scaling for ATAR calculations – Tasmania: 2005–2013

Reg Allen

November 2014
Tasmanian Qualifications Authority

Contents

List of Figures

List of Tables

Note

All analyses were performed using the R statistical software

- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
- James Honaker, Gary King, Matthew Blackwell (2011). Amelia II: A Program for Missing Data. Journal of Statistical Software, 45(7), 1-47. URL http://www.jstatsoft.org/v45/i07/ http://CRAN.R-project.org/package=softImpute
- Thomas Kiefer, Alexander Robitzsch and Margaret Wu (2014). TAM: Test Analysis Modules. R package version 1.0-3. http://CRAN.R-project.org/package=TAM
- Diedenhofen, B. (2013). cocor: Comparing correlations (Version 1.0-0). Available from http://r.birkdiedenhofen.de/p .

Some of the graphs in this document make meaningful use of colour. This means that important differences may not be visible if the report is printed out in grey scale rather than viewed on screen.

This report was produced in LaTeX, using Sweave to process R code embedded in the source files.

Section 1

Background

At present, the Tasmanian Qualifications Authority produces Australian Tertiary Admission Ranks (ATARs) as part of its productive partnership with the University of Tasmania. This partnership brings the Authority and Tasmanian senior secondary education both benefits and challenges.

ATARs are used by Australian universities

- to make high-stakes selection decisions for admission to highly competitive university courses
- to describe basic entry requirements
- to report on the minimum academic standard of persons admitted to university courses.

An ATAR is an estimate of a student's overall academic achievement in senior secondary studies, regardless of the particular subjects a student has studied.[1]

ATARs are useful in making decisions about admission to tertiary studies because it is generally true that the best single predictor of academic achievement at one stage of education, such as first-year university, is academic achievement the previous year.

The generic nature of ATARs is both a strength and a weakness. It is a significant strength because provides students with an opportunity to select studies without committing to a particular post-secondary school destination – a study pattern based on an initial intention to seek entry to medicine can be used to seek entry to law. Under an ATAR system, students can delay closing off options. This very strength is also in some sense a weakness in that students entering a particular university course may have had very different patterns of previous study. The attached report (see page 30) from a national discussion expands on these issues.

1.1 Purpose of this report

This report identifies fundamental assumptions of a key aspect of the processes used in Tasmania to determine ATARs – scaling – and examines the extent to which the data support these assumptions.

1.2 Introduction – overall achievement in senior secondary studies – ATAR

An Australian Tertiary Admissions Rank (ATAR, for short) provides a comparison of relative overall "academic" achievement, based on a student's performance in a set of courses [2]at "Year 12"[3] level, scaled, added up and turned into a percentile rank within the total age-cohort.[4]

[1] The term subject is used in this document in the generic sense. It is used whenever the interests of clarity and accuracy require it. Tasmanian senior secondary students study TQA accredited courses. The term "course" is used in this document whenever this particular sense of a subject is required. Students in Tasmania can also include in the calculation of their ATARs subjects studied at University as part of their senior secondary studies.

[2] at TQA level 3 or level 4

[3] The term "Year 12" is familiar but has multiple meanings. It can mean the number of years a student has been in schooling. It can imply a level of cognitive demand. It can mean studies typically done by students in their twelfth year of schooling, regardless of the level of demand or difficulty

[4] Note, the total age-cohort, not the group that completed Year 12. This means that a policy goal to have 40 per cent of the age group complete a bachelor degree will decrease the minimum ATAR to 60, assuming that all the top 40 per cent go on to complete degrees. A target like this is incompatible with expressions of concern about falling ATARs in university entry.

So an ATAR of 80 means an estimated overall achievement that is better than 80 percent of the *total* group of persons of the relevant age who could have participated.

Note that ATAR

- measures *overall achievement*, not achievement in a particular subject or group of subjects
- ATAR is a kind of rank (a percentile rank), not a score in the common sense meaning of that word: the difference in achievement represented by an ATAR of 90 and an ATAR of 95 is about three times greater than the difference in achievement represented by an ATAR of 50 and an ATAR of 55.

The calculation of an ATAR includes a step where results in different subjects are put onto a common scale.

While excellent guides have been produced on the purpose, principles and nature of scaling [5], there are many misstatements and misunderstandings in widespread and continuing circulation amongst students, parents, teachers and educators [6].

The "black box" of scaling appears resistant to key mis-beliefs including that

- a scaled score represents my achievement in a subject: so a lower scaled score for my achievement in this subject than you got for your achievement in another subject means that my achievement has been "scaled down" – unfairly and inappropriately
- students who choose some subjects are unfairly advantaged because these subjects are deliberately and intentionally "scaled up".

Attempts to demystify the mathematics of the "black box" seem fated to be shipwrecked on the rocks of these beliefs.

1.3 limitations of this study

This study focuses only on the key elements of scaling as used for ATAR purposes. This means it leaves out a lot of important issues, including

- special procedures to deal with anomalous situations (odd cases and cases with very small numbers)
- the inclusion of results from University courses
- other aspects of ATAR calculations, including the conversion to percentile ranks intended to be comparable across Australia.

1.4 how to read this report

The reader who wants to know about and understand the details of scaling and the extent to which it works as it should will read the entire report. This has been written to provide a largely non-technical explanation of the technical issues, analyses and data supporting the report's conclusions.

Unavoidably, an account of technical issues must go into matters that may not be familiar to some readers.

The reader who wants only an overview and the conclusions demonstrated by the analyses in this report should read Section 2 on scaling, the introductions to Sections 3, 4, 5 and the conclusions in section 6.

[5] See, for example, https://satac.edu.au/SACE_NTCET/Scaling.htm

[6] See, for example, the discussion at http://parentscouncil.nsw.edu.au/announcements/scary-scaling-choosing-your-subjects-for-the-hsc

Section 2

Scaling

The relativities of subject scores[1] when scaled are relativities in terms of overall achievement, not subject- or group-specific achievement, such as literacy or numeracy.

Put simply, calculating an ATAR is like adding apples, oranges and bananas – the result is fruit, not information about specific varieties of apple.

Put more technically, the scaling processes used in the determination of ATARs are one-factor models of the set of achievements of students across a wide variety of subjects.

This single factor id not and cannot be a subject-specific factor or a factor common to a subset of subjects. Scaled scores in Physics and French are on a common scale of overall achievement, a scale that measures neither Physics nor French.

D.J.Daley (1989) in "Determining Relative Academic Achievement for Fair Admission to Higher Education" provided a thorough exploration of the underlying mathematical issues and the requirements for a "fair" system. Dr Daley demonstrates that any moderately positively correlated set of scores in different subjects will be dominated by a factor that, in the simple case of a common curriculum, is aligned with the first principal component of these sets of scores. The situation becomes more complicated when a student's overall score is based on a *best-n* subset of scores and when students choose different sets of subjects. These complications, however, while they make the technical aspects of scaling more complex, do not take away from the fundamentals of the idea of overall academic achievement.

2.1 why do we have scaling

Scaling puts results from different subjects/studies/courses onto a common scale.

We do this when we calculate a grade point average (GPA), but GPA is a bad measure of overall achievement, since it introduces systematic bias in favour of some students and against others, biases that reflect systematic differences in the groups of students taking different courses, in the grading practices in those courses and in the courses themselves [2]

The purpose of scaling is to create a "no systematic advantage/disadvantage situation", to avoid the unfairness you get if you just add up scores, grades or results from different courses/subjects. Scaling is designed to *create* fairness, to create a "level playing field" before scores in different subjects are added up.[3]

[1] results in courses

[2] See, for example: Caulkins, J. P., Larkey, P. D., Wei, J. (1996). Adjusting GPA to reflect course difficulty. Working paper, Heinz School of Public Policy and Management, Carnegie Mellon University. Lei P., Bassiri, D., Schultz, E.M. (2001), Alternatives to the Grade Point Average as a Measure of Academic Achievement in College. ACT Research Report Series

[3] There are some analogous aspects in, for example, rankings in sports, especially where players or teams do not or can not take part in every possible event. For example, the tables of points for professionals playing in contests of different difficulties and for reaching the later rounds – see http://www.atpworldtour.com/Rankings/Rankings-FAQ.aspxpoints – clearly reflect the idea that some tournaments are harder than others (attract more of the better players) and that winning a hard tournament is a greater achievement than winning an easier tournament. Association Football and chess provide examples where the Elo method of calculating rankings is used – see http://en.wikipedia.org/wiki/World$_{Football_{E}lo_{R}atingsB}asic_{c}alculation_{p}rinciples$.

2.2 how scaling is done in Tasmania

The scaling process used in Tasmania applies Item Response Theory (IRT) to each year's data set of student results in TQA level 3 courses onto a single common scale of overall academic achievement.

The idea of using IRT to put results from different courses onto a common scale seems to have first appeared in a study described by J. W. Young in 1990.[4]

The process we use treats the results in each TQA level 3 course (a result is one of five "awards" – known as PA, SA, CA, HA, EA[5]) as if each were an item in a test reported on a five step scale (from 0 to 4). Most students, of course, do not do more than five or so of these items – that is, most of the 320000 plus possible data points (the ones we would have if every student did every subject) are missing.

Data sets (deidentified to protect student privacy and omitting courses with fewer than 30 results) from 2010 through to 2013 can be found at http://www.tqa.tas.gov.au/32416 . Readers are encouraged to analyse these data sets.

IRT modelling (using a variant called "partial credit") is used to establish values for each award in each course. This involves calculating an estimate for each course and each award in each course so-called "item and item-step difficulties" and, as an estimate for each student, so-called "person abilities". A basic IRT model, such as we use, assumes that a data set of results in a test is best explained by the idea that the chance of a person getting a particular result on a particular item is a function of some characteristic of the result (its difficulty) and some characteristic of the person (the person's ability). In our case, the test is the set of TQA level 3 courses and each item in the test is a TQA accredited course.

The values in this model are then used to calculate "thresholds" for each award in each course – these are the person abilities that, within the model, indicate that a person with this ability is equally likely to have this award or one of the awards below it.

These thresholds are then adjusted so that the average threshold for a CA is 7 and for an EA is 20. [6] These values were chosen for historical reasons. Values for SA that are less than 1.0 are, arbitrarily, raised to 1.0. The value for a result of PA or lower is set as zero. There are conceptual weaknesses in this set of decisions, but it appears difficult to make changes in these sorts of details without the impetus of a major review of the system.

These limits are then used to assign a range of scaled scores for each award in each course (note that there is some element of extrapolation in estimating the range for EAs). Each student with an award in a course is then, using the pattern of ratings within each award, assigned a scaled value for this course within these ranges.

