each rater is considered to have interpreted the response categories in Part B in the same manner as they did for Part A of the survey.

categories. The Partial Credit and Rating Scale models differ in the way the response categories for an item are treated for analysis. Using a Partial Credit model, the relative differences in difficulty between response categories are allowed to vary for each item. Under a Rating Scale model, the relative differences in difficulty between response categories are the same across all items. In more technical terms, a Partial Credit model will estimate 2 parameters for each item. One parameter for each adjacent pair of response categories (commonly referred to as an item 'step'). That is, one parameter for the 'step' between the 'Not Very Likely' and 'Somewhat Likely' categories and one parameter for the 'step' between the 'Somewhat Likely' and the 'Very Likely' categories. In the Rating Scale model, the difference between the two step parameters is assumed to be constant across all items. That is, each item has a parameter for an overall difficulty, but there is only one step parameter estimated for all items.

The Partial Credit model is useful as it allows for the analysis of a wide range of performance assessments that might contain items with different numbers of response categories between items. It does not require an assumption that items with the same number of response categories have to have the same 'step' size as is needed for the Rating Scale model. In the present study, as few raters rated high ability learners, there were few ratings against Level 4 and Level 5 items. By using a Rating Scale model for analysis, the 'step' parameters for higher level items are somewhat better estimated as the step parameters are estimated using information from all items.

Items were initially analysed using both the Partial Credit and Rating Scale models. Similar results were found using both approaches although there were some differences between item estimates for higher level items. In general, it is recommended that results at higher levels are treated with some caution due to the low number of ratings made to these items. The remainder of the results presented in this report is based on the Rating Scale modeling of the data.

The survey design ensured that there were items common to multiple survey forms (e.g. some of the items on Form A also appeared on Form B). The IRT analysis enabled all items on Forms A to C to be mapped onto the same scale using common item equating (as detailed in Chapter 3).

The outcome of the analyses was that the complexity of each item and the ability of each learner were estimated on a scale measured in logit units (analogous to the unit of measurement in centimetres on a ruler used to measure length). A linear transformation was then applied to the logit values for each item and learner so that all values were expressed as a positive number to aid readability of the results. That is, items were assigned complexity estimates ranging from 100 (low complexity) to 200 (high complexity).

2.7 APPROVAL PROCESSES

Two types of approval were required to undertake this research: ethics approval to administer the survey and in one state/territory approval from the lottery commission was also required.

An application to gain ethical approval of the study was submitted to the Arts, Education and Human Development Research Ethics Subcommittee at Victoria University and approval was subsequently granted.

As the study design also included a draw to win one of six IPad (3rd generation 32 GB WiFi) as part of its recruitment strategy, it was necessary to administer the draw in accordance with the Lottery ACT 1964 in which a permit was required for 1 of the 8 states/territories, namely the Australian Capital Territory (ACT). In addition, the Western Australia Department of Racing, Gaming and Liquor required the Terms and Conditions to be lodged. The remaining states/territory did not require the research team to officially lodge its Terms and Conditions or apply for a permit, as participation in the study was to be free, undertaken for educational research purposes, and/or the prize was to be valued at less than $5000. A summary of each state/territories requirement for permits has been displayed next.

Table 5: Lottery Requirements per state/territory

Location	Permit	Lodge Terms and Conditions	Comments
VIC	No	No	Although neither permit nor lodgment was required, all advertising must clearly specify, when and where the draw will take place, as well as a clear description of the prize and value.
NSW	No	No	Permit not required due to the nonprofit nature of the draw and its attempt to advance developments in education.
SA	No	No	As less than $5000, no need to gain permit nor lodge anything, just ensure terms and conditions have been met
WA	No	Yes	When lodging Terms and Conditions, terms and conditions must be specified. Statement must be specific and concise. See www.rgl.wa.gov.au gaming form and applications and trade promotions. See www.rgl.wa.gov.au gaming form and applications and trade promotions
QLD	No	No	As entry is fee, there is no need to gain a permit nor lodge terms and conditions. However, terms and conditions must be satisfied. For terms and conditions see www.olgr.qld.gov.au under promotion game.
TAS	No	No	As this is not regulated in Tasmania, there was no need to lodge terms and conditions
ACT	Yes	Yes	Paper –based application to be completed. Permit Received: ACT TP 11/04754
NT	No	No	As total draw prize is under $5000, no permit is required. Note that even if the prize value was greater than $5000, another state's permit is acceptable. No need for a new or additional application in the NT.

A copy of the Terms and Conditions as well as the ACT's permit to conduct a trade promotion lottery can be found in Appendices 4 and 14 respectively.

At the completion of data collection, six winners were randomly drawn and notified of their prize in accordance with the Terms and Conditions specified in Appendix 4.

CHAPTER 3 SURVEY DESIGN

This chapter has six sections. Section 3.1 describes the process implemented to determine the content of the survey. Section 3.2 describes the process and outcomes of the paneling exercise undertaken to review item content, instructions and format. Section 3.3 describes the functionality features of the survey whilst Section 3.4 outlines the design of the forms and linkages. Section 3.5 describes how forms were allocated and finally, Section 3.6 presents the survey format.

3.1 DETERMINING THE ITEM CONTENT

ACSF Statements

The Australian Core Skills Framework (ACSF) describes five levels of performance in five core skills of Learning, Reading, Writing, Oral Communication and Numeracy. Within each core skill, there are four key components:

- *Indicators* – statements that provide an overview of performance at each level of each core skill.
- *Support, Context, Text and Task Complexity - statements* which describe factors that may influence a person's performance at each level.
- *Performance Features* – statements which provide detailed descriptors of what a person operating at a level is able to do.
- *Sample Activities* – statements that provide specific examples of what a person may be able to do at each level in each of the 6 Aspects of Communication (which represent different contexts in which an individual may use the core skills).

Each of the components within the ACSF was examined for potential use within the study as stand-alone prompts to form the focus of each item within the survey. An example of a statement within each of the key components of the ACSF has been displayed in Table 6 to illustrate its potential contribution to the development of the item pool.

Table 6: Key components of the ACSF and implications for survey design.

Key Component	Example with Reading Level 1	Implications for survey design
Indicator	1.03 Identifies personally relevant information and ideas within highly familiar contexts	Major component within the ACSF to describe typical performance at each level of a core skill. Provides a direct measure of the constructs.
Task Complexity	Concrete tasks of 1 or 2 processes, e.g. locating, recognising	To be stand-alone statements within the survey, the statements would need to be contextualised with examples to enhance consistency in interpretation across raters.
Performance Features	Identifies personally relevant reasons for reading	Provides more specific information about the Indicators through detailed descriptions of what a person *who is fully competent in an ACSF level is able to do.* These statements could readily be converted into survey items but would need to be piloted first and completed by those with expertise in adult literacy and/or numeracy. They are also a direct measure of the constructs.
Sample Activities	Locates particular information of personal relevance from a familiar source, e.g. expiry date on a driving license, name on shift roster.	Concrete descriptions of activities that could typically be performed by individuals at varying levels which could readily be converted to individual prompts within a survey. They are an indirect measure of the construct. Furthermore, as some raters may not have had the opportunity to observe an individual using the core skills in all contexts, some sample activities may not appear to be applicable (i.e., they may be too context dependent) and therefore, may result in missing data and/or guessing.

As displayed in Table 6, statements could possibly be drawn from the Level Indicators, Performance Features and Sample Activities of the ACSF. However, practical constraints related to sample size, rater fatigue, and the desired number of ratings per item meant that it was not feasible to include all components within the survey. Hence, for the purposes of this study, only the Level Indicators and Performance Features were included in the survey design. Performance Features were included as they have been established to provide a direct measure of the indicators by providing detailed descriptions of what a person who is fully competent in an ACSF level would be able to do[14]. Sample Activities were illustrative of the core skills and had been drawn from a variety of contexts (e.g. personal, public, technical) and hence may not have been as generally applicable to the different raters surveyed.

The number of items from within the Performance Features and Indicators of the revised ACSF Reading and Numeracy core skills were determined to meet all constraints of the study. It can be seen in Table 7 and Table 8 respectively, that within the ACSF Reading core skill there were 123 statements that could be used to operationalise the construct, whereas with the ACSF Numeracy core skill, there were a potential pool of 96 statements.

14 DEEWR (2012). Australian Core Skills Framework, Commonwealth of Australia, Canberra.

Table 7: Summary of the pool of items within the ACSF 2012 Reading core skill

ACSF Level	Indicators	Performance Features	Sub-total
5	2	16	18
4	2	27	29
3	2	29	31
2	2	24	26
1	2	17	19
Sub-total	10	113	**123**

Table 8: Summary of the pool of items within the ACSF 2012 Numeracy core skill

ACSF Level	Indicators	Performance Features	Sub-total
5	3	16	19
4	3	17	20
3	3	18	21
2	3	16	19
1	3	14	17
Sub-total	**15**	**81**	**96**

ALLS

The ALLS had been designed to measure knowledge and skills of the adult population in four areas: Prose Literacy, Document Literacy, Numeracy and Problem Solving[15]. T he survey comprised a set of psychometrically scaled test items as well as descriptors for five levels of performance used for purposes of reporting achievement levels. Three potential sources of items for the current survey were identified:

1. Level Descriptors
2. Publicly available scaled Items
3. Numeracy Complexity Statements

Each will be considered next.

1. *ALLS Level Descriptors* for each skill area. T hese were publicly available statements that defined each level on the scale for Prose Literacy, Document Literacy and Numeracy. These included both the original statements published by Statistics Canada as well those modified by

[15] Note, as recommended by the NCVER, problem solving will be excluded from this study.

the expert panel of this (see Section 3.2). A summary of the number of items per level has been displayed below.

Table 9: Prose and Document Literacy ALLS Level Descriptions

ALLS Level	Number of Performance Level Descriptors		
	Literacy (Prose + Document)		Sub-total
	Original	Panel Modified	
5	2 (1+1)	4 (1+3)	6
4	2 (1+1)	5 (3+2)	7
3	2 (1+1)	5 (3+2)	7
2	2 (1+1)	2 (1+1)	4
1	2 (1+1)	2 (1+1)	4
Total	10	18	28

Table 10: Numeracy ALLS Level Descriptions

ALLS Level	Number of Performance Level Descriptors		
	Numeracy		Sub-total
	Original[16]	Panel Modified	
5	1	3	4
4	1	3	4
3	1	3	4
2	1	2	3
1	1	2	3
Total	5	13	18

It total, there were 28 ALLS Prose/Document Literacy Level statements and a further 18 Numeracy Level statements to be included in the current study.

2. *Publicly available scaled items* for Prose Literacy, Document Literacy and Numeracy/Quantitative Literacy. A number of scaled items, which were no longer in use (referred to as 'retired' items), have been published in publicly available reports[17] and therefore, could be used to measure Reading and Numeracy skills in the current study. A summary of the number of publicly available ALLS scaled items in which there were complete stimulus materials and instructions has been displayed in Table 11.

[16] Please note that the original ALLS Level Descriptors were slightly edited by the research team to remove any specific reference to the assessment task and/or level within the framework. Note that such edits were required to ensure that the statements from within the ALLS and ACSF could be displayed concurrently and interchangeably within the same section of the survey. That is, the revisions helped ensure that the descriptors made sense to survey respondents in the context of the survey. The descriptors were consistent with the original description in all other respects. Refer to Appendix 1 for the editorial changes made to the presentation of the ALLS original statements in the survey.

[17] See OECD (2005). Learning a Living: first results of the Adult Literacy and Life Skills Survey, Statistics Canada, Ottawa and Paris. and see also Murray, T.S., Clermon, Y., & Binkley, M. (2005). International Adult Literacy Survey. Measuring Adult Literacy and Life Skills: New Frameworks for Assessment. Statistics Canada, Catalogue no. 89-552-MIE, no. 13.

Table 11: Number of publicly available ALLS scaled items per skill area.

ALLS level	Number of publicly available ALLS scaled items	
	Literacy (Prose + Document)	Numeracy[18]
5	2 (1+1)[19]	3
4	2 (1+1)	1
3	3 (1+2)	2
2	4 (2+2)	3
1	2 (1+1)	2
Total	13 (6+7)	11

Note that some of the Numeracy ALLS scaled items reported in Table 11 have been sourced from the Quantitative Literacy skill of the International Adult Literacy Survey (IALS) which was the predecessor to the ALLS Numeracy skill. It can however be assumed that the two surveys are measuring the same construct (i.e., 'Numeracy') as both the IALS and ALLS have been equated on to the same scale using common items across both testing situations.

3. *Statements of Numeracy Complexity* have been used to underpin the design and development of Numeracy items in the ALLS. These statements express the complexity of the mathematical information/data needed to be manipulated (referred to as Factor 3) as well the mathematical action required (referred to as Factor 4 (see Murray, Clermont & Brinkley, 2005, p.190-191)). It should be acknowledged that although these statements had not been previously empirically validated, they were theoretically constructed to guide item writing. Furthermore, while such statements had not been mapped onto the five ALLS levels, the set of statements had been designed to cover the full range of complexity of the ALLS. Such statements were then used to score items in terms of perceived complexity during the item development phase. Thirty statements were selected for inclusion in the current study (see Appendix 2) to supplement the publicly available scaled items. The breakdown of such statements in terms of their complexity score has been displayed in Table 12.

[19] Note that one of the publicly available items (titled MEDCO ASPRIN) requires copyright permission if it was to be used in the current study.

Table 12: The number of selected ALLS Complexity Statements for Numeracy

ALLS Complexity Score	No. of statements selected
5 (hardest)	6
4	6
3	9
2	5
1 (easiest)	3
Sub-total	29

In summary, a potential pool of 58 ALLS Numeracy statements/items and a potential pool of 41 ALLS literacy statements/items drawn from a variety of sources were identified. Table 13 displays the breakdown of statements/items per domain and ALLS level.

Table 13: Summary of ALLS statements/items for each domain and level

Total:	Level 1	Level 2	Level 3	Level 4	Level 5	Total
Prose	3	4	5	5	3	20
Document	3	4	5	4	5	21
Literacy (P+D)	*6*	*8*	*10*	*9*	*8*	*41*
Numeracy	*8*	*11*	*15*	*11*	*13*	*58*
Total	*14*	*19*	*25*	*20*	*21*	*99*

3.2 PANELLING THE ITEM CONTENT

A one day workshop was held to panel the draft statements that were to underpin the design of the survey. Five adult literacy and numeracy specialists participated in the workshop. The panelists were selected based upon their familiarity and expertise in:

- Adult literacy and numeracy in general;
- The development, design features, purpose and/or use of the ACSF;
- The administration, design and/or calibration of the ALLS (and its predecessors); and/or
- Previous involvement in Stage 1 of the Mapping of Adult Literacy Performance Study[20].

[20] See Circelli, M., Curtis, D., & Perkins, K. (2011). Mapping Adult Literacy Performance, NCVER: Adelaide.

The ultimate aim of the workshop was to review the statements within the ACSF and ALLS prior to transferring the statements into an on-line survey format for testing in the pilot study. The panel reviewed the draft content that was to underpin the on-line survey, namely:

- The appropriateness of the ALLS descriptor statements as stand-alone items within the survey;
- Items to be used in the background questionnaire (in terms of relevance, coverage etc);
- Survey instructions (e.g., the ease of interpretation of the proposed item format for rating each statement); and the
- Marketing strategies and documentation.

As the ALLS level descriptors were primarily designed to describe the nature of the tasks to be performed, there was general agreement by the expert panel that these descriptions could be modified slightly to be expressed in terms of competency statements without compromising the original intent of each description. Accordingly, within small groups, the panel reviewed the original ALLS descriptors and where necessary, re-wrote the statements in terms of the specific skills and knowledge that underpinned each description. Once drafted, the revised statements were then reviewed by the entire group (both during and post workshop via a series of email exchanges), and where necessary, amended until group consensus was reached. The final modified ALLS descriptor statements used in the study have been presented in Appendix 1. Note that the original ALLS descriptors as well as the modified statements were also to be included in both the pilot study and final survey.

The draft items to be used in the background questionnaire were also reviewed by the panel in terms of relevance, coverage, clarity etc. The panel agreed that the background questions should be kept to a minimum. It was therefore suggested that the background questions be limited to the following:

- Where do you work?
- Where are you located? (state/territory)
- What type of program do you manage and/or teach within?
- Who are your learners?
- How much experience have you had in Adult Language, Literacy and/or Numeracy (years)
- In terms of the ACSF, what level of reading/numeracy do you think the learner is typically performing at?

There was general agreement that the background questions would provide sufficient information to assist with monitoring the characteristics of participants for targeted follow-up recruitment as well as to provide a description of the survey sample for subsequent reporting purposes.

The panel also reviewed the draft survey instructions and rating scales to be used to record raters' ratings against each statement within the survey. Such content and presentation were reviewed in terms of clarity, succinctness, appropriateness, ease of responding and potential for engagement. In general, the proposed item format was well received with a small number of amendments suggested to improve the clarity and succinctness of the instructions and rating scales.

The panel also undertook the following activities:

- Consideration of the similarities and differences in the design and use of the ACSF and ALLS;
- Identification of potential raters for the pilot study;
- Review of the appropriateness of the target population and proposed recruitment strategies for the main data collection; as well as
- Recommendations for identifying organisations and/or individuals who could be directly approached to participate in the field study.

At the completion of the workshop, the panel also developed a set of summary statements for each Reading and Numeracy Level of the ACSF. The summary ACSF level statements were required in the current study to assist the raters to make an initial holistic judgement of the learner's reading/numeracy level. The survey was designed to link this initial judgement to the most appropriate form. Each form was to contain items (drawn from both the ACSF and ALLS) that were at or around the level of ability of the learner, as determined by the initial holistic judgement of the rater. See Appendix 8 for the final version of the ACSF Level Summary Statements for both Reading and Numeracy.

3.3 DESIGNING THE SURVEY FUNCTIONALITY

The on-line survey was designed to:

1. Collect data separately for the Reading and Numeracy domains.

2. Collect information on the complexity of each statement within the ALLS (i.e., Level Descriptors – original and modified as well as Numeracy Complexity Statements) and ACSF (i.e., Performance Features and Indictors). Each statement formed a separate item on the survey. The diagnostic information collected on each item may inform future revisions to the ACSF.

3. Randomly present items to:

    - Ensure that the ordered nature of the items *could not* be easily determined by the rater (i.e., which level on the framework the statement refers to);

- Avoid an item positioning effect as there is a tendency for items positioned toward the end of the survey to have more missing data or guessing responses due to raters' fatigue and/or boredom; and,

- Validate the measures used within the study to measure Reading and Numeracy for each framework.

4. Restrict raters from moving onto the next survey web page until the previous page has been completed to minimise missing data on some items that may be more difficult to rate.

5. Minimise rater workload by designing a number of alternative forms so that not all individuals have to complete the same set of items for each domain, nor does any one individual have to rate all items for each domain. This required designing alternative forms that had:

    o Common items across forms (to allow each form to be linked to another form so that all items within a domain could be calibrated onto a single scale); and

    o Statements/items that were likely to be at appropriate levels for the learner being rated.

6. Randomly present one of several forms based on the rater's holistic judgement.

7. Enable a rater to return to a partially completed form at a later time (i.e., save and resume functions).

3.4 DESIGING THE FORMS AND LINKS

3.4.1 LINKING FORMS

A suitable link design was developed to ensure an adequate sample of items appeared in multiple forms. This made it possible for all forms to be 'equated' onto a single scale of measurement. Separate forms were established for the Reading and Numeracy domain. Items within the forms were drawn from both the ACSF and the ALLS framework. The alternative forms contained common items (those that appear in more than one form) sourced from both the ACSF and ALLS. Furthermore, the forms were limited in the range of difficulty of the ACSF and ALLS statements to help ensure that raters rated against items that were likely to be relevant to their particular learner. By matching the range of item complexity within a particular form to the level of the estimated ability of the learner (as determined from the initial holistic judgement of the rater), the ratings would be more informative about the learners and about the items.

The study comprised 6 survey forms. That is, three forms related to the Numeracy domain and three forms related to the Reading domain. The general form design for each domain has been

illustrated in Table 14. Within each form, there were two sections: Part A and Part B. Part A contained statements drawn from both the ACSF and ALLS. Part B contained ALLS scaled items that were publicly available. Within Table 14, the 'X' represents the set of items that were within that particular form and survey section.

Table 14: General form design for a single domain

Survey Section	Component	Level represented	Item Pool		Form A	Form B	Form C
			Total	Sub-set			
Part A: Mapping statements of adult language, literacy and numeracy	ALLS Level Descriptors	5	X				X
		4	X			X	X
		3	X		X	X	X
		2	X		X	X	
		1	X		X		
	ALLS Complexity Statements (Numeracy only)	5		X			X
		4		X		X	X
		3		X	X	X	X
		2		X	X	X	
		1		X	X		
	ACSF Levels Indicators	5	X				X
		4	X			X	X
		3	X		X	X	X
		2	X		X	X	
		1	X		X		
	ACSF Performance Features	5		X			X
		4		X		X	X
		3		X	X	X	X
		2		X	X	X	
		1		X	X		
Part B: Mapping adult language, literacy and numeracy test items	ALLS Scaled Items	5		X			X
		4		X		X	X
		3		X	X	X	X
		2		X	X	X	
		1		X	X		

For example, it can be seen in Table 14 that for each domain, Form C comprised statements that were drawn from:

- A random sample of ACSF Performance Features at Levels 3, 4 and 5;
- The entire pool of ACSF level indicators at Levels 3, 4 & 5; and
- The entire pool of ALLS level descriptors (both original and modified by the panel) that were at aligned at Levels 3, 4 and 5 (see Appendix 1)
- A sub-set of the ALLS Complexity Statements (numeracy only) for Levels 3, 4 & 5.
- A sub set of ALLS scaled items (publicly available) aligned to Level 3, 4 or 5.

3.4.2 PART A: MAPPING STATEMENTS OF ADULT LANGUAGE, LITERACY AND NUMERACY

This section of the survey contained:

- The entire pool of ALLS level descriptors (both original and modified) as displayed in Appendix 1);
- a sample of the ALLS Complexity Statements (for the Numeracy domain only) (as displayed in Appendix 2);
- the entire pool of ACSF Levels Indicators; and
- a sample of ACSF Performance Features.

As previously noted, there were a relatively large number of ACSF Performance Features (see Table 7 and Table 8). To achieve the aim of mapping the ACSF to the ALLS without overburdening raters it was proposed that only a sample of all ACSF Performance Features be included in the survey forms. To include all Performance Features would result in either the forms becoming too long for any individual rater to complete or that insufficient ratings would be made against each Performance Feature. A more suitable approach was to randomly sample a set of statements from the complete list of Performance Features for inclusion in the study. According to this design, whilst not all Performance Features were included, a sizeable proportion was still calibrated. Furthermore, as the items have been selected via random sampling, inferences about the complexity of the Performance Features not included in the current study could still be made.

It should also be acknowledged that the Performance Features that appeared in one form at a particular level were the same as those that appeared at the same level in another form. For example, the set of ACSF Level 4 Performance Features that were included in Form B were exactly the same as the set of ACSF Level 4 Performance Features included in Form C. This was undertaken to ensure there were sufficient common items between forms for linking purposes and also to ensure that any Performance Feature in the study would have sufficient number of ratings made against it for analysis.

Table 15 and Table 16 display the number of items by component and form for Reading and Numeracy, respectively.

Table 15: Reading: Number of items per component by form.

Component	Level represented	Item Pool		Form A	Form B	Form C
		Total	Sampled			
ALLS Level Descriptors	5	4	4			4
	4	4	4		4	4
	3	7	7	7	7	7
	2	7	7	7	7	
	1	6	6	6		
ACSF Levels Indicators	5	2	2			2
	4	2	2		2	2
	3	2	2	2	2	2
	2	2	2	2	2	
	1	2	2	2		
ACSF Performance Features	5	13	7			7
	4	17	7		7	7
	3	24	7	7	7	7
	2	22	7	7	7	
	1	17	7	7		
Sub-total		131	73	47	45	42

Table 16: Numeracy: Number of items per component by form.

Component	Level represented	Item Pool		Form A	Form B	Form C
		Total	Sampled			
ALLS Level Descriptors	5	4	4			4
	4	4	4		4	4
	3	4	4	4	4	4
	2	3	3	3	3	
	1	3	3	3		
ALLS Complexity Statements	5	6	2			2
	4	6	2		2	2
	3	9	2	2	2	2
	2	5	2	2	2	
	1	3	2	2		
ACSF Levels Indicators	5	3	3	0	0	3
	4	3	3	0	3	3
	3	3	3	3	3	3
	2	3	3	3	3	0
	1	3	3	3	0	0
ACSF Performance Features	5	20	7	0	0	7
	4	16	7	0	7	7
	3	16	7	7	7	7
	2	15	7	7	7	0
	1	13	7	7	0	0
Sub-total		142	78	46	47	48

3.4.2.1 Part A: Number of links per form for the statements

A robust link design was created for Part A of the survey to ensure the ACSF and ALLS statements presented within different forms could be mapped onto a s ingle scale with confidence. As shown in Table 17, a total of 32 reading statements were common between Forms A and B. Similarly 29 statements in Form B also appeared in form C. In Form C, 16 of the statements also appeared in Form A. For example, the two ACSF Level 3 Indicators appeared in all 3 Forms while 7 sampled ACSF Level 4 Performance Features were only common to Forms B and C (and hence do not appear in Form A).

Table 17: Part A: Number of common/link statements for the Reading domain.

Component	Level represented	Common statements between forms		
		A to B	B to C	A to C
ALLS Level Descriptors	5	0	0	0
	4	0	4	0
	3	7	7	7
	2	7	0	0
	1	0	0	0
ACSF Levels Indicators	5	0	0	0
	4	0	2	0
	3	2	2	2
	2	2	0	0
	1	0	0	0
ACSF Performance Features	5	0	0	0
	4	0	7	0
	3	7	7	7
	2	7	0	0
	1	0	0	0
SUB-TOTAL		32	29	16

Table 18: Part A: Number of common/link statements for the Numeracy domain.

Component	Level represented	Common Statements between forms		
		A to B	B to C	A to C
ALLS Level Descriptors	5	0	0	0
	4	0	3	0
	3	4	4	4
	2	4	0	0
	1	0	0	0
ACSF Levels Indicators	5	0	0	0
	4	0	3	0
	3	3	3	3
	2	3	0	0
	1	0	0	0
ACSF Performance Features	5	0	0	0
	4	0	7	0
	3	7	7	7
	2	7	0	0
	1	0	0	0
ALLS Complexity Statements	5	0	0	0
	4	0	2	0
	3	2	2	2
	2	2	0	0
	1	0	0	0
SUB-TOTAL		32	31	16

3.4.3 PART B: MAPPING ADULT LANGUAGE, LITERACY AND NUMERACY TEST ITEMS

In Part B of the Survey, a subset of the ALLS publicly available scaled items was included in the design of the forms. A subset was selected to:

- minimise rater workload;
- maintain rater engagement;
- present those items that could be administered on-line (e.g., did not require access to any other resources such as newspapers etc); and to
- overcome difficulties experienced with sizing the images for the display of the items on-line.

As previously discussed, each of the ALLS scaled items had a difficulty value (expressed in terms of a scale score ranging from 0-500) and a band level (i.e., 1 to 5). In this section of the survey, the publicly available ALLS Scaled Items were incorporated and targeted at the ability level of the learner. This was achieved by matching the scaled item to the rater's initial holistic judgement of the

learner using the summary statements of the ACSF levels. Please note that although it has been assumed that the five levels on the ACSF and ALLS are roughly aligned for purposes of form design (at least within one band level either side), the actual relationship between the levels will be determined through the analysis of the survey data.

It can be seen in Table 19 that each form contains three to four items collected from similar and/or adjacent levels, as determined by their scaled scores and accompanying band levels. For example, Form A for Reading targets learners at ACSF Level 1, 2 or 3, thus four ALLS items were included on this form. Within the pool of items on each form, two items were randomly selected for the rater to make ratings.

Table 19: Part B: Allocation of ALLS scaled Items to each form.

Domain	Level	Image	Stem	Form A	Form B	Form C
Reading	1	FEW DUTCH WOMEN AT THE BLACKBOARD	Identify from the chart the percentage of teachers from Greece who are women.	Y		
Reading	2	IMPATIENS	What does the smooth leaf and stem suggest about the plant?	Y	Y	
Reading	2	IMPATIENS	What happens when the impatiens plant is exposed to temperatures of 14 degrees C or below?	Y	Y	
Reading	3	FIREWORKS IN THE NETHERLANDS	Write a brief description of the relationship between sales and injuries based on the information shown	Y	Y	Y
Reading	4	The Hiring Interview	Write in your own words one difference between the panel and the group interview		Y	Y
Reading	5	CANCO	List two ways in which CIEM (an employee support initiative within a company) helps people who lose their jobs because of departmental reorganization.			Y
Numeracy	1	COCA COLA BOTTLES	Find the total number of bottles in the two full cases shown in the picture.	Y		
Numeracy	1	Election Results	Determine the total number of votes cast.	Y		
Numeracy	2	FEW DUTCH WOMEN AT THE BLACKBOARD	Calculate the percentage of men in the teaching profession in Italy.	Y	Y	
Numeracy	2	GAS (PETROL) Gauge	The tank holds 48 gallons. How many gallons remain in the tank?	Y	Y	
Numeracy	3	FIREWORKS IN THE NETHERLANDS	Calculate how many more people were injured in 1989 than in 1988	Y	Y	Y
Numeracy	4	COMPOUND INTEREST	Calculate the total amount of money you will have if you invest $100 at a rate of 6% for 10 years.		Y	Y
Numeracy	5	DOUBLE YOUR INVESTMENT	Is it possible to double $1000 invested at this rate after seven years? Support your answer with calculations.			Y
Numeracy	5	IS BREAST MILK SAFE?	Compare the change in Dioxin level from 1975 to 1985 to the percent of change in Dioxin level from 1985 to 1995. Determine which percent of change is larger, and explain your answer.			Y

Appendix 3 displays the final set of publicly available ALLS Scaled Items that were included in the survey.

3.5 ALLOCATING FORMS

A rater's initial holistic judgement of the ACSF level of the learner (see Section 3.6 Background Information) was used to determine the appropriate form for a rater to complete. Table 20 shows

the relationship between the ACSF holistic judgement for a domain and the form that raters completed. For example, raters who rated their learner holistically at ACSF Level 4 had a 50% chance of completing Form B, a 50% chance of completing Form C, and were never asked to complete Form A.

Table 20: Relationship between rater holistic ACSF level judgement and administered form.

Holistic ACSF judgement	Form A	Form B	Form C
5			100%
4		50%	50%
3	33%	34%	33%
2	50%	50%	
1	100%		

3.6 SURVEY FORMAT

There were five main sections to the survey:

1. Welcome
2. Background Information
3. Part A: Mapping statements of adult literacy, language and numeracy
4. Part B: Mapping ALLS test questions
5. Entering the Draw

Each section is described below. It should also be acknowledged that two versions of the survey were created: a tablet/mobile/iPad version as well as a general version for personal computers, notebooks and desktop computers etc.

Welcome

In accordance with Victoria University's Human Research Ethics Committee (HREC) requirements, the first section of the survey was designed to provide an overview of the study and to gain voluntary consent to participate in the study (see Appendix 5). Information about the study, presented in plain English text was approved by the HREC in November 2011. The information was presented according to the following five key questions:

- What is this study about?
- What will I be asked to do?
- How will the information I give be used?
- What are the potential risks of participating in this project?
- Who is conducting the study?

Before proceeding to the next section of the survey, the raters then had to indicate (by ticking the boxes) that they had:

- Read and understood the information; and
- Voluntarily consented to participate in the study.

Background Information

The second section of the survey was designed to collect non-identifiable background information about the learner being rated as well as about the rater. Each is considered next.

Question	Response Formats
Which of the following best describes your workplace?	☐ Education &/or Training Institution o TAFE o Private RTO o Community based RTO o School Sector o University o Dual/multi sector ☐ Not for profit organisation o With RTO status o Without RTO status ☐ Private enterprise o With RTO status o Without RTO status ☐ Private consultant ☐ Public Service o Federal o State ☐ Other o Please list_____
Where are you located?	☐ VIC ☐ NSW ☐ ACT ☐ QLD ☐ TAS ☐ NT ☐ SA ☐ WA
How much experience have you had in Adult Language, Literacy and/or Numeracy (years)?	☐ Less than 1 year ☐ 1-5 years ☐ 6-10 years ☐ More than 10 years

Figure 1: Background Questions about the Rater

Question	Response Formats
In this survey we ask you to rate a learner against a number of skills descriptions. Think of a learner whose reading or numeracy ability you are familiar with. Please enter the first name of the learner in the box below.	☐ ** This is optional and does not need to be the learner's real name.*
In relation to <<The LEARNER>>[21], which program best matches the type of program s/he was or is currently undertaking?	☐ English As a Second Language (ESL) ☐ English Language Intensive Courses for Overseas Students (ELICOS) ☐ Workplace English Language & Literacy (WELL) ☐ Language, Literacy and Numeracy Program (LLNP) ☐ Adult Literacy and Numeracy Course ☐ Course in Applied Vocational Study Skills (CAVSS) ☐ VET Training Package qualification (with literacy &/or numeracy unit(s)) ☐ Senior Secondary Education qualification ☐ Higher Education qualification ☐ Other_____please list
Do you want to rate against <<the learner's**>> Reading or Numeracy skills?**	☐ Reading ☐ Numeracy
In terms of the Australian Core Skills Framework (ACSF), which of the following statements best describes the reading level of <<the learner>>?	☐ ACSF 1 (note a summary description has been included) ☐ ACSF 2 (note a summary description has been included) ☐ ACSF 3 (note a summary description has been included) ☐ ACSF 4 (note a summary description has been included) ☐ ACSF 5 (note a summary description has been included)

Figure 2: Background Questions about the Learner

Note, that the 'type of survey to be completed' in combination with the rater's holistic judgement of the learner's ACSF Level was used to assign an appropriate form to the participant.