2.2.1 Using scaled scores in ATAR calculations

Oversimplified, the next stages of the ATAR calculations are:

- assign eligible students an aggregate of their scaled scores, including those from previous years
- rank these students in terms of these aggregate scores
- convert these ranks into estimated ranks in the total age-group.

The details of these further steps in the ATAR calculations are not relevant to this discussion of scaling.

It may be worth noting, however, that the uncertainties in ATARs reflect

- uncertainties in the assessment and standards-setting processes for results in subjects – each student could possibly have had a different result
- uncertainties in the scaled values attributable to applying the scaling process to the Tasmanian data set – students might have taken different subjects or performed differently in the subjects they did take.

[4] Young, J.W., (1990) Are validity coefficients understated due to correctable defects in the GPA? *Research in Higher Education*,vol. 31, No.4, pp. 319-325

[5] NN is also included and considered, for these purposes, as the same as PA

[6] In recent years, these averages are weighted averages - weighted by the number of persons with each award.

We have estimated the second set of uncertainties[7] and concluded that, on the basis of using the scaling process we currently use on the sorts of data sets we have in Tasmania:

- a student with five EAs (at the bottom) can, on average, expect an ATAR of 96.0 plus or minus 1.8 [8]
- a student with five CAs (at the bottom) can, on average, expect an ATAR of 57.5 plus or minus 2.25.

There do not appear to be any published studies of uncertainties in ATARs in other jurisdictions. The precision with which ATARs are used (four significant figures) is probably more than is justified by the properties of the input data and the scaling processes.

2.3 key assumptions of the scaling process

There are some critical assumptions underpinning the scaling process:

1. the set of awards in TQA level 3 courses defines a single common scale of overall achievement sufficiently well to make the calculation meaningful
2. the groups of students completing two or more TQA level 3 courses each year are sufficiently similar "academically" that the tables of scaled values are sufficiently comparable from year to year
3. the IRT process produces values that do not show a systematic advantage/disadvantage for groups of students who choose different combinations of subjects

Each of the following sections explores one of these assumptions.

[7]We used a bootstrap (resampling) process, running the scaling process on 700 random samples (with replacement) of the 2014 data set. We then repeatedly (300 times) drew a random sample of five subjects and each time calculated the sum of the EA and CA scaled values for these subjects in each of these 700 sets of scaled values. This gave us a mean and a spread for these totals. These were then converted to ATARs. The distribution of these values tells us the likely ATAR and likely range of ATARs.

[8]This is a 75 per cent confidence interval.

Section 3

Assumption one: a meaningful common scale of overall achievement can be derived from combining results in TQA level 3 courses

3.1 Introduction

The construction of an ATAR rests on the assumption that it makes sense to speak of "doing well at school". So does calculating a grade point average. So does awarding an academic prize or a scholarship medal. Sometimes it makes sense to add up very different things. Sometimes it does not. For example, if you are selecting a person to accompany you on a mountaineering expedition as your emergency medical support you want a person who can do well two very different things: medical help and climbing. You wouldn't want someone who was wonderful at one and no good at the other.

When we construct an ATAR we are adding up achievement in different subjects. Whether that makes any sense depends on whether the subjects involved have enough in common.[1]

Whether the set of TQA level 3 courses defines sufficiently well a single common construct that we can identify as "overall academic achievement" presents some technical challenges.

The following sections describe how we have approached these challenges to show that we can construct from results in TQA level 3 and 4 courses a meaningful single measure of overall academic achievement.

3.2 the data we have

In 2013, the data set included 19386 results across a total of 60 courses. If we omit the 11 courses with fewer than 21 results, there are 49 courses with an average count of 394.3 results per course.

There are 6937 students with one or more results in these courses in 2013. Table 3.1 shows the number of students with one, two, three or more awards.

1	2	3	4	5	6	7
1102	1588	2085	2033	120	8	1

Table 3.1: 2013 data – numbers of students with 1 or more awards in with at least 20 awards at NN/PA or above

In analysing the relationships between results in these courses there is no value in including those students with a result in only one course[2]. These single cases are, accordingly, removed and the following analyses conducted on a data set with 5835 students with two or more results in subjects with, originally, 20 results.

In 2013 there were 1176 pairs of TQA level 3 courses. In 33 per cent of these cases there was no observed relationship between results.

[1]And none of them have so much in common that they are really the same subject

[2]Any scaling process uses information about relationships between results. A student with only one result provides no information about these relationships. For example, in analysing multiple choice tests, a question is said to be hard if it usually got right mostly by people who get lots of other questions right and got wrong by people who get wrong lots of other questions.Somebody who answers only one question doesn't add any information about this question.

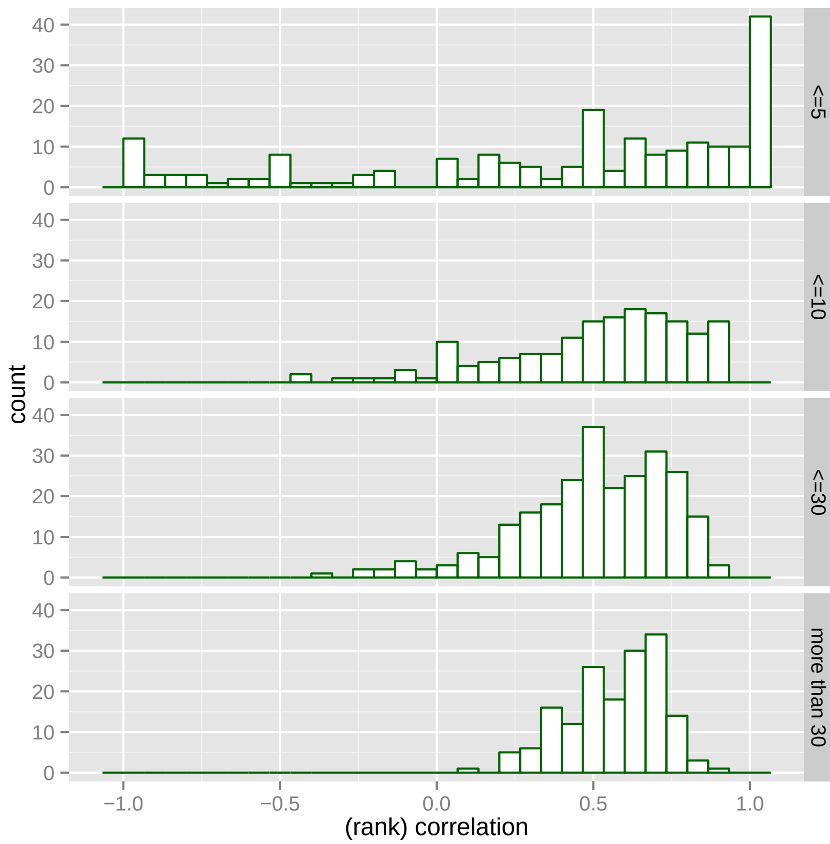

Figure 3.1: Observed correlations of results in pairs of TQA level 3 courses in 2013 by count category

Figure 3.1 shows, for the 2013 data, the observed correlations [3]of results in pairs of subjects. As we would expect, the correlations for cases where there are few data points in common can vary wildly between -1.0 and 1.0. [4] With more than 30 data points, the correlations are mostly moderately positive – an essential feature if we are to make any sense of the idea of scaling and adding up results to get an indicator of overall achievement. For example, if we included five categories of students' height we would expect to see correlations close to zero for the relationship between these and results in most if not all courses.

For about 76 per cent of the possible pairs the actual data gives us a missing or not very reliable estimate of the correlation.

3.3 the data we need

If we had complete data (good estimates of the correlations between every possible pair of subject results), it would be easy to work out whether the set of TQA level 3 subject results provides a good estimate of a single common factor, one we can identify as overall academic achievement.

Using some modern statistical methods (in this case those provided by the Amelia package in R) we can derive a series of estimates for these correlations. The basic idea is to fill in the missing values using the observed relationships as a guide to the likely values. This process ("imputation") is done

[3]A correlation is a number intended to capture the (linear) relationship between two sets of numbers. A correlation varies between -1.0 and 1.0. A correlation around zero says that there is no relationship (or it is not a simple relationship). A correlation of nearly 1.0 says that the two sets of numbers line up well together (or there are a few odd cases giving a misleading result)

[4]It's helpful to remember that awards (PA, SA, CA, HA, EA) have **ordinal** properties, have a coarse grain (vary by one whole result at a time) and that there are serious ceiling and floor effects (you can't get higher than EA and you can't get a result below the lowest).

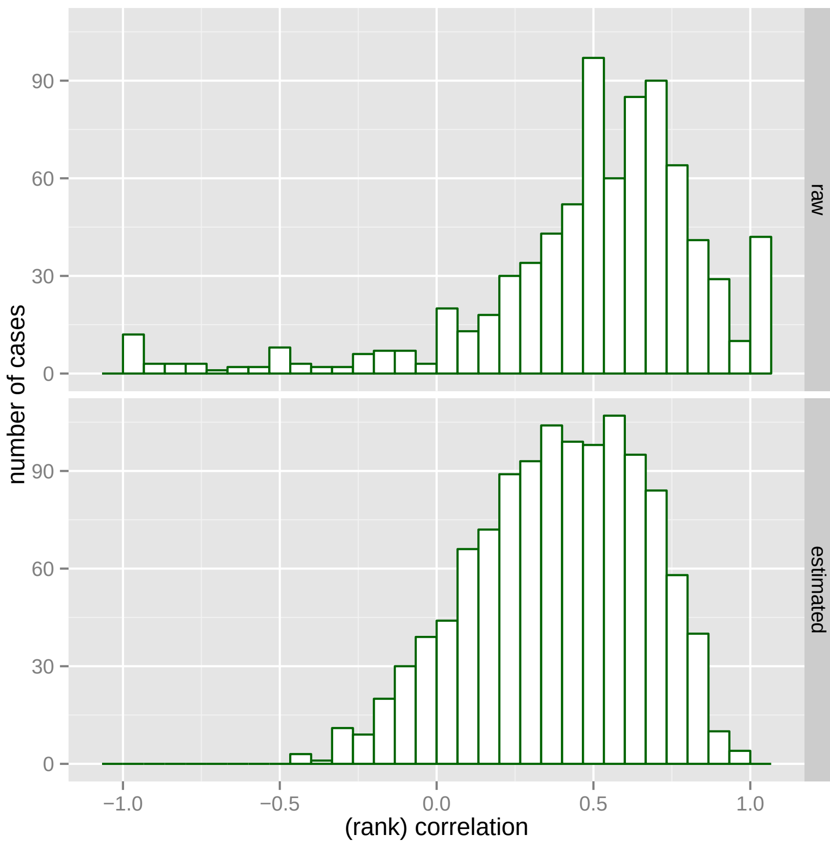

Figure 3.2: Distributions of observed and estimated correlations of results in pairs of TQA level 3 courses in 2013

many times to get a sense of the likely variability in the missing values. Having done this, we can then get a sense of what the correlations are like (and how much they are likely to vary). From these sets of correlations [5] we can get an idea of how well the data set is dominated by a single common factor (the first principal component).

We have done this.

We derived fifty sets of complete data by treating the raw results as being one of five ordinal [6]values between 0 and 4 and using the relationships between these results to estimate plausible values for the missing data points[7].

For each of these fifty data sets we calculated the complete set of correlations. Are these values indeed plausible (along the lines we should expect)?.

Figure 3.2 compares the distributions of these correlations for one of these fifty data sets. Figure 3.3 compares the distributions of these two sets of correlations for the five categories of the numbers of pairs of results (including the cases where there was no original correlation).

We can see that, as we should expect, the estimated correlations are less extreme than the raw values.

Putting together the fifty sets of estimated data, we get for each pair of TQA level 3 course an average estimated correlation. Figure 3.4 plots these average estimated correlations against the observed values. The straight line on each plot is a line of equal value to make it easy to see the types

[5]There is an argument that we should use the covariances. In practice, this does not significantly alter the results. The term correlation is used here as being more familiar to more readers.

[6]five discrete values where only the order – this value is bigger than that value – has any meaning

[7]We are not trying to work out what an individual student "would have" got. We are exploring the relationships in the original data. Each set of estimated data is based on the real relationships – and so the variability between each set of estimated data reflects the variability in the relationships in the original data.

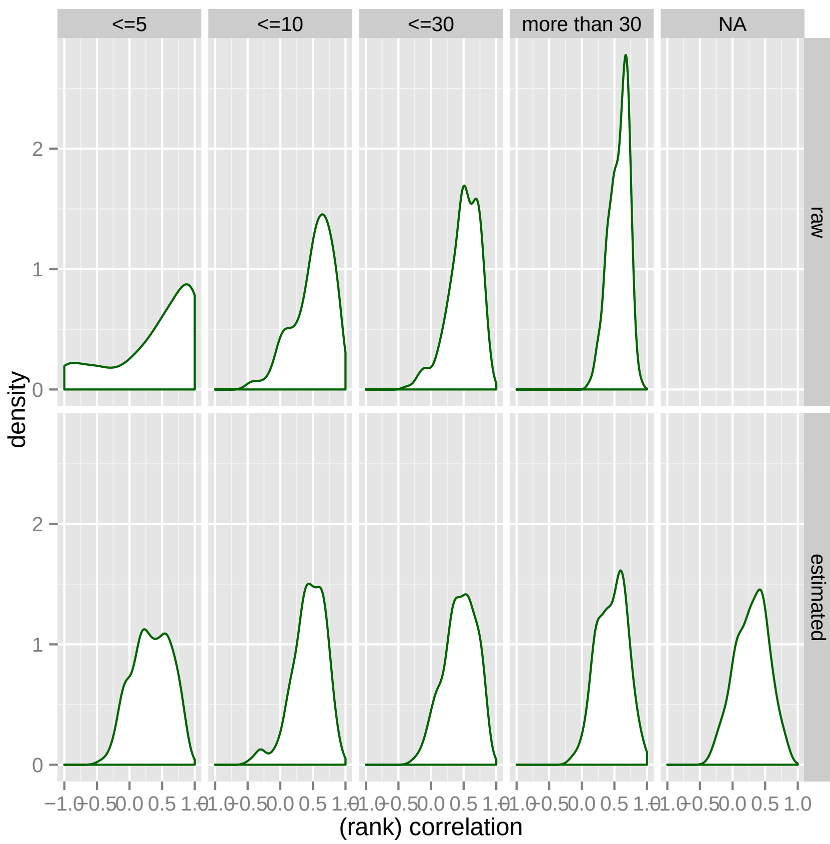

Figure 3.3: Distributions of values of observed and estimated correlations of results in pairs of TQA level 3 courses in 2013 by count category

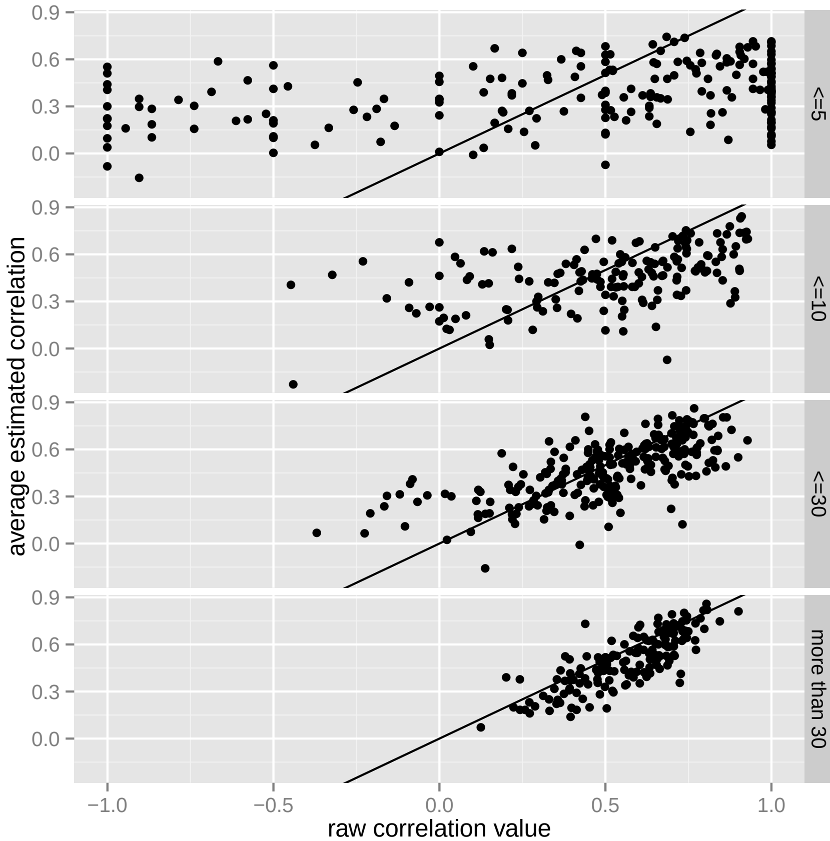

Figure 3.4: Values of average estimated correlations against observed values for TQA level 3 courses in 2013 by count category

of differences.

Figure 3.5 shows the sorts of variations in the estimated correlations. As we should expect, the estimated correlations for larger groups are generally more stable and the higher the original value the more stable the estimates.

We can see from all this that the estimated values are

- plausible – they make sense
- certainly not an overestimate of the relationships in the data set (the correlations tend to be lower and less spread than the observed values).

This means that we can apply standard techniques to each of the estimated data sets to get a sense of how well each defines a single common factor and whether these common factors are similar. If they are, we can infer that the original data set, if it were complete, would have similar properties.

3.4 a single factor common across achievement in a wide range of subjects?

A set of results in over forty subjects (just like a test with over forty items) has a lot of different things going on in it – it is multi-dimensional, with over forty dimensions (there are differences between subjects just as there are differences between different questions on a test). Principal Component Analysis (PCA) is a well-known technique for exploring multi-dimensional data sets. In PCA we find a sets of (linear) combinations of all the variables. Each set is independent of all the others. The sets are put in order from the one that captures the largest part of the variation in the original data set (the first principal component) to the one that captures the smallest part of the variation. The second

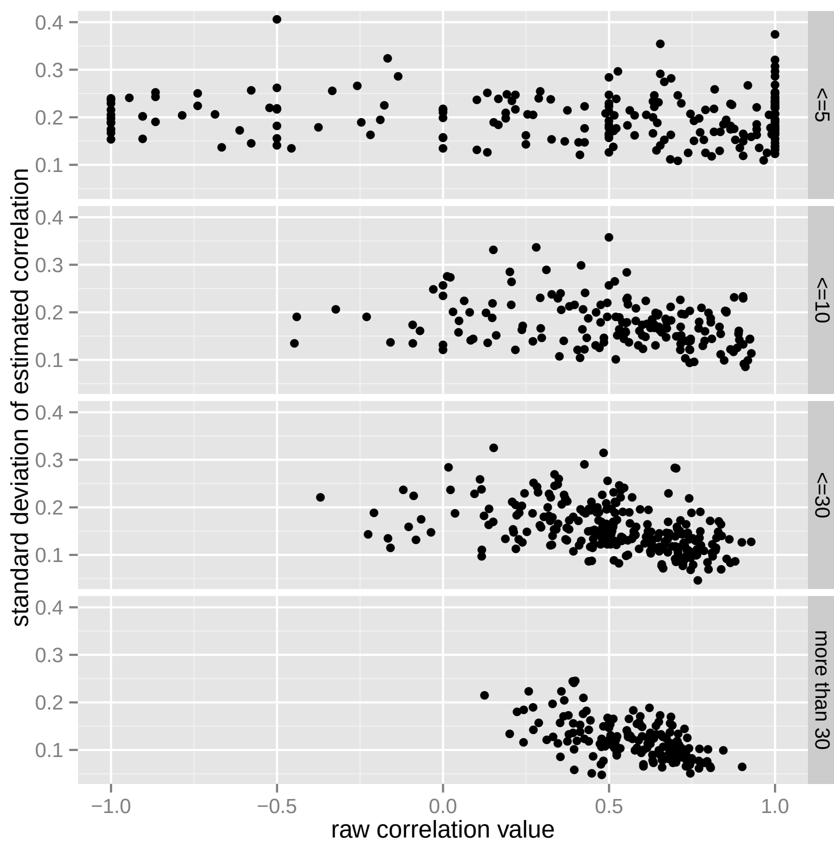

Figure 3.5: Standard deviations of the estimated correlations against observed values TQA level 3 courses in 2013 by count category