PART A: Mapping statements of adult literacy, language and numeracy

Each of the statements drawn from both the ACSF (Indicators and Performance Features) and ALLS (original and modified Level Descriptors as well as Complexity Statements for Numeracy)

[21] Note: that if a name is supplied earlier (e.g. Bob), then this name will appear instead of the term "the learner".

formed a separate item within the survey. An illustration of Part A items has been shown in Figure 3.

Part B: Mapping ALLS test questions

Similar to the previous section of the survey, raters were asked to judge how likely the learner would be able to independently answer each question successfully. In this section of the survey, the original ALLS Scaled Item was included (i.e., stimulus and instructions) and the rater was instructed to use the scroll features to view the entire image (if required). An example of how each item was displayed in the survey has been provided in Figure 4 and Appendix 3 displays all the scaled items that were included in the survey.

A full listing of all items included in the survey (both Part A and Part B) can be found in Appendix 10.

Entering the Draw

In the final section of the survey, raters were presented with the Terms and Conditions for the Draw (in accordance with the Permit No ACT TP 11/04754.3), and were invited to enter the draw to win one of six iPad 3rd generation (32GB WiFi valued at $649 each). If they agreed to the Terms and Conditions (as indicated by ticking a box), they were asked to supply the following contact details:

- First name (required)
- Last Name (required)
- Email address (required)
- Mobile (optional - for notifying winners via SMS)

At the completion of the survey, raters were then encouraged to complete another survey for another learner, or the same learner in a different domain by clicking on a direct link to the survey's homepage.

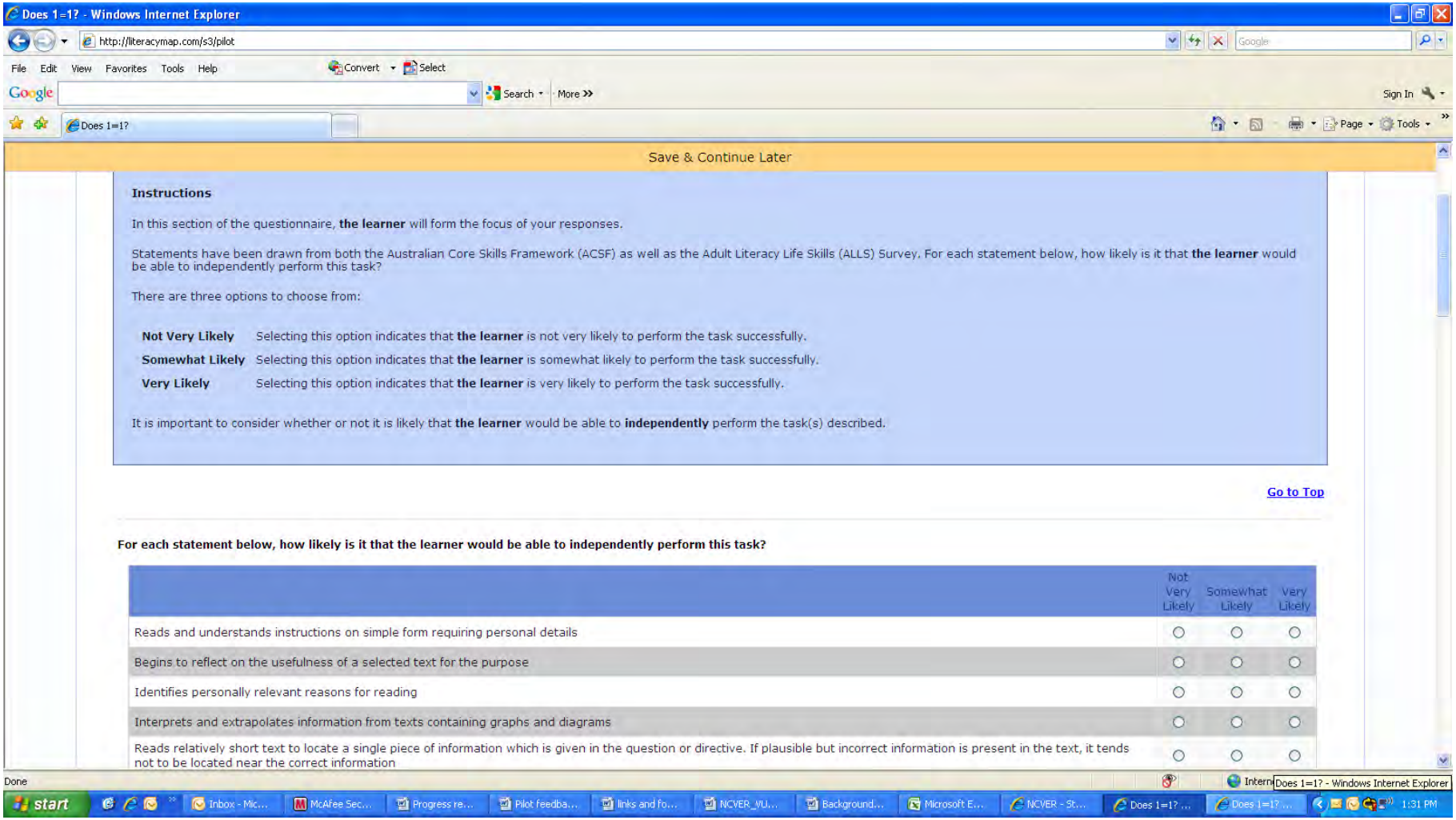


Figure 3: Item Format for the ACSF and ALLS statements: Reading Example

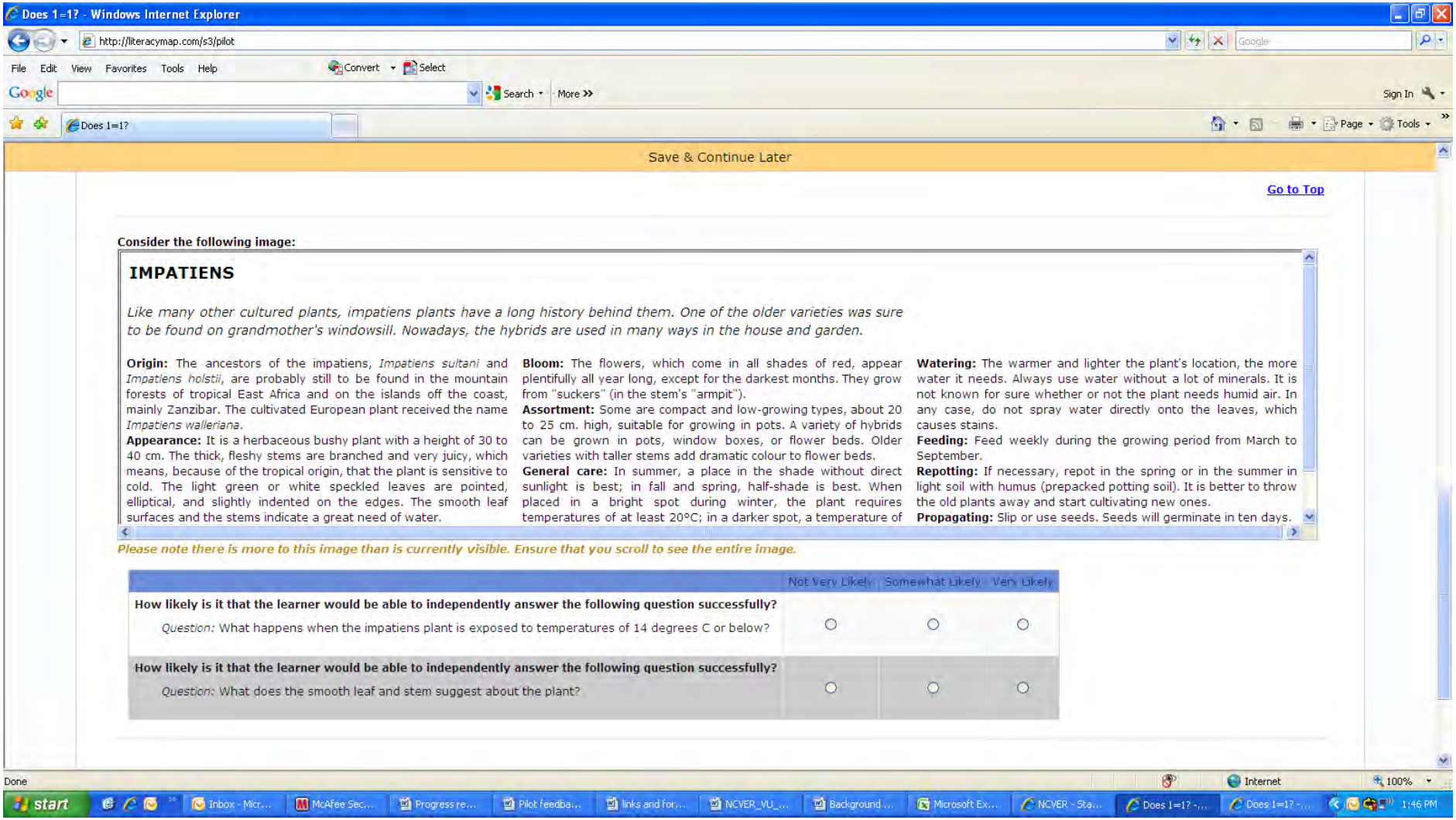


Figure 4: Item Format for the publicly released ALLS Scaled Items: Reading Example

CHAPTER 4: RESULTS

This chapter firstly describes the background characteristics of the raters who participated in the survey, and the characteristics of the learners rated by the raters. The second part of this chapter, Section 4.2, presents the findings from the Item Response Theory analysis. Firstly, the internal reliability and validity of the measures for each framework are explored (Section 4.2.1 and 4.2.2, respectively). Next, the relationship between the levels within and across each framework is examined for each domain (Section 4.2.3). The final section of this chapter reports on two supplementary analyses that were undertaken to provide further evidence of the validity of the study.

4.1 THE SAMPLE

4.1.1 The Raters

Four hundred and eleven (411) raters submitted completed surveys.[22] Each of the 411 raters focused on either the Reading or Numeracy skills of a learner they were acquainted with. The breakdown of the number of raters per domain is displayed in Table 21.

Table 21: Number of raters per domain

Domain	Frequency	Percent
Numeracy	117	28
Reading	294	72
Total	411	100

It can be seen in Table 21 that 72% of raters focused on Reading and 28% of raters focused on Numeracy.

Figure 5 is a bar chart that displays the location of the 411 raters according to each Australian state/territory. The figure presents percentages (%) and the number of raters (n) for each location. The majority of raters were from Victoria and New South Wales (40% and 23%, respectively). Very few raters were located in the Australian Capital Territory (1%).

[22] 411 completed surveys were received. For convenience, this is referred to as 411 raters. However, it is important to note that the 411 raters need not be 411 different people. Rather, in this context a rater refers to a unique individual / learner / domain combination.

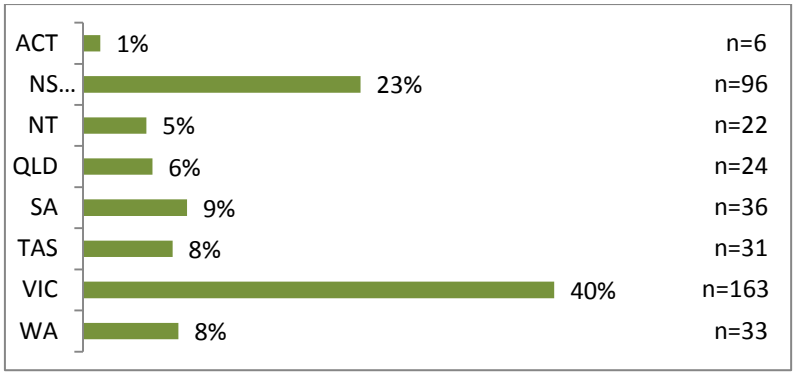


Figure 5: Location of raters

The raters were also required to describe their workplace. Figure 6 displays the percentage (%) of raters as well as the number of raters (n) who selected each role category.

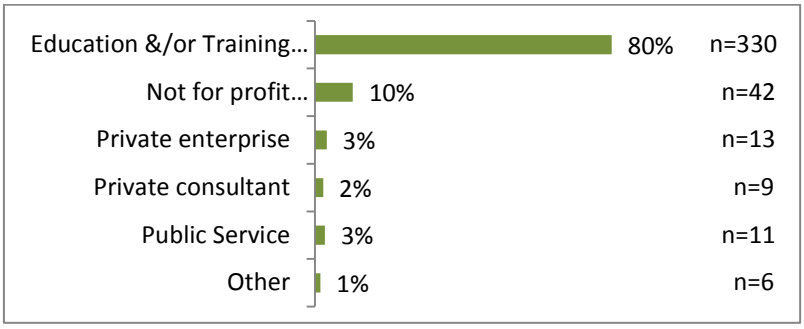


Figure 6: Description of workplace

The majority of raters worked in an Education and/or Training Institution (80%), followed by a Not for Profit Organisation (10%). Eight percent of raters identified themselves as working in a private enterprise or Public Service or working as a private consultant. One percent of raters indicated that their workplace was outside those listed in the survey. The specific roles identified by these raters can be found in Table 22.

Table 22: Description of 'Other' Type of Workplace

Workplace Type	Frequency	Percent
Adult Literacy	1	17%
CSNSW (Gaol)	1	17%
Retired TAFE LLN teacher and NRS assessor, now a volunteer adult literacy/numeracy volunteer	1	17%
Student	2	33%
Volunteer Read Write Now	1	17%
Total	6	100%

For particular workplaces, raters were asked further questions to describe their workplace. For example, raters who indicated that they worked in an education and/or training institution were then

asked to specify the type of institution. A summary of workplace sub-categories is displayed in Table 23.

Table 23: Number of raters in workplace sub-categories

Workplace	Total Number of raters	Workplace category	Number of raters for each category
Education &/or Training Institution	330	TAFE	185
		Community Based RTO	21
		Dual/Multi sector	28
		Private RTO	32
		School Sector	9
		University	55
Not for profit organisation	42	With RTO status	37
		Without RTO status	5
Private enterprise	13	With RTO status	10
		Without RTO status	3
Public Service	11	Federal	1
		State	10

In addition to identifying their workplace, each rater was asked to specify their experience in Adult Language, Literacy and/or Numeracy. Figure 7 displays the overall results for all raters, and Figure 8 displays results according to each domain, Reading and Numeracy.

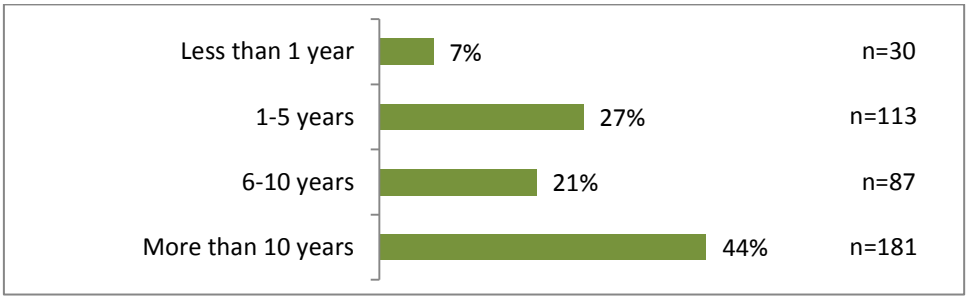


Figure 7: Experience in Adult Language, Literacy and/or Numeracy

The majority of raters had more than 10 years experience (44%). Twenty one percent of raters had between 6 to 10 years experience in Adult Language, Literacy and/or Numeracy and 27% of raters had between 1 to 5 years experience. Seven percent of raters indicated that they had less than one year experience.

The level of experience of raters within each domain was similar to the results displayed above. Figure 8 displays the level of experience for raters according to each domain.

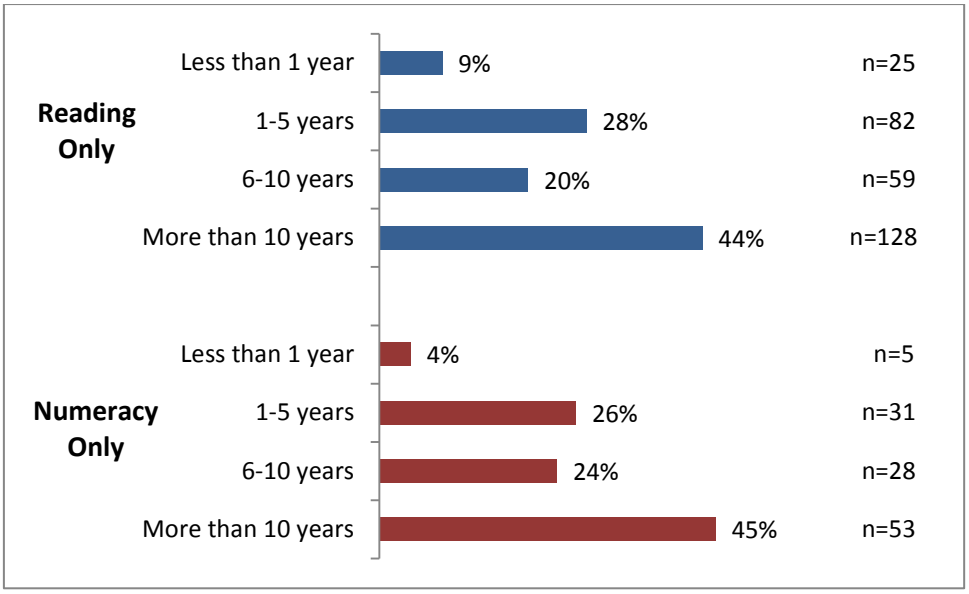


Figure 8: Experience in Adult Language, Literacy and/or Numeracy within each domain

It can be seen that in both domains, the majority of raters had more than 10 years of experience in the adult language, literacy and/or numeracy field (44% for the Reading domain and 45% for the Numeracy domain).

4.1.2 The Learners

Raters were also asked to think of a learner who they were familiar with in terms of their Reading or Numeracy ability. Raters then selected a program that best matched the type of program the learner had completed or was currently undertaking. The results of this question are displayed in Table 24.

Table 24: Programs undertaken by the learners

Program	Reading	Numeracy	Total
Adult Literacy and Numeracy Course	44	37	81
Course in Applied Vocational Study Skills (CAVSS)	1	4	5
English As a Second Language (ESL)	76	9	85
English Language Intensive Courses for Overseas Students (ELICOS)	4	0	4
Higher Education qualification	34	9	43
Language, Literacy and Numeracy Program (LLNP)	64	26	90
Senior Secondary Education qualification	13	4	17
VET Training Package qualification (with literacy &/or numeracy unit(s))	32	17	49
Workplace English Language & Literacy (WELL)	15	9	24
Other	11	2	13
Total	**294**	**117**	**411**

As seen in Table 24, most raters who chose to focus on the Reading domain selected a familiar learner who was undertaking, or had previously undertaken, an ESL program or an LLNP program (76 raters and 64 raters respectively). Alternatively, most raters who chose to focus on Numeracy, nominated a familiar learner who had undertaken (or was currently undertaking) either an Adult Literacy and Numeracy Course or an LLNP program (37 raters and 26 raters respectively). Thirteen raters indicated that the program in which his/her learner had undertaken was outside those listed in the survey. The specific programs identified by these raters can be found in Table 25.

Table 25: Description of 'Other Type of Learning Program' not listed in Survey.

Other Type of Program	Frequency	Percent
10697	1	8%
Access10 program	1	8%
Certificate I	1	8%
Diploma in Electrotechnology	1	8%
Foundation Skills	1	8%
General Education	1	8%
IELTS	1	8%
Make a presentation	1	8%
Pathway course in humanities	1	8%
Prevocational course	1	8%
Skilled migrant program	1	8%
Skills Tasmania's Employer Pledge Program	1	8%
Volunteer Literacy Tutor program	1	8%
Total	13	100%

Each rater was presented with summary descriptions of each of the five levels of proficiency within the ACSF, and asked to pick the level that best described the ability of their learner. Table 26 presents the results of the initial holistic judgement made by the raters for each domain.

Table 26: ACSF Level of learners: Initial Holistic Judgement per domain

	Total	Percent
Reading		
Level1	98	33%
Level2	98	33%
Level3	69	23%
Level4	23	8%
Level5	6	2%
Subtotal	***294***	***100%***
Numeracy		
Level1	40	34%
Level2	39	33%
Level3	27	23%
Level4	7	6%
Level5	4	3%
Subtotal	***117***	***100%***
Total	**411**	

It can be seen in the table above that most raters chose learners whose ability was thought to be located at ACSF Levels 1 to 3, irrespective of the domain. Approximately 10% of raters in each domain specified that their learner was in the upper two levels.

Most Reading raters stated that their learner was at ACSF Level 1 or Level 2 (33% for each level) and twenty-three percent of Reading raters selected learners at ACSF Level 3. The remaining Reading raters selected learners thought to be at ACSF Level 4 or Level 5 (8% and 2%, respectively). Similar to the Reading raters, most Numeracy raters indicated that their learner was of ACSF Level 1 or Level 2 (34% in Level 1 and 33% in Level 2) and twenty-three percent of Numeracy raters chose learners thought to be at ACSF Level 3. Likewise, a small proportion of raters selected learners at ACSF Level 4 or Level 5 within the Numeracy domain (6% in Level 4 and 3% in Level 5).

As discussed in Chapter 3, the holistic judgement of the ACSF level of the learner was used to allocate an appropriate form within each domain. Table 27 presents the distribution of learners across each level and form per domain.

Table 27: Distribution of learners across levels and forms

	Frequency				Percentage			
	Form A	Form B	Form C	Total	Form A	Form B	Form C	Total
Reading								
Level1	98			98	100%			100%
Level2	46	52		98	47%	53%		100%
Level3	21	22	26	69	30%	32%	38%	100%
Level4		16	7	23		70%	30%	100%
Level5			6	6			100%	100%
Subtotal	***165***	***90***	***39***	***294***				
Numeracy								
Level1	40			40	100%			100%
Level2	18	21		39	46%	54%		100%
Level3	10	8	9	27	37%	30%	33%	100%
Level4		5	2	7		71%	29%	100%
Level5			4	4			100%	100%
Subtotal	***68***	***34***	***15***	***117***				

In Table 27 we can see that for each domain, raters who selected learners thought to be at ACSF Level 3 were evenly distributed across the three alternative forms (Form A, Form B, or Form C). Similarly, raters who selected learners thought to be at ACSF Level 2 were evenly distributed across Form A and Form B. For example, of the 98 Reading Level 2 learners rated, 47% were rated using Form A and 53% of learners were rated using Form B. Similarly, the Reading Level 3 learners were spread across the three forms (30% were rated using Form A, 32% were rated using Form B and 38% were rated using Form C).

In both domains the split of Level 4 learners was not equal across Form B and Form C. Approximately 70% of Level 4 learners were assigned Form B and 30% were assigned Form C. This skewed distribution was due to the small number of learners at this level (i.e., n=23 Reading Level 4 learners and n=7 Numeracy Level 4 learners).

4.1.3 The Time to Complete the Survey

It was estimated that raters should take between 10 and 20 minutes to complete the survey. Figure 9 displays the median time taken to complete the survey according to the perceived ACSF level of the learner (as determined by the initial holistic judgement of the rater).

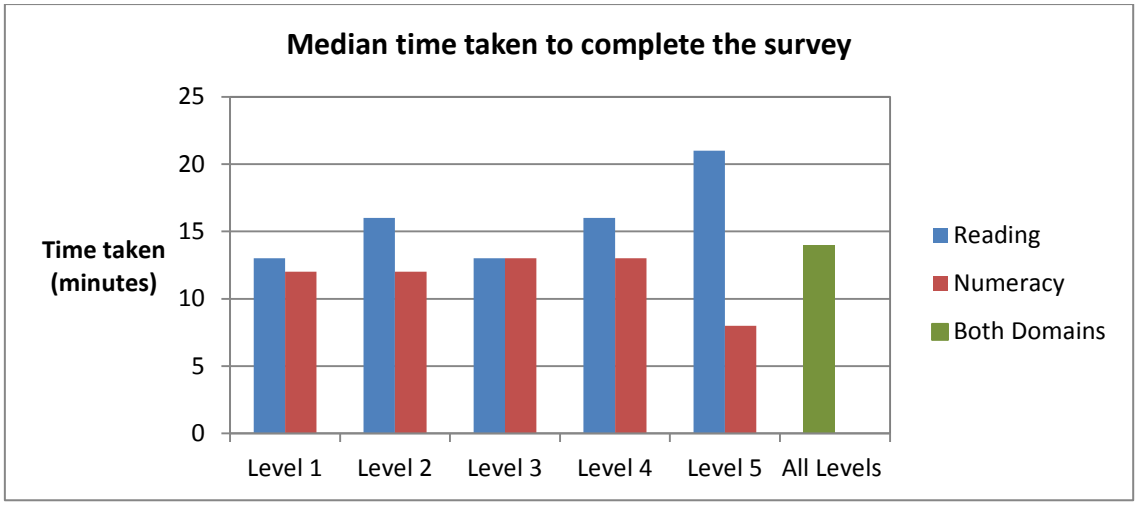


Figure 9: Median time to complete the survey according to the perceived ACSF Level of the Learner

The median time taken to complete the survey for all raters was 14 minutes. In general, the Reading survey took raters as long or slightly longer to complete than the Numeracy survey in each level. Given the small number of raters in Level 5, caution should be exercised when comparing median times at this level.

4.2 EMPIRICAL ALIGNMENT OF THE ACSF AND ALLS

The Reading items from both frameworks were analysed as a single set using an IRT analysis by fitting a Rating Scale model. Similarly, the Numeracy items from both frameworks were also analysed as a single set using the same approach as that used for the Reading domain (See Section 2.6 for further information about IRT).

4.2.1 Item statistics

A total of 165 items were originally included in the survey. The number of ratings made for each item per level in Part A of the survey has been shown in Table 28.

Table 28: Part A: Number of item ratings per level for each domain.

Level	Numeracy	Reading
1	**68**	**165**
2	**102**	**255**
3	**117**	**294**
4	**49**	**129**
5	**15**	**39**

It can be seen that the majority of the ratings were made against items aligned to Levels 2 and 3 for both the Reading and Numeracy domain, with very few ratings made against the items at Level 5

on both domains. Caution therefore should be exercised when comparing the highest complexity levels of both frameworks.

As outlined in Section 3.4.3, for Part B of the survey, two ALLS scaled items were randomly presented to each rater. The number of ratings varied from as low as 4 ratings for one ALLS Numeracy Level 5 item up to 199 ratings for ALLS Reading Level 3 items. Prior to IRT analysis, the two ALLS calibrated items in the Numeracy domain at Level 5 (as shown on pages 102 and 103) were excluded from analysis due to the very low number of ratings for these two items (4 and 7 ratings, respectively). Once these items had been excluded the next lowest number of ratings made on an ALLS calibrated item was 21 ratings. As such, 79 items were retained for the Reading scale and 84 items were calibrated for the Numeracy scale (163 items in total) using Item Response Theory.

For each of the 163 items, the relevant characteristics and IRT statistics have been presented in Table 29 (for Reading) and Table 30 (for Numeracy). The values presented in these tables largely form the basis of the remaining results and discussion. Hence, a brief description has been provided about each type of value/statistic below, along with a reference to later sections containing more detailed discussion of the results.

The first five columns in both Table 29 and Table 30 summarise the characteristics of each of the items. The column titled **'Item No.'** refers to the code that was assigned to each item by the Conquest Software program when undertaking the item response theory analysis for each domain (i.e., 1 to 79 for Reading and 1 to 84 for Numeracy). This column is used to identify the items presented in Figure 10 and Figure 11. The second column, labeled **'ID'** refers to the unique identification code that was assigned to the item by the research team to track items (refer to Appendix 10 for the description of the statement that relates to each unique identification code). The next three columns state the **'framework'** from which the item's content originated from (i.e, ACSF or ALLS); the expected **'level'** of the item (i.e., Level 1 to 5); and the **'component'** within the relevant framework in which the content of the item was drawn from (e.g., Performance Feature, Scaled Item).

The right hand side of the two tables presents the item statistics. The column labeled **'N'** refers to the number of ratings that were made per item. The next column, labeled **'Logit'** refers to the item's logit value and is a measure of item complexity (with higher values indicating higher levels of complexity for learners to demonstrate). By convention, the average learner ability within each domain was set equal to a logit value of zero. See Section 4.2.1.1 for more detail on the relationship between logit values, items and learners. The column labeled **'Error'** refers to the standard error of measurement associated with the logit value; and is an indication of the level of uncertainty associated with the complexity estimate of the item.

The column titled **'Fit'** refers to how well the responses to the item fit the theoretical IRT model used for analysis. Good fit to the model suggests that the items are measuring the same dominant

construct (in this case reading/numeracy). As the purposes of this study was to compare the complexity estimates of statements drawn directly from both frameworks, items were not excluded solely on the basis of their fit values. However, it is recommended that ACSF statements with large fit values in particular are reviewed in future revisions to the ACSF to ensure they are describing the skills and knowledge as originally intended. Mis-fitting items (those with large fit values) are discussed in detail in Appendix 11.

The last three columns within each table refer to transformed measures which have been included to aid readability of the results. That is, the column labeled '**Complexity'** is the linear transformation of the logit value displayed in the 'Logit' column. The lowest logit value for an item was set equal to 100 and the highest item logit value was set equal to 200, with higher scores indicating higher complexity, whilst the last two columns in the tables refer to the '**95% Confidence Interval'** of the complexity estimate (which were calculated using the logit values and error estimates reported in the 'Logit' and 'Error' columns). The general formula used to convert logit values into complexity estimates was:

$$Complexity = ((Est - MinEst)/(MaxEst - MinEst)))*(MaxComplexity - MinComplexity) + MinComplexity. \quad \text{(i)}$$

Specifically, the formulas used to create the Complexity values for each domain were:

$$Reading\ Complexity = ((Estimate - (-2.76))/(2.87 - (-2.76)))*(200 - 100) + 100. \quad \text{(ii)}$$

$$Numeracy\ Complexity = ((Estimate - (-3.54))/(3.25 - (-3.54)))*(200 - 100) + 100. \quad \text{(iii)}$$

Lower and Upper bound logit values for the 95% confidence interval were estimated using the formula:

$$LowerEstimate/UpperEstimate = Estimate^{-/+} 1.96*Error \quad \text{(iv)}$$

LowerEstimates and UpperEstimates produced using equation (iv) were then converted to Lower and Upper complexity values using the Complexity-Estimate conversion formulas as shown in equations (ii) and (iii).

Finally, the blue shading in the last several rows of Table 29 and Table 30 indicates the ALLS scaled items that were located in Part B of the survey. Those non-shaded refer to items contained within Part A of the survey. Those six items highlighted red (2 for Reading and 4 for Numeracy) indicate unexpected complexity estimates and are discussed in detail in *Section 4.2.2 Empirically validating the measures used for each framework.*

Table 29: Reading Item Characteristics and Statistics.