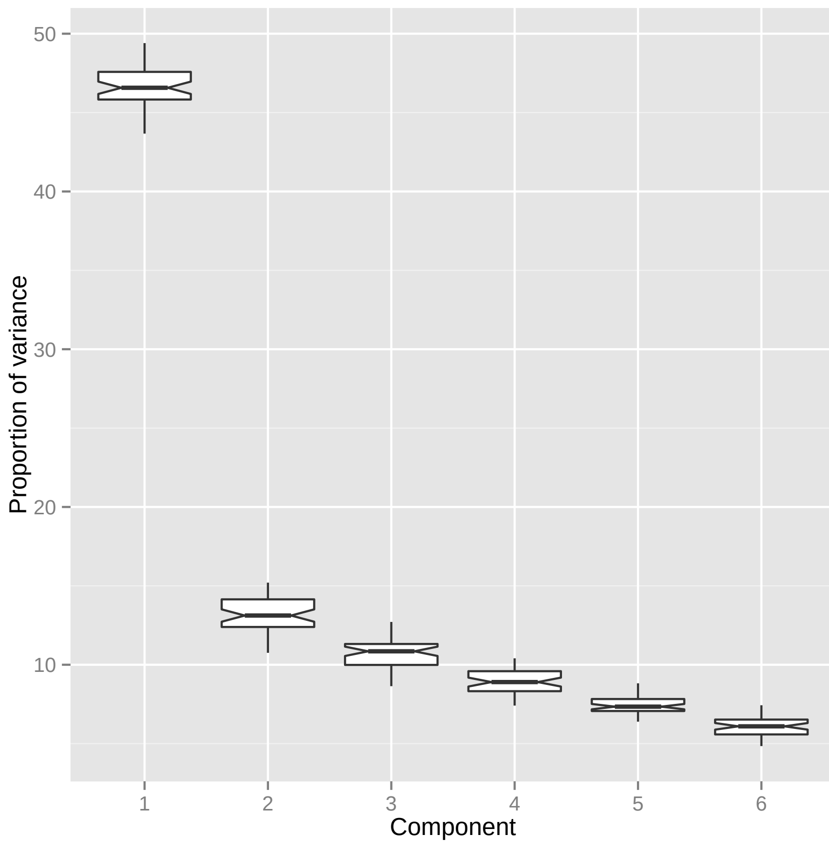

Figure 3.6: Boxplots of the proportion of variance captured by the first six principal components in each of the estimated complete data sets

principal component is the combination that best captures the variation that is left after we have taken out the first principal component.

Figure 3.6 shows how well the first six principal components capture the variation. The first principal component typically captures nearly half the total variation in the data set, a rate that is much greater than that captured by the second and third.

As we can see from the small variations shown in these first principal components are similar across all the estimated data sets (even though the estimated individual values vary widely, the relationships are similar). Table 3.2 lists all subjects in this study and the average loading of each on the first principal component together with the variation across all the estimated data sets. This table also includes the average observed correlation of results in each subject with other subjects.

Taking all these aspects together, our data set has the required properties for calculating an overall index and, given that the subjects are mostly academic subjects, we can interpret this as an index of overall academic achievement in year 12 academic studies.

However, as is clear from figure 3.6 where the next three components capture about one-third of the variation, there are other specific aspects to achievement that are not captured by the first principal component (and hence by ATAR). If you want to know whether someone is better at mathematics or science than they are at writing humanities essays then there is no point in looking at overall academic achievement: you have to look at results in particular subjects or groups of subjects. Of course.

The data in table 3.2 has been ordered from those subjects those that line up better with most other subjects (greater loadings on the first principal component) down to those that align less well overall.

Figure 3.7 shows the relationship between these loadings on the first principal component and the observed correlations between results in a subject and results in other subjects. As should be expected,

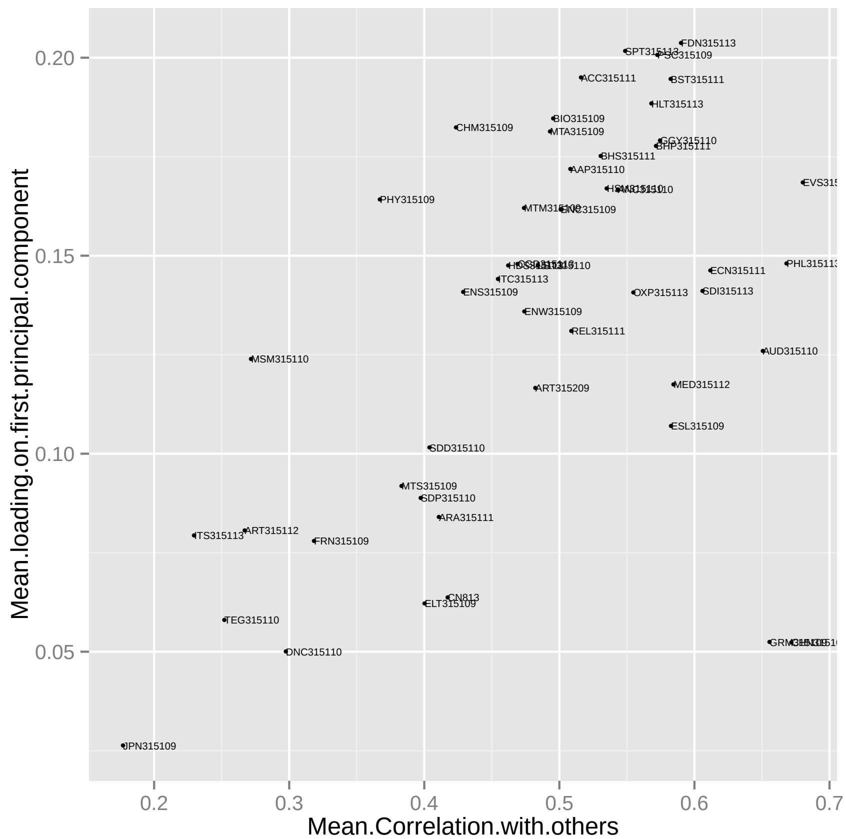

Figure 3.7: Relationship between loadings on first principal component and average observed relationship with results in other subjects

the subjects that tend to correlate better with most of the others are those that relate more strongly with overall achievement.

That is, mostly[8], the subjects with smaller loadings on the first principal component tend to be those where the results are less well related with results in other subjects.

As discussed earlier, we should not expect subjects where there is a less clear relationship with results in other subjects to give us as much information about students' overall academic achievement. A technically sound argument can be made that subjects that don't relate to overall academic achievement (as defined by the majority) should not be included. However, this technical issue should be balanced against the educational value of having such subjects "count" for an ATAR.

We know from the results of these analyses that we can make sense of calculating an index of overall achievement from results in TQA level 3 courses. The next step is to ask how well our current processes work in practice.

[8] the exceptions are two languages with very small numbers in common with other subjects – German and Chinese

Subject	Mean loading on first principal component	SD loading on first principal component	Mean Correlation with others
FDN315113	0.204	0.012	0.590
SPT315113	0.202	0.012	0.549
PSC315109	0.201	0.013	0.573
ACC315111	0.195	0.014	0.516
BST315111	0.195	0.013	0.583
HLT315113	0.188	0.010	0.568
BIO315109	0.185	0.012	0.496
CHM315109	0.182	0.014	0.424
MTA315109	0.181	0.012	0.493
GGY315110	0.179	0.009	0.575
BHP315111	0.178	0.011	0.572
BHS315111	0.175	0.010	0.531
AAP315110	0.172	0.013	0.508
EVS315109	0.168	0.010	0.680
HSM315110	0.167	0.010	0.535
ANC315110	0.167	0.010	0.543
PHY315109	0.164	0.019	0.367
MTM315109	0.162	0.018	0.474
ENC315109	0.162	0.006	0.501
PHL315113	0.148	0.012	0.668
CGD315113	0.148	0.014	0.469
LST315110	0.148	0.010	0.484
HDS315113	0.148	0.011	0.462
ECN315111	0.146	0.011	0.612
ITC315113	0.144	0.012	0.455
SDI315113	0.141	0.015	0.606
ENS315109	0.141	0.014	0.429
OXP315113	0.141	0.009	0.555
ENW315109	0.136	0.014	0.474
REL315111	0.131	0.011	0.509
AUD315110	0.126	0.018	0.651
MSM315110	0.124	0.026	0.272
MED315112	0.117	0.019	0.585
ART315209	0.117	0.013	0.482
ESL315109	0.107	0.018	0.583
SDD315110	0.102	0.014	0.404
MTS315109	0.092	0.021	0.383
SDP315110	0.089	0.023	0.397
ARA315111	0.084	0.020	0.411
ART315112	0.081	0.016	0.267
ITS315113	0.079	0.019	0.230
FRN315109	0.078	0.023	0.318
CN813	0.064	0.021	0.417
ELT315109	0.062	0.014	0.400
TEG315110	0.058	0.015	0.252
GRM315109	0.052	0.020	0.656
CHN315109	0.052	0.013	0.672
DNC315110	0.050	0.019	0.298
JPN315109	0.026	0.027	0.177

Table 3.2: Loadings on first principal component by subject: average, variation together with average observed pairwise correlation

Section 4

Assumption two: sufficiently comparable groups of students – year to year

4.1 Introduction - causes of changes in scaled values

The results of IRT analyses of the awards in TQA level 3 courses are adjusted each year so that the average value for an EA (the highest award) is 20 and the average value for a CA is 7.0.

This means that the scaled values for EAs in some subjects are above 20 and for EAs in other subjects must be below 20. This must be so – an increase in one is a decrease in another.

It also means that changes from year to year in the scaled values for EAs (and other awards) are interpreted as if they were deliberate (rather than the outcome of a data-driven process) and, more significantly, interpreted as signposts to students choosing subjects. Some will choose a subject that was, they think, "scaled up this year" and then feel let down by the system when their hoped for bonus doesn't arrive. Others will argue that students avoid subjects that are "scaled down".

What happens? There are changes from year to year, sometimes small (about half are plus or minus 1.0) and sometimes large.

Why are there such changes? Aren't these the same subjects with the same standards studied by similar groups of students? Isn't it reasonable to expect the scaled values not to change from year to year?

Changes in the scaled values for EA [1] largely depend on

- changes in the proportion of students getting an EA (more EAs lowers the scaled value)
- changes in the average academic standing of the group of students doing a subject (a less academic group means a lower scaled EA)
- changes in the range of students doing a subject (a group with a wider range of academic achievement means a higher scaled value for an EA)
- changes in the relationship between results in this subject and overall academic achievement (an increase in this relationship means a higher scaled EA).

4.2 Year to year changes in scaled values for awards

To see why this is so we start with the data for each year from 2005 to 2013. We use an IRT program [2] to derive scaled values for each award in each subject.[3] We can then look at the year-on-year changes (a simple matter of subtraction) and explore the relationships between these changes and other information we have.

[1] We are looking at EA since we get a similar picture if we study the scaled values for HA and since it is the EA scaled value that attracts most attention.

[2] R package TAM

[3] The values got this way aren't exactly the same as the published values on TQA website since we are not making any of the adjustments for small groups and for anomalous cases and since we are using the final data set rather than the interim one used at the time of the year when we do the calculations. However, the differences are small and don't matter in this context.

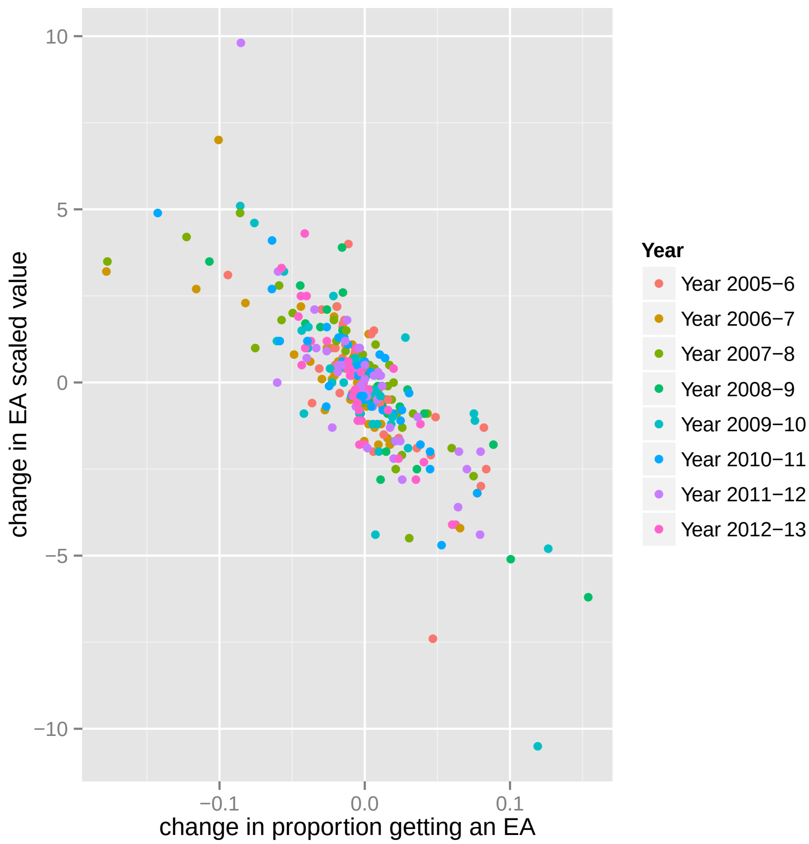

Figure 4.1: Relationship between changes in minimum EA scaled value and changes in the proportion of students with an EA

The first and most important factor is clearly apparent in figure 4.1: a change in the proportion of students getting an EA [4]. A change in proportion of 0.1 means that the percentage of students getting an EA has gone up from, say, 14 per cent of the students doing this subject to 24 per cent.

The proportion of students getting an EA each year in a subject depends on

- how well the students do on the examination in that subject that year
- judgments by assessment panels about how to apply the standards defined in the course document to these students results in the examination
- the overall academic capacity and subject specific capacities (skills, knowledge and willingness to work) of the students doing the subject that year

From the evidence of figure 4.1 you would not advise assessment panels to be too generous in assigning EAs. However, you would do well to remember that this is a zero-sum process – an increase in one area is a decrease in the others. If all courses change in the same way at the same time there will be no changes.

It is also very clear from figure 4.1 that changes in the proportion of students getting an EA are not the only changes that matter. Figure 4.1 includes cases of obvious changes in the scaled values for an EA without any corresponding change in the proportion of students getting an EA.

What other factors are there? We explored a range of possibilities. We found a significant relationship for the ones shown in table 4.1. This summarises an analysis of the relationship – linear regression [5] – between a change in EA scaled value and :

[4] Multiply a proportion by 100 to get a percentage.

[5] Because there are some extreme values, we also did robust regression. This showed that the relationships are there even without the extreme cases

- EA.prop.change: the change in proportion of students getting an EA
- two cohort measures reflecting changes in the groups of students doing this subject
  - Poly.change: a change in the average overall academic achievement of the group of students doing this subject
  - Poly.SD.change: a change in the spread of overall academic achievements of the group of students doing this subject

The measures of overall academic achievement used here are described in more detail in the next section. At this stage, it is enough to note that on average the scaled value for an EA goes up

- with a decrease in the proportion of EAs: a drop from 15 per cent to 10 per cent on average raises the EA scaled value by 0.05*44 – about 2 (if nothing else changes)
- with an increase in the academic standing of the subject group: the kind of difference seen in the Mathematics Applied cohort between 2009 and 2013 on average raises the EA scaled value by about half a point – but only if this is the only change: if the proportion of EAs goes up at the same time ...
- with an increase in the spread of academic standing of the subject group: the sort of change seen in the Ancient Civilisations group between 2008 and 2009, if positive, raises the EA scaled value by about half a point (if nothing else changes)

These factors, taken together, account for about 73 per cent of the variance [6] shown in figure ??

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.1004	0.0643	-1.56	0.1191
EA.prop.change	-43.8048	1.6020	-27.34	0.0000
Poly.change	4.2638	0.5397	7.90	0.0000
Poly.SD.change	4.6017	0.7627	6.03	0.0000

Table 4.1: modelling change in the EA scaling value against change in proportion getting EA, cohort measures

There are real changes in the scaled values from year to year. Much of this can be understood in terms of changes in standards and changes in the participating groups.

A fundamental assumption behind using the scaled scores from several years – an essential feature of the Tasmanian system – is that the participating cohorts are sufficiently comparable from year to year.

To know if the cohorts are comparable we need some measures that we can get from the data we hold.

4.3 Measures for comparing cohorts

In this study, the total cohort is the group of students with one or more results in a TQA level 3 course in a given year. A subject cohort is the group of students with a result in a given year in one of these courses.

Cohorts can usefully be compared in terms of:

- mean age
- gender balance (proportion of females and comparison with the total student cohort balance[7])
- relative overall achievement
  - mean ATAR of those students in a cohort who have an ATAR in the current year

[6] Leaving out odd values – outliers – does not materially alter this summary, as does including a range of other variables.

[7] Note that including gender (or age) is *not* a claim of a causal link. Changes in the gender balance are, however, a useful proxy indicator of changes in composition of a cohort. So are changes in average age.

- mean transformed ATAR – "ATAR score" (see below)
- mean overall achievement – "polyscore" (see below)
- the number of students with only one result.

Note that we could include measures reflecting the balance across sectors or measures reflecting estimated Socio-Economic Advantage/Disadvantage. Such measures would give us additional perspectives on the composition of a cohort and of changes over time.This is a matter for future studies.

4.3.1 Relative overall achievement

As discussed earlier, an ATAR provides a comparison of relative overall achievement, based on a student's performance in a set of subjects, scaled, and turned into a percentile rank within the total age-cohort.

ATAR provides a measure that is comparable across jurisdictions and across years, on the assumptions that:

- the underlying populations are comparable in terms of potential academic achievement
- the overall achievement measures used in each year in each jurisdiction reflect potential academic achievement in comparable ways.

Since ATAR is a percentile rank it can be used to derive an "ATAR score" – a score on a normal distribution[8]. A percentile rank treats the difference between 50 and 51 as equivalent to the difference between 90 and 91, where the corresponding difference in a normal score is about 2.3 times bigger.

Cohort comparisons should seek to take into account all the students with a result, not just those who also have an ATAR. This requires an overall achievement score for each student.

4.3.2 Deriving an overall achievement score for each student

The method used is a modified version of the algorithm described by Sympson and Haladyna [9] that uses all subject results to derive an "achievement score" for students (weighted average of subject scores) and a set of subject result scores (average of achievement scores of students with this result).

This algorithm assumes that students' achievement scores are the weighted average of their subject scores and that subject scores are the average of the achievement scores of students with this result.

Estimates are updated until convergence.

The student scores from this process are a percentile rank. This was turned into a normally distributed score (a "polyscore"). These scores were then scaled so that the scores of those with an ATAR had the same mean and standard deviation as their normed ATAR scores.

This process produces normal scores that correlate reasonably well [10] on an individual basis with the normed ATAR scores for these students. Correlations of this order on an individual basis means that it is appropriate to use these measures for comparing groups. The advantage of the polyscore measure is that it uses information about all the subjects included in a cohort data set – a difference that matters when a subtantial proportion of the cohort is not eligible for an ATAR that year.

It is important to note that different subjects are **not** directly comparable and that the mean ATAR (and its related ATAR score) in for a subject group reflects participation – a group with a lower mean ATAR is one with a higher participation rate (for ATAR eligible students).

4.3.3 similarity of groups of students from year to year – overall

Table ?? shows these measures for the total cohort (all those with one or more TQA level results in a given year).

NULL It is obvious that there are continuing statistically significant changes in the age-structure

[8]

$$ATARscore = \Phi^{-1}(ATAR \div 100) \mathit{Thisis, indeed, ascore.AnATARitselfisarank, notascoreintheusualsenseofthatterm.}$$

[9] Sympson, J. B., and Haladyna, T. M. (1988, April). *An evaluation of "polyweighting" in domain-referenced testing*. Paper presented at the annual meeting of the American Educational Research Association. New Orleans, LA

[10] 0.91 to 0.98 in studies using data from interstate – 0.91 is a typical figure for Tasmanian data, where the data set includes a lot of students with at ATAR in a given year with only three results

Year	Total.cohort	Total.age	Total.age.sd	Total.Prop.ATAR	Total.Prop.Female	Total.meanPoly	Total.meanATAR.score	Total.Prop.one.only
2005	6270	18.50	4.37	0.33	0.55	0.44	0.88	0.16
2006	6554	17.69	2.69	0.34	0.56	0.38	0.86	0.16
2007	6598	17.76	3.30	0.33	0.54	0.39	0.87	0.16
2008	6589	17.80	3.44	0.34	0.55	0.40	0.86	0.15
2009	6677	17.70	2.87	0.32	0.55	0.40	0.87	0.15
2010	6690	17.68	2.72	0.35	0.54	0.48	0.89	0.15
2011	6775	17.71	3.00	0.34	0.54	0.47	0.91	0.15
2012	6833	17.56	1.62	0.34	0.54	0.49	0.91	0.15
2013	6945	17.50	1.11	0.33	0.53	0.45	0.90	0.16

Table 4.2: Total cohort characteristics - counts, mean age, proportion with an ATAR, proportion female

of the total cohort – lower average age and smaller spread – indicating a reduction in participation of older learners. At the same time, there is over time a significant shift in gender balance (less female).

The proportion completing only one subject at this level falls and rises by modest amounts, although the value at its lowest (2010) is significantly lower than it was in 2005.

The proportion with an ATAR changes from year to year, showing moderate to signficant differences.

The average academic standing (whether estimated with our measure of overall achievement or with a score derived from the ATARs of the one-third who have them) also shows significant changes.