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
					N					Lower	Higher
1	R1I10301	ACSF	1	Indicator	165	-2.318	0.158	0.93	108	102.35	113.35
2	R1P10301	ACSF	1	Performance Feature	165	-1.958	0.153	1.03	114	108.92	119.57
3	R1P10306	ACSF	1	Performance Feature	165	-0.667	0.146	1.26	137	132.09	142.26
4	R1I10401	ACSF	1	Indicator	165	-1.403	0.147	1.10	124	118.99	129.22
5	R1P10405	ACSF	1	Performance Feature	165	-1.600	0.149	1.16	121	115.42	125.79
6	R1P10407	ACSF	1	Performance Feature	165	-1.060	0.146	1.16	130	125.11	135.28
7	R1P10408	ACSF	1	Performance Feature	165	-2.342	0.157	1.02	107	101.96	112.89
8	R1P10409	ACSF	1	Performance Feature	165	-2.293	0.156	1.12	108	102.86	113.73
9	R1P10411	ACSF	1	Performance Feature	165	-1.801	0.150	1.32	117	111.81	122.26
10	R2I20301	ACSF	2	Indicator	255	-1.336	0.121	0.87	125	121.08	129.51
11	R2P20301	ACSF	2	Performance Feature	255	-1.803	0.125	1.04	117	112.65	121.35
12	R2P20302	ACSF	2	Performance Feature	255	-0.530	0.118	0.92	140	135.50	143.72
13	R2P20311	ACSF	2	Performance Feature	255	0.012	0.118	0.98	149	145.13	153.34
14	R2P20312	ACSF	2	Performance Feature	255	-0.392	0.118	0.93	142	137.95	146.17
15	R2I20401	ACSF	2	Indicator	255	-0.405	0.118	0.88	142	137.72	145.94
16	R2P20402	ACSF	2	Performance Feature	255	-0.963	0.119	1.03	132	127.78	136.06
17	R2P20403	ACSF	2	Performance Feature	255	-1.249	0.120	0.90	127	122.66	131.02
18	R2P20411	ACSF	2	Performance Feature	255	-0.503	0.118	1.74	140	135.98	144.20
19	R3I30301	ACSF	3	Indicator	294	0.683	0.114	0.75	161	157.19	165.12
20	R3P30302	ACSF	3	Performance Feature	294	0.003	0.111	0.94	149	145.21	152.94
21	R3P30303	ACSF	3	Performance Feature	294	-0.603	0.112	0.92	138	134.41	142.21

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
					N					Lower	Higher
22	R3P30309	ACSF	3	Performance Feature	294	0.189	0.112	0.90	152	148.48	156.28
23	R3P30313	ACSF	3	Performance Feature	294	0.541	0.113	0.87	159	154.70	162.57
24	R3I30401	ACSF	3	Indicator	294	0.415	0.112	0.82	156	152.50	160.29
25	R3P30405	ACSF	3	Performance Feature	294	0.985	0.116	1.05	167	162.48	170.56
26	R3P30413	ACSF	3	Performance Feature	294	0.289	0.112	0.95	154	150.26	158.06
27	R3P30414	ACSF	3	Performance Feature	294	-0.232	0.111	1.64	145	141.04	148.77
28	R4I40301	ACSF	4	Indicator	129	2.689	0.190	0.71	197	190.17	203.40
29	**R4P40301**	**ACSF**	**4**	**Performance Feature**	**129**	**-0.312**	**0.169**	**0.89**	**143**	**137.60**	**149.36**
30	R4P40304	ACSF	4	Performance Feature	129	1.720	0.172	0.86	180	173.59	185.56
31	R4P40309	ACSF	4	Performance Feature	129	2.146	0.179	0.99	187	180.91	193.37
32	R4I40401	ACSF	4	Indicator	129	1.036	0.166	0.71	167	161.65	173.20
33	R4P40401	ACSF	4	Performance Feature	129	1.147	0.167	0.96	169	163.58	175.21
34	R4P40405	ACSF	4	Performance Feature	129	1.777	0.173	0.89	181	174.56	186.61
35	R4P40409	ACSF	4	Performance Feature	129	0.625	0.165	0.87	160	154.38	165.87
36	R4P40410	ACSF	4	Performance Feature	129	1.119	0.167	0.82	169	163.08	174.71
37	R5I50301	ACSF	5	Indicator	39	1.778	0.308	0.63	181	169.88	191.33
38	R5P50302	ACSF	5	Performance Feature	39	2.871	0.327	0.96	200	188.63	211.40
39	R5P50304	ACSF	5	Performance Feature	39	0.737	0.310	0.86	162	151.32	172.91
40	R5P50306	ACSF	5	Performance Feature	39	0.444	0.315	0.91	157	145.94	167.88
41	R5P50308	ACSF	5	Performance Feature	39	2.358	0.315	0.94	191	179.94	201.87
42	R5I50401	ACSF	5	Indicator	39	0.737	0.310	0.68	162	151.32	172.91
43	R5P50401	ACSF	5	Performance Feature	39	1.213	0.307	0.76	171	159.88	181.26
44	R5P50402	ACSF	5	Performance Feature	39	2.162	0.312	1.02	187	176.56	198.29

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings N	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
										Lower	Higher
45	R5P50407	ACSF	5	Performance Feature	39	0.929	0.309	0.70	166	154.77	176.28
46	R1LOrg01	ALLS	1	Level Description (Original)	165	-1.158	0.147	1.03	128	123.34	133.57
47	R1LOrg02	ALLS	1	Level Description (Original)	165	-1.866	0.152	1.05	116	110.59	121.17
48	R1LMod01	ALLS	1	Level Description (Modified)	165	-2.148	0.155	0.92	111	105.47	116.27
49	R1LMod02	ALLS	1	Level Description (Modified)	165	-2.519	0.160	1.04	104	98.71	109.85
50	R1LMod03	ALLS	1	Level Description (Modified)	165	-2.756	0.165	0.80	100	94.33	105.82
51	R1LMod04	ALLS	1	Level Description (Modified)	165	-2.269	0.156	0.82	109	103.29	114.15
52	R2LOrg01	ALLS	2	Level Description (Original)	255	0.468	0.121	0.96	157	153.12	161.55
53	R2LOrg02	ALLS	2	Level Description (Original)	255	0.630	0.122	0.97	160	155.97	164.46
54	R2LMod01	ALLS	2	Level Description (Modified)	255	-0.405	0.118	1.02	142	137.72	145.94
55	R2LMod02	ALLS	2	Level Description (Modified)	255	-0.308	0.118	0.75	144	139.44	147.66
56	R2LMod03	ALLS	2	Level Description (Modified)	255	-0.893	0.119	0.92	133	129.02	137.30
57	R2LMod04	ALLS	2	Level Description (Modified)	255	0.917	0.124	1.01	165	160.99	169.63
58	R2LMod05	ALLS	2	Level Description (Modified)	255	0.353	0.120	0.70	155	151.12	159.47
59	R3LOrg01	ALLS	3	Level Description (Original)	294	1.358	0.119	1.17	173	169.00	177.29
60	R3LOrg02	ALLS	3	Level Description (Original)	294	2.260	0.134	0.87	189	184.50	193.83
61	**R3LMod01**	**ALLS**	**3**	**Level Description (Modified)**	**294**	**-1.816**	**0.120**	**1.08**	**117**	**112.59**	**120.94**
62	R3LMod02	ALLS	3	Level Description (Modified)	294	2.137	0.131	1.07	187	182.42	191.54
63	R3LMod03	ALLS	3	Level Description (Modified)	294	1.855	0.126	1.08	182	177.59	186.36
64	R3LMod04	ALLS	3	Level Description (Modified)	294	0.892	0.115	0.86	165	160.86	168.87
65	R3LMod05	ALLS	3	Level Description (Modified)	294	1.430	0.120	1.03	174	170.25	178.60
66	R4LOrg01	ALLS	4	Level Description (Original)	129	2.547	0.187	0.90	194	187.75	200.77
67	R4LOrg02	ALLS	4	Level Description (Original)	129	2.310	0.182	0.86	190	183.72	196.39

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings N	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval Lower	Higher
68	R4LMod01	ALLS	4	Level Description (Modified)	129	2.310	0.182	1.21	190	183.72	196.39
69	R4LMod02	ALLS	4	Level Description (Modified)	129	2.342	0.182	0.84	191	184.29	196.96
70	R5LOrg01	ALLS	5	Level Description (Original)	39	2.358	0.315	0.75	191	179.94	201.87
71	R5LOrg02	ALLS	5	Level Description (Original)	39	1.969	0.310	0.83	184	173.20	194.79
72	R5LMod01	ALLS	5	Level Description (Modified)	39	2.458	0.317	0.67	193	181.65	203.72
73	R5LMod02	ALLS	5	Level Description (Modified)	39	1.873	0.309	0.70	182	171.53	193.05
74	ALLS_R_Blac..	ALLS	1	Question (FEW DUTCH WOMEN)	100	-1.472	0.193	1.39	123	116.16	129.60
75	ALLS_R_Imp1..	ALLS	2	Question (IMPATIENS1)	186	0.594	0.141	1.15	160	154.67	164.48
76	ALLS_R_Imp2..	ALLS	2	Question (IMPATIENS2)	186	-0.044	0.138	1.14	148	143.44	153.05
77	ALLS_R_Fire..	ALLS	3	Question (FIREWORKS)	199	1.550	0.148	1.15	177	171.40	181.71
78	ALLS_R_Hiri..	ALLS	4	Question (The Hiring Interview)	82	-0.191	0.215	1.60	146	138.15	153.12
79	ALLS_R_CANC..	ALLS	5	Question (CANCO)	21	0.181	0.444	1.51	152	136.78	167.70

Table 30: Numeracy Item Characteristics and Statistics

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
					N					Lower	Higher
1	N1I10901	ACSF	1	Indicator	68	-2.419	0.248	0.89	117	109.35	123.67
2	N1P10901	ACSF	1	Performance Feature	68	-2.606	0.252	0.84	114	106.48	121.03
3	N1I11001	ACSF	1	Indicator	68	-2.543	0.251	1.15	115	107.44	121.93
4	N1P11005	ACSF	1	Performance Feature	68	-3.018	0.264	0.93	108	100.07	115.31
5	N1P11006	ACSF	1	Performance Feature	68	-2.814	0.257	1.33	111	103.27	118.11
6	**N1P11007**	**ACSF**	**1**	**Performance Feature**	**68**	**-1.422**	**0.231**	**0.82**	**131**	**124.52**	**137.86**
7	N1P11008	ACSF	1	Performance Feature	68	-2.496	0.248	0.91	115	108.22	122.53
8	**N1P11009**	**ACSF**	**1**	**Performance Feature**	**68**	**-0.997**	**0.230**	**1.08**	**137**	**130.81**	**144.09**
9	N1I11101	ACSF	1	Indicator	68	-2.255	0.242	1.00	119	111.94	125.91
10	N1P11101	ACSF	1	Performance Feature	68	-3.465	0.282	1.03	101	92.96	109.24
11	N2I20901	ACSF	2	Indicator	102	-1.683	0.198	0.85	127	121.63	133.06
12	N2P20901	ACSF	2	Performance Feature	102	-1.300	0.194	1.17	133	127.39	138.59
13	N2I21001	ACSF	2	Indicator	102	-0.966	0.191	0.87	138	132.40	143.42
14	N2P21002	ACSF	2	Performance Feature	102	-1.375	0.195	0.97	132	126.26	137.51
15	N2P21003	ACSF	2	Performance Feature	102	-1.842	0.201	0.79	125	119.21	130.81
16	N2P21004	ACSF	2	Performance Feature	102	-0.675	0.190	0.98	142	136.71	147.68
17	N2P21008	ACSF	2	Performance Feature	102	-1.528	0.196	0.99	130	123.97	135.29
18	N2P21009	ACSF	2	Performance Feature	102	-0.747	0.191	1.00	141	135.62	146.65
19	N2I21101	ACSF	2	Indicator	102	-0.747	0.191	1.06	141	135.62	146.65
20	N2P21102	ACSF	2	Performance Feature	102	-0.602	0.190	0.97	143	137.78	148.75

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
					N					Lower	Higher
21	N3I30901	ACSF	3	Indicator	117	0.244	0.180	0.95	156	150.53	160.92
22	N3P30902	ACSF	3	Performance Feature	117	-0.529	0.179	0.79	144	139.18	149.51
23	N3I31001	ACSF	3	Indicator	117	0.408	0.182	0.93	158	152.89	163.40
24	N3P31002	ACSF	3	Performance Feature	117	-1.116	0.183	0.79	136	130.42	140.98
25	N3P31003	ACSF	3	Performance Feature	117	-1.017	0.182	1.02	137	131.90	142.41
26	N3P31010	ACSF	3	Performance Feature	117	-0.143	0.179	0.93	150	144.86	155.20
27	N3P31012	ACSF	3	Performance Feature	117	0.607	0.183	1.12	161	155.79	166.36
28	N3P31013	ACSF	3	Performance Feature	117	0.674	0.184	1.45	162	156.75	167.37
29	N3I31101	ACSF	3	Indicator	117	-0.208	0.179	1.10	149	143.91	154.24
30	N3P31103	ACSF	3	Performance Feature	117	0.147	0.180	0.94	154	149.10	159.50
31	N4I40901	ACSF	4	Indicator	49	0.441	0.270	0.65	159	150.84	166.42
32	N4P40902	ACSF	4	Performance Feature	49	1.628	0.279	0.70	176	168.06	184.17
33	N4I41001	ACSF	4	Indicator	49	1.322	0.274	1.11	172	163.70	179.51
34	N4P41002	ACSF	4	Performance Feature	49	0.900	0.274	0.90	165	157.48	173.30
35	N4P41005	ACSF	4	Performance Feature	49	1.832	0.286	0.91	179	170.86	187.37
36	N4P41006	ACSF	4	Performance Feature	49	1.512	0.280	1.02	174	166.32	182.49
37	N4P41008	ACSF	4	Performance Feature	49	0.077	0.275	0.95	153	145.33	161.21
38	N4P41012	ACSF	4	Performance Feature	49	1.591	0.281	1.01	176	167.46	183.68
39	N4I41101	ACSF	4	Indicator	49	0.228	0.274	1.01	155	147.58	163.40
40	N4P41101	ACSF	4	Performance Feature	49	0.750	0.273	1.33	163	155.30	171.06
41	N5I50901	ACSF	5	Indicator	15	2.071	0.515	0.76	183	167.77	197.50
42	N5P50902	ACSF	5	Performance Feature	15	3.253	0.583	1.97	200	183.22	216.87
43	N5I51001	ACSF	5	Indicator	15	3.253	0.583	0.82	200	183.22	216.87

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings N	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval Lower	Higher
44	N5P51005	ACSF	5	Performance Feature	15	2.928	0.559	1.22	195	179.12	211.39
45	N5P51006	ACSF	5	Performance Feature	15	3.253	0.583	0.83	200	183.22	216.87
46	N5P51007	ACSF	5	Performance Feature	15	1.056	0.499	1.40	168	153.28	182.09
47	N5P51008	ACSF	5	Performance Feature	15	2.071	0.515	0.83	183	167.77	197.50
48	N5P51011	ACSF	5	Performance Feature	15	2.928	0.559	0.80	195	179.12	211.39
49	N5I51101	ACSF	5	Indicator	15	2.071	0.515	0.64	183	167.77	197.50
50	N5P51102	ACSF	5	Performance Feature	15	1.556	0.502	0.65	175	160.56	189.54
51	N1LOrg01	ALLS	1	Level Description (Original)	68	-2.297	0.245	0.99	118	111.23	125.38
52	N1LMod01	ALLS	1	Level Description (Modified)	68	-2.670	0.254	0.78	113	105.48	120.14
53	N1LMod02	ALLS	1	Level Description (Modified)	68	-3.538	0.287	0.70	100	91.74	108.31
54	N1LMod03	ALLS	1	Level Description (Modified)	68	-2.082	0.239	0.98	121	114.57	128.37
55	N2LOrg01	ALLS	2	Level Description (Original)	102	-0.784	0.191	0.87	141	135.08	146.10
56	**N2LMod01**	**ALLS**	**2**	**Level Description (Modified)**	**102**	**-2.174**	**0.207**	**1.19**	**120**	**114.14**	**126.09**
57	N2LMod02	ALLS	2	Level Description (Modified)	102	-0.276	0.190	1.16	148	142.59	153.56
58	N2LMod03	ALLS	2	Level Description (Modified)	102	-0.893	0.191	1.18	139	133.47	144.50
59	N3LOrg01	ALLS	3	Level Description (Original)	117	0.742	0.184	0.85	163	157.75	168.37
60	N3LMod01	ALLS	3	Level Description (Modified)	117	-0.852	0.181	0.82	140	134.36	144.81
61	N3LMod02	ALLS	3	Level Description (Modified)	117	0.811	0.185	0.90	164	158.74	169.42
62	N3LMod03	ALLS	3	Level Description (Modified)	117	0.375	0.181	0.73	158	152.43	162.88
63	N4LOrg01	ALLS	4	Level Description (Original)	49	2.640	0.306	0.93	191	182.18	199.85
64	N4LMod01	ALLS	4	Level Description (Modified)	49	1.946	0.285	1.11	181	172.57	189.02
65	N4LMod02	ALLS	4	Level Description (Modified)	49	1.551	0.278	1.08	175	166.95	183.00
66	N5LOrg01	ALLS	5	Level Description (Original)	15	1.810	0.507	1.32	179	164.16	193.43

Item Characteristics					Item Estimates						
Item No.	ID	Framework	Level	Component	No. of Ratings N	IRT Estimate			Linear Transformation		
						Logit	Error	Fit	Complexity	95% Confidence Interval	
										Lower	Higher
67	N5LMod01	ALLS	5	Level Description (Modified)	15	2.625	0.540	1.50	191	175.21	206.38
68	N5LMod02	ALLS	5	Level Description (Modified)	15	2.342	0.526	0.91	187	171.44	201.81
69	N1CCom01	ALLS	1	Complexity Statement	68	-1.339	0.233	1.11	132	125.69	139.14
70	N1CCom02	ALLS	1	Complexity Statement	68	-2.120	0.241	0.77	121	113.96	127.87
71	N2CCom01	ALLS	2	Complexity Statement	102	0.015	0.193	0.94	152	146.79	157.93
72	N2CCom02	ALLS	2	Complexity Statement	102	-0.856	0.191	1.30	140	134.02	145.04
73	N3CCom01	ALLS	3	Complexity Statement	117	1.054	0.188	1.27	168	162.23	173.09
74	N3CCom03	ALLS	3	Complexity Statement	117	0.708	0.184	1.02	163	157.25	167.87
75	N4CCom01	ALLS	4	Complexity Statement	49	0.805	0.270	1.23	164	156.20	171.78
76	N4CCom02	ALLS	4	Complexity Statement	49	1.322	0.274	1.19	172	163.70	179.51
77	N5CCom04	ALLS	5	Complexity Statement	15	2.158	0.502	1.01	184	169.43	198.41
78	N5CCom06	ALLS	5	Complexity Statement	15	1.305	0.500	0.88	171	156.92	185.79
79	ALLS_N_Cola	ALLS	1	Question (COCA COLA BOTTLES)	28	-2.380	0.376	1.15	117	106.23	127.94
80	ALLS_N_Elec..	ALLS	1	Question (Election Results)	25	-2.198	0.386	1.78	120	108.62	130.91
81	ALLS_N_Blac..	ALLS	2	Question (FEW DUTCH WOMEN)	43	0.100	0.296	1.44	154	145.06	162.15
82	ALLS_N_Gas	ALLS	2	Question (GAS (PETROL) Gauge)	48	-0.227	0.279	0.90	149	140.74	156.85
83	ALLS_N_Fire..	ALLS	3	Question (FIREWORKS)	52	-0.231	0.267	0.77	149	141.03	156.44
84	ALLS_N_Comp..	ALLS	4	Question (COMPOUND INTEREST)	27	0.378	0.349	1.34	158	147.63	167.78

4.2.1.1 RELATIONSHIP BETWEEN LEARNER ABILITY AND ITEM COMPLEXITY

Figure 10 and Figure 11 illustrate the relationship between the estimated ability of learners rated in the study and the complexity of the items for the Reading and Numeracy domains respectively. The vertical scale is shown in logit values. Learner ability estimates are distributed in the left hand panel (indicated with an 'x'). By convention, the average learner ability estimate was set equal to zero on the scale (Wu et.al., 2007)[23]. Each item complexity is represented by an item number on the right hand side (the number each item relates to can be found in column 1 (Item No.) of Table 29 and Table 30, whilst the actual logit value per item can be found in column 7 (Logit)).

More able learners and more difficult items are located towards the top of the chart, whilst less able learners and less complex items are found toward the bottom of the chart. For example, in relation to the Reading domain, it can be seen in Figure 10 (as well as Table 29) that the most complex Reading item within the survey was item no. 38 which referred to an ACSF Level 5 Performance Feature statement with a logit value of 2.871 (see Table 29). This indicates that the competencies described within this task would typically be more difficult for a learner to successfully complete independently compared to any of the other items presented within the survey. The least complex Reading item for a learner to independently succeed on was item no. 50 which referred to a panel modified ALLS Level 1 descriptor which had a logit value of -2.756. Note that a more detailed comparison of empirically calibrated item complexity measures with their designated theoretical levels has been addressed in Section 4.2.2 of this chapter.

Figure 10 and Figure 11 show that the items and learners cover similar ranges of the ability distribution, indicating an overall good match between learner ability and item difficulty. Overall, this suggests that the items used within the survey, for both domains, were suitably distributed across the learner ability distribution.

[23] Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. (2007). ConQuest (Version 2.0) [Computer Software]. Camberwell, Australia: ACER.

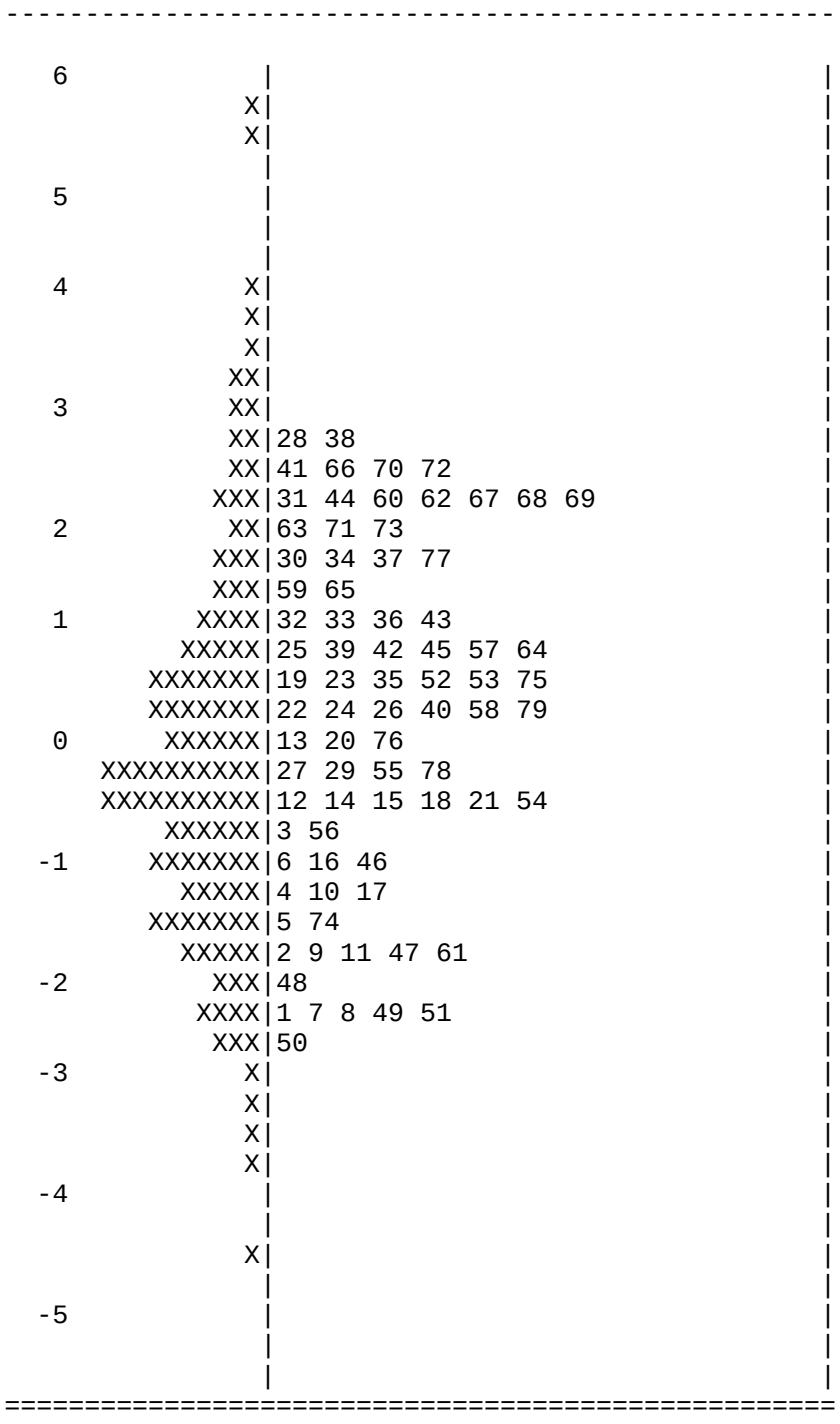


Figure 10: Reading domain Item-Learner map

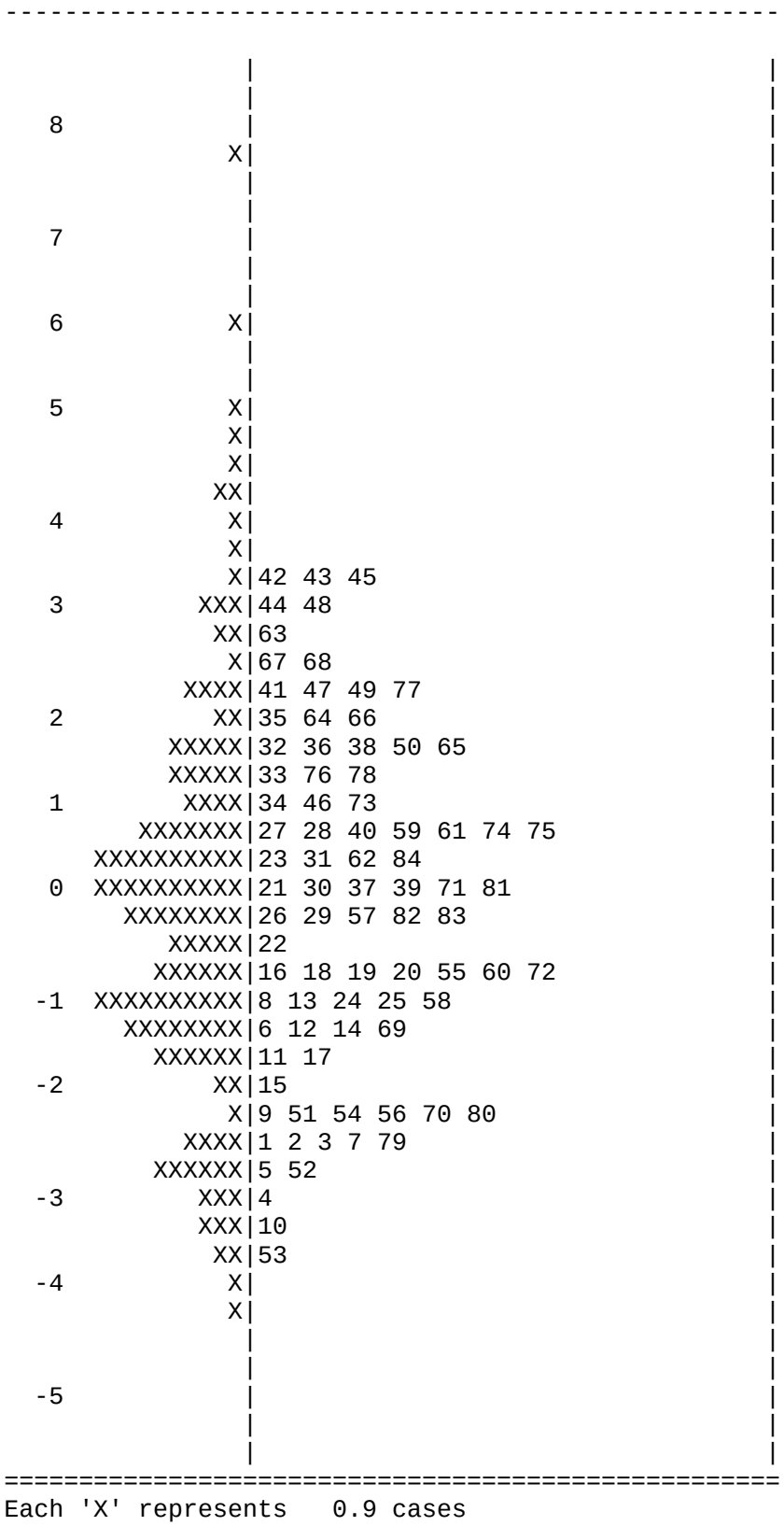


Figure 11: Numeracy domain Item-Learner map

4.2.1.2 RELIABILITY

The reliability of the learner ability estimates was very high across both the Numeracy and Reading domain. For the Numeracy domain, the survey had an overall mean reliability of 0.971. In the Reading domain, the survey had an overall mean reliability of 0.977.

Another indication of the reliability of the measures used within this study was the standard errors of measurement for each item. It can be seen in Table 29, that the measurement error for the reading items ranged from 0.111 logits (for items R3P30302 and R3P30414, both of which were Performance Features at ACSF Level 3) to 0.327 logits (for item ID R5P50302 which was an ACSF Level 5 Performance Feature). The measurement error for items below Level 5 was less than 0.2 logits, which indicates that one can be reasonably confident in the estimates obtained. However, in relation to the Reading items at Level 5, the standard errors of measurement were nearly twice the size (irrespective of the domain) of those at the lower levels. This indicates that one should be more cautious in making any comparisons of the complexity estimates at Reading Level 5.

In relation to Numeracy, the 84 items tended to have larger standard errors of measurement for the complexity estimates than those found for the 79 Reading items (see Table 30). This was due to the lower number of ratings made for the Numeracy items in comparison to the Reading items. Table 30 shows that the standard errors of measurement for the Numeracy items ranged from 0.179 logits (for Item ID N3P30902 which was an ACSF Level 3 item) to 0.583 logits (for Item ID N5P51006 which was an ACSF Level 5 Performance Feature Item). The large standard errors of measurement (i.e., ≥ 0.5 logits) were found at Numeracy Level 5, in which there were very few ratings made against such items. Given such findings, extreme caution should be exercised when making any comparisons within and across the frameworks at Level 5 Numeracy.

4.2.2 Empirically validating the measures used for each framework

Prior to examining how each of the ACSF levels related to the levels on the ALLS, it was first necessary to empirically validate the measures used within each framework by checking their expected level structures[24]. For example, to what extent did those items that were thought to be measuring ACSF Level 1 have lower complexity estimates than those thought to be at Level 2? And were Level 2 items found to be less complex than Level 3 and so on? Matching the expected sequencing of the items with the empirically calibrated complexity measures within each level for each framework would provide evidence of the validity of the measures used within the current study.

[24] Note that the expected level for each item was determined by its source and positioning within each of the frameworks (see Chapter 3)

The following four figures (Figure 12 to Figure 15) display the item sets for the Numeracy domain ACSF items, Numeracy domain ALLS items, Reading domain ACSF items and Reading domain ALLS items, respectively. The exact values used to produce these figures are displayed in Table 29 and Table 30 (i.e., columns labeled 'Complexity' and '95% Confidence Intervals').

Within each chart, the horizontal (x) axis represents the expected Level within each framework; the vertical (y) axis denotes the complexity estimate. Each item is represented as a shaded bar. The value range of the bar (as indicated by bar length) represents the 95% confidence interval surrounding the complexity estimate for the item. For example, at the 95% confidence level, the first Level 1 item shown for the ACSF Numeracy had a true complexity estimate that is between 93-109 on the complexity scale (see Table 30 for item number 10, unique identifier code N1P11101). A longer bar indicates more uncertainty about the complexity measure of the item. This increase in range is prevalent in items at higher levels, where the response rate was particularly low. For example, the size of the bar for Level 5 items was much larger than items at Level 2 and Level 3.

All four figures exhibit a general pattern of positive association between an item's complexity estimate and its expected level (as determined by the level it was originally assigned within each framework). That is, in general, items thought to be at higher levels were associated with higher empirically calibrated complexity (and items thought to be at lower levels were associated with lower calibrated complexity), indicating that the frameworks were developmental and hierarchical within themselves (as intended). In other words, the empirically derived complexity estimates were consistent with those hypothesised. Some overlap between levels should be expected, especially for frameworks designed to provide measurements against a continuous developmental scale. Due to the limited number of raters rating items at Level 4 and in particular Level 5, extra caution is needed when making observations at these levels.

There was however a small number of items with complexity estimates that were not consistent with their designated levels in the framework. These items were highlighted red in Table 29 and Table 30 for Reading and Numeracy, respectively. These unexpected outliers have also been highlighted within each of the charts displayed in Figure 12 to Figure 15. The following table (Table 31) summarises the list of highlighted items and their details.

Table 31: Complexity Estimate Outliers

Domain	Framework	Level	ID	Unexpected complexity estimate
Reading	ALLS	3	R3LMod01	Y – Easier than expected
	ALLS	4	ALLS_R_Hiring	Y – Easier than expected
	ALLS	5	ALLS_R_CANCO	Y – Easier than expected
	ACSF	4	R4P40301	Y – Easier than expected
Numeracy	ALLS	2	N2LMod01	Y – Easier than expected
	ACSF	1	N1P11007	Y – Harder than expected
	ACSF	1	N1P11009	Y – Harder than expected

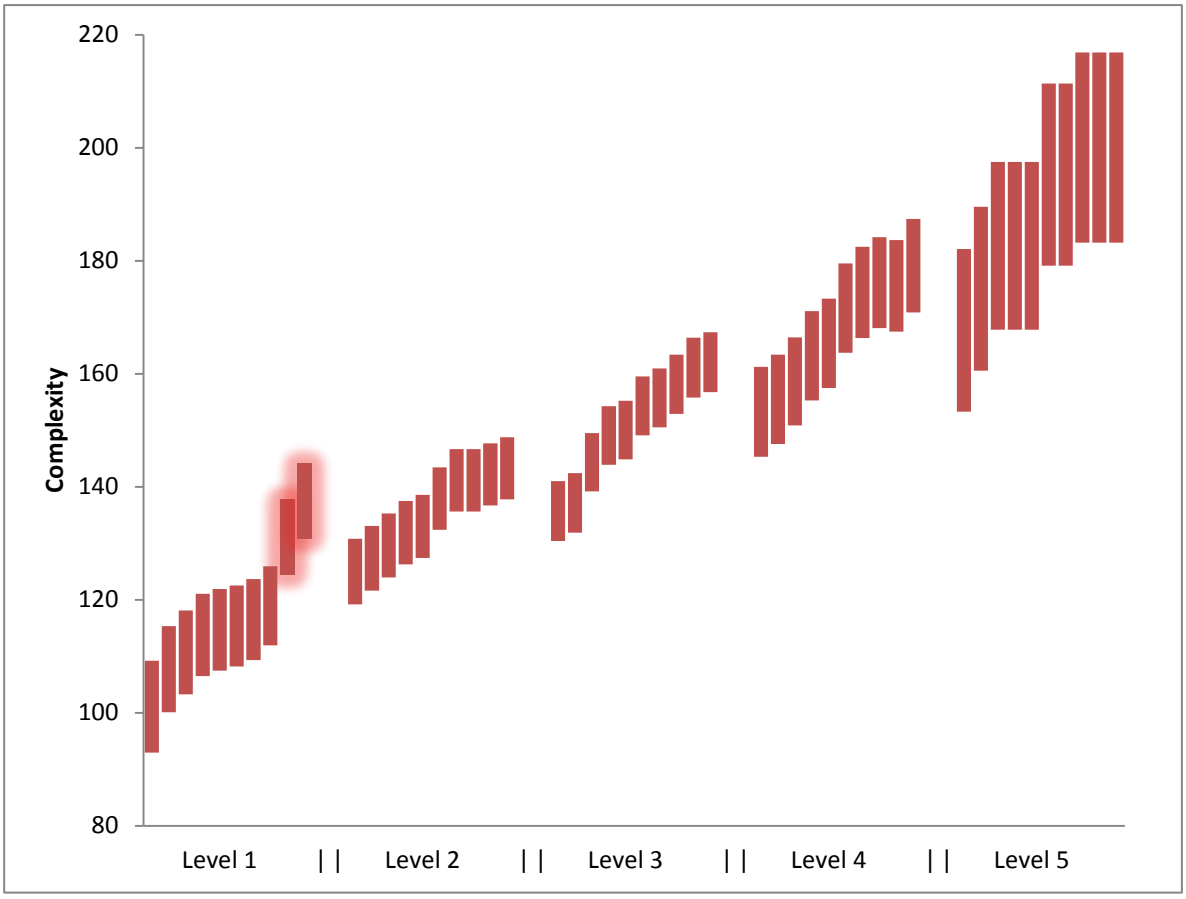


Figure 12: ACSF Numeracy Items according to Expected Level

Figure 12 illustrates that within the Numeracy domain, the ACSF items showed a general progression from low complexity to high complexity as the levels increased. Possible exceptions to

this were the two most complex items in Level 1 (namely N1P11007, N1P11009) which appeared to have complexity estimates similar to high Level 2 / low Level 3 items. These two items have been highlighted in the figure.

Note that the length of the bars for the items displayed in Level 5 (i.e. the complexity value range) were quite large, indicating that there was considerable uncertainty associated with the complexity estimates for these items. This is the result of fewer ratings made against these items. Conversely, the bar length of the items within the other levels was considerably less, indicating greater measurement precision.

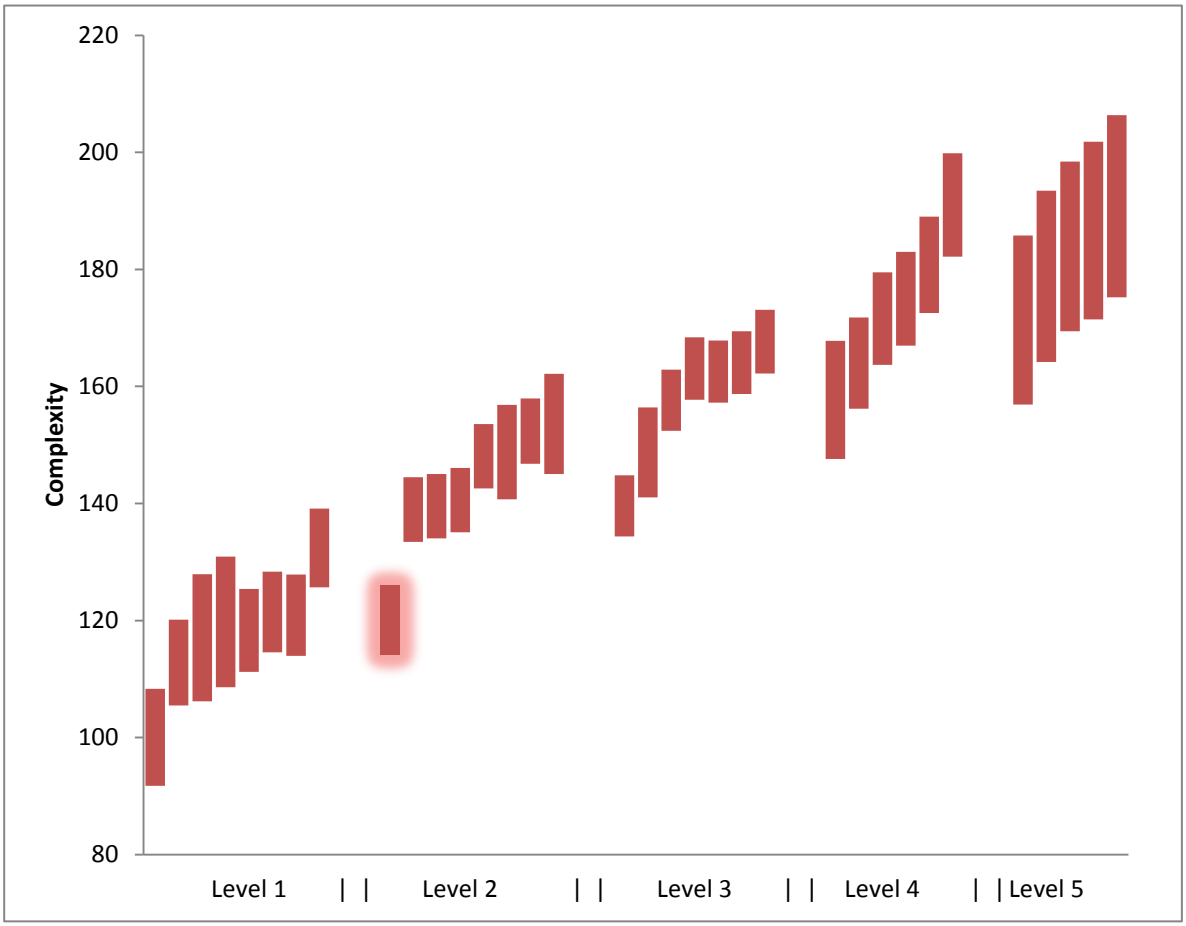


Figure 13: ALLS Numeracy Items according to Expected Level

Similar to the ACSF items, the ALLS items within the Numeracy domain also showed a general progression from low to high complexity as the levels increased. Note that the complexity scale displayed on the vertical axis in Figure 13 is the same as that used in Figure 12, as both frameworks were placed on the same scale of measurement. Hence, direct comparisons can be made. It can be seen in Figure 13 that one item at Level 2 appeared to be unexpectedly low in complexity. This item, N2LMOD01, is a modified version of the original ALLS descriptors.

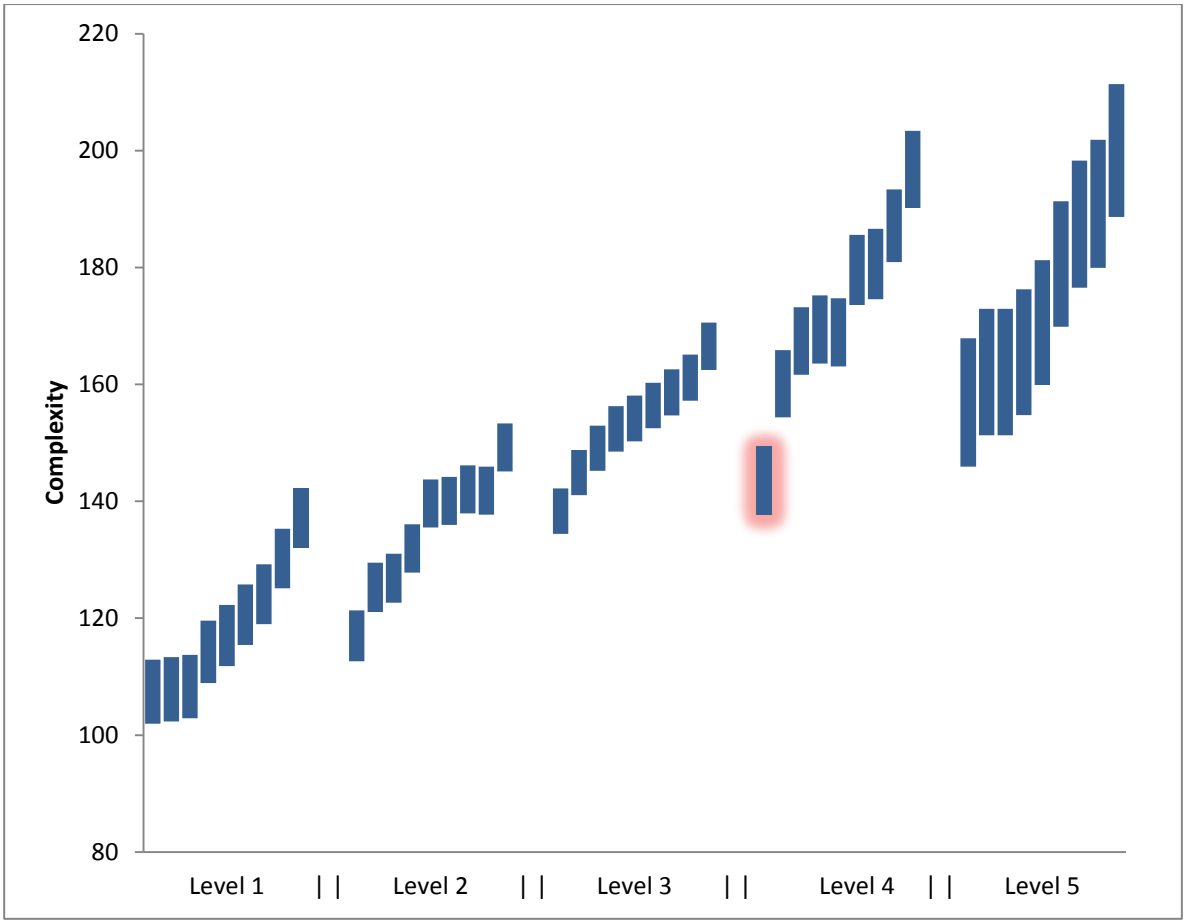


Figure 14: ACSF Reading Items according to Expected Level

The ACSF items within the Reading domain displayed in Figure 14 showed a similar pattern to the Numeracy items. Level 1 and Level 2 appeared to have considerable overlap, as was the case for items at Level 4 and Level 5. Within Level 4, one item (R4P40301) appeared to be lower in complexity than expected.

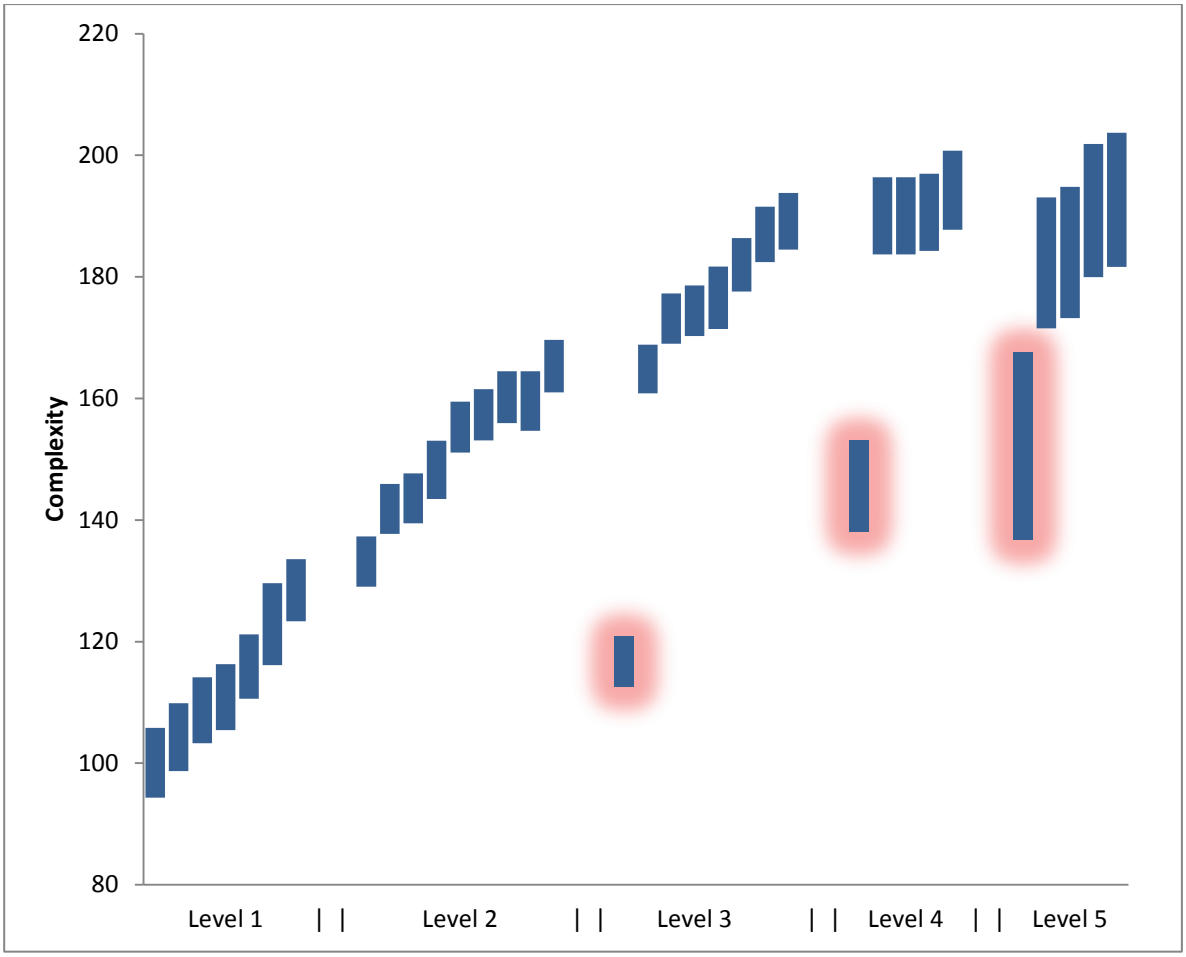


Figure 15: ALLS Reading Items according to Expected Level

As can be seen in Figure 15, the ALLS items within the Reading domain showed a steady increase in complexity across the lower levels (i.e., Levels 1 to 3). Note that there does not appear to be the overlap between Levels 1 and 2 as for ACSF Reading items. However, the progression of complexity from the top of Level 3 across to Level 5 appeared to be quite flat. This may be due to the language used to describe the complexity of the scale at these levels.

The most striking feature within Figure 15 was the three items which had much lower complexity estimates than was expected given their assigned ALLS level. The two outlying items at Levels 4 and 5 were scaled items (calibrated test questions) ALLS_R_Hiring and ALLS_R_CANCO respectively. In similar exercises for other large scale assessments we have found that it is sometimes challenging for subject matter experts to accurately judge the difficulty of test items. This may be one possible explanation for these findings in that raters underestimated the difficulty posed to learners within these questions. The outlying item at Level 3 was a modified ALLS descriptor (R3LMod01).

Overall, items within each domain and framework exhibit the expected characteristics of a continuous scale that increases in complexity across the levels. A small number of items (i.e., 7 of

the 165 items), however, had complexity estimates that were exceedingly lower (or higher) than the level the item was thought to represent (see Table 31 for a summary of these items). As the overall purpose of the study was to compare levels across frameworks the outliers were excluded from further analysis. The removal of such items ensured that the measures used within this study were valid measures of the Reading and Numeracy domains within both the ACSF and ALLS. Further evidence of the validity of the measures used within this study can be found in Section 4.4.

4.2.3 Determining the relationship across the frameworks - Does 1 = 1?

To compare the levels across the two frameworks, complexity estimates (as displayed in Table 29 and Table 30) were summarised by level into boxplots after removing extreme outliers detailed in Table 31. Boxplots for both frameworks have been shown in Figure 16 for the Numeracy domain and Figure 17 for Reading.

Each boxplot contains several pieces of statistical information. The 'box' indicates the middle 50% of items (known as the interquartile range). The top 'whisker' indicates the distance between the upper quartile and the most complex item (excluding any outliers). Similarly, the bottom 'whisker' indicates the distance between the lower quartile (bottom of the 'box') and the least complex item (excluding any outliers). The median is displayed as a black line within the box.

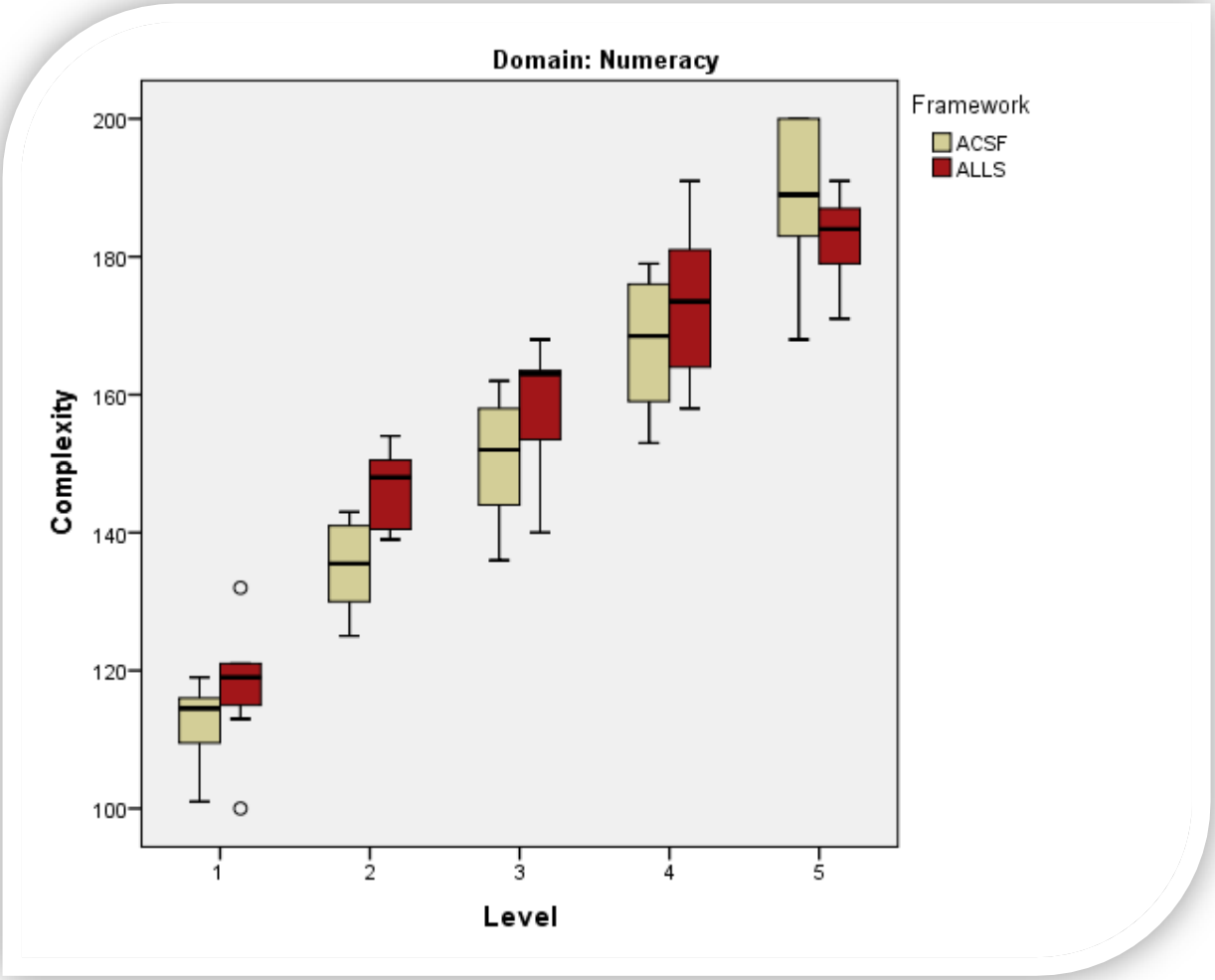


Figure 16: Numeracy domain boxplots by level and framework

In relation to the Numeracy domain, Figure 16 illustrates that ACSF Level 1 appeared to be similar to ALLS Level 1. However, ALLS Level 2 tended to be more complex than ACSF Level 2. Similarly, ACSF Level 3 was found to be distributed across a lower complexity range compared to ALLS Level 3, again indicating that ACSF Level 3 was less complex than that of ALLS Level 3. At Level 4 there was more overlap found between the two frameworks. Finally, the relationship between the two frameworks at Level 5 was not as clear. The large uncertainty surrounding the estimates at this level means we are unable to determine if this pattern is a property of the frameworks themselves or possibly an issue associated with the measurement uncertainty of the complexity estimates at this level (as a result of the low number of ratings (n=15) made against Numeracy items at Level 5, see Table 28).

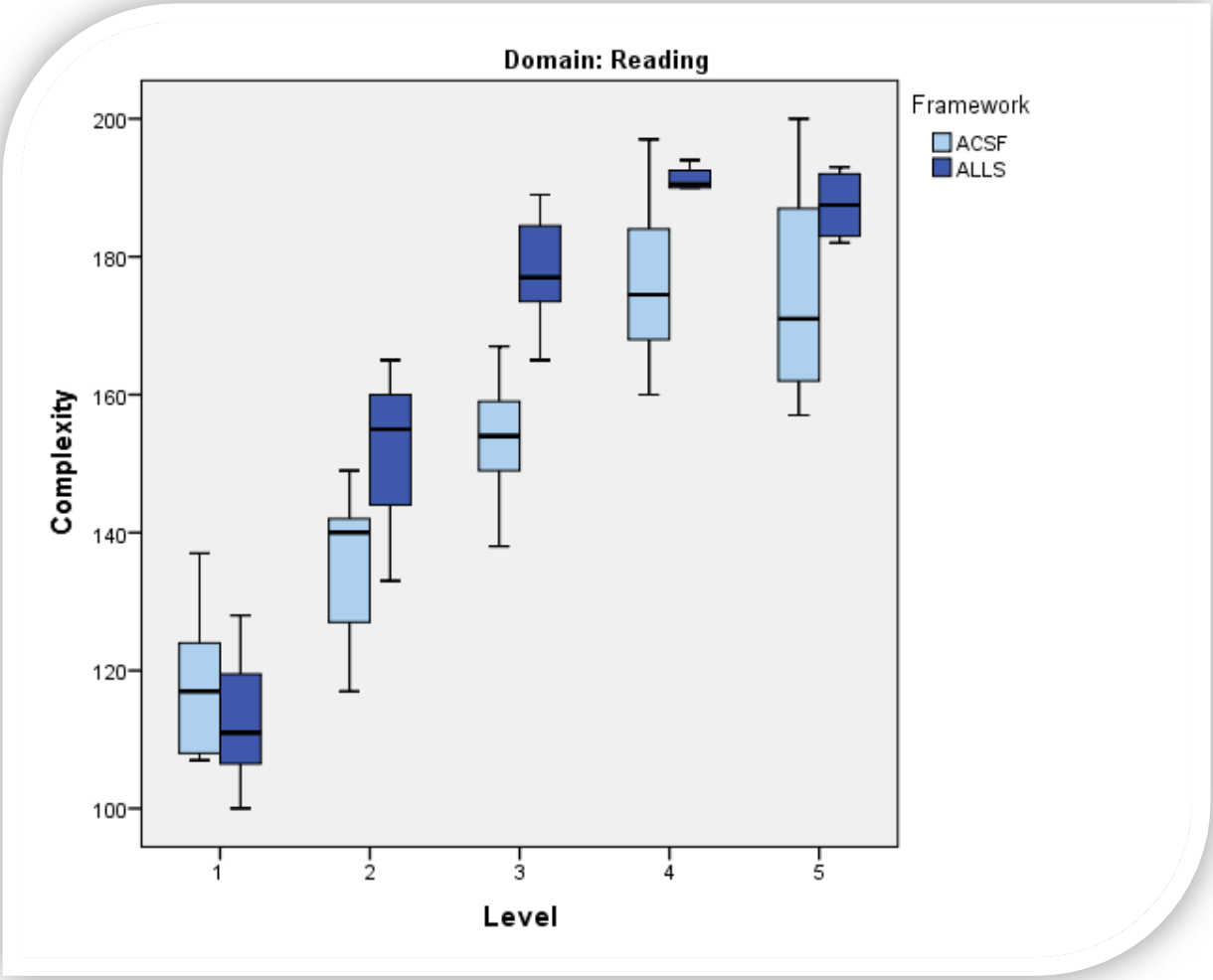


Figure 17: Reading domain boxplots by level and framework

As with the Numeracy domain, it can be seen in Figure 17 that ACSF Reading Level 1 appeared to be similar to ALLS Level 1 whilst ALLS Level 2 was found to be considerably more complex than ACSF Level 2. Similarly ALLS Level 3 was estimated to cover a higher range of complexity than ACSF Level 3. The differences between Levels 2 and 3 across the frameworks appeared to be more pronounced for the Reading domain than the Numeracy domain (see Figure 16 and Figure 17). At Reading Levels 4 and 5, it was more difficult to determine the relationship between the frameworks due to the low response rate for items at this level (see Table 28).

In summary, a similar relationship between frameworks was evident for both domains. Firstly, it appeared that at Level 1 both frameworks for each domain were similar in their complexity estimates. Secondly, and perhaps most importantly, it was evident that ALLS Level 2 and ALLS Level 3 appeared to be more complex than ACSF Level 2 and ACSF Level 3 respectively in both Reading and Numeracy, with the difference more pronounced for the Reading domain. In relation to Reading, it appeared as though ACSF Level 3 was more similar to ALLS Level 2 than ALLS

Level 3; and that ALLS Level 3 was more closely aligned to ACSF Level 4. A table summarising the indicative empirical relationship between the levels based on the results of this study has been shown in Table 32. Please note that as there were very few ratings made at Level 5 Numeracy (and thus producing large measurement error for each of the level 5 item's complexity estimates), the table below has deliberately omitted making any comparisons at Level 5 Numeracy.

Table 32: Empirical alignment of ACSF to ALLS by ACSF Level

Reading		Numeracy	
ACSF Level	**ALLS Level**	**ACSF Level**	**ALLS Level**
1	1	1	1
2	1-2	2	1-2
3	2	3	2-3
4	3	4	3-4
5	4-5	5	Uncertain

4.3. SUPPLEMENTARY EVIDENCE OF VALIDITY

In *Section 4.2.2 Empirically validating the measures used for each framework*, the replication of the original levels structure of each framework (for each domain) demonstrated that the items used in the survey were valid measures of Reading and Numeracy for both the ACSF and the ALLS. Such a replication of each framework's level structure indicates that the measures used within this study matched the real world (existing) structure of each framework.

Additional approaches to checking the validity of the measures used within this study was to compare the results to other external measures. In this case we have two external measures that could be used. The first was the holistic judgement of learner ability made by each rater when commencing the survey. The second was the published Scaled Score for the ALLS items used within this study.

4.3.1 Association between rater holistic judgement and learner ability estimate

At the start of the survey, raters were asked to estimate the ACSF level of the learner they had selected. The purpose of this initial question was to assist in identifying the most appropriate survey form to administer to the rater. Without making such a judgement, raters would be more likely to be presented with statements not appropriate to the ability of the learner they were rating. For example, there is not much point presenting ACSF Level 5 or ALLS Level 5 statements to a rater who is intending to rate a learner expected to be at about Level 1. The descriptions used for the holistic judgement were developed by the research panel (see Chapter 3). Figure 18 and Figure 19 illustrate the association between the holistic descriptor rating (ranging from Level 1 to Level 5) and the average learner ability as estimated using the IRT analysis.

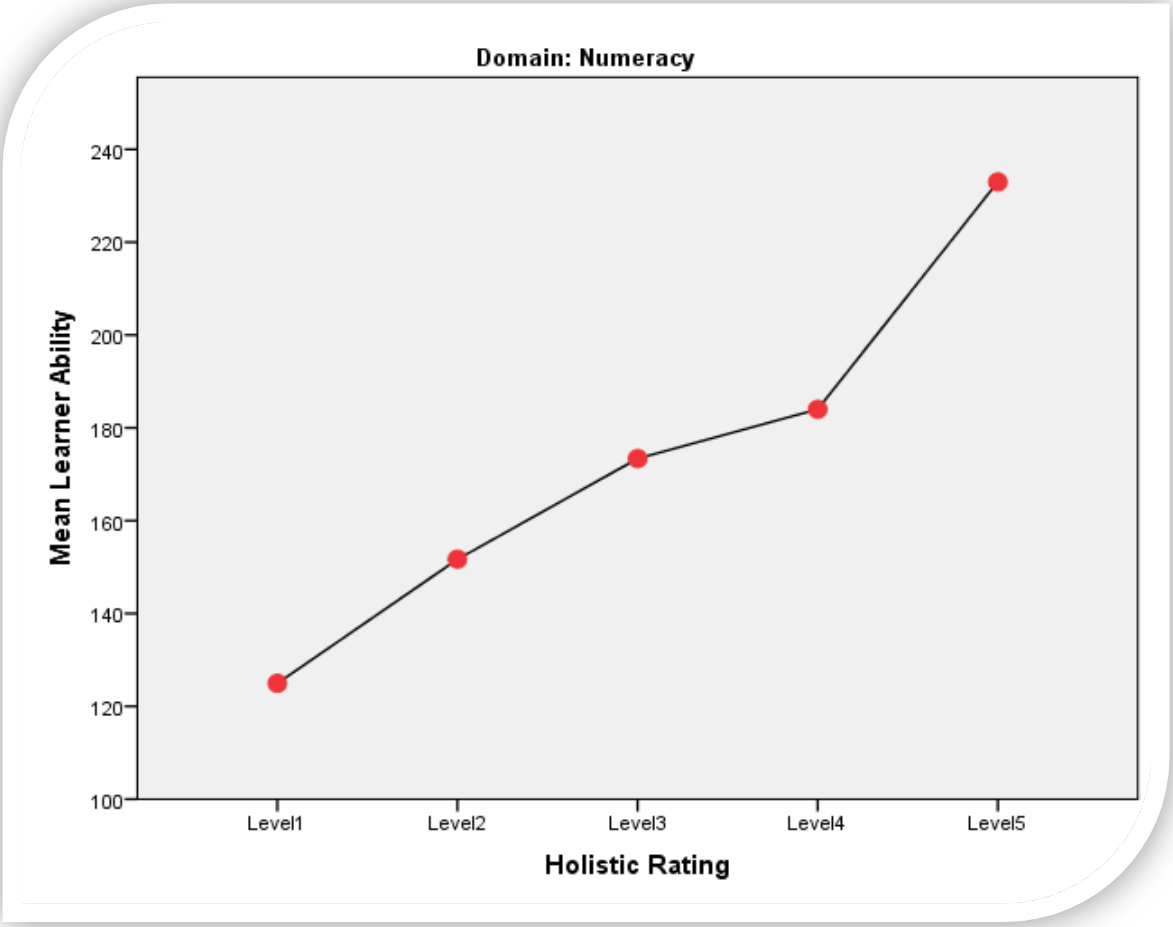


Figure 18: Relationship between Numeracy Holistic Judgement and Mean Level of Numeracy Ability of the learner

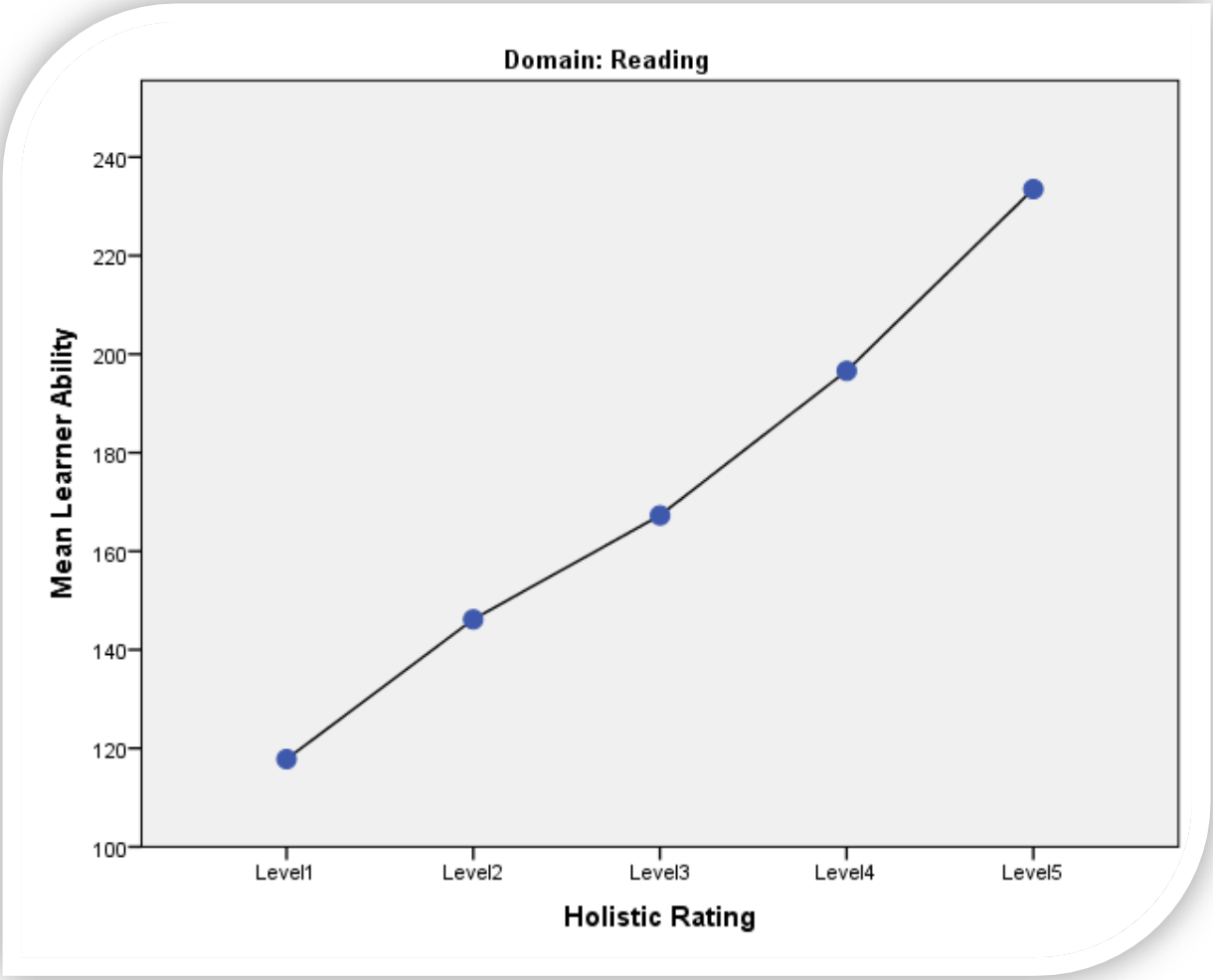


Figure 19: Relationship between the Holistic Judgement and the Mean Level of Reading Ability of the Learner

The findings presented in Figure 18 and Figure 19 illustrate that the assumption used for allocating forms to raters was appropriate for both domains; and that average ability estimates increased in line with holistic ratings even though ability estimates were based on rating a random set of items comprising three adjacent levels from two different frameworks (i.e. a particular survey form as described in Chapter 3).