All this means that we should not assume that each year the group of students completing one or more TQA level 3 courses is sufficiently similar to the previous year's to make the scaling scores from different years automatically comparable. The changes should be actively monitored.

However, the impact of these changes is smaller than the sizes of the changes in scaled scores for subjects and the factors directly associated with these changes discussed in section 4.2, in particular both the changes in the standards applied in subjects and changes in patterns of participation in subjects, which are discussed in more detail below.

4.4 similarity of groups of students from year to year – subject by subject

Appendix B (see page 8) provides tables for courses with, on average, thirty or more students showing average and range of

- number of students
- proportion of female students[11]
- ages of students
- relative proportion [12] of students with an ATAR (that is, leaving out students in their first year since completing year 11 or those who do not complete at least four TQA level 3 courses over two years
- proportion of students with only one result in a given year
- average overall achievement (the columns headed MeanPoly)
- the relationship between achievement in the course (awards) and overall achievement (the columns headed Spearman.cor).

The tables are first sorted by course title, then by size, average proportion female and average overall achievement. We can see from these tables that there are lots of differences – courses with, on average, few students, courses with hardly any female students, courses taken by, on average, high achieving students and so on. It is obvious from these tables that courses with similar patterns of participation on one characteristic can vary widely on another: any single perspective is likely to mislead.

There are also courses where there is considerable variation from year to year.

There are clearly courses that are taken by students of higher overall academic achievement and courses that are taken by students who do not do well overall.

[11]a proportion greater than 0.54 is a course with, relative to the total cohort, a higher proportion of female students

[12]This is the actual proportion divided by the proportion in the total cohort. A value greater than 1.0 is a group with relatively more students with an ATAR than the total cohort

Subject	Prop.Female.Average	Prop.Female:variability	Age.Average	Age:variability	Study.count.Average	Study.count:variability
Computer Science	0.05	more unstable 25%	17.54	middle 50%	158.78	middle 50%
Computer Graphics and Design Stage 6	0.11	more stable 25%	17.78	middle 50%	257.67	more stable 25%
Audio Design	0.14	more unstable 25%	17.76	middle 50%	132.22	middle 50%
Computer Graphics and Design	0.16	more unstable 25%	17.69	middle 50%	277.20	more stable 25%
Physics	0.21	more unstable 25%	17.95	middle 50%	360.00	middle 50%
Outdoor Leadership	0.34	more unstable 25%	17.91	more stable 25%	126.83	more unstable 25%
Economics	0.38	more unstable 25%	17.84	middle 50%	254.56	middle 50%
Chinese (Specialist Level)	0.39	more unstable 25%	18.12	more unstable 25%	69.11	more unstable 25%
Media Production	0.44	more unstable 25%	17.62	middle 50%	151.57	more unstable 25%
Music	0.49	more unstable 25%	17.48	more unstable 25%	135.56	middle 50%
Housing and Design	0.49	more unstable 25%	17.67	middle 50%	210.00	more unstable 25%
Student Directed Inquiry	0.50	more unstable 25%	17.89	more stable 25%	61.75	more unstable 25%
Mathematics Applied	0.50	more stable 25%	17.49	more unstable 25%	1599.33	middle 50%
Australian Studies	0.59	more unstable 25%	18.81	more unstable 25%	37.38	more unstable 25%
Biology	0.67	more stable 25%	17.86	middle 50%	784.56	middle 50%
Studies of Religion	0.67	more stable 25%	17.66	more stable 25%	335.67	middle 50%
Art, Craft & Design - Production	0.68	more stable 25%	18.05	middle 50%	534.50	more stable 25%
German	0.68	more unstable 25%	18.79	more unstable 25%	43.11	more unstable 25%
Performance Stage 5	0.68	more stable 25%	17.98	more stable 25%	188.50	more stable 25%
Drama Stage 5	0.69	more stable 25%	17.77	more unstable 25%	307.50	more stable 25%
Sociology	0.78	more stable 25%	17.70	middle 50%	542.44	middle 50%
Art, Craft & Design - Appreciation	0.80	more stable 25%	17.99	more stable 25%	34.50	middle 50%
Choreography and Dance Performance	0.90	more stable 25%	17.45	middle 50%	80.33	middle 50%

Table 4.3: Significant variability proportion female cohort data for subjects 2005-2013)

Subject	MeanPoly.Average	MeanPoly:variability	Spearman.cor.Average	Spearman.cor:variability
Information Systems	0.12	more unstable 25%	0.80	middle 50%
Dance Stage 5	0.14	more stable 25%	0.83	more stable 25%
Information Technology and Systems	0.15	more unstable 25%	0.78	middle 50%
Chinese (Specialist Level)	0.18	more unstable 25%	0.68	more unstable 25%
Health Studies	0.20	more unstable 25%	0.89	middle 50%
Food Studies	0.24	more unstable 25%	0.86	middle 50%
Australian Studies	0.25	more unstable 25%	0.88	middle 50%
Audio Design	0.27	more unstable 25%	0.79	more unstable 25%
English as a Second Language	0.27	more unstable 25%	0.77	middle 50%
Australia in Asia and the Pacific	0.27	more unstable 25%	0.88	middle 50%
Food and Nutrition	0.33	more unstable 25%	0.87	middle 50%
Dance Choreography & Performance	0.35	more unstable 25%	0.60	more unstable 25%
Science of Natural Resources	0.38	more unstable 25%	0.87	middle 50%
Media Production	0.41	more unstable 25%	0.71	middle 50%
Art, Craft & Design - Production	0.49	more stable 25%	0.57	middle 50%
Performance Stage 5	0.63	more stable 25%	0.69	middle 50%
Studies of Religion	0.80	more stable 25%	0.82	middle 50%
Art, Craft & Design - Appreciation	0.90	more stable 25%	0.78	middle 50%
Music (Solo Performance)	0.96	more stable 25%	0.61	more unstable 25%
Chemistry	1.40	more stable 25%	0.85	middle 50%
Physics	1.43	more stable 25%	0.86	middle 50%
Mathematics Specialised	1.72	more stable 25%	0.66	more unstable 25%

Table 4.4: Significant variability overall achievement cohort data for subjects 2005-2013)

Table 4.3 shows the data for the courses where the variability[13] of the proportion of female students is outside (above or below) the middle range. There are clearly large and small courses with high and low female participation and with high and low variability in this rate of participation.

Table 4.4 shows the data for the courses where the variability of the overall achievement of the students is outside (above or below) the middle range. There is a tendency for the courses that are taken by higher achieving students to show lower variability over time, to have more stable participation. For example, the data confirms what many would expect: Mathematics Specalised is taken by an academically select group of students and there isn't much variation in this over time. Participation in Information Systems is, in contrast, changing over time.

The data given here and in appendix B demonstrates that no-one should ever assume that students doing different courses are essentially similar groups – a key assumption behind grade point averages or any process that treats the awards in different courses as being on the same five point scale. This is why some form of scaling is essential.

However, the data also show that there are often considerable variations over time: a key assumption of the current scaling and combining model is that there is little important variation from one year to the next.

In recent years, the Authority has provided assessment panels with information about cohort variations as a perspective that may explain variations in standards that they see. This process should continue.

[13]To estimate the variability, we divided the observed average by the range of values. We then grouped these ratios into the middle 50 per cent, the lower quartile – more stable – and the upper quartile – less stable.

It is also important to monitor actively changes in participation patterns. It will be necessary to alter the scaling processes if the data shows that these changes in participation are having are a more significant impact on students' ATARs than changes in the standards applied each year to student work.

Section 5

Assumption three: no systematic advantage/disadvantage for groups of students who choose different combinations of subjects

5.1 Introduction

It is common to hear people say that the ATAR system advantages the maths/science students, disadvantages the language students, the music students and so on.

Where we can show a systematic advantage/disadvantage to a group of students **by virtue of their membership of that group and no other** we should intervene in the scaling process to correct it (whether an unfair advantage or an unfair disadvantage).

The data do not support strong claims about widespread and obvious automatic advantage/disadvantage.

5.2 Patterns of study and ATAR outcomes

Generalisations about the relationship between pattern of study and ATAR results from the data (rather than from supposition or anecdote) are, as we can see from the following, difficult to make.

In 2013, there were 6937 students with one or more results in TQA level 3 courses. There were results in 60 different courses.

We simplified things a bit by taking into account *only* the thirty most common courses.

Across the 6937 students there were 2185 *different* combinations of these courses, Sexprdim(Pattern.summary)[1] different patterns of study. There are 1426 patterns where there is only one student with this choice of study. That is, two out every three of the patterns of study we saw in 2013 were chosen by one student. Nine out of ten of the patterns of study were chosen by four or fewer students. The most frequent pattern (across the thirty courses with the highest enrolment) in 2013 was the 190 students who did English Communications alone (from this set). The second most common was the 153 students who did English Communications, Mathematics Applied and Physical Sciences.

Most of the time, therefore, claims of the form "a person who does a particular combination of subjects gets a lower ATAR because of this choice" are not grounded in enough data to allow us to evaluate the claim. If there are only a few people who have actually studied this particular combination the data almost certainly won't tell us anything useful.

What about associations between your choosing particular subjects and the ATAR you get?

5.3 Subject choice and ATAR outcomes

The reason that the awards in some courses have higher scaled scores than the awards in others is that the students with, say, an EA in Chemistry have done better at the *other* things that they do in comparison with how well the students who have an EA in English Communications do in the other things that *they* do.

This is what the scaling system is intended to do. By and large, it is what it does.

There are exceptions – small numbers can cause problems. A more subtle problem arises because the approach we use enforces a requirement that the scaled score for an EA in a course must be higher than the scaled score for an HA in the same course. Sometimes this doesn't fit the data very well: there are cases where the students with exceptional achievement in a course have done less well at the other things they do than the students who haven't done as well in this course. It is easy to think up ways in which this can happen.

We can get some idea of the differences in achievement in other courses for students with an EA by working out an average GPA for each such student[1]. As discussed earlier(page 2.1 and page 4.4, GPA is a bad measure, with systematic bias programmed into it. However, it has some intuitive appeal in checking the plausibility of claims about relative performance.

Table 5.1 lists courses where there were at least ten EAs in 2013 in the order of students' average result in the other courses they did that year. A high average means that, on average, the students getting an EA in this course did well in their other studies.

The courses with students doing well (even on the broken measure provided by GPA) in their other studies include, as expected, the courses with higher scaled values for EAs: Mathematics Specialised, Physics, Economics and so on. The ones where students do less well in the other courses that they do include, as expected, the courses with lower scaled values for EAs: Japanese, Art Studio Practice, Computer Science. Exceptions, such as Art Production[2], illustrate the difference between the scaling model we use and a model – GPA – that asserts, *wrongly*, that the awards in different courses are already on a common five-point scale.