4.3.2 Relative differences in the complexity of ALLS scaled items

Another approach to confirming the validity of the design of the study was to compare the complexity estimates for the ALLS scaled items obtained in this study against the published location of the ALLS scaled items on the ALLS scale. Ideally, there should be a strong correlation between the ALLS scaled scores and the calibrated complexity measures obtained in this study. A high correlation of over 0.9 was observed for both the Reading domain and Numeracy domain between

ALLS proficiency scale estimates and the complexity estimates independently obtained from the current study (as shown in Table 33). Note that the correlation for the Reading domain was achieved after two ALLS scaled items that were previously identified as complexity estimate outliers (see Section 4.2.2) were excluded from analysis.

Table 33: ALLS Items Comparison

Domain	UniqueID	ALLS proficiency scale	Complexity
Reading	ALLS_R_Blackboard	188	123
	ALLS_R_Impatiens_1	254	160
	ALLS_R_Impatiens_2	230	148
	ALLS_R_Fireworks	295	177
	ALLS_R_Hiring*	338	146
	ALLS_R_CANCO*	377	152
Numeracy	ALLS_N_Cola	174	117
	ALLS_N_Election	192	120
	ALLS_N_Blackboard	268	154
	ALLS_N_Gas	248	149
	ALLS_N_Fireworks	293	149
	ALLS_N_CompoundInterest	348	158

*item was previously identified as an outlier (unexpected complexity estimate) and was not included in estimation of correlation.

Based on their responses, raters were overall able to replicate the relative differences in complexity between the ALLS scaled items. However, two of the Reading ALLS scaled items were rated at a much lower complexity than expected (see Section 4.2.2).

CHAPTER 5: CONCLUSION

This chapter firstly summarises the major findings of the study and then discusses the implications of the findings for further research and development activities.

5.1 MAJOR FINDINGS

This study set out to empirically determine the relationship between Levels of the ACSF and the ALLS for the Reading and Numeracy domains using a complex survey design methodology. Within this research design, survey participants familiar with adult literacy and numeracy concepts rated one or more anonymous learners, whose reading and/or numeracy skills were familiar to the rater, against statements drawn from both frameworks. The survey featured multiple forms (i.e., three forms per domain) and link items (i.e., items common across forms to enable the forms to be equated onto a single scale) which not only minimised rater workload, but also enabled data to be collected on all 79 items for Reading and 86 items for Numeracy. Each form had approximately 50 items comprising statements from both frameworks across three adjacent levels. The items were presented randomly so the raters were not able to obtain external cues about the level of an item (other than the wording of the item itself), and also to avoid any item positioning effect. As such, the complexity of the item could be determined solely upon the language contained within each item, as opposed to making an a priori assumption about the relative complexity of the statements according to its original positioning within the framework.

Four hundred and eleven surveys were completed, with the majority of ratings made against learners' reading ability (72%). Although all states and territories were represented by the raters, the majority were located in Victoria (40%) and New South Wales (23%), with very few raters located in the Australian Capital Territory (1%). As expected, the raters tended to be employed within an Education and/or Training Institute (80%) and had more than 5 years' experience in Adult, Language, Literacy and/or Numeracy (65%). Only 7% of the raters had less than one year experience, indicating that majority of raters who participated in this study were very experienced within this field.

When rating the Reading domain, raters tended to select learners who were (or had previously) undertaken an ESL program (26%) or an LLN Program (22%); whereas with Numeracy, raters tended to select learners from an adult literacy and numeracy course (32%) or an LLN Program (22%). In both instances, raters tended to select learners who were at Levels 1 to 3 on the ACSF (as determined by the rater's initial holistic judgement), with less than 10% of raters selecting learners thought to be at Levels 4 & 5 in either Reading or Numeracy.

Subsequently, the majority of ratings were made against Levels 2 and 3 items for both Reading and Numeracy (ranging from 102 to 294 ratings per item), with very few ratings made against the items that were at Level 5 on both frameworks (15 ratings per item for Numeracy and 29 ratings per item for Reading). Given the large standard errors of measurement for the Level 5 items, it was

recommended that caution be exercised when comparing the highest complexity levels on both frameworks, particularly for the Numeracy domain.

Estimates of learners' ability were highly reliable in both Reading and Numeracy, with reliability estimates greater than 0.97 for both domains. Furthermore, there was also strong evidence of the construct validity of the measures as indicated by the match between the empirically calibrated measures against theoretical levels of the items within each framework. That is, even though the 50 or so items presented to raters were presented randomly, and drawn from two frameworks spanning three levels each, the calibration of the items (using Item Response Theory) replicated the expected sequencing of the set of items within each level. In fact, only 7 of the 165 items were found to produce unexpected complexity estimates, and these items were subsequently excluded from further analysis that compared the complexity levels across frameworks.

When comparing the levels across frameworks, a similar pattern was evident for both Reading and Numeracy. For example, Level 1 on both frameworks appeared to be similar in their complexity, whereas Level 2 & 3 on the ALLS was found to be more complex than ACSF Level 2 and ACSF Level 3, respectively on both domains; with the difference more pronounced for Reading. That is, ACSF Reading Level 3 was closely aligned to ALLS Reading Level 2; and ALLS Reading Level 3 was closely aligned to ACSF Reading Level 4. Whereas with Numeracy, although ALLS Numeracy Level 2, 3 & 4 was found to be more complex than ACSF Numeracy Levels 2, 3 & 4, respectively, the differences in the level complexities was not as great as that found for the Reading domain. For example, whilst ACSF Numeracy Level 2 was found to be less complex than ALLS Numeracy Level 2, it was still higher in complexity than ALLS Level 1 Numeracy. Similarly, ACSF Numeracy Level 3 was found to be more complex than ALLS Level 2 but not as complex as ALLS Level 3. Similar findings were found for the complexity of ACSF Numeracy Level 4, in which it was found to be located somewhere between ALLS Numeracy Level 3 & 4. Given the very few ratings made at Level 5 Numeracy, the relationship between the two frameworks at this level could not be reliably determined.

5.2 IMPLICATIONS FOR FUTURE RESEARCH AND DEVELOPMENT

The findings of the current study had direct implications for future use and refinement of the ACSF, as well as future mapping to other similar programs (e.g., AMEP) and frameworks (e.g., the Core Skills for Employment Framework).

5.2.1 Empirical Mapping of the ACSF and ALLS at Levels 4 & 5

As there were very few ratings of learners at Levels 4 & 5 in both Reading and Numeracy, there was less certainty about the true complexity estimates of statements drawn from the higher levels of both the ACSF and ALLS. Low number of ratings may have been due to raters being less familiar with the reading/numeracy skills of learners at the higher levels, as it is less likely that higher able learners (in terms of reading and/or numeracy) would undertake specific adult literacy and/or

numeracy programs. If it is desirable to obtain more accurate measures at the higher levels on the ACSF and to compare these to the ALLS, a future study could be employed in which raters judge their own level of ability directly against the statements (hence, the survey would be designed to resemble a self-assessment as opposed to performance judgement of a learner). A similar methodology was employed to empirically validate the Developmental Learning Framework for School Leaders[25] in Victoria in which aspiring School Principals were surveyed using a self-assessment tool. If the survey instructions for the current study were re-designed to reflect a self-assessment, the target population for calibrating the ACSF and ALLS at Levels 4 and 5, could include the current raters used within this study (i.e., adult literacy and numeracy educators and specialists), as well as students undertaking undergraduate and post graduate studies, majoring in English/Mathematics.

5.2.2 Improvements to the structure of the ACSF

This study has clearly demonstrated that the Performance Features within the ACSF have varying levels of complexity, even within each level of the framework. The implications of such findings have potential for further improvements to the current structure of the ACSF. For example, three statements drawn from the ACSF were found to have unexpected complexity estimates which should be reviewed when future revisions are made to the ACSF. The study has also shown that it is possible to recognise sub-levels within each of the five levels. That is, the empirical positioning of Performance Features in the current study illustrates that the empirical results can be used to assist in describing the typical developmental pathway of learners **within** an ACSF level. If a similar follow up study was employed, with a larger pool of Performance Features sampled, then it would be possible to describe what it means to be at a higher or lower level within each of the existing five levels of the ACSF (e.g., a profile description could be developed to describe a *typical* high level ACSF 2 learner as opposed to a *typical* lower level ACSF 2 learner). Hence, it is possible that the ACSF could be restructured to have profile descriptions within each level (e.g. two sub-level profiles per level).

5.2.3 Future Roll-Out to Other Frameworks

This study has successfully demonstrated that developmental learning frameworks, thought to be measuring the same overarching constructs, can be empirically positioned onto the same measurement scale using a c ombination of complex survey design methodology and Item Response Theory analyses. In addition, this study has demonstrated that such a methodology can also be used to empirically validate developmental frameworks (such as the ACSF) to ensure that they are hierarchical, developmental and cumulative, as intended. A similar study was undertaken

[25] See http://www.eduweb.vic.gov.au/edulibrary/public/teachlearn/leader/developmental_learning_framework_20070418.pdf

for the Australian Qualifications Framework (AQF)[26]. Prior to conducting the current study, it was only assumed that the ACSF had five levels of increasing complexity in Reading and Numeracy.

Furthermore, the results from this study could be used to map other similar frameworks or programs onto the ACSF and/or the ALLS. For example, the Adult Migrant English Program (AMEP), if thought to have similar constructs in terms of reading/numeracy, could also be mapped onto the Reading and/or Numeracy complexity scales that were developed in this particular study. Under such circumstances, only a sample of the calibrated statements used within this study would need to be included in the follow-up study, in addition to a sample of additional statements (e.g., learning outcomes) drawn directly from the AMEP.

Similarly, the new, yet to be released Core Skills for Employment Framework (CSFE)[27], which has been designed to have five developmental levels across 10 skill areas (to complement the ACSF), could also be empirically validated using a similar methodology to that employed in the current study. The empirical validation of the CSFE would enable the developmental, hierarchical and cumulative nature of the five levels within the framework to be tested per skill area, as well as minimise any redundancies and/or ambiguity in the Performance Features which may exist within and across the 10 skills sets. In addition to empirically validating the CSFE in terms of its architectural structure etc, it may also be desirable to map certain skills sets within its framework to the ACSF.

[26] see Gillis, S., Wu., M., Dulhunty, M., Calvitto, L., & Bateman, A. (2010). Empirical Validation of the strengthened Australian Qualifications Framework using Item Response Theory. Conducted for the AQF Council. www.aqf.edu.au

[27] ITHACA Group (2012) Core Skills For Employment, Draft Version 3; DIISRTE & DEEWR, Canberra.

APPENDICES

Appendix 1: Modified ALLS Level Descriptions

Prose Literacy Domain

Level	Prose Literacy Descriptor (original)	Research Team Edits (to remove specific reference to task and/or level)	Panel Modified Version
1	Most of the tasks in this level require the respondent to read relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information.	Reads relatively short text to locate a single piece of information which is given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information.	*Reads relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or instruction.*
2	Some tasks in this level require respondents to locate a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present, or low-level inferences may be required. Other tasks require the respondent to integrate two or more pieces of information or to compare and contrast easily identifiable information based on a criterion provided in the question or directive.	Locates a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present, or low-level inferences may be required. Integrates two or more pieces of information and can compare and contrast easily identifiable information based on the criterion provided in the question or directive.	*Locates a single piece of information in a text containing some distracting, yet plausible information.* *Integrates two or more pieces of information* *Compares and contrasts easily identifiable information*
3	Tasks in this level tend to require respondents to make literal or synonymous matches between the text and information given in the task, or to make matches that require low-level inferences. Other tasks ask respondents to integrate information from dense or lengthy text that contains no organizational aids such as headings. Respondents may also be asked to generate a response based on information that can be easily identified in the text. Distracting information is present, but is not located near the correct information.	Makes literal of synonymous matches between the text and information given in the task. Makes matches that require low-level inferences. Integrates information from dense or lengthy text that contains no organisational aids such as headings. Can also generate a response based on information that can be easily identified in the text where distracting information is present, but not located near the correct information.	*Locates relevant information in a text where the information is literal or requires only low level inference.* *Integrates information from dense text with no organisational aids such as heading.* *Integrates information from lengthy text with no organisational aids such as heading.*

Level	Prose Literacy Descriptor (original)	Research Team Edits (to remove specific reference to task and/or level)	Panel Modified Version
4	These tasks require respondents to perform multiple-feature matches and to integrate or synthesize information from complex or lengthy passages. More complex inferences are needed to perform successfully. Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent.	Performs multiple-feature matches and integrates or synthesizes information from complex or lengthy passages. To perform these tasks successfully, more complex inferences are needed and conditional information must be taken into consideration.	*Reads complex lengthy text, integrating or synthesising information that may require high level inferencing.*
5	Some tasks in this level require the respondent to search for information in dense text which contains a number of plausible distractors. Others ask respondents to make high-level inferences or use specialized background knowledge. Some tasks ask respondents to contrast complex information.	Searches for information in dense text which contains a number of plausible distractors. Can typically complete tasks that require high-level inferences or specialized background knowledge. Can also complete tasks that ask the learner to contrast complex information.	*Locates and contrasts complex information and ideas in dense texts which may require high level inferencing and specialised background knowledge.*

Document Literacy Domain

	Document Literacy Descriptor	Research Team Edits (to remove specific reference to task and/or level)	Panel Modified Version
1	Tasks in this level tend to require the respondent either to locate a piece of information based on a literal match or to enter information from personal knowledge onto a document. Little, if any, distracting information is present.	Locates a piece of information based on a literal match and can enter information from personal knowledge onto a document in which there is little, if any, distracting information present.	*Locates specific information from a simple unambiguous text* *Reads and understands instructions on simple form requiring personal details* *Enters information from personal knowledge onto a document that contains little, if any, distracting information.*
2	Tasks in this level are more varied than those in Level 1. Some require the respondents to match a single piece of information; however, several distractors may be present, or the match may require low-level inferences. Tasks in this level may also ask the respondent to cycle through information in a document or to integrate information from various parts of a document.	Match a single piece of information; however, several distractors may be present, or the match may require low-level inferences. Cycles through information in a document or to integrate information from various parts of a document.	*Cycles through a text to distinguish relevant information from distracting information* *Selects information from a text which may require integration from various parts of the document or low level inferencing*
3	Some tasks in this level require the respondent to integrate multiple pieces of information from one or more documents. Others ask respondents to cycle through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.	Integrates multiple pieces of information from one or more documents. Cycles through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.	*Integrates multiple pieces of information from one or more documents.* *Selects information from complex tables or graphs, where irrelevant or inappropriate information may be present.*
4	Tasks in this level, like those at the previous levels, ask respondents to perform multiple-feature matches, cycle through documents, and integrate information; however, they require a greater degree of inferencing. Many of these tasks require respondents to provide numerous responses but do not designate how many responses are needed. Conditional information is also present in the document tasks at this level and must be taken into account by the respondent.	Complete tasks that require multiple-feature matches, cycling through documents, and integrating information with some degree of inference. Can typically complete tasks that require numerous responses but the tasks do not designate how many responses are needed. Conditional information must be taken into account to complete these tasks.	*Reads complex texts, requiring:* • *Integration of multiple pieces of information* • *Inferencing at high level* • *Selection of appropriate responses where the number of responses is not designated*
5	Tasks in this level require the respondent to search through complex displays that contain multiple distractors, to make high-level text-based inferences, and to use specialized knowledge.	Searches through complex displays that contain multiple distractors, to make high-level text-based inferences, and use specialised knowledge to complete the task.	*Searches through complex texts containing several pieces of distracting information which may require high level thinking and specialised background knowledge.*

Numeracy domain

level	Numeracy Descriptor	Research Team Edits (to remove specific reference to task and/or level)	Panel Modified Version
1	Tasks in this level require the respondent to show an understanding of basic numerical ideas by completing simple tasks in concrete, familiar contexts where the mathematical content is explicit with little text. Tasks consist of simple, one-step operations such as counting, sorting dates, performing simple arithmetic operations or understanding common and simple percents such as 50%.	Understands basic numerical ideas as demonstrated by completing simple tasks in concrete, familiar contexts where the mathematical content is explicit with little text. Can typically complete tasks that are simple, one-step operations such as counting, sorting dates, performing simple arithmetic operations or understanding common and simple percents such as 50%.	*Understands basic and explicit numerical ideas in order to complete a simple task in a concrete, familiar context.* *Can complete simple, one-step arithmetical operations (e.g., counting, sorting dates)* *Understands common and simple numerical ideas (e.g., 50%) where the mathematical context is explicit with little text.*
2	Tasks in this level are fairly simple and relate to identifying and understanding basic mathematical concepts embedded in a range of familiar contexts where the mathematical content is quite explicit and visual with few distractors. Tasks tend to include one-step or two-step processes and estimations involving whole numbers, benchmark percents and fractions, interpreting simple graphical or spatial representations, and performing simple measurements.	Identifies and understands basic mathematical concepts that are embedded in a range of familiar contexts, where the mathematical content is quite explicit and visual with few distractors. Can typically complete tasks that are fairly simple, one step or two step processes and estimations involving whole numbers, benchmark percents and fractions, interpreting simple graphical or spatial representation, and performing simple measurements.	*Identifies and understands basic mathematical concepts in a range of familiar contexts where the mathematics content is quite explicit and visual with little distracting information.* *Can complete tasks with one-step or two-step processes involving whole numbers and common percents and fractions.* *Can interpret simple graphical or spatial representations where the mathematical context is quite explicit and visual with little distracting information.*
3	Tasks in this level require the respondent to demonstrate understanding of mathematical information represented in a range of different forms, such as in numbers, symbols, maps, graphs, texts, and drawings. Skills required involve number and spatial sense, knowledge of mathematical patterns and relationships and the ability to interpret proportions, data and statistics embedded in relatively simple texts where there may be distractors. Tasks commonly involve undertaking a number of processes to solve problems.	Understands mathematical information represented in a range of different forms, such as in numbers, symbols, maps, graphs, texts, and drawings. Skills include number and spatial sense, knowledge of mathematical patterns and relationships and the ability to interpret proportions, data and statistics embedded in relatively simple texts where there may be distractors. Can typically complete tasks that involve undertaking a number of processes to solve problems.	*Can locate and use mathematical information in a range of different forms, e.g., numbers, symbols, maps, graphs, texts, drawings.* *Can solve problems involving a number of processes with skills related to mathematics patterns and relationships.* *Interpret mathematical information (e.g., data and statistics) embedded in relatively simple texts where there may be distracting information.*

level	Numeracy Descriptor	Research Team Edits (to remove specific reference to task and/or level)	Panel Modified Version
4	Tasks at this level require respondents to understand a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts. These tasks involve undertaking multiple steps to find solutions to problems and require more complex reasoning and interpretation skills, including comprehending and working with proportions and formulas or offering explanations for answers.	Understands a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts. Can typically complete tasks that involve undertaking multiple steps to find solutions to problems which require more complex reasoning and interpretation skills, including comprehending and working with proportions and formulas or offering explanations for answers.	*Understands a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts.* *Can solve problems involving multiple steps requiring complex mathematical reasoning and interpretation skills, e.g., working with proportions and formulas.*
5	Tasks in this level require respondents to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Respondents may have to integrate multiple types of mathematical information, draw inferences, or generate mathematical justification for answers.	Understands complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Can typically integrate multiple types of mathematical information, draw inferences, or generate mathematical justification for answers.	*Understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts.* *Can interpret multiple types of mathematics information to draw inferences or generate mathematical justification for answers.*

Appendix 2: Sampled ALLS Numeracy Complexity Statements

Source: Murray, T.S., Clermont, Y & Brinkley, H. (2005). Measuring adult literacy and life skills: New frameworks for assessments, Statistics Canada, p.190-191.

Complexity Score	Numeracy Complexity Statement
Complexity Score 1 (easiest)	Can carry out one-step arithmetical operations (+, -, ×,÷) with whole numbers to 1,000. Can complete simple tasks involving everyday measures in whole units (e.g., kg, m, dates, hours, minutes). Can recognise simple, common 2D shapes.
Complexity Score 2	Can carry out arithmetical operations (+, -, ×,÷) with large whole numbers including millions. Can evaluate a given formula involving common operations (+, -, ×,÷). Can carry out computations with everyday standard measures (e.g., length, weight) involving common fractions (e.g., 25%, 10%) and common decimal units (e.g., 0.1, 0.25). Can recognise common 3D shapes and their representation via diagrams or photos. Can use common measuring instrument to read off marked units on a scale, but not necessarily able to interpolate between gradations.
Complexity Score 3	Is familiar with area and volume formulae. Can relate 2D and 3D shapes in combination. Can work with decimals to 3 decimal places. Can compute rates and ratios, NOT limited to just the common ones. Can round numbers off to a requested number of decimal places. Can carry out simple probability calculations. Can use angle and symmetry properties to describe shapes or objects. Can interpolate values between gradations on scales. Can calculate distances from scales on maps.
Complexity Score 4	Can carry out operations involving negative numbers. Can carry out operations involving squares, square roots, etc. Can create formulae based on the problem descriptions. Can calculate mean and standard deviation from data. Can convert between measuring units (e.g., km to m, kg to g) within the same system. Can use strategies such as working backwards or backtracking (e.g., 15% of ? = $255).
Complexity Score 5 (hardest)	Can complete tasks with abstract ideas or in unfamiliar contexts. Can communicate complex and abstract reasoning. Can use advanced mathematical techniques (e.g., trigonometry). Can use algebraic conventions and techniques. Can convert between measurements across different systems (e.g., kg to lb, km to miles, Fahrenheit to Celsius). Can work with formal mathematics involving formulae and relationships between variables.

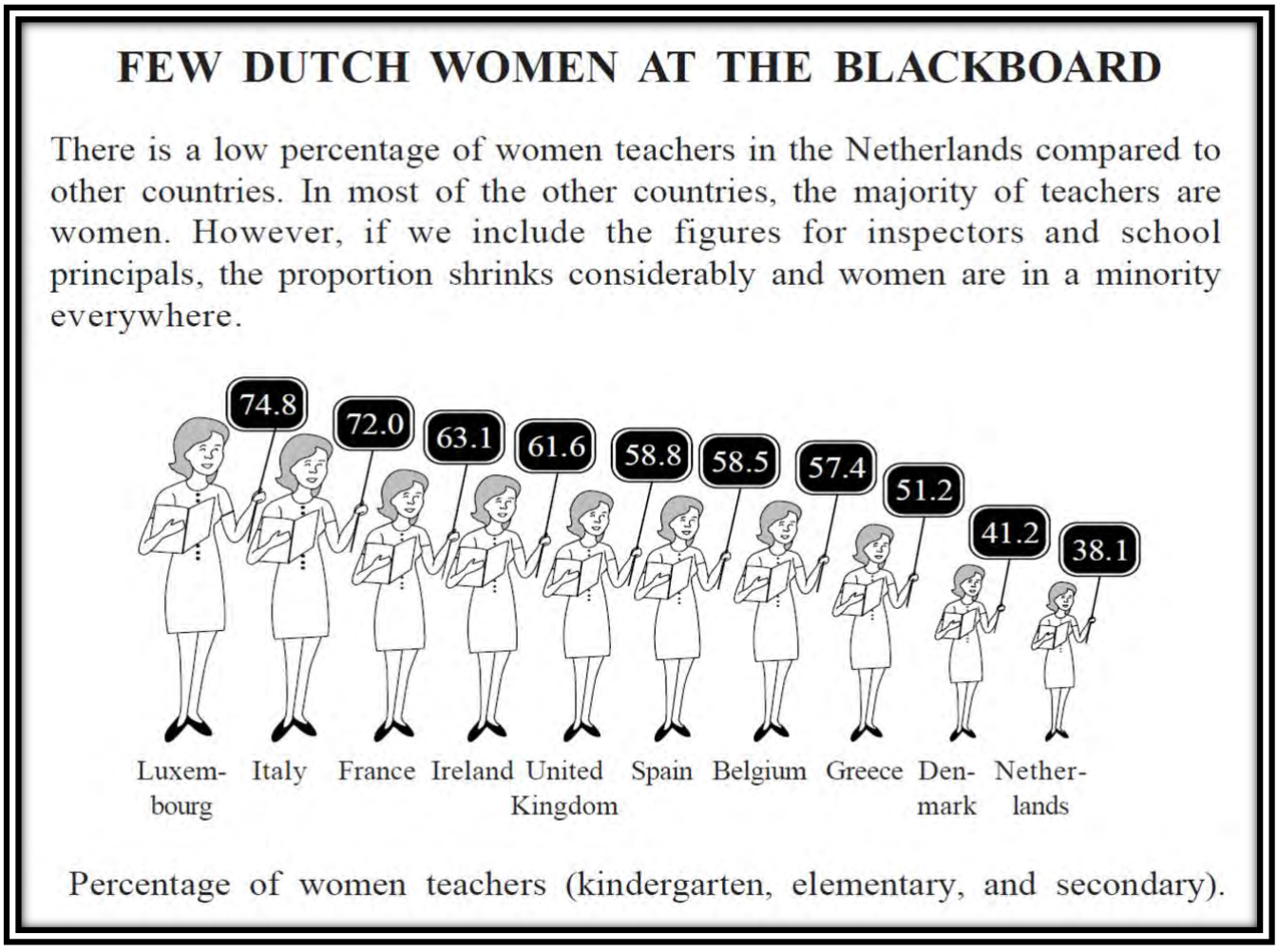
Appendix 3: Selected publicly available ALLS Scaled Items

ALLS READING ITEM - LEVEL 1 - DUTCH WOMEN AT THE BLACKBOARD

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 112 (**used this source mainly**)
- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.288
- ABS 2006, Adult literacy and life skills survey: User guide, Catalogue No. 4228.0.55.002, ABS.

Consider the following image:



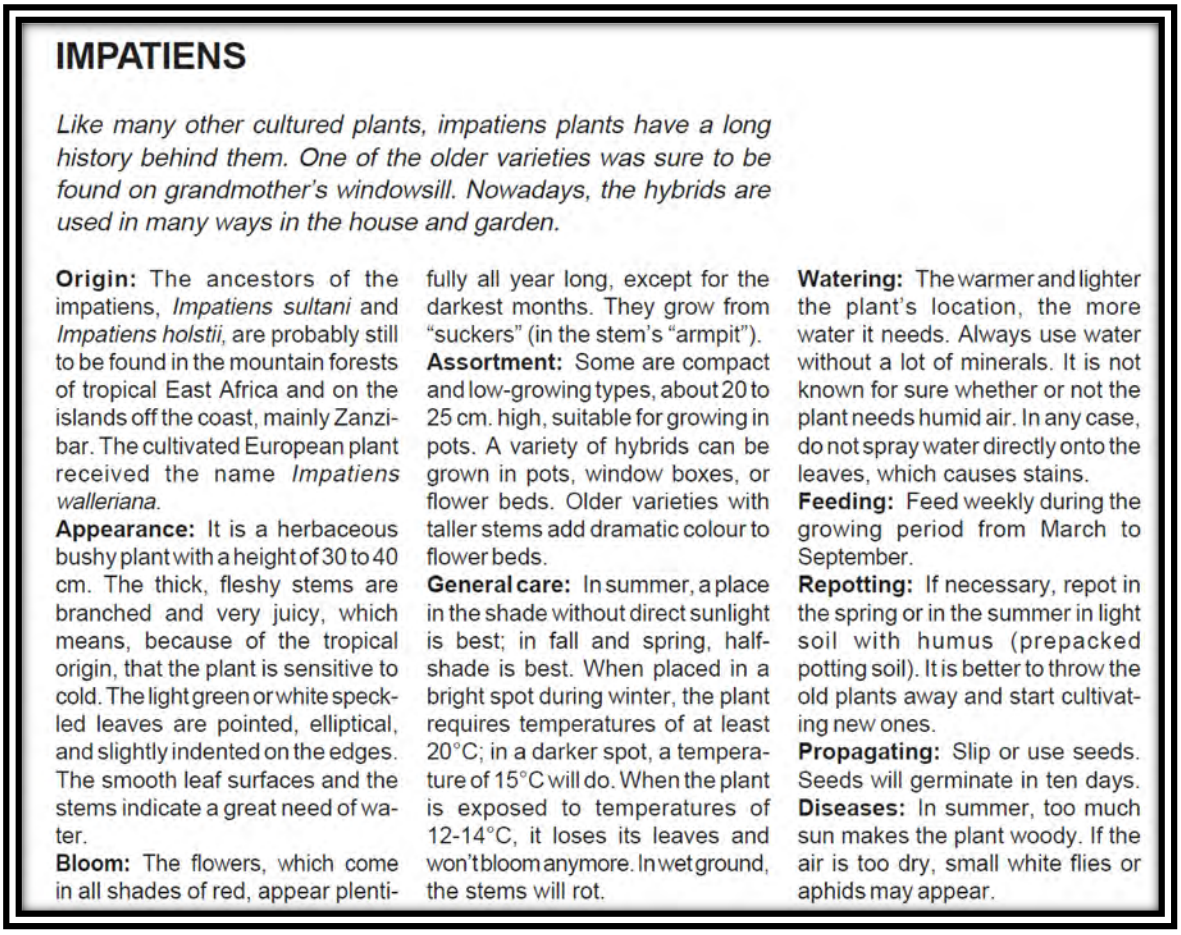
Question: Identify from a chart the percentage of teachers from Greece who are women.

ALLS READING ITEM - LEVEL 2 - IMPATIENS

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 106 & 107 (**used this source mainly**)
- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.285
- Murray TS, Clermont Y & Binkley H 2005, Measuring adult literacy and life skills: New frameworks for assessment, Statistics Canada, p.106;

Consider the following image:



IMPATIENS

Like many other cultured plants, impatiens plants have a long history behind them. One of the older varieties was sure to be found on grandmother's windowsill. Nowadays, the hybrids are used in many ways in the house and garden.

Origin: The ancestors of the impatiens, *Impatiens sultani* and *Impatiens holstii*, are probably still to be found in the mountain forests of tropical East Africa and on the islands off the coast, mainly Zanzibar. The cultivated European plant received the name *Impatiens walleriana*.

Appearance: It is a herbaceous bushy plant with a height of 30 to 40 cm. The thick, fleshy stems are branched and very juicy, which means, because of the tropical origin, that the plant is sensitive to cold. The light green or white speckled leaves are pointed, elliptical, and slightly indented on the edges. The smooth leaf surfaces and the stems indicate a great need of water.

Bloom: The flowers, which come in all shades of red, appear plentifully all year long, except for the darkest months. They grow from "suckers" (in the stem's "armpit").

Assortment: Some are compact and low-growing types, about 20 to 25 cm. high, suitable for growing in pots. A variety of hybrids can be grown in pots, window boxes, or flower beds. Older varieties with taller stems add dramatic colour to flower beds.

General care: In summer, a place in the shade without direct sunlight is best; in fall and spring, half-shade is best. When placed in a bright spot during winter, the plant requires temperatures of at least 20°C; in a darker spot, a temperature of 15°C will do. When the plant is exposed to temperatures of 12-14°C, it loses its leaves and won't bloom anymore. In wet ground, the stems will rot.

Watering: The warmer and lighter the plant's location, the more water it needs. Always use water without a lot of minerals. It is not known for sure whether or not the plant needs humid air. In any case, do not spray water directly onto the leaves, which causes stains.

Feeding: Feed weekly during the growing period from March to September.

Repotting: If necessary, repot in the spring or in the summer in light soil with humus (prepacked potting soil). It is better to throw the old plants away and start cultivating new ones.

Propagating: Slip or use seeds. Seeds will germinate in ten days.

Diseases: In summer, too much sun makes the plant woody. If the air is too dry, small white flies or aphids may appear.

Question: What happens when the impatiens plant is exposed to temperatures of 14 degrees C or below?

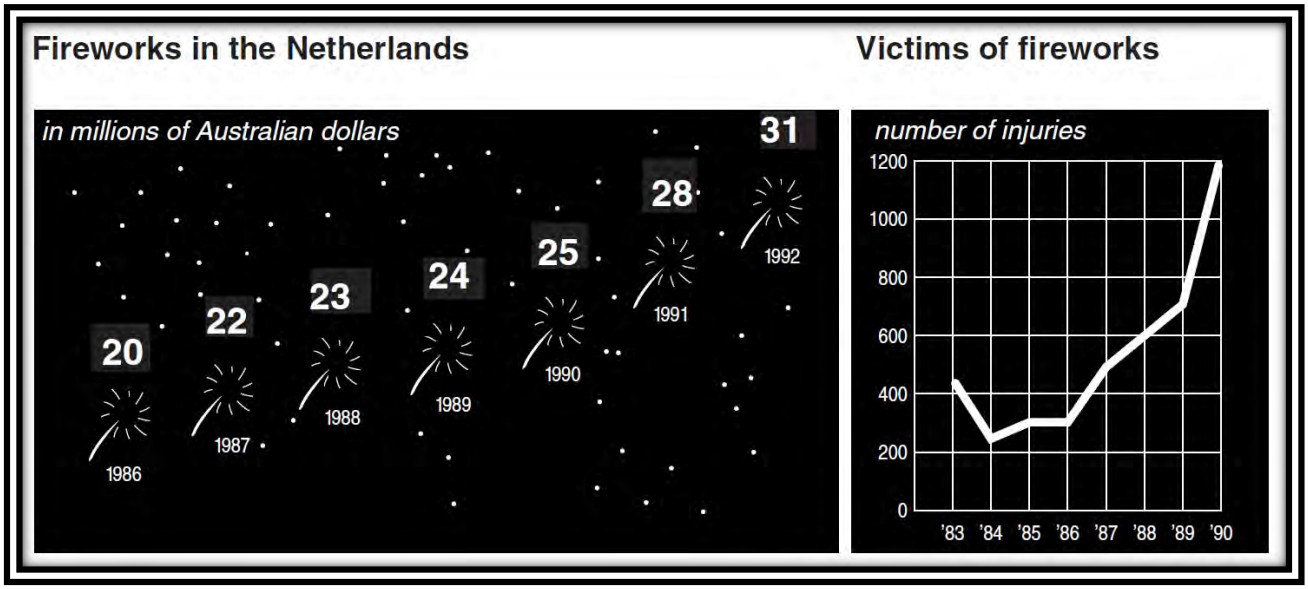
Question: What does the smooth leaf and stem suggest about the plant?