Table 5.2 provides a similar picture for students with an award of SA. The information in these tables is consistent with the intended purpose of scaling (and reflects common perceptions about the "difficulty" of more demanding mathematics and science courses). The information is, however, also consistent with a claim that the higher scaled results in these courses reflects higher inter-subject correlations (students who are good at one of these are good at others) because, and this is where the problem arises, these subjects are more-or-less the same subject. This feature, it is said, is not so for other areas of study to the same extent.

This issue is not easily resolved. Further extensive study of the data set over the years and of examinations is needed. It is gaining in importance with the increase in students' including in their ATAR calculations results from studies at the University.

[1] PA is assigned 0, SA is assigned 1, CA is assigned 2, HA is assigned 3 and EA is assigned 4

[2] In 2013, an EA in Art Production had a minimum scaled value of 21.2

Subject	Award	Mean GPA in other TQA level 3 courses
Art Production	EA	2.19
Japanese	EA	2.35
Computer Science	EA	2.58
Art Studio Practice	EA	2.59
Theatre Performance	EA	2.61
Ancient Civilisations	EA	2.66
Sport Science	EA	2.66
Food & Nutrition	EA	2.69
Business Studies	EA	2.72
Accounting	EA	2.73
English Writing	EA	2.74
Student Directed Inquiry	EA	2.77
Mathematics Applied	EA	2.78
English as a Second Language	EA	2.78
Music	EA	2.79
German	EA	2.83
Sociology	EA	2.89
French	EA	2.90
English Communications	EA	2.90
Psychology	EA	2.96
English Studies	EA	2.97
Media Production	EA	2.98
Geography	EA	3.00
Health Studies	EA	3.02
Physical Sciences	EA	3.03
Studies of Religion	EA	3.10
Australia in Asia and the Pacific	EA	3.11
Biology	EA	3.20
Legal Studies	EA	3.26
Environmental Science	EA	3.28
Mathematics Methods	EA	3.32
Chemistry	EA	3.38
Modern World History	EA	3.39
Economics	EA	3.46
Philosophy	EA	3.47
Physics	EA	3.48
Mathematics Specialised	EA	3.59

Table 5.1: Average award in other courses (2013 data) for courses with at least 10 EAs

	Subject	Award	Mean GPA in other TQA level 3 courses
86	Chinese (Specialist Level)	SA	0.68
128	English as a Second Language	SA	0.86
235	Outdoor Leadership	SA	0.94
140	Food & Nutrition	SA	1.05
14	Ancient Civilisations	SA	1.11
62	Business Studies	SA	1.11
265	Drama	SA	1.13
68	Computer Graphics & Design	SA	1.15
8	Accounting	SA	1.15
134	Environmental Science	SA	1.15
181	Computer Science	SA	1.17
110	English Communications	SA	1.18
169	Health Studies	SA	1.21
283	Sport Science	SA	1.21
152	Geography	SA	1.22
44	Psychology	SA	1.22
50	Sociology	SA	1.24
122	English Writing	SA	1.24
32	Art Studio Practice	SA	1.25
241	Philosophy	SA	1.25
104	Advanced Electronics	SA	1.26
2	Australia in Asia and the Pacific	SA	1.26
289	Technical Graphics	SA	1.27
187	Information Systems & Digital Technologies	SA	1.30
175	Modern World History	SA	1.32
211	Music	SA	1.34
116	English Studies	SA	1.35
163	Housing and Design	SA	1.36
92	Dance Choreography & Performance	SA	1.37
205	Media Production	SA	1.38
38	Audio Design	SA	1.38
26	Art Production	SA	1.40
277	Theatre Performance	SA	1.41
199	Legal Studies	SA	1.42
217	Mathematics Applied	SA	1.42
98	Economics	SA	1.48
259	Studies of Religion	SA	1.48
271	Student Directed Inquiry	SA	1.51
193	Japanese	SA	1.59
56	Biology	SA	1.64
253	Physical Sciences	SA	1.67
20	Art Appreciation	SA	1.72
229	Mathematics Specialised	SA	1.88
223	Mathematics Methods	SA	1.94
247	Physics	SA	2.00
74	Chemistry	SA	2.20

Table 5.2: Average award in other courses (2013 data) for courses with at least 10 SAs

Section 6

Conclusions and implications

It is clear that scaling, in some form, is essential. The data do not support an assumption that awards in different courses are already on the same scale and so can just be added up.

Scaling – having a data-driven[1] way of putting results in different courses on a common scale – is fundamental to fairness. This applies to all kinds of results that are to be included in ATAR calculations, whether TQA accredited courses or results in University courses.

The inclusion of results from a wide range of courses in ATAR calculations requires a balance across competing considerations. The extent to which results in a course provide an adequate estimate of overall academic achievement has sometimes to be balanced against policy reasons for allowing results to count towards an ATAR. That is, an argument for dropping a course on the grounds that the results are not a good estimate of overall academic achievement has to be set against an argument that including the course has sufficient merit on other grounds, despite the anomalies that this causes.

Students can reasonably expect that the value they get from an EA in a course taken one year should be similar to the value they would have got had they taken the same course in another year. This is an expectation for year-to-year comparability of scaled scores.

The year-to-year comparability of scaled scores (essential given the way that Tasmanian students usually count scaled scores from two years' of studies) rests on assumptions that are only partly met with in practice.

Year-to-year comparability of standards applied to external assessment in each course is typically the largest single factor affecting year-to-year comparability of scaled scores. The Authority has had an external audit of its procedures for year-to-year comparability of standards (see http://www.tqa.tas.gov.au/2621) and has made some enhancements to its processes. Further developments are needed.

There are other sources of uncertainty in students' ATARs – assessment and standards setting, the scaling processes that are used (see section 2.2.1, page 7), the alignment of ATARs with population parameters, a process intended to make ATARs comparable across jurisdictions. Further studies are needed.

Continuing monitoring and review of the technical adequacy of the procedures used for ATARs is clearly essential.

[1]Data-driven means using data rather than *a priori* beliefs, assertions or guesses based on experience. There is a lot to be said for using data.

Section 7

Appendix A

AUSTRALASIAN CURRICULUM, ASSESSMENT AND CERTIFICATION AUTHORITIES

ATAR – policy and technical issues: implications for ACACA agencies

Report to September 2014 CEOs meeting in Wellington, New Zealand:
summary of discussions held at meeting in Melbourne August 2014 and recommendations

Background

At their May 2014 meeting, ACACA members discussed the changing environment of university admissions and the continuing implications for ACACA agencies of the determination, publication and use of Australian Tertiary Admissions Ranks (ATAR). Member felt that it would be timely to revisit both policy and technical issues in depth. It was agreed that each agency would sponsor attendance at a full day's workshop for senior policy and technical officers from each jurisdiction.

A well-attended meeting was held August 13 2014 in Melbourne

Summary of discussions

1. National picture

1.1. There is a changing balance of selection and recruitment into tertiary education under pressures of market dynamics, funding and demographic changes.

1.2. Currently, selection for students from year 12 into a wide range of institutions and courses uses ATAR – a simple measure of overall achievement in senior secondary studies.

1.3. There are long-standing arrangements that put ATARs onto a national comparable scale. A review of these arrangements is in progress.

1.4. There are differences in the detail of ATAR arrangements across jurisdictions in terms of

- scaling algorithms
- slight variations in the number and weight of inputs
- variations in the role of the ACACA agency, whether providing subject level data, doing ATAR calculations as a service or working in partnership with the universities. These variations sit within the different policy contexts and circumstances of ACACA agencies

1.5. There are a variety of bonus schemes in operation, adding considerable complications without necessarily having the intended substantive effect. Bonus schemes

- may involve changes to the input data before or after scaling but before aggregation and ranking
- may in effect reduce an admission cut-off for a course through an addition to a student's ATAR to bring the student up to the required minimum ATAR
- may lead to bonuses for nearly all students

1.6. There is an increasing variety of schemes for 'alternative entry' (applicants who do not have an ATAR). The evidence-basis of these schemes is not clear and publicly available. There are some cases of clearly evident large-scale anomalies.

2. Selection processes for high prestige courses and institutions have consequences for the senior secondary sector far beyond the students immediately involved.

3. ACACA agencies' collective experience over time indicates that

3.1.1. student behaviour in senior secondary studies in response to selection processes can usefully be modelled in terms of 'maximising return on effort'

3.1.2. some aspects of school behaviours can be understood in terms of the impact of selection processes and reporting of school data about student results

3.1.3. there is clear value to students in selection systems being as clear and simple as possible – multiple options and complexities restrict student pathways in senior secondary education

3.1.4. It is important that the mechanics of ATAR calculations do not have perverse outcomes

3.1.5. The general nature of an ATAR as an indicator of overall academic achievement in year 12 provides greater flexibility and choice to students and fewer restrictions on pathways during and after year 12. Destination specific measures increase restrictions on pathways.

3.1.6. There are greater pressures on quality assurance and integrity of results that are used in ATAR calculations.

3.1.7. The more value the processes and practices of an assessment and certification system place on the content and standards of a subject course the more students will value this content and the standards.

4. ATAR provides a simple and widely used selection mechanism, although its details and internal workings are not widely understood.

5. 'Small groups' (data from subjects where there are not many students) and ATAR calculations present significant technical and policy challenges. At least three classes of methods are in use, including

- aligning with cognate subjects
- administrative decision, informed by inspection of the data and prior experience
- combining several small subject groups to create a single, larger group

6. While there are excellent materials providing clear explanations of the ATAR process, it is clear that

6.1. much of the community sees this as 'black box'

6.2. there are persistent myths about the purpose, nature and outcomes of scaling.