ALLS READING ITEM - LEVEL 3 - FIREWORKS IN THE NETHERLANDS

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 114 & 118 (**used this source mainly**)
- Murray TS, Clermont Y & Binkley H 2005, Measuring adult literacy and life skills: New frameworks for assessment, Statistics Canada, p.113-114;

Consider the following image:



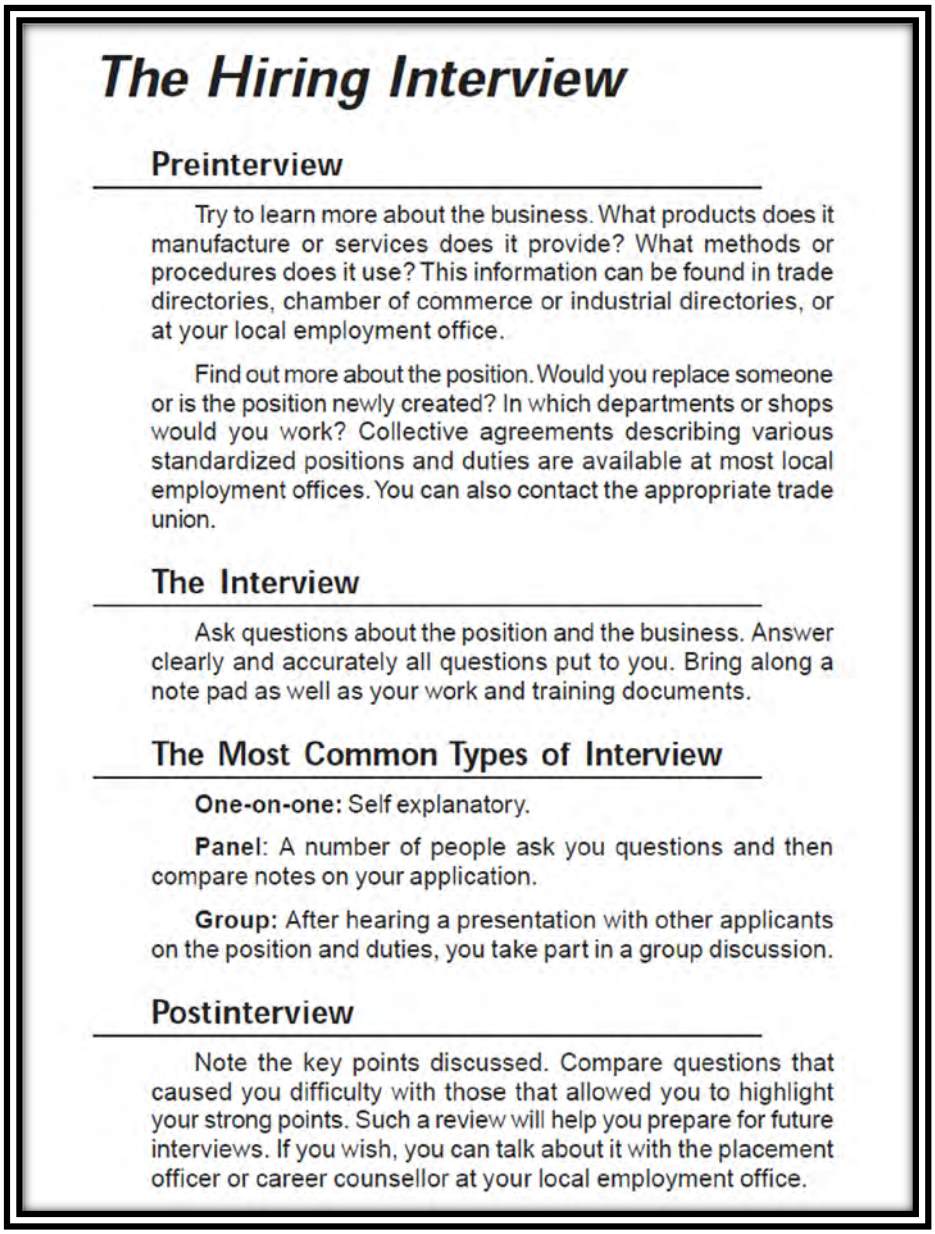
Question: Write a brief description of the relationship between sales and injuries based on the information shown

ALLS READING ITEM - LEVEL 4 - THE HIRING INTERVIEW

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 109 & 118 (**used this source mainly**)
- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.286-87

Consider the following image:



The Hiring Interview

Preinterview

Try to learn more about the business. What products does it manufacture or services does it provide? What methods or procedures does it use? This information can be found in trade directories, chamber of commerce or industrial directories, or at your local employment office.

Find out more about the position. Would you replace someone or is the position newly created? In which departments or shops would you work? Collective agreements describing various standardized positions and duties are available at most local employment offices. You can also contact the appropriate trade union.

The Interview

Ask questions about the position and the business. Answer clearly and accurately all questions put to you. Bring along a note pad as well as your work and training documents.

The Most Common Types of Interview

One-on-one: Self explanatory.

Panel: A number of people ask you questions and then compare notes on your application.

Group: After hearing a presentation with other applicants on the position and duties, you take part in a group discussion.

Postinterview

Note the key points discussed. Compare questions that caused you difficulty with those that allowed you to highlight your strong points. Such a review will help you prepare for future interviews. If you wish, you can talk about it with the placement officer or career counsellor at your local employment office.

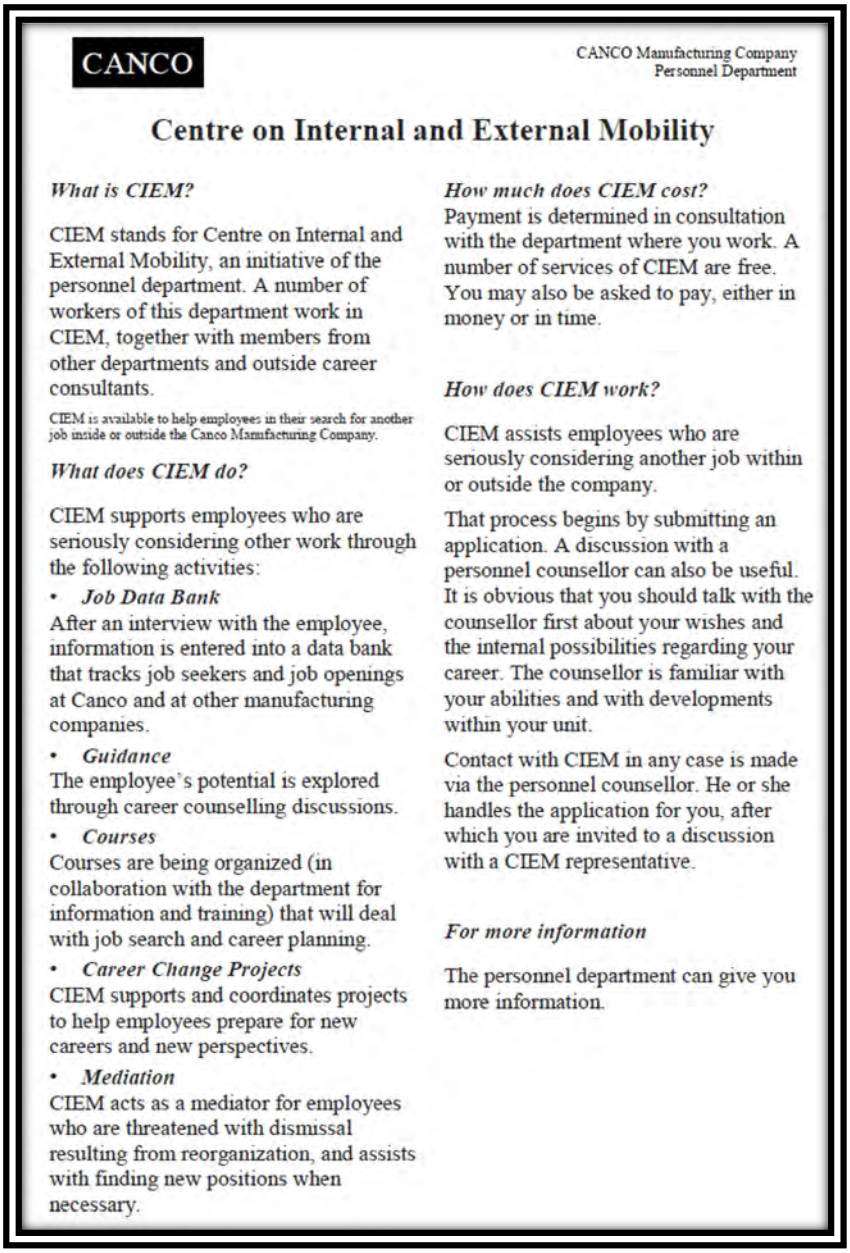
Question: Write in your own words one difference between the panel and the group interview

ALLS READING ITEM - LEVEL 5 - CANCO

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 110 & 111
- Murray TS, Clermont Y & Binkley H 2005, Measuring adult literacy and life skills: New frameworks for assessment, Statistics Canada, p.110-111
- ABS 2006, Adult literacy and life skills survey: User guide, Catalogue No. 4228.0.55.002, ABS.p.34 (**used this source mainly**)

Consider the following image:



CANCO

CANCO Manufacturing Company
Personnel Department

Centre on Internal and External Mobility

What is CIEM?

CIEM stands for Centre on Internal and External Mobility, an initiative of the personnel department. A number of workers of this department work in CIEM, together with members from other departments and outside career consultants.

CIEM is available to help employees in their search for another job inside or outside the Canco Manufacturing Company.

What does CIEM do?

CIEM supports employees who are seriously considering other work through the following activities:

- ***Job Data Bank***

After an interview with the employee, information is entered into a data bank that tracks job seekers and job openings at Canco and at other manufacturing companies.

- ***Guidance***

The employee's potential is explored through career counselling discussions.

- ***Courses***

Courses are being organized (in collaboration with the department for information and training) that will deal with job search and career planning.

- ***Career Change Projects***

CIEM supports and coordinates projects to help employees prepare for new careers and new perspectives.

- ***Mediation***

CIEM acts as a mediator for employees who are threatened with dismissal resulting from reorganization, and assists with finding new positions when necessary.

How much does CIEM cost?

Payment is determined in consultation with the department where you work. A number of services of CIEM are free. You may also be asked to pay, either in money or in time.

How does CIEM work?

CIEM assists employees who are seriously considering another job within or outside the company.

That process begins by submitting an application. A discussion with a personnel counsellor can also be useful. It is obvious that you should talk with the counsellor first about your wishes and the internal possibilities regarding your career. The counsellor is familiar with your abilities and with developments within your unit.

Contact with CIEM in any case is made via the personnel counsellor. He or she handles the application for you, after which you are invited to a discussion with a CIEM representative.

For more information

The personnel department can give you more information.

Question: List two ways in which CIEM (an employee support initiative within a company) helps people who lose their jobs because of departmental reorganization.

ALLS Numeracy item - Level 1 - Coca Cola Bottles

Sources:

- Stats Can/OECD 2005, Learning a living: First results of the ALLS Numeracy item - Level 1 - Coca Cola Bottles
- Adult Literacy and Life Skills survey, OECD, p.299

Consider the following image:

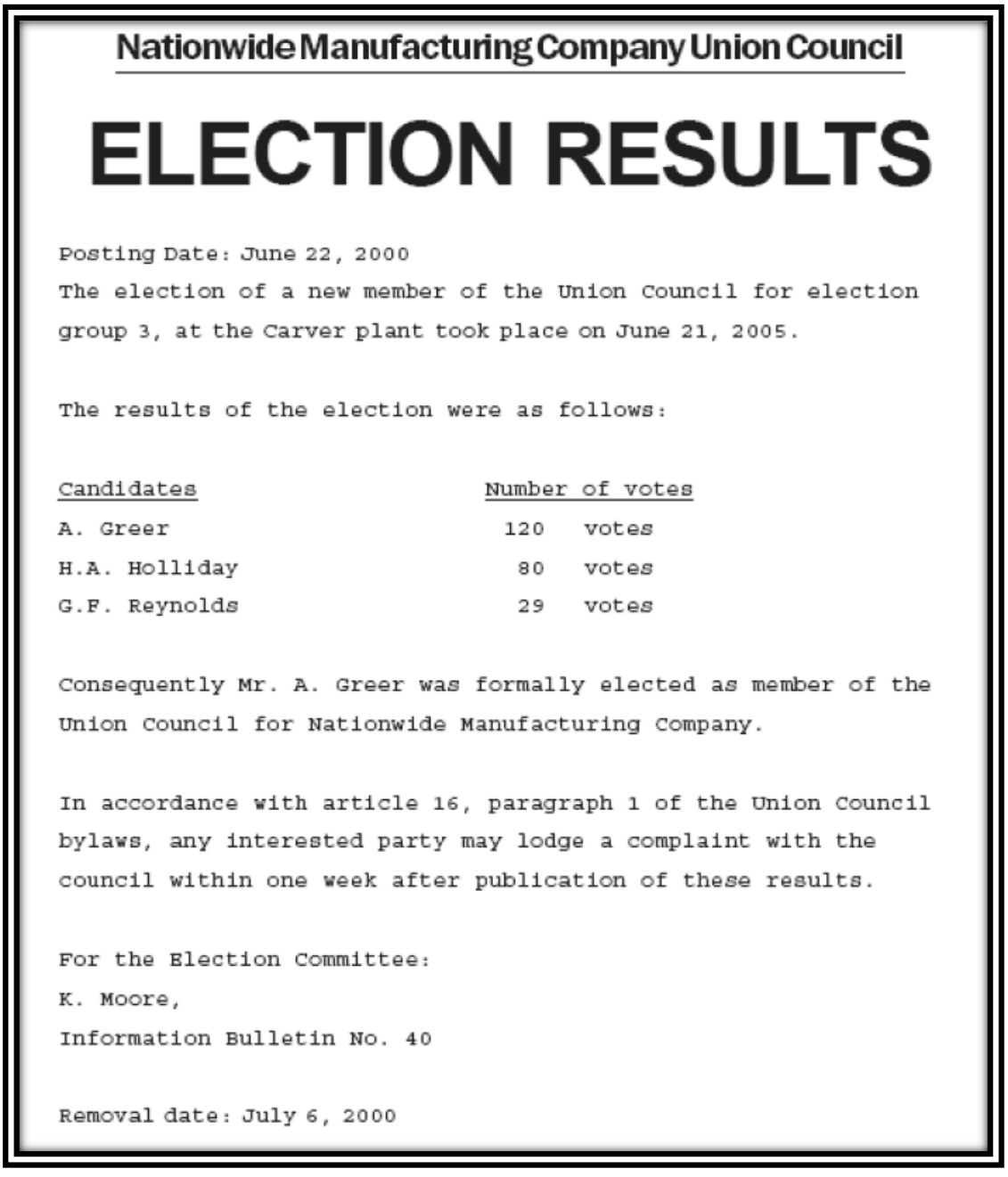
Question: Find the total number of bottles in the two full cases shown in the picture.

ALLS NUMERACY ITEM - LEVEL 1 - ELECTION RESULTS

Sources:

- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.299
- ABS 2006 Adult literacy and life skills survey: User Guide, Catalogue No. 4228.0.55.002. ABS, p.37 (**used this source mainly**)

Consider the following image:



Nationwide Manufacturing Company Union Council

ELECTION RESULTS

Posting Date: June 22, 2000

The election of a new member of the Union Council for election group 3, at the Carver plant took place on June 21, 2005.

The results of the election were as follows:

Candidates	Number of votes
A. Greer	120 votes
H.A. Holliday	80 votes
G.F. Reynolds	29 votes

Consequently Mr. A. Greer was formally elected as member of the Union Council for Nationwide Manufacturing Company.

In accordance with article 16, paragraph 1 of the Union Council bylaws, any interested party may lodge a complaint with the council within one week after publication of these results.

For the Election Committee:
K. Moore,
Information Bulletin No. 40

Removal date: July 6, 2000

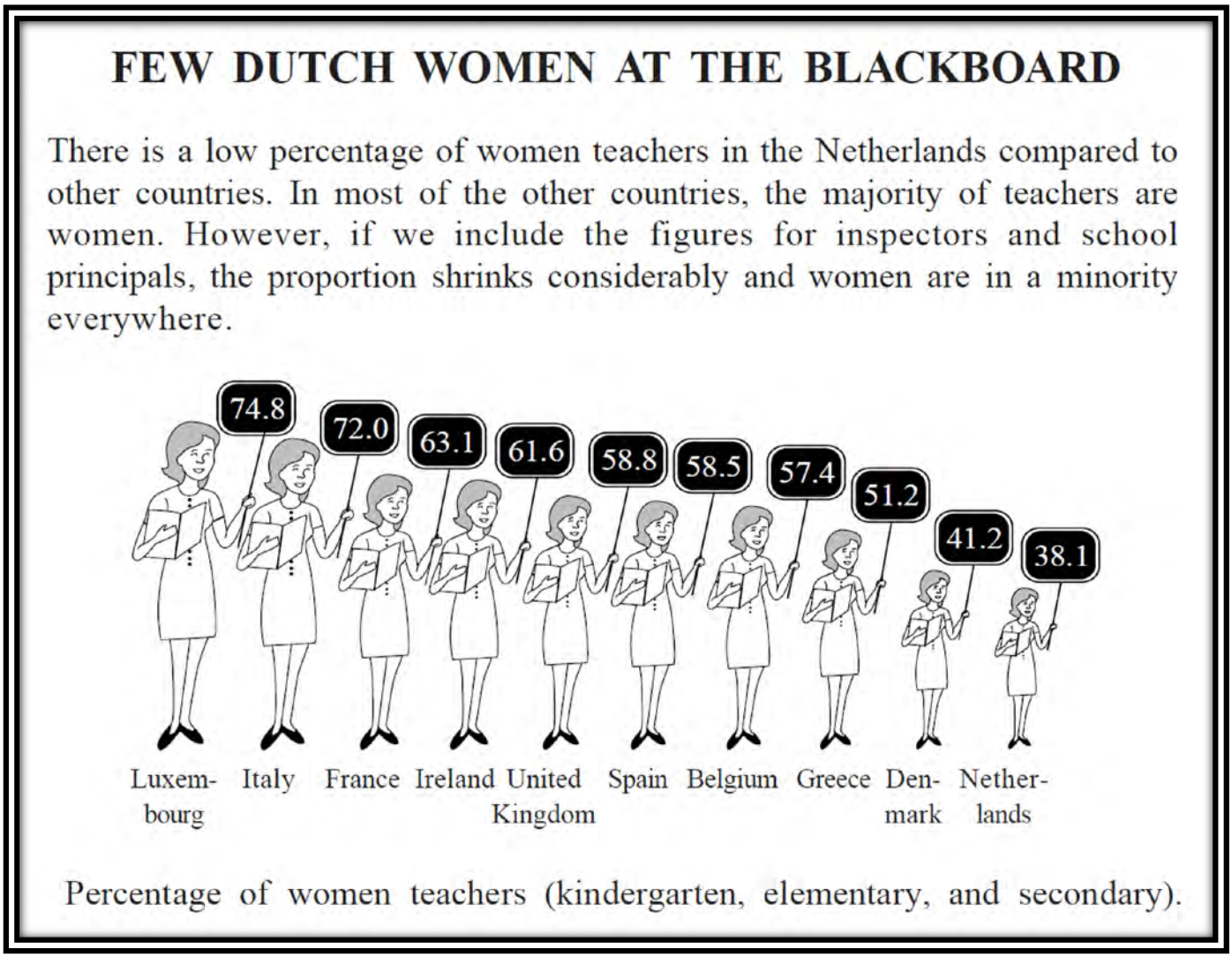
Question: Determine the total number of votes cast.

ALLS NUMERACY ITEM - LEVEL 2 - FEW DUTCH WOMEN AT THE BLACKBOARD

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 113 & 118

Consider the following image:



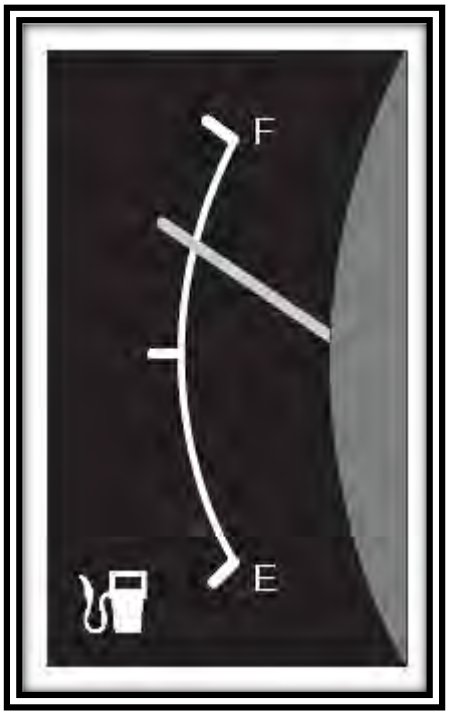
Question: Calculate the percentage of men in the teaching profession in Italy.

ALLS NUMERACY ITEM - LEVEL 2 - GAS GAUGE

Sources:

- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.299

Consider the following image:



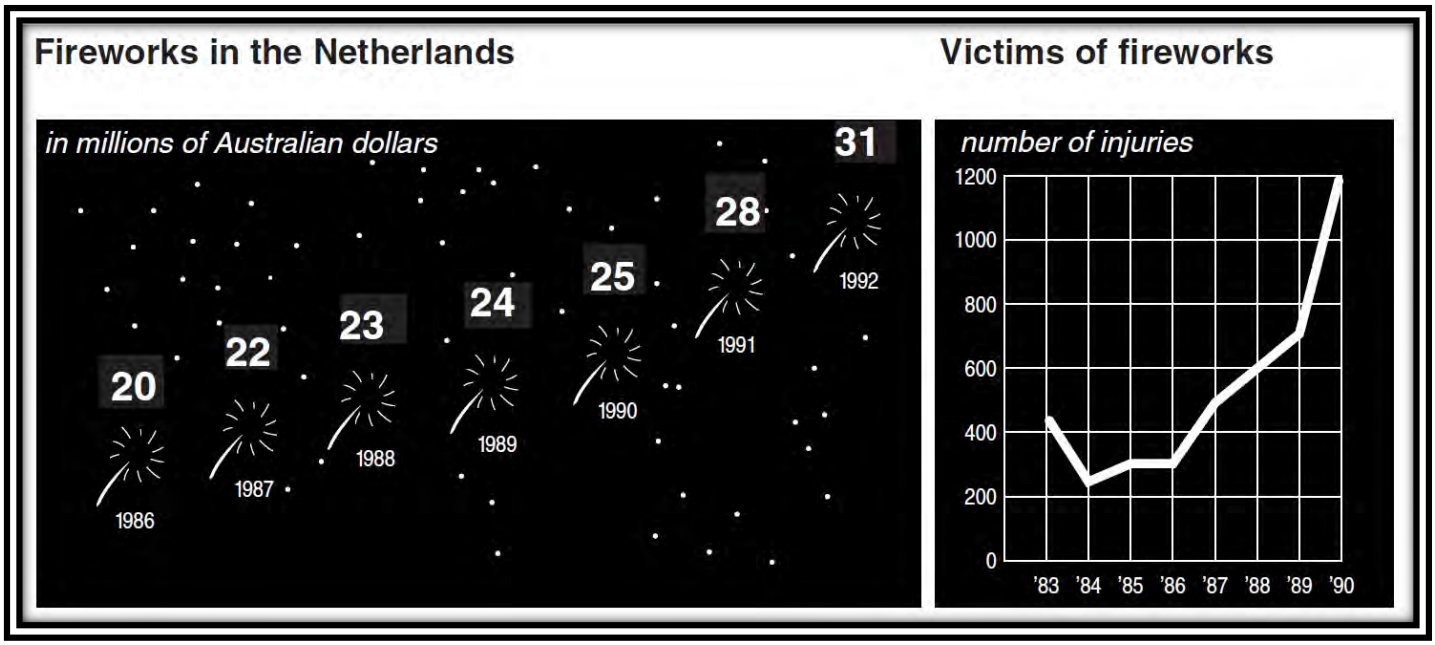
Question: The tank holds 48 gallons. How many gallons remain in the tank?

ALLS NUMERACY ITEM - LEVEL 3 - FIREWORKS IN THE NETHERLANDS

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 114 & 118

Consider the following image:



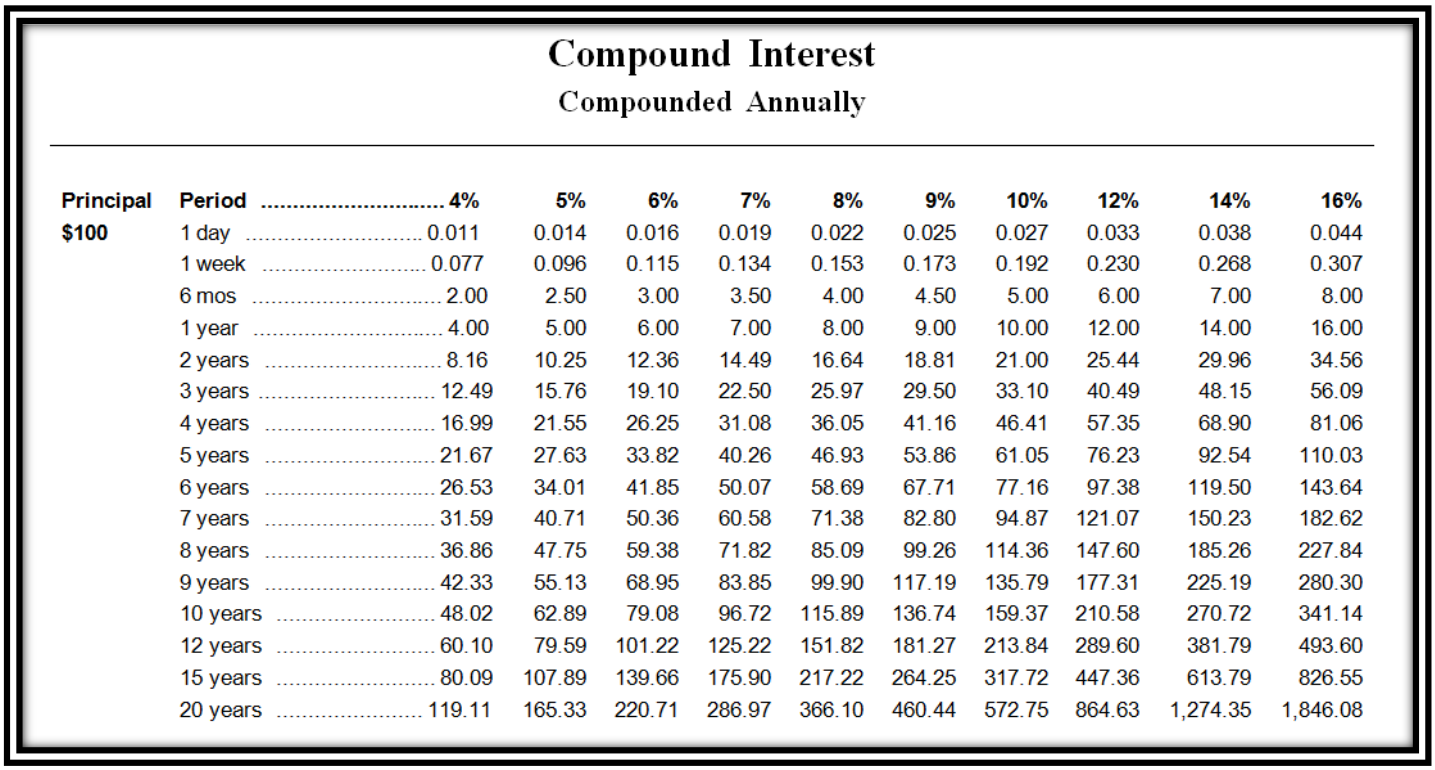
Question: Calculate how many more people were injured in 1989 than in 1988

ALLS NUMERACY ITEM - LEVEL 4 - COMPOUND INTEREST

Sources:

- International Adult Literacy Survey. Statistics Canada (2006), p. 118 & 119

Consider the following image:



Compound Interest

Compounded Annually

Principal	Period	4%	5%	6%	7%	8%	9%	10%	12%	14%	16%
$100	1 day	0.011	0.014	0.016	0.019	0.022	0.025	0.027	0.033	0.038	0.044
	1 week	0.077	0.096	0.115	0.134	0.153	0.173	0.192	0.230	0.268	0.307
	6 mos	2.00	2.50	3.00	3.50	4.00	4.50	5.00	6.00	7.00	8.00
	1 year	4.00	5.00	6.00	7.00	8.00	9.00	10.00	12.00	14.00	16.00
	2 years	8.16	10.25	12.36	14.49	16.64	18.81	21.00	25.44	29.96	34.56
	3 years	12.49	15.76	19.10	22.50	25.97	29.50	33.10	40.49	48.15	56.09
	4 years	16.99	21.55	26.25	31.08	36.05	41.16	46.41	57.35	68.90	81.06
	5 years	21.67	27.63	33.82	40.26	46.93	53.86	61.05	76.23	92.54	110.03
	6 years	26.53	34.01	41.85	50.07	58.69	67.71	77.16	97.38	119.50	143.64
	7 years	31.59	40.71	50.36	60.58	71.38	82.80	94.87	121.07	150.23	182.62
	8 years	36.86	47.75	59.38	71.82	85.09	99.26	114.36	147.60	185.26	227.84
	9 years	42.33	55.13	68.95	83.85	99.90	117.19	135.79	177.31	225.19	280.30
	10 years	48.02	62.89	79.08	96.72	115.89	136.74	159.37	210.58	270.72	341.14
	12 years	60.10	79.59	101.22	125.22	151.82	181.27	213.84	289.60	381.79	493.60
	15 years	80.09	107.89	139.66	175.90	217.22	264.25	317.72	447.36	613.79	826.55
	20 years	119.11	165.33	220.71	286.97	366.10	460.44	572.75	864.63	1,274.35	1,846.08

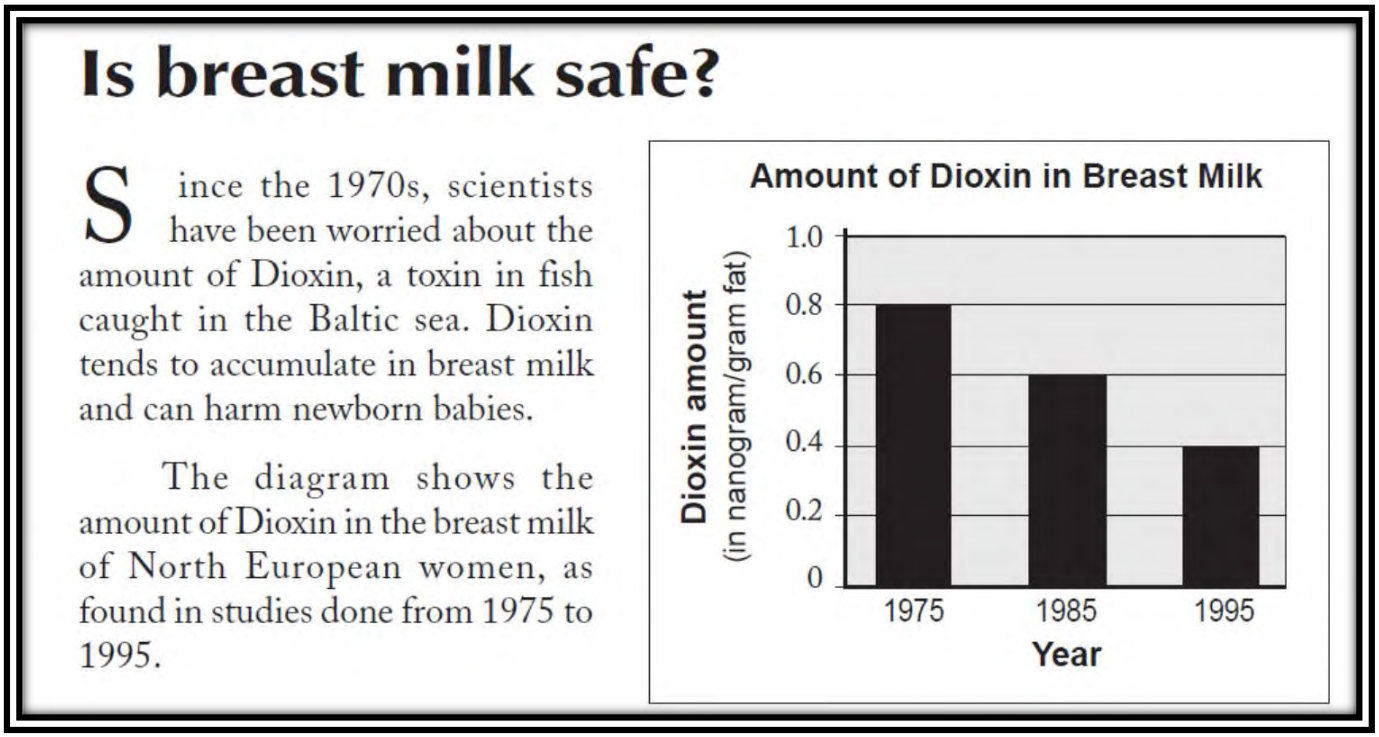
Question: Calculate the total amount of money you will have if you invest $100 at a rate of 6% for 10 years.

ALLS NUMERACY ITEM - LEVEL 5 - IS BREAST MILK SAFE?

Sources:

- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.301 & 302

Consider the following image:



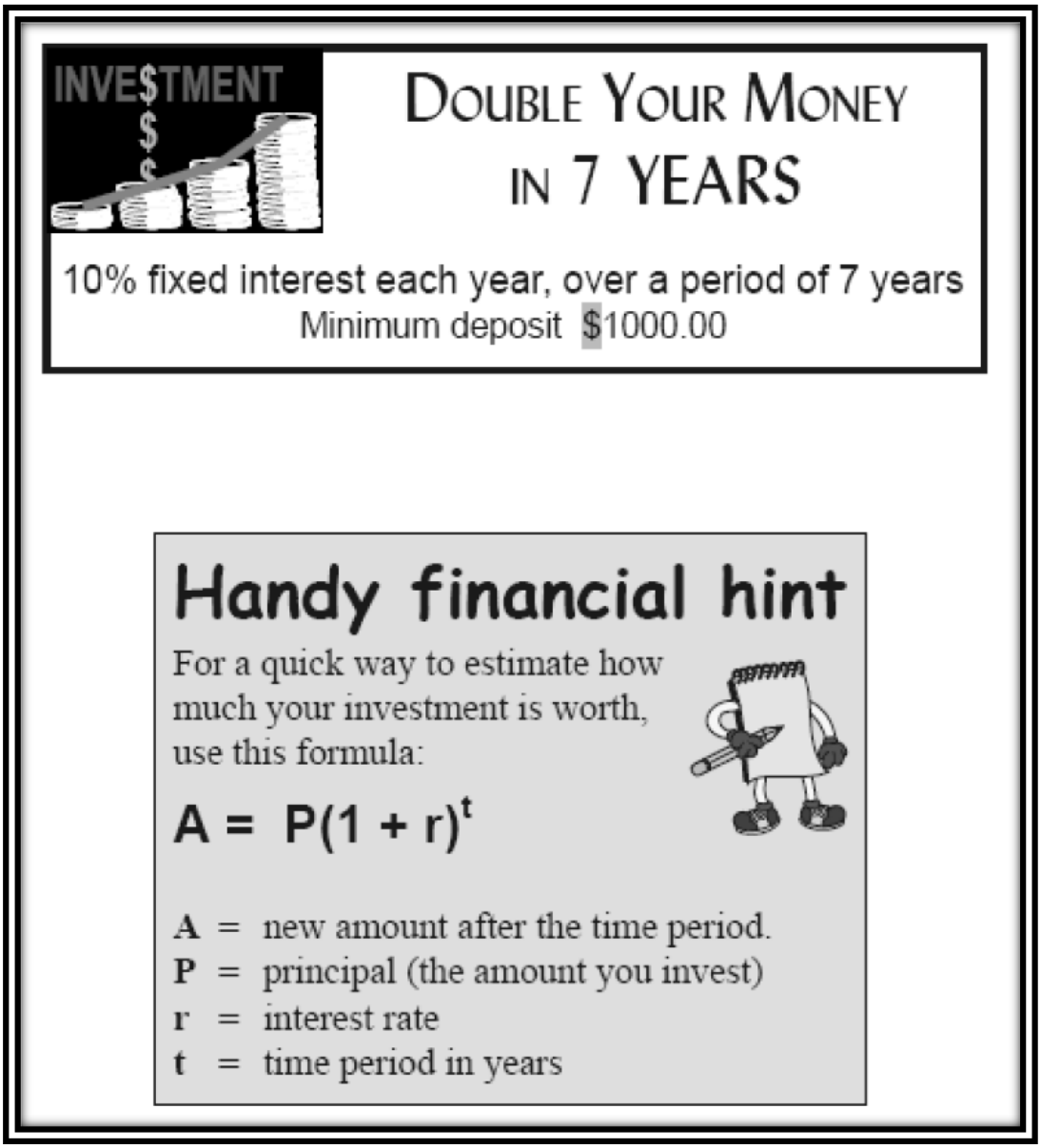
Question: Compare the change in Dioxin level from 1975 to 1985 to the percent of change in Dioxin level from 1985 to 1995. Determine which percent of change is larger, and explain your answer.

ALLS NUMERACY ITEM - LEVEL 5 - DOUBLE YOUR INVESTMENT

Sources:

- Stats Can/OECD 2005, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, p.302 (**used this source mainly**)
- ABS 2006 Adult literacy and life Skills Survey, User Guide, Catalogue No. 4228.0.55.002, ABS, p.38

Consider the following image:



Question: Is it possible to double $1000 invested at this rate after seven years? Support your answer with calculations.

Appendix 4: Terms and Conditions

Terms and Conditions: Mapping of Adult Literacy Performance Survey – Victoria University

Respondents who have completed the Mapping of Adult Literacy Performance Survey during March to May 2012 will have the option to enter into a prize draw to win one of six Apple iPads with Wi-Fi 32GB (3rd generation). Respondents will be given an opportunity to opt into the prize draw after they have completed the on-line survey. Survey respondents can enter the prize draw by answering 'Yes' to the prize draw opt-in question (this question will appear on the last survey page completed).

1. Information on how to enter and prizes form part of these Terms & Conditions. Any entry not complying with these Terms & Conditions is invalid.
2. The Promoter is Victoria University (ABN 8377694731) of 8-18 Whitehall Street, Footscray, VIC, 3011.
3. Entry is open to all respondents who complete the Mapping of Adult Literacy Performance Survey.
4. Entrants must complete the Mapping of Adult Literacy Performance Survey in order to be eligible to enter the prize draw.
5. For each survey completed, a respondent can enter into the prize draw. There is no restriction of how many times a respondent can complete a survey and enter into the prize draw.
6. Victoria University reserves the right to verify the validity of entries and to disqualify any entrant who subverts or attempts to subvert the entry process or who submits an entry not in accordance with these Terms & Conditions.
7. The survey and associated prize draw commences on the 26th March 2012 and entry will be closed 7th May, 2012. Six winners will be randomly drawn on 8th May 2012 at Victoria University, 8-18 Whitehall Street, Footscray, VIC 3011 at 10am (AEST).
8. Six Prizes will be drawn. Each 'Prize' is an Apple iPad with Wi-Fi 32GB (3rd generation) valued at AUD $649 (RRP) each.
9. Prize is not transferable and cannot be taken as cash.
10. Prize value is correct at time of publication, but no responsibility is taken for any variation in the value of the prize.
11. Victoria University will not be held responsible for the loss, theft or damage to any prize after it has been awarded, or for any injury that results directly or indirectly from this promotion.
12. Victoria University is not responsible for receipt of incorrect, inaccurate or incomplete information, caused by an entrant or occurring during transmission.
13. Prize winners will be notified by email to the email address supplied during entry. The winners will also be published on the National Centre for Vocational Education Research's website (www.ncver.edu.au) on the 10th May 2012. The prize will be forwarded to their home address within 15 working days of the draw. Victoria University will make reasonable attempts to contact the prize winners by email and telephone.
14. Victoria University's decision is final and no correspondence will be entered in to.
15. Victoria University cannot amend the terms & conditions of the prize draw without approval of all the relevant state and territory lottery departments.
16. By entering the prize draw, entrants will be deemed to have accepted these terms & conditions.
17. Unclaimed prizes will be kept for three months after the winners are drawn. If unclaimed after that time another winner will be drawn on the 9th August, 2012 at 10am at Victoria University, 8-18 Whitehall Street, Footscray, VIC 3011. The redrawn winner(s) will be notified by email and telephone. Prizes will not jackpot. The redrawn winners will also be published on the National Centre for Vocational Education Research (NCVER) website (www.ncver.edu.au) on the 13th August, 2012
18. Any queries arising from the interpretation of these Terms & Conditions may be raised by emailing shelley.gillis@vu.edu.au.

INFORMATION TO PARTICIPANTS INVOLVED IN RESEARCH

You are invited to participate

You are invited to participate in a research project entitled *Mapping of Adult Literacy Performance.*

This project is being undertaken by the National Centre for Vocational Education Research, Victoria University and Educational Measurement Solutions (Pty Ltd). It is funded by the Department of Education, Employment and Workplace Relations.

Project explanation

Language, literacy and numeracy are critical for greater workforce participation, productivity and social inclusion with research demonstrating the relationship between increasing levels of language, literacy and numeracy proficiency and positive outcomes for individuals, communities and the economy. Being able to measure how skilled people are, and any changes in their level of skill, is important for getting a sense of how well language, literacy and numeracy programs are working for learners.

In Australia, among the tools used to determine the literacy and numeracy skill of learners are the Australian Core Skills Framework (ACSF) and the Adult Literacy and Life Skills (ALLS) survey. The ALLS is used internationally and is done every 10 years of so. The ACSF contextualises learning and provides evidence of progress so that at any point in time, a learner's performance in a core skill can be assessed and their strengths and weaknesses identified. Both have five performance levels and it is sometimes assumed that these levels are equal. But are they?

Given the ALLS levels are used by the federal government as a benchmark for one of the goals in the 2008 National Skills and Workforce Development Agreement, the Department of Employment, Education and Workplace Relations (DEEWR) wants to know whether performance levels on the ALLS literacy and numeracy scales can be reliably mapped to the performance levels of the ACSF, essentially meaning that ACSF performance levels could be used as a proxy for ALLS performance levels. In essence, does 1 = 1?

If this can be done, we can then get more frequent information on the literacy and numeracy progression of particular groups of adult learners against national goals.

We are not evaluating the usefulness of either the ALLS or the ACSF.

What will I be asked to do?

We've developed an online survey and we're asking teachers/tutors/lecturers familiar with adult literacy and numeracy concepts to anonymously rate a student, whose literacy and/or numeracy levels are most familiar to them, against statements and sample tasks drawn directly from both the ACSF and ALLS frameworks. These anonymous ratings will be analysed and placed onto the same scale of measurement using Item Response Theory. This is the most direct method for determining and comparing the complexity of the two frameworks.

If you wish to voluntarily complete the survey, it should only take about 20 minutes to do. Please head to [insert link] to begin the survey. The survey will be available until <<insert date>>.....

Please help us by encouraging your colleagues to participate.

What will I gain from participating?

By participating, not only will you be helping us find out whether it is possible to map the ACSF to the ALLS, but you'll also go in the draw to win one Apple iPad 2 – 32GB Wi-Fi valued at AUD $689 (RRP) each. You will be given the option at the end of the survey to enter the draw by selecting the "Yes" button to the prize draw-opt in question. If they choose to enter the draw, you will be asked to supply your contact details (e.g. name, address and email) so that if you win, you can be notified. If you elect to participate in the draw, your responses to the survey will be kept separate to your contact details so that your survey responses will remain anonymous during data storage, data analysis and reporting. Your contact details will only be used for purposes of the draw.

If you win one of the six prizes, you will be notified directly by email. The winners will also be published on the NCVER's website (www.ncver.edu.au). See the Terms and Conditions for more information about the draw.

How will the information I give be used?

Information from the survey will be analysed to see how the five literacy and numeracy ACSF levels relate to those of the ALLS. The information will be used for research purposes only. The findings of the study will only be reported at the aggregate level and no individual participants will be identified in any reports that arise from the study.

What are the potential risks of participating in this project?

The findings will be reported at the aggregate level which means that no individual participants will be identified in any reports that arise from the research. There are also no wrong or right answers to the survey questions, we are just seeking your professional judgement to assist us with determining how the statements from within the ACSF and ALLS relate to each other. You can postpone or end the on-line survey at any time whilst maintaining anonymity. If you do elect to enter the draw to win one of six IPADs 32GP WiFi, then your contact details will be kept separate to your survey responses, so that we will not be able to identify who said what in the survey.

Who is conducting the study?

This project is being conducted by Associate Professor Shelley Gillis and Geri Pancini from the Work-based Education Research Centre at Victoria University, together with Dr Margaret Wu, Director, Educational Measurement Solutions and Mark Dulhunty, Director, Educational Measurement Solutions.

The project is being undertaken jointly with the National Centre for Vocational Education Research. It has been funded by the Department of Education, Employment and Workplace Relations (DEEWR)>

Any queries about your participation in this project may be directed to the Chief Investigator, Associate Professor, Shelley Gillis on the following email: Shelley.Gillis@vu.edu.au

If you have any queries or complaints about the way you have been treated, you may contact the Research Ethics and Biosafety Manager, Victoria University Human Research Ethics Committee, Victoria University, PO Box 14428, Melbourne, VIC, 8001 or phone (03) 9919 4148.

Appendix 6: Pilot participants

fname	lname	organisation
Jenni	Oldfield	Precision Consultancy
Chris	Tully	Kangan TAFE
Lindee	Conway	Community West
Dave	Tout	ACER
Kate	Perkins	Private Consultant
Philippa	McLean	Private Consultant
Linda	Wyse	Private Consultant
Juliette	Mendelovits	ACER
Michelle	Circelli	NCVER
Geri	Pancini	VU
Rob	McCormack	VU

Appendix 7: ACSF Level Summary Statements

Performance level	ACSF	
	Reading	**Numeracy**
1	Identifies the main idea in short simple texts of personal interest with a highly explicit purpose and highly familiar vocabulary Locates one or two pieces of information and makes simple connections, e.g. matches, groups	A person at this level can find and recognise highly explicit numerical information in short and simple activities or texts, and use simple mathematical (e.g. sorting, ordering, adding, subtracting) and personal problem-solving strategies in highly familiar contexts. Everyday, informal and mainly oral language is used to express mathematical concepts. Appropriate numeracy tasks will involve 1 or 2 similar processes, e.g. locating, adding or subtracting, recognising numerical information.
2	Begins to skim and scan to identify relevant information and ideas in short, unambiguous texts with simple structures and syntax and familiar vocabulary Compares and contrasts information and draws low level inferences	A person at Level 2 can identify and understand the relevant and simple mathematical information in familiar texts or activities and can use an appropriate mathematical (including simple operations with +, –, x and ÷) and problem-solving strategies in familiar contexts. A combination of mainly informal and some formal oral and written mathematical and general language is used to express mathematical concepts. Appropriate numeracy tasks will involve a limited number of familiar processes, e.g. identifying, simple calculating and measuring, comparing and contrasting.
3	Identifies the main messages in longer routine texts requiring integration of a number of ideas and pieces of information and containing some specialised vocabulary. Understands texts that may include some unfamiliar elements, embedded information and abstraction, and contain simple diagrams and charts Applies several steps and processes such as sequencing, interpreting, simple extrapolating, inferencing and abstracting.	Someone at Level 3 can find and interpret mathematical information that may be partly embedded in both familiar and less familiar tasks and texts, and use a variety of routine mathematical and problem-solving strategies. A combination of both informal and formal mathematical oral and written mathematical language is used to express mathematical concepts. Appropriate numeracy tasks include a number of steps within one task, such as sequencing, basic inferencing, simple extrapolation and integration, calculating with whole numbers and routine fractions, decimals and percentages, measuring, comparing and contrasting.
4	Understands complex texts that may contain embedded information, abstraction and symbolism and incorporate information presented in graphic, diagrammatic or visual form Texts may contain technical specificity and	A person at Level 4 can extract and evaluate mathematical information embedded in a range of tasks and texts and selects and applies a range of mathematical and problem-solving strategies. Appropriately uses a range of oral and written informal and formal mathematical language to express mathematical concepts. Appropriate numeracy tasks may involve more

Performance level	ACSF	
	Reading	**Numeracy**
	use specialised vocabulary Reading involves complex task analysis and application of a number of processes, such as extracting, extrapolating, inferencing, reflecting and abstracting	complex task analysis involving application of a number of processes such as collecting, organising and representing data, extracting, comparing, interpreting mathematical information, undertaking a range of calculations including using routine and familiar formulae.
5	Understands complex, lexically dense texts containing highly embedded information, symbolism, cultural references and technical vocabulary. Reading involves sophisticated task conceptualisation, organisation and analysis e.g. selecting, synthesising and critically reflecting on and evaluating evidence, arguments and ideas	A person at Level 5 understands, analyses and synthesises highly embedded mathematical information across a broad range of tasks or texts and chooses and applies highly developed mathematical and problem-solving strategies and communicates mathematical concepts and outcomes using a wide range of informal and specialised oral and written mathematical language. Appropriate mathematical tasks are complex including interpretation, analysis, reflection, synthesis, evaluation and recommendations when applying statistical techniques to analyse data, when calculating with rational numbers and formulae, when measuring and calculating quantities.

Appendix 8: Pilot Participants Invitation

Dear XXX

As you may be aware, NCVER has commissioned Victoria University (Shelley Gillis, Geri Pancini) in conjunction with Educational Measurement Solutions (Margaret Wu and Mark Dulhunty) to empirically align the Australian Core Skills Framework (ACSF) to the international Adult Literacy and Life Skills (ALLS) Survey. This study is a follow up to a preliminary study undertaken last year in which a small number of adult literacy and numeracy experts used their professional judgement to qualitatively align a sample of completed ALLS items to the ACSF performance descriptors. If you are not familiar with the preliminary study, a copy of the report can be found at:
http://www.ncver.edu.au/publications/2463.html

The current study involves a national survey of the adult literacy, language and numeracy teaching profession. We will be asking individuals familiar with teaching adult literacy and numeracy concepts to anonymously rate a learner that they are familiar with against statements and sample tasks drawn directly from both the ACSF and ALLS frameworks. The ratings of statements drawn from both frameworks will then be analyzed and placed onto the same scale of measurement using Item Response Theory.

Given your expertise in LL&N, we need your help to pilot the survey before we begin data collection in March 2012. The purpose of the pilot study is to identify any current and/or potential problems with the draft survey. This is where we need your help.

If you agree to participate, we will be asking you to firstly complete the draft survey as though it was part of the real data collection. This should take around 20 minutes. Please note that your responses to the survey will not be kept or used by the research team in anyway. We are only asking you to complete the survey as though you are a participant so that you can explore both the content and the functions of the survey through the eyes of a potential respondent.

After you have had a chance to familiarise yourself with the features of the survey, we will then ask you a few questions about your experience. This should take around 5 minutes to complete. For example, we would like to know whether you experienced any problems with completing the survey and whether you have any recommendations for improving the survey.

As you complete the survey, you will see that we are trying to encourage as many people as possible to complete the survey by giving participants the opportunity to enter a draw to win 1 of 6 iPad 2s. Unfortunately, there are no prizes to be won for participating in the pilot study. However, we will be encouraging you to complete the survey again once it does go live in March. We hope that you participate in both the pilot study and the main study so that you will be given the opportunity to win an iPad!

Please also note that as the ACSF is still under embargo, this email has been sent to you **in-confidence**. Please do not distribute this email or forward the link to the survey until after the public release of the ACSF.

Please also note that we need your feedback by COB ***Monday, the 20th February, 2012.*** If you have any queries or concerns, please do not hesitate to contact me.

If you agree to participate in the pilot study, please click on the link below:
http://literacymap.com/s3/pilot

Your feedback is extremely important to us.

Thank you

Shelley

Associate Professor Shelley Gillis
Deputy Director, Work-based Education Research Centre
Unit 11, Level 2, 8-18 Whitehall Street
Footscray VIC 8001
Victoria University
Ph: 61 3 9919 1615
Mobile: 0432 756 638
email: shelley.gillis@vu.edu.au
web: www.werc.vu.edu.au

Appendix 9: Pilot Report

Pilot Aim

A small pilot study was undertaken to examine the usability and functionality of the on-line survey prior to the site being launched in March to empirically map the levels of the ACSF to the ALLS. Specifically, the pilot study sought to examine the:

- Appropriateness of the workload of each participant;
- Appropriateness of the background questions
- Appropriateness and ease of completion of the item formats;
- Clarity of the instructions to complete each section of the survey;
- Ease of navigation throughout the on-line survey; and
- Ways in which the survey could be improved.

The Pilot Participants

Twenty-eight individuals were nominated by the panel workshop participants as potential participants in the pilot study due to their expertise within adult literacy, language and/or numeracy. However, as the ACSF was still under embargo at the time of piloting, only 3 of the 28 individuals nominated were approached to participate in the pilot study. In addition to the three nominated individuals, each panel member was also invited to participate in the pilot. As such, 11 individuals in total were invited to voluntarily participate in the pilot study. See Appendix 6 for a listing of the names of the individuals invited to participate in the pilot study. Of the 11 individuals invited to participate, 9 completed the on-line feedback form and 1 provided general feedback. Hence, there was a 90% response rate for the pilot study.

The Feedback Questionnaire

A series of closed and open ended questions were developed to gather feedback from the participants on how the survey design and functions could be improved. This feedback questionnaire was administered on-line. A listing of the questions, item format and response values for each feedback question has been displayed in Table 34.

Table 34: Pilot Feedback Questionnaire: Stimulus, response values and item format.

Stimulus	Item Format	Response Value
1. Approximately how long did you take to complete the survey?	Multiple choice	Less than 10 minutes
		10-19 minutes
		20-29 minutes
		30-39 minutes
		> 40 minutes
2. Did you find the background questions adequately captured the range of background characteristics of those teaching in the adult literacy and numeracy field (i.e., the potential pool of respondents for the survey)?	Dichotomous	Yes/No
If no, can you please indicate which questions need to be revised and how they could be improved?	Open ended	
3. Did you find the survey instructions easy and clear to understand?	Dichotomous	Yes/No
If no, Which instructions were not easy and/or clear to understand? How would you like to see them improved?	Open ended	
4. Did you find some parts of the survey difficult to complete?	Dichotomous	Yes/No
If no, can you please explain which parts of the survey were difficult and why? How would you like to see this/these part(s) of the survey improved?	Open ended	
5. Did you experience any technical/IT related problems when completing the survey (eg formatting issues, image resolution)?	Dichotomous	Yes/No
If no, can you please provide more details of: • The nature of the problem • The internet browser you were using (eg Internet Explorer 8) • The device (eg iPad, PC, Netbook).	Open ended	
6. Did you experience any difficulties navigating your way around the survey? For example, using the scroll features or the 'save and continue later' features of the survey.	Dichotomous	Yes/No
If no, can you please provide details of the problem(s) you experienced.	Open ended	
7. Do you have any suggestions of how the survey could be improved to be more engaging for the survey participants?	Open ended	
8. Any other comments?		Open ended

DATA COLLECTION PROCEDURES

Each potential participate was invited to participate in the pilot study via an email invitation. A copy of the invitation sent to each individual can be found in Appendix 8. In summary, the invitation outlined the purpose of the pilot study, the in-confidence nature of the survey (due to the embargo on the ACSF), the process for undertaking the pilot study, as well as the voluntary nature of participation. It also included a direct link to the survey in which the following instructions were included:

> Before beginning the survey, please read all the instructions below.
>
> 1. Work your way through the survey and try to answer each question honestly and accurately as possible. Answer each question as though you were doing it as part of the real study (note your responses will be confidential).
> 2. Whilst completing the survey, attempt to use as many of the survey features as possible (eg roll-over your mouse, try to use the hyper text links as well as the 'Save & Continue Later' function).
> 3. Note how long it takes you to complete the survey.
> 4. Keep notes on any questions or features of the survey that you think should be revised or improved.
> 5. Once you have had sufficient opportunity to interrogate the features of the survey, please complete a short feedback form (which takes less than five minutes to complete) by visiting http://literacymap.com/s3/feedback

Data collection occurred over a one week period.

RESULTS

Following are the results from each question administered in the pilot feedback survey.

It can be seen in Figure 20 that the majority of pilot participants completed the survey within 10-19 minutes.

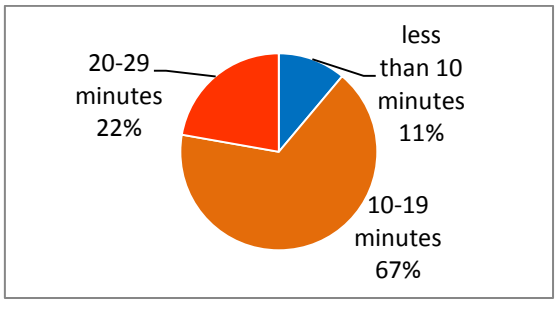


Figure 20: Pilot Q.1 Approximately how long did you take to complete the survey?

One respondent commented….

"The survey was much shorter than I expected. I was surprised when I realised that I had completed the questions in Part A and then again when I finished Part B."

Recommendation 1: The marketing and survey instructions in terms of the approximate amount of time for completing the survey be amended from '*20 minutes'* to '*10 to 20 minutes'* to more accurately reflect the pilot findings.

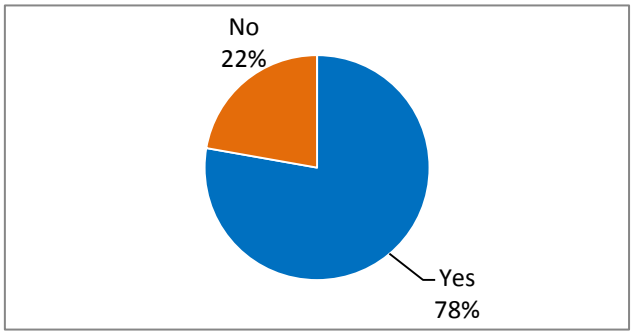


Figure 21: Pilot Q.2 Did you find the background questions adequately captured the range of background characteristics of those teaching in the adult literacy and numeracy field (i.e., the potential pool of respondents for the survey)?

The majority of participants found the background questions adequately captured the range of background characteristics of those teaching in the adult literacy and numeracy field. Of the two respondents who disagreed with this statement, both raised queries with the item that required them to make a holistic judgement of the ACSF level of the learner. In both instances, the respondents indicated that their learner was "operating at a level considerably below Level 1" and wondered whether it would be possible to include ACSF descriptions below Level 1. However, as the current study was limited to the level statements and publicly available scaled items from within the ALLS (as opposed to the yet to be released PIAC), future equating studies could be undertaken to link the pre-levels within both the PIAC to the newly developed pre-levels within the ACSF. Such an equating exercise will be feasible and cost effective as only a sample of items that have been calibrated in this study (i.e., from levels 1 to 5 on both frameworks) would need to be included in a subsequent equating study, in addition to a sample of pre-level statements/items from both frameworks.

Recommendation 2: At the completion of stage 2, further research be undertaken to equate the PIAC and ACSF pre-levels on the same scale using common item equating procedures.

To overcome any potential difficulties respondents may experience with judging learners at or below Level 1 of the ACSF within the current study, it is recommended that the stem of the questionnaire be slightly reworded.

Recommendation 3: In relation to the background question that required the respondents to make a holistic judgement of the ACSF level of the learner, it is recommended that the stem be revised to reflect the notion of the "*level that best describes the learner*" (as opposed to the level which the "*learner is typically performing at*").

As can be seen in Figure 22, all the pilot participants found the survey instructions easy and clear to understand.

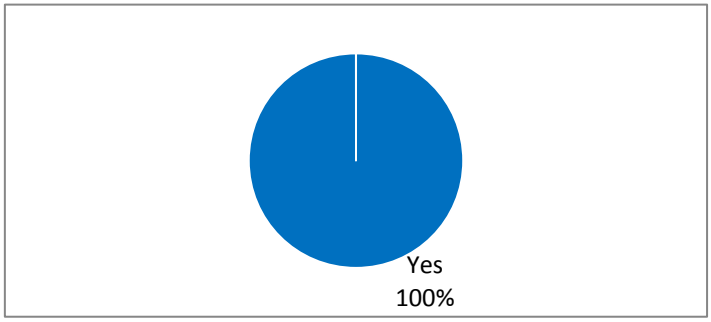


Figure 22: Pilot Q.3 Did you find the survey instructions easy and clear to understand?

However, one of the 9 participants who completed this feedback questionnaire found some parts of the survey difficult to complete (see Figure 23).

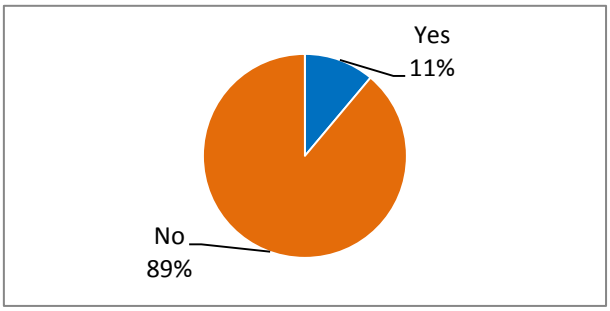


Figure 23: Pilot Q.4 Did you find some parts of the survey difficult to complete?

In particular, the respondent found some of the statements within the frameworks difficult to judge. That is, s/he stated that….

"Some of the questions where too complex. ie "Can locate and use mathematical information in a range of different forms eg numbers, symbols, maps, graphs. I know <<Learner>> is comfortable with some of these but not all so I was forced to make a decision as to what to click. This was also the case where the question did not specify degree of difficulty. For example "uses knowledge about chance and probability to estimate and interpret the outcomes". <<Learner>> can do this with simple concepts that are familiar to her but not at a higher level, so again I made a decision based on what I thought the question was asking."

Although this issue could not be dealt with as part of the survey design, as the statements have been sourced directly from the ACSF and the ALLS, it should be acknowledged that as each statement formed a separate item on the survey, this will enable diagnostic information to be gathered on each item which may inform future revisions to the ACSF in terms of clarity etc.

No technical/IT related problems were experienced by the pilot participants, as displayed in Figure 24.

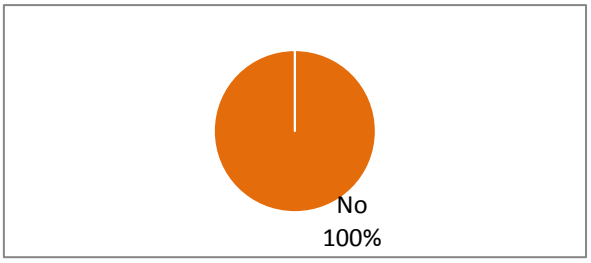


Figure 24: Pilot Q.5 Did you experience any technical/IT related problems when completing the survey (eg formatting issues, image resolution)?

Similar findings were found in terms of navigation, with only one respondent indicating s/he experienced some problems (see Figure 25). This respondent suggested that future respondents should be explicitly instructed to scroll to see all the scaled item presented in Part B of the survey (similar to the instructions that were presented in the tablet/mobile version of the survey).

Recommendation 4: In Part B of the survey in which the ALLS scaled items are presented, the statement "You will have to scroll" should be added to the main version survey (as was the case for the ipad/tablet friendly version).

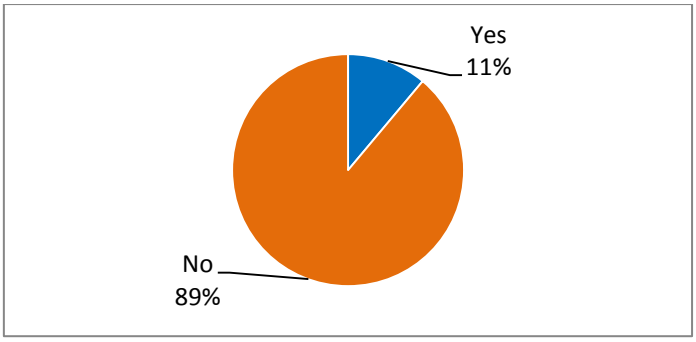


Figure 25: Pilot Q.6 Did you experience any difficulties navigating your way around the survey? For example, using the scroll features or the 'save and continue later' features of the survey

A number of suggestions for improving the design and layout of the survey was also recommended by the respondents. These have been summarised in the table below, alongside the resolutions that were made by the research team. In addition to these resolutions, all recommendations proposed in this study, with the exception of Recommendation 2, have been completed.

	Feedback	Resolution
F1	*Suggest that when people have finished doing the survey/s, can they just exit, and not go to NCVER website. They will not necessarily be entering the survey from NCVER website so may be somewhat confused to finish there.*	Link removed
F2	*For Part A categories, is there a need for an 'every time' category, ie, more than just 'almost always'. This applies to Part B as well*	Change Part A instructions, question stem and rating scale to be similar to Part B in that they refer to "how likely is it that the learner would be able to independently perform this task?" • Not Very Likely • Somewhat Likely • Very Likely
F3	*When someone enters the prize draw and enters their contact details, do you think we need to expand the assurance to say that the details will be deleted following the prize draw (or something sufficiently vague to cover the prize draw and potential re-draw.*	Change made
F4	*When I did another survey focusing on numeracy, there were some statements that were essentially repeated (up to 3 times) with just minor differences in wording. Is this intentional?*	No change necessary as this was intentional and due to the fact that we have included the original ALLS level statements plus the panel modified statements to ensure that the original intent of the statements have not been compromised.
F5	*With the group of 'reading' statements I looked at in Part A, one was contradictory with the instructions (i.e, consider that the learner would be able to independently perform etc), with it stating "With assistance, identifies one or two questions that reading a text may answer".*	Replace with another performance feature at the same level and focus area and within the same domain. Original: R2P20303 _ With assistance, identifies one or two questions that reading a text may answer Replacement: R2P20301 _ Identifies texts in the immediate environment that are relevant to own needs and interests

	Feedback	Resolution
F6	*With 'Part A – Statements', any possibility this header could be a little more descriptive?.* *Will the "Does = 1" Mapping Measures of adult language, literacy and numeracy" header be used or what is in the pilot?*	Headings changed throughout survey. For example, in part A, the heading is now… Does 1=1? Part A - Mapping statements of adult language, literacy and numeracy
F7	*Participants may not like that they cannot view the whole image without scrolling down (this applies to all the ALLS scaled items).*	No change necessary No other issues reported related to this. The scroll bar was the best way to present the items that varied in image quality, size and resolution.
F8	*When rating the ACSF statement, there a couple of performance features that use the term 'limited' – this led to some confusion when rating the student against that in terms of rarely, sometimes always etc (as do they always have limited texts etc...).*	This issue has been overcome with the change to the instructions, stem and rating scale for Part A of the survey (as specified in F2), which now refers to how likely would the candidate be able to perform this task, as opposed to how often they perform this task.
F9	*Perhaps a brief explanation to differentiate the terms 'numeracy' and 'mathematics' would help non-Numeracy practitioners to get a clear picture of the question...*	No change undertaken
F10	*I thought part B was more interesting and therefore easier to engage with. I could imagine the learner better with an actual problem in front of me .I found this easier than judging abstract statements.*	No change necessary
F11	*As I am not currently working with anyone with reading issues, I deliberately selected a 4th year uni student I know very well to see how the high levels might look. Not surprisingly, as she is a strong performer, I know she is likely to be able to do all that was described . Not sure if this helps or not really, but I guess it supports the thrust of the general descriptors, whatever the source.... As this survey is being pitched more at people working with learners who may need assistance with reading, I'm sure other respondents will be able to provide more useful input around the lower levels.*	No change necessary
F12	*Diagram mentions Canadian $$ while question asks for Australian $$*	Changes made to reflect Australian $$.
F13	*Note a couple of typos- 'inferences'. 'complete tasks' instead of 'completes tasks' several American z/s spellings - e.g. synthezises.* *There's a typo in the Intro section under "What is the study about?", see 2nd para, 2nd sentence: "...every 10 years or so"*	Typographical errors corrected but American spelling from ALLS statements may still be present (although every attempt has been made to identify these)
F14	*Can we bold, "Please help us by encouraging your colleagues to participate" (in the 'What will I be asked to do' bit)*	Change made
F15	*In the Part A Instructions box, can we bold references to 'independently'. Did this for the Part B instructions too.*	Changes made
F16	*Change dates in Terms of Conditions*	Amendment lodged with appropriate lottery commission and Terms and Conditions updated.

Appendix 10: Unique Identification Codes and Statement Descriptors

Domain	Framework	Level	ID	Type	Statement
Reading	ACSF	1	R1I10301	Indicator	Identifies personally relevant information and ideas from texts on highly familiar topics
Reading	ACSF	1	R1P10301	Performance Feature	Identifies personally relevant reasons for reading
Reading	ACSF	1	R1P10306	Performance Feature	Recognises some simple ways in which visual features like layout are used to send a message and how this may influence interpretation, e.g. the placement of a photo or heading in a newspaper
Reading	ACSF	1	R1I10401	Indicator	Uses a limited range of strategies to locate specific information and construct meaning from explicit and highly familiar texts
Reading	ACSF	1	R1P10405	Performance Feature	Uses a limited range of decoding strategies for unfamiliar words, e.g. sounding out letters and syllables
Reading	ACSF	1	R1P10407	Performance Feature	Uses knowledge of familiar phrases to predict the next word, e.g. Once upon a …
Reading	ACSF	1	R1P10408	Performance Feature	Recognises some basic punctuation and understands its use in meaning-making, e.g. a full stop
Reading	ACSF	1	R1P10409	Performance Feature	Recognises a number of high frequency words/basic sight words and common phrases, e.g. down, would, have, little, come, when
Reading	ACSF	1	R1P10411	Performance Feature	Uses simple strategies to assist with word identification and extend vocabulary, e.g. a pictorial or bilingual dictionary, or a personal word list
Reading	ACSF	2	R2I20301	Indicator	Identifies and interprets relevant information and ideas from texts on familiar topics
Reading	ACSF	2	R2P20301	Performance Feature	Identifies texts in the immediate environment that are relevant to own needs and interests
Reading	ACSF	2	R2P20302	Performance Feature	Recognises that some texts are more appropriate for a purpose than others
Reading	ACSF	2	R2P20311	Performance Feature	Recognises the difference between formal and informal registers in simple familiar texts
Reading	ACSF	2	R2P20312	Performance Feature	Understands that the meaning of some familiar words and phrases may change in different contexts
Reading	ACSF	2	R2I20401	Indicator	Uses a number of reading strategies to identify and interpret relevant information within familiar text types
Reading	ACSF	2	R2P20402	Performance Feature	Recognises some features of diagrammatic texts, e.g. grid references, dot points or arrows
Reading	ACSF	2	R2P20403	Performance Feature	Begins to skim and scan familiar texts, using pictures and graphics to help locate specific information
Reading	ACSF	2	R2P20411	Performance Feature	Uses a dictionary or online resource to check word meanings
Reading	ACSF	3	R3I30301	Indicator	Evaluates and integrates information and ideas to construct meaning from a range of familiar, and some unfamiliar, texts and text types
Reading	ACSF	3	R3P30302	Performance Feature	Begins to reflect on the usefulness of a selected text for the purpose

Domain	Framework	Level	ID	Type	Statement
Reading	ACSF	3	R3P30303	Performance Feature	Understands familiar texts of limited complexity that may incorporate graphs, tables and charts
Reading	ACSF	3	R3P30309	Performance Feature	Identifies the purpose and intended audiences of a range of familiar, and some unfamiliar, text types
Reading	ACSF	3	R3P30313	Performance Feature	Interprets and extrapolates information from texts containing graphs and diagrams
Reading	ACSF	3	R3I30401	Indicator	Selects and applies a range of reading strategies as appropriate to purpose and text type
Reading	ACSF	3	R3P30405	Performance Feature	Uses explicit strategies to make connections between information and ideas while reading, e.g. margin notes or simple diagrams
Reading	ACSF	3	R3P30413	Performance Feature	Understands that some words and phrases have figurative meanings
Reading	ACSF	3	R3P30414	Performance Feature	Routinely uses dictionaries and other references to determine the meaning of unknown words
Reading	ACSF	4	R4I40301	Indicator	Interprets and critically analyses complex texts
Reading	ACSF	4	R4P40301	Performance Feature	Actively identifies an explicit purpose for reading, e.g. to gather background information, identify specific facts or understand a concept
Reading	ACSF	4	R4P40304	Performance Feature	Understands texts with complex syntactic structures that may incorporate some technical specificity and information presented in graphic, diagrammatic or visual form
Reading	ACSF	4	R4P40309	Performance Feature	Explicitly identifies some ways in which an author uses structure, language and tone to create an impression and explain or reinforce a message, e.g. through choice of text structure, use of rhetorical questions, repetition, simile and metaphor or figures of speech
Reading	ACSF	4	R4I40401	Indicator	Applies appropriate strategies to construct meaning from complex texts
Reading	ACSF	4	R4P40401	Performance Feature	Selects an appropriate reading approach according to text structure and purpose, e.g. skims and scans to identify areas of interest, reads closely to identify explicit and implicit information and ideas, understands that reports do not necessarily need to be read sequentially and uses contents page and headings to find relevant sections
Reading	ACSF	4	R4P40405	Performance Feature	Uses a range of strategies to synthesise ideas and information from several texts, e.g. constructs mind maps to show connections between ideas
Reading	ACSF	4	R4P40409	Performance Feature	Responds to transitional words and phrases that signal important information or a shift in focus
Reading	ACSF	4	R4P40410	Performance Feature	Understands how linking devices are used to demonstrate conceptual connections and/or causal relationships
Reading	ACSF	5	R5I50301	Indicator	Organises, evaluates and critiques ideas and information from a range of complex texts
Reading	ACSF	5	R5P50302	Performance Feature	Understands highly complex, lexically dense texts, including those incorporating a high level of technical specificity
Reading	ACSF	5	R5P50304	Performance Feature	Draws on broad general knowledge to aid understanding of texts on a wide range of subjects and within specialised disciplines

Domain	Framework	Level	ID	Type	Statement
Reading	ACSF	5	R5P50306	Performance Feature	Builds breadth and depth of understanding by integrating prior knowledge with ideas and information from multiple texts
Reading	ACSF	5	R5P50308	Performance Feature	Identifies how social relations, register and audience influence an author's choice of text type, structure and language and how they may be used to express or hide attitudes and bias
Reading	ACSF	5	R5I50401	Indicator	Draws on a broad range of strategies to build and maintain understanding throughout complex texts
Reading	ACSF	5	R5P50401	Performance Feature	Recognises the distinguishing structures, layout, features and conventions of a broad range of complex text types and understands how to use these as an aid to locating information, developing understanding and focusing reading effort
Reading	ACSF	5	R5P50402	Performance Feature	Sustains reading activity for long periods when required using a range of strategies to maintain comprehension and integrate concepts, particularly when moving between texts, e.g. reference cards, matrices of key points or flow diagrams
Reading	ACSF	5	R5P50407	Performance Feature	Uses a range of support resources to extend understanding, and to investigate, research and reflect on the use of specific words and phrases
Reading	ALLS	1	R1LOrg01	Level Description (Original)	Reads relatively short text to locate a single piece of information which is given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information
Reading	ALLS	1	R1LOrg02	Level Description (Original)	Locates a piece of information based on a literal match and can enter information from personal knowledge onto a document in which there is little, if any, distracting information present.
Reading	ALLS	1	R1LMod01	Level Description (Modified)	Reads relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or instruction.
Reading	ALLS	1	R1LMod02	Level Description (Modified)	Locates specific information from a simple unambiguous text
Reading	ALLS	1	R1LMod03	Level Description (Modified)	Reads and understands instructions on simple form requiring personal details
Reading	ALLS	1	R1LMod04	Level Description (Modified)	Enters information from personal knowledge onto a document that contains little, if any, distracting information.
Reading	ALLS	2	R2LOrg01	Level Description (Original)	Locates a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present, or low-level inferences may be required. Integrates two or more pieces of information and can compare and contrast easily identifiable information based on the criterion provided in the question or directive.
Reading	ALLS	2	R2LOrg02	Level Description (Original)	Match a single piece of information; however, several distractors may be present, or the match may require low-level inferences. Cycles through information in a document or to integrate information from various parts of a document.
Reading	ALLS	2	R2LMod01	Level Description (Modified)	Locates a single piece of information in a text containing some distracting, yet plausible information.

Domain	Framework	Level	ID	Type	Statement
Reading	ALLS	2	R2LMod02	Level Description (Modified)	Integrates two or more pieces of information
Reading	ALLS	2	R2LMod03	Level Description (Modified)	Compares and contrasts easily identifiable information
Reading	ALLS	2	R2LMod04	Level Description (Modified)	Cycles through a text to distinguish relevant information from distracting information
Reading	ALLS	2	R2LMod05	Level Description (Modified)	Selects information from a text which may require integration from various parts of the document or low level inferencing
Reading	ALLS	3	R3LOrg01	Level Description (Original)	Makes literal of synonymous matches between the text and information given in the task. Makes matches that require low-level inferences. Integrates information from dense or lengthy text that contains no organisational aids such as headings. Can also generate a response based on information that can be easily identified in the text where distracting information is present, but not located near the correct information.
Reading	ALLS	3	R3LOrg02	Level Description (Original)	Integrates multiple pieces of information from one or more documents. Cycles through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.
Reading	ALLS	3	R3LMod01	Level Description (Modified)	Locates relevant information in a text where the information is literal or requires only low level inference.
Reading	ALLS	3	R3LMod02	Level Description (Modified)	Integrates information from dense text with no organisational aids such as heading.
Reading	ALLS	3	R3LMod03	Level Description (Modified)	Integrates information from lengthy text with no organisational aids such as heading.
Reading	ALLS	3	R3LMod04	Level Description (Modified)	Integrates multiple pieces of information from one or more documents.
Reading	ALLS	3	R3LMod05	Level Description (Modified)	Selects information from complex tables or graphs, where irrelevant or inappropriate information may be present.
Reading	ALLS	4	R4LOrg01	Level Description (Original)	Performs multiple-feature matches and integrates or synthesizes information from complex or lengthy passages. To perform these tasks successfully, more complex inferences are needed and conditional information must be taken into consideration.
Reading	ALLS	4	R4LOrg02	Level Description (Original)	Complete tasks that require multiple-feature matches, cycling through documents, and integrating information with some degree of inference. Can typically complete tasks that require numerous responses but the tasks do not designate how many responses are needed. Conditional information must be taken into account to complete these tasks.
Reading	ALLS	4	R4LMod01	Level Description (Modified)	Reads complex lengthy text, integrating or synthesising information that may require high level inferencing.
Reading	ALLS	4	R4LMod02	Level Description (Modified)	Reads complex texts, requiring: o Integration of multiple pieces of information o Inferencing at high level o Selection of appropriate responses where the number of responses is not designated

Domain	Framework	Level	ID	Type	Statement
Reading	ALLS	5	R5LOrg01	Level Description (Original)	Searches for information in dense text which contains a number of plausible distractors. Can typically complete tasks that require high-level inferences or specialized background knowledge. Can also complete tasks that ask the learner to contrast complex information.
Reading	ALLS	5	R5LOrg02	Level Description (Original)	Searches through complex displays that contain multiple distractors, to make high-level text-based inferences, and use specialised knowledge to complete the task.
Reading	ALLS	5	R5LMod01	Level Description (Modified)	Locates and contrasts complex information and ideas in dense texts which may require high level inferencing and specialised background knowledge.
Reading	ALLS	5	R5LMod02	Level Description (Modified)	Searches through complex texts containing several pieces of distracting information which may require high level thinking and specialised background knowledge.
Reading	ALLS	1	ALLS_R_Blac..	Question (FEW DUTCH WOMEN)	Identify from the chart the percentage of teachers from Greece who are women.
Reading	ALLS	2	ALLS_R_Imp1..	Question (IMPATIENS1)	What does the smooth leaf and stem suggest about the plant?
Reading	ALLS	2	ALLS_R_Imp2..	Question (IMPATIENS2)	What happens when the impatiens plant is exposed to temperatures of 14 degrees C or below?
Reading	ALLS	3	ALLS_R_Fire..	Question (FIREWORKS)	Write a brief description of the relationship between sales and injuries based on the information shown
Reading	ALLS	4	ALLS_R_Hiri..	Question (The Hiring Interview)	Write in your own words one difference between the panel and the group interview
Reading	ALLS	5	ALLS_R_CANC..	Question (CANCO)	List two ways in which CIEM (an employee support initiative within a company) helps people who lose their jobs because of departmental reorganization.
Numeracy	ACSF	1	N1I10901	Indicator	Locates and recognises key mathematical information in simple activities or texts
Numeracy	ACSF	1	N1P10901	Performance Feature	Locates and recognises simple, everyday mathematical information in highly familiar short and simple oral and/or written materials where the mathematics is highly explicit
Numeracy	ACSF	1	N1I11001	Indicator	Uses simple mathematical and personal problem solving strategies in highly familiar contexts
Numeracy	ACSF	1	N1P11005	Performance Feature	Adds and subtracts simple whole number amounts (into the 100s) and familiar monetary amounts in personally relevant contexts
Numeracy	ACSF	1	N1P11006	Performance Feature	Recognises and compares familiar shapes and objects in relation to size and shape
Numeracy	ACSF	1	N1P11007	Performance Feature	Recognises and compares familiar basic metric measurements and quantities such as length, mass, capacity/volume, time, temperature, e.g. personal height and weight, a litre of milk or vehicle height clearances
Numeracy	ACSF	1	N1P11008	Performance Feature	Gives and follows simple and familiar oral directions, including using highly familiar maps/diagrams
Numeracy	ACSF	1	N1P11009	Performance Feature	Compares information and data within highly familiar simple texts, lists, charts, diagrams and tables
Numeracy	ACSF	1	N1I11101	Indicator	Uses everyday informal oral language or highly familiar and simple written representation to communicate simple

Domain	Framework	Level	ID	Type	Statement
					mathematical information
Numeracy	ACSF	1	N1P11101	Performance Feature	Writes numbers and monetary amounts into the 100s
Numeracy	ACSF	2	N2I20901	Indicator	Identifies and comprehends relevant mathematical information in familiar activities or texts
Numeracy	ACSF	2	N2P20901	Performance Feature	Identifies and interprets simple mathematical information in familiar and simple oral instructions and written texts where the mathematics is partially embedded
Numeracy	ACSF	2	N2I21001	Indicator	Selects and uses appropriate familiar mathematical problem solving strategies to solve problems in familiar contexts
Numeracy	ACSF	2	N2P21002	Performance Feature	Uses personal and informal 'in-the-head' methods and pen and paper methods to calculate or uses calculator/technological processes and tools to calculate
Numeracy	ACSF	2	N2P21003	Performance Feature	Identifies appropriate tools and uses them in familiar applications, e.g. uses a familiar measuring instrument, such as a tape measure, to measure length in cm or records workplace data on a simple hand-held device
Numeracy	ACSF	2	N2P21004	Performance Feature	Identifies and uses whole numbers, including numbers into the 1000s, money and simple everyday fractions, decimals and percentages, e.g. 1/4, 1/10, 50% or 0.25
Numeracy	ACSF	2	N2P21008	Performance Feature	Identifies, draws and describes common 2D shapes and some common 3D shapes, e.g. sphere, cube or cylinder
Numeracy	ACSF	2	N2P21009	Performance Feature	Measures and estimates length, mass, capacity/volume, time and temperature, using simple instruments graduated in familiar units, e.g. cm, m, ml, °C or hours/min/sec
Numeracy	ACSF	2	N2I21101	Indicator	Uses informal and some formal oral and written mathematical language and representation to communicate mathematically
Numeracy	ACSF	2	N2P21102	Performance Feature	Uses a combination of mainly informal and some formal oral mathematical and general language to report on and discuss the mathematical and problem solving process
Numeracy	ACSF	3	N3I30901	Indicator	Selects and interprets mathematical information that may be partly embedded in a range of familiar, and some less familiar, tasks and texts
Numeracy	ACSF	3	N3P30902	Performance Feature	Interprets and comprehends: o whole numbers and familiar or routine fractions, decimals and percentages o dates and time, including 24 hour times o familiar and routine 2D and 3D shapes, including pyramids and cylinders o familiar and routine length, mass, volume/capacity, temperature and simple area measures o familiar and routine maps and plans o familiar and routine data, tables, graphs and charts, and common chance events

Domain	Framework	Level	ID	Type	Statement
Numeracy	ACSF	3	N3I31001	Indicator	Selects from and uses a variety of developing mathematical and problem solving strategies in a range of familiar and some less familiar contexts
Numeracy	ACSF	3	N3P31002	Performance Feature	Uses a blend of personal 'in-the-head' methods and formal pen and paper methods to calculate and uses calculator/technological processes and tools to undertake the problem solving process
Numeracy	ACSF	3	N3P31003	Performance Feature	Selects and uses appropriate tools, hand-held devices, computers and technological processes, e.g. uses a tape measure to measure the dimensions of a window in mm or creates a personal weekly budget in a spreadsheet
Numeracy	ACSF	3	N3P31010	Performance Feature	Converts between routine metric units by applying understanding of common prefixes, e.g. milli, centi or kilo
Numeracy	ACSF	3	N3P31012	Performance Feature	Collects and organises familiar data and constructs tables, graphs and charts, manually or with spreadsheets, using simple and familiar or routine scales and axes
Numeracy	ACSF	3	N3P31013	Performance Feature	Describes, compares and interprets the likelihood of everyday chance events (e.g. rolling a six on a dice or the chance of rain) using qualitative terms such as certain, likely, impossible and relates these to everyday or routine fractions, decimals or percentages
Numeracy	ACSF	3	N3I31101	Indicator	Uses a combination of both informal and formal oral and written mathematical language and representation to communicate mathematically
Numeracy	ACSF	3	N3P31103	Performance Feature	Uses a combination of both formal and informal symbolism, diagrams, graphs and conventions relevant to the mathematical knowledge of the level, e.g.: o 1/100, 12.5% o km/hr, $/kg o 1.25 m = 1250 mm
Numeracy	ACSF	4	N4I40901	Indicator	Extracts and evaluates the mathematical information embedded in a range of tasks and texts
Numeracy	ACSF	4	N4P40902	Performance Feature	Extracts, interprets and comprehends: o fractions, decimals and percentages, including their equivalent values o ratio, rates and proportions o positive and negative numbers o numbers expressed as powers, e.g. 2³ or 3.6 x 10³ o routine formulae and algebraic representations and conventions o 2D and 3D shapes, including compound shapes o detailed maps and plans

Domain	Framework	Level	ID	Type	Statement
					o statistical data in complex tables and spreadsheets, graphs, measures of central tendency, simple measures of spread and common chance events
Numeracy	ACSF	4	N4I41001	Indicator	Selects from, and applies, an expanding range of mathematical and problem solving strategies in a range of contexts
Numeracy	ACSF	4	N4P41002	Performance Feature	Flexibly uses both 'in-the-head' methods and formal pen and paper methods to calculate and uses technological processes and tools, including a range of calculator or spreadsheet functions, e.g. memory function on a calculator, formulae in a spreadsheet or software to undertake a problem solving process
Numeracy	ACSF	4	N4P41005	Performance Feature	Calculates with fractions, decimals and percentages and flexibly uses equivalent forms; calculates with relevant positive and negative numbers; and uses numbers expressed as roots and powers, e.g. 2³ = 8, √4 = 2 or 3.6 x 10³ = 3,600
Numeracy	ACSF	4	N4P41006	Performance Feature	Develops, interprets and uses routine formulae and algebraic representations and conventions that describe relationships between variables in relevant contexts, e.g. in sport, when considering the cost of repairs, in calculating routine area and volume, using Pythagoras's theorem or in using workplace formulae
Numeracy	ACSF	4	N4P41008	Performance Feature	Estimates, accurately measures and calculates quantities, including areas and volumes, using relevant routine formulae
Numeracy	ACSF	4	N4P41012	Performance Feature	Uses knowledge about chance and probability to estimate and interpret the outcomes of common chance events in both numerical and qualitative terms
Numeracy	ACSF	4	N4I41101	Indicator	Uses a range of informal and formal oral and written mathematical language and symbols to communicate mathematically
Numeracy	ACSF	4	N4P41101	Performance Feature	Uses a combination of informal, but mostly formal, written mathematical and general language, including some specialised mathematical symbolism, abbreviations, terminology and representation to document, interpret and communicate the processes, results and implications of the mathematical activities or tasks
Numeracy	ACSF	5	N5I50901	Indicator	Analyses and synthesises highly embedded mathematical information in a broad range of tasks and texts
Numeracy	ACSF	5	N5P50902	Performance Feature	Extracts, comprehends and analyses a wide range of mathematical information related to number and algebra, measurement and geometry, and statistics and probability, including: o rational and relevant irrational numbers o selected appropriate concepts and information from specialist areas of mathematics relevant to personal, study or workplace needs, e.g. trigonometry, statistics, geometry, linear and non-linear relationships, including parabolas, hyperbolas, circles and exponential functions, introductory calculus, matrices or vectors
Numeracy	ACSF	5	N5I51001	Indicator	Selects from, and flexibly applies, a wide range of highly developed mathematical and problem solving strategies and

Domain	Framework	Level	ID	Type	Statement
					techniques in a broad range of contexts
Numeracy	ACSF	5	N5P51005	Performance Feature	Uses and solves a range of equations using a variety of algebraic techniques
Numeracy	ACSF	5	N5P51006	Performance Feature	Applies graphical techniques to analyse and solve algebraic relationships and equations, including the connections between formulae, their graphical representations and the situations they represent, e.g. linear, quadratic, exponential or inverse relationships
Numeracy	ACSF	5	N5P51007	Performance Feature	Uses and applies knowledge about space and shape, including angle properties, symmetry and similarity to describe, draw or construct accurate 2D and 3D shapes and scale plans and drawings
Numeracy	ACSF	5	N5P51008	Performance Feature	Estimates, accurately measures and calculates quantities, including for complex areas and volumes using measurement formulae
Numeracy	ACSF	5	N5P51011	Performance Feature	Uses and applies knowledge about probability to a range of relevant contexts (e.g. sporting events), calculates theoretical probabilities and uses tree diagrams to investigate the probability of outcomes in simple multiple event trials
Numeracy	ACSF	5	N5I51101	Indicator	Uses a wide range of mainly formal, and some informal, oral and written mathematical language and representation to communicate mathematically
Numeracy	ACSF	5	N5P51102	Performance Feature	Uses a combination of oral specialised mathematical and general language to discuss, explain and interpret the processes, results and implications of the mathematical investigation
Numeracy	ALLS	1	N1LOrg01	Level Description (Original)	Understands basic numerical ideas as demonstrated by completing simple tasks in concrete, familiar contexts where the mathematical content is explicit with little text. Can typically complete tasks that are simple, one-step operations such as counting, sorting dates, performing simple arithmetic operations or understanding common and simple percents such as 50%.
Numeracy	ALLS	1	N1LMod01	Level Description (Modified)	Understands basic and explicit numerical ideas in order to complete a simple task in a concrete, familiar context.
Numeracy	ALLS	1	N1LMod02	Level Description (Modified)	Can complete simple, one-step arithmetical operations (e.g., counting, sorting dates)
Numeracy	ALLS	1	N1LMod03	Level Description (Modified)	Understands common and simple numerical ideas (e.g., 50%) where the mathematical context is explicit with little text.
Numeracy	ALLS	2	N2LOrg01	Level Description (Original)	Identifies and understands basic mathematical concepts that are embedded in a range of familiar contexts, where the mathematical content is quite explicit and visual with few distractors. Can typically complete tasks that are fairly simple, one step or two step processes and estimations involving whole numbers, benchmark percents and fractions, interpreting simple graphical or spatial representation, and performing simple measurements.
Numeracy	ALLS	2	N2LMod01	Level Description (Modified)	Identifies and understands basic mathematical concepts in a range of familiar contexts where the mathematics content

Domain	Framework	Level	ID	Type	Statement
					is quite explicit and visual with little distracting information.
Numeracy	ALLS	2	N2LMod02	Level Description (Modified)	Can complete tasks with one-step or two-step processes involving whole numbers and common percents and fractions.
Numeracy	ALLS	2	N2LMod03	Level Description (Modified)	Can interpret simple graphical or spatial representations where the mathematical context is quite explicit and visual with little distracting information.
Numeracy	ALLS	3	N3LOrg01	Level Description (Original)	Understands mathematical information represented in a range of different forms, such as in numbers, symbols, maps, graphs, texts, and drawings. Skills include number and spatial sense, knowledge of mathematical patterns and relationships and the ability to interpret proportions, data and statistics embedded in relatively simple texts where there may be distractors. Can typically complete tasks that involve undertaking a number of processes to solve problems.
Numeracy	ALLS	3	N3LMod01	Level Description (Modified)	Can locate and use mathematical information in a range of different forms, e.g., numbers, symbols, maps, graphs, texts, drawings.
Numeracy	ALLS	3	N3LMod02	Level Description (Modified)	Can solve problems involving a number of processes with skills related to mathematics patterns and relationships.
Numeracy	ALLS	3	N3LMod03	Level Description (Modified)	Interpret mathematical information (e.g., data and statistics) embedded in relatively simple texts where there may be distracting information.
Numeracy	ALLS	4	N4LOrg01	Level Description (Original)	Understands a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts. Can typically complete tasks that involve undertaking multiple steps to find solutions to problems which require more complex reasoning and interpretation skills, including comprehending and working with proportions and formulas or offering explanations for answers.
Numeracy	ALLS	4	N4LMod01	Level Description (Modified)	Understands a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts.
Numeracy	ALLS	4	N4LMod02	Level Description (Modified)	Can solve problems involving multiple steps requiring complex mathematical reasoning and interpretation skills, e.g., working with proportions and formulas.
Numeracy	ALLS	5	N5LOrg01	Level Description (Original)	Understands complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Can typically integrate multiple types of mathematical information, draw inferences, or generate mathematical justification for answers.
Numeracy	ALLS	5	N5LMod01	Level Description (Modified)	Understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts.
Numeracy	ALLS	5	N5LMod02	Level Description (Modified)	Can interpret multiple types of mathematics information to draw inferences or generate mathematical justification for

Domain	Framework	Level	ID	Type	Statement
					answers.
Numeracy	ALLS	1	N1CCom01	Complexity Satement	Can carry out one-step arithmetical operations (+, -, ×, ÷) with whole numbers to 1,000.
Numeracy	ALLS	1	N1CCom02	Complexity Satement	Can complete simple tasks involving everyday measures in whole units (e.g., kg, m, dates, hours, minutes).
Numeracy	ALLS	2	N2CCom01	Complexity Satement	Can carry out arithmetical operations (+, -, ×, ÷) with large whole numbers including millions.
Numeracy	ALLS	2	N2CCom02	Complexity Satement	Can evaluate a given formula involving common operations (+, -, ×, ÷).
Numeracy	ALLS	3	N3CCom01	Complexity Satement	Is familiar with area and volume formulae.
Numeracy	ALLS	3	N3CCom03	Complexity Satement	Can work with decimals to 3 decimal places.
Numeracy	ALLS	4	N4CCom01	Complexity Satement	Can carry out operations involving negative numbers.
Numeracy	ALLS	4	N4CCom02	Complexity Satement	Can carry out operations involving squares, square roots, etc.
Numeracy	ALLS	5	N5CCom04	Complexity Satement	Can use algebraic conventions and techniques.
Numeracy	ALLS	5	N5CCom06	Complexity Satement	Can work with formal mathematics involving formulae and relationships between variables.
Numeracy	ALLS	1	ALLS_N_Cola	Question (COCA COLA BOTTLES)	Find the total number of bottles in the two full cases shown in the picture.
Numeracy	ALLS	1	ALLS_N_Elec..	Question (Election Results)	Determine the total number of votes cast.
Numeracy	ALLS	2	ALLS_N_Blac..	Question (FEW DUTCH WOMEN)	Calculate the percentage of men in the teaching profession in Italy.
Numeracy	ALLS	2	ALLS_N_Gas	Question (GAS (PETROL) Gauge)	The tank holds 48 gallons. How many gallons remain in the tank?
Numeracy	ALLS	3	ALLS_N_Fire..	Question (FIREWORKS)	Calculate how many more people were injured in 1989 than in 1988
Numeracy	ALLS	4	ALLS_N_Comp..	Question (COMPOUND INTEREST)	Calculate the total amount of money you will have if you invest $100 at a rate of 6% for 10 years.

Appendix 11: Supplementary results – FIT statistics

Part of the results obtained using an IRT analysis was the estimation of 'fit' statistics for each item. Fit statistics provide a measure of 'how well' responses to an item fit the data model relative to other items contained within the same analysis. Items with very large fit values relative to other items in the analysis set may be classified as misfitting the data model used for analysis. In other words, these items do not separate learners into different ability groups in a similar way to other items. For example, this can be due to items assessing against different construct(s) to the other items within the same set or perhaps because the item was ambiguous in some way.

Table 29 and Table 30 contain the 4 items for Reading and 3 items for Numeracy with the largest fit statistics within each domain. The selection of only a few items each is somewhat arbitrary. For this study, the main purpose of analysis was alignment of complexity along the developmental scale along the two frameworks. Hence, it is important to remove complexity estimate outliers from analysis (which has already been done – see Chapter 4). The main purpose was not to identify the inherent unidimensionality underlying the combination of the two frameworks (this has already been validated previously during the stage 1 study which found the frameworks were similar in underlying constructs)[28]. The selected items presented here are more to illustrate the types of items with the largest fit values and what implications this might have for the relevant frameworks. Details relating to the misfitting items for Reading have been presented in Table 35 whilst the misfitting Numeracy items have been displayed in Table 35.

Table 35: Description of Misfitting Reading Items

Item No.	ID Framework and Level	Description	No. of Ratings	Fit Estimate	IRT Complexity Estimate
18	R2P20411 ACSF Reading Level 2 Vocabulary	Uses a dictionary or online resource to check word meanings	255	1.74	-0.503
27	R3P30414 ACSF Reading Level 3 Vocabulary	Routinely uses dictionaries and other references to determine the meaning of unknown words	294	1.64	-0.232
78	ALLS_R_Hiri.. ALLS Level 4 Scaled Test Item	The Hiring Interview	82	1.60	-0.191
79	ALLS_R_CANC..ALLS Level 5 Scaled Test Item	CANCO	21	1.51	0.181

The large fit statistic for these items indicates that raters rated these items in a way that was not consistent with the ratings on other items in the survey. In other words, these items may be associated with constructs that are different to the construct measured by the other survey items. In relation to the two Reading items, one possible explanation might be due to both items containing

[28] See Circelli, M., Curtis, D., & Perkins, K. (2011). Mapping Adult Literacy Performance, NCVER: Adelaide.

the word 'dictionary'. There was only one other statement within the survey containing the term 'dictionary': That is, item number R1P10411 ACSF Level 1 Vocabulary which stated *"uses simple strategies to assist with word identification and extend vocabulary, e.g. a pictorial or bilingual dictionary, or a personal word list*". The fit value for this item was also relatively large (1.32). In relation to the Reading items included in Part B of the Survey (i.e., the ALLS scaled items), two items had large fit values. These items were ALLS_R_Hiring and ALLS_R_Canco. These two items were also identified as outliers in terms of their complexity estimates and hence were excluded from analysis comparing the two frameworks.

Table 36: Description of Misfitting Numeracy Items

Item No.	ID Framework and Level	Description	No. of Ratings	Fit Estimate	IRT Complexity Estimate
42	N5P50902 ACSF Numeracy Level 5	Extracts, comprehends and analyses a wide range of mathematical information related to number and algebra, measurement and geometry, and statistics and probability, including: o rational and relevant irrational numbers o selected appropriate concepts and information from specialist areas of mathematics relevant to personal, study or workplace needs, e.g. trigonometry, statistics, geometry, linear and non-linear relationships, including parabolas, hyperbolas, circles and exponential functions, introductory calculus, matrices or vectors	15	1.97	3.253
67	N5LMod01 ALLS Numeracy Level 5	Understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts.	15	1.50	2.625
80	ALLS_N_Elec.. ALLS Level 1, Scaled Test Item	Election Results Question: Determine the total number of votes cast.	25	1.78	-2.198

It is unclear why the two Numeracy items (item numbers 42 and 67) from Part A of the survey had large misfit values. Item number 80 was an ALLS scale item at Level 1 that was also identified as a complexity estimate outlier.

Items were excluded from additional analysis if they were identified as complexity estimate outliers. Three out of the seven misfitting items discussed here fell into this category and therefore were excluded from further analysis. As the purposes of this study was to compare the complexity estimates of statements drawn directly from both frameworks, items were not excluded solely on the basis of their fit values. However, it is recommended that ACSF statements with large fit values in particular are reviewed in future revisions to the ACSF to ensure they are describing the skills and knowledge as originally intended.