Section 8

Appendix B

Subject	Study.count.Average	Study.count.Range	Prop.Female.Average	Prop.Female.Range	Age.Average	Age.Range
Accounting	332.89	153	0.45	0.09	17.84	0.51
Ancient Civilisations	311.00	126	0.60	0.13	17.75	0.49
Art Appreciation	56.57	44	0.80	0.09	18.01	0.86
Art Production	672.14	169	0.71	0.06	17.64	0.42
Art Studio Practice	132.80	137	0.70	0.07	17.98	0.04
Art, Craft & Design - Appreciation	34.50	7	0.80	0.02	17.99	0.09
Art, Craft & Design - Production	534.50	91	0.68	0.02	18.05	0.23
Audio Design	132.22	66	0.14	0.13	17.76	0.43
Australia in Asia and the Pacific	368.67	359	0.62	0.10	17.71	0.51
Australian Studies	37.38	76	0.59	0.54	18.81	4.61
Biology	784.56	154	0.67	0.04	17.86	0.30
Business Studies	313.33	311	0.50	0.06	17.68	0.19
Chemistry	616.67	157	0.47	0.07	17.97	0.16
Chinese (Specialist Level)	69.11	60	0.39	0.19	18.12	0.79
Choreography and Dance Performance	80.33	44	0.90	0.01	17.45	0.19
Computer Graphics and Design	277.20	42	0.16	0.06	17.69	0.30
Computer Graphics and Design Stage 6	257.67	44	0.11	0.01	17.78	0.57
Computer Science	158.78	35	0.05	0.03	17.54	0.42
Dance Choreography & Performance	83.75	31	0.83	0.06	17.50	0.12
Dance Stage 5	74.50	13	0.89	0.07	17.60	0.26
Design Graphics Stage 6	40.50	36	0.19	0.07	17.73	0.50
Drama	252.57	44	0.65	0.11	17.07	0.15
Drama Stage 5	307.50	3	0.69	0.01	17.77	1.42
Economics	254.56	113	0.38	0.17	17.84	0.24
English as a Second Language	171.89	70	0.45	0.14	18.50	1.05
English Communications	2007.00	309	0.55	0.04	17.35	0.74
English Studies	662.78	138	0.70	0.11	17.39	0.89
English Writing	708.11	256	0.64	0.05	17.78	0.62
Environmental Science	235.78	81	0.54	0.09	17.69	0.41
Food and Nutrition	279.40	336	0.78	0.11	17.77	0.19
Food Studies	88.33	52	0.81	0.18	17.80	0.16
French	118.89	58	0.66	0.11	20.69	11.95
Geography	236.00	116	0.56	0.12	17.84	0.26
German	43.11	39	0.68	0.32	18.79	8.88
Health Studies	996.44	193	0.71	0.07	17.63	0.58
Housing and Design	210.00	229	0.49	0.23	17.67	0.50
Information Systems	218.33	137	0.16	0.06	17.69	0.43
Information Technology and Systems	103.60	41	0.14	0.03	17.66	0.21
Japanese	111.67	62	0.66	0.11	17.62	1.46
Legal Studies	479.44	217	0.60	0.11	17.83	0.27
Mathematics Applied	1599.33	476	0.50	0.03	17.49	0.65
Mathematics Methods	793.44	126	0.41	0.07	17.43	0.88
Mathematics Specialised	234.78	61	0.29	0.08	17.98	0.27
Media Production	151.57	124	0.44	0.18	17.62	0.28
Modern World History	443.67	185	0.56	0.06	17.80	0.59
Music	135.56	56	0.49	0.23	17.48	0.62
Music (Solo Performance)	63.00	28	0.61	0.12	17.76	0.05
Music Performance	84.67	16	0.47	0.03	17.81	0.14
Outdoor Leadership	126.83	124	0.34	0.15	17.91	0.04
Performance Stage 5	188.50	7	0.68	0.01	17.98	0.02
Physical Sciences	1329.56	154	0.44	0.05	17.26	1.35
Physics	360.00	73	0.21	0.09	17.95	0.18
Psychology	826.78	138	0.74	0.06	17.80	0.14
Religion and Philosophy	433.67	470	0.63	0.16	17.83	0.19
Science of Natural Resources	45.00	34	0.38	0.10	17.64	0.50
Sociology	542.44	169	0.78	0.04	17.70	0.44
Sport Science	496.75	114	0.50	0.06	17.63	0.51
Sports Science	548.00	85	0.46	0.08	17.56	0.07
Student Directed Inquiry	61.75	72	0.50	0.33	17.89	0.13
Studies of Religion	335.67	97	0.67	0.02	17.66	0.09
Theatre Performance	185.71	37	0.63	0.09	17.96	0.31

Table 8.1: Variations in cohort data for subjects 2005-2013: counts, proportion female, age

Subject	Rel.Prop.ATAR.Average	Rel.Prop.ATAR.Range	Prop.One.result.Average	Prop.One.result.Range
Accounting	1.61	0.37	0.05	0.04
Ancient Civilisations	1.48	0.24	0.07	0.05
Art Appreciation	2.37	0.67	0.01	0.04
Art Production	1.22	0.47	0.11	0.06
Art Studio Practice	2.30	0.30	0.05	0.05
Art, Craft & Design - Appreciation	2.28	0.39	0.00	0.00
Art, Craft & Design - Production	1.50	0.04	0.10	0.02
Audio Design	1.16	0.41	0.15	0.10
Australia in Asia and the Pacific	1.47	0.77	0.08	0.10
Australian Studies	1.70	0.87	0.08	0.22
Biology	2.12	0.28	0.01	0.01
Business Studies	1.57	0.45	0.08	0.07
Chemistry	2.72	0.30	0.00	0.01
Chinese (Specialist Level)	0.87	0.56	0.11	0.11
Choreography and Dance Performance	1.32	0.18	0.11	0.03
Computer Graphics and Design	1.21	0.19	0.10	0.04
Computer Graphics and Design Stage 6	1.22	0.25	0.14	0.04
Computer Science	1.12	0.49	0.09	0.05
Dance Choreography & Performance	1.25	0.19	0.09	0.06
Dance Stage 5	0.78	0.11	0.15	0.00
Design Graphics Stage 6	1.18	0.39	0.08	0.17
Drama	0.25	0.17	0.11	0.05
Drama Stage 5	0.19	0.03	0.11	0.02
Economics	2.27	0.41	0.01	0.03
English as a Second Language	0.95	0.38	0.14	0.11
English Communications	0.65	0.31	0.07	0.03
English Studies	0.83	0.16	0.04	0.04
English Writing	1.58	0.18	0.08	0.04
Environmental Science	1.70	0.59	0.04	0.04
Food and Nutrition	1.82	0.28	0.05	0.03
Food Studies	1.74	0.78	0.09	0.14
French	0.76	0.55	0.12	0.17
Geography	2.03	0.25	0.04	0.03
German	0.91	0.87	0.06	0.14
Health Studies	1.30	0.15	0.10	0.02
Housing and Design	1.40	0.75	0.09	0.07
Information Systems	1.29	0.43	0.13	0.09
Information Technology and Systems	1.23	0.41	0.09	0.08
Japanese	0.84	0.49	0.08	0.10
Legal Studies	1.94	0.30	0.04	0.04
Mathematics Applied	0.92	0.27	0.05	0.02
Mathematics Methods	0.73	0.23	0.01	0.01
Mathematics Specialised	2.82	0.30	0.00	0.01
Media Production	1.54	0.61	0.04	0.04
Modern World History	1.70	0.29	0.05	0.04
Music	0.83	0.58	0.04	0.07
Music (Solo Performance)	2.25	0.39	0.01	0.03
Music Performance	2.37	0.28	0.04	0.05
Outdoor Leadership	1.93	0.42	0.07	0.04
Performance Stage 5	2.37	0.03	0.07	0.02
Physical Sciences	0.24	0.07	0.04	0.03
Physics	2.60	0.20	0.00	0.01
Psychology	2.02	0.32	0.03	0.02
Religion and Philosophy	2.31	0.56	0.02	0.02
Science of Natural Resources	1.31	0.44	0.09	0.07
Sociology	1.70	0.34	0.03	0.03
Sport Science	1.43	0.16	0.06	0.03
Sports Science	1.38	0.05	0.07	0.03
Student Directed Inquiry	2.39	0.77	0.02	0.04
Studies of Religion	1.82	0.08	0.01	0.01
Theatre Performance	2.27	0.29	0.07	0.08

Table 8.2: Variations in cohort data for subjects 2005-2013: relative proportion with an ATAR, proportion with only one result

Subject	MeanPoly.Average	MeanPoly.Range	Spearman.cor.Average	Spearman.cor.Range
Accounting	0.51	0.21	0.86	0.10
Ancient Civilisations	0.43	0.25	0.86	0.07
Art Appreciation	0.80	0.30	0.67	0.30
Art Production	0.46	0.16	0.57	0.20
Art Studio Practice	0.57	0.14	0.68	0.09
Art, Craft & Design - Appreciation	0.90	0.00	0.78	0.12
Art, Craft & Design - Production	0.49	0.06	0.57	0.08
Audio Design	0.27	0.25	0.79	0.23
Australia in Asia and the Pacific	0.27	0.26	0.88	0.05
Australian Studies	0.25	0.55	0.88	0.14
Biology	0.93	0.17	0.89	0.02
Business Studies	0.39	0.09	0.86	0.05
Chemistry	1.40	0.09	0.85	0.06
Chinese (Specialist Level)	0.18	0.24	0.68	0.21
Choreography and Dance Performance	0.30	0.12	0.43	0.29
Computer Graphics and Design	0.30	0.14	0.81	0.13
Computer Graphics and Design Stage 6	0.22	0.13	0.78	0.11
Computer Science	0.51	0.33	0.88	0.07
Dance Choreography & Performance	0.35	0.27	0.60	0.34
Dance Stage 5	0.14	0.01	0.83	0.03
Design Graphics Stage 6	0.56	0.17	0.75	0.27
Drama	0.33	0.17	0.81	0.12
Drama Stage 5	0.35	0.25	0.81	0.02
Economics	0.87	0.39	0.88	0.05
English as a Second Language	0.27	0.34	0.77	0.15
English Communications	0.43	0.16	0.82	0.02
English Studies	0.74	0.19	0.82	0.06
English Writing	0.52	0.13	0.78	0.06
Environmental Science	0.57	0.22	0.88	0.05
Food and Nutrition	0.33	0.25	0.87	0.05
Food Studies	0.24	0.45	0.86	0.10
French	1.13	0.30	0.71	0.21
Geography	0.55	0.23	0.86	0.06
German	1.06	0.54	0.47	0.72
Health Studies	0.20	0.17	0.89	0.05
Housing and Design	0.31	0.22	0.80	0.08
Information Systems	0.12	0.35	0.80	0.08
Information Technology and Systems	0.15	0.24	0.78	0.14
Japanese	0.78	0.27	0.61	0.45
Legal Studies	0.58	0.26	0.89	0.03
Mathematics Applied	0.46	0.15	0.84	0.05
Mathematics Methods	1.04	0.15	0.76	0.09
Mathematics Specialised	1.72	0.19	0.66	0.20
Media Production	0.41	0.32	0.71	0.18
Modern World History	0.63	0.17	0.87	0.04
Music	0.67	0.23	0.75	0.15
Music (Solo Performance)	0.96	0.02	0.61	0.20
Music Performance	0.90	0.18	0.53	0.12
Outdoor Leadership	0.36	0.23	0.83	0.18
Performance Stage 5	0.63	0.04	0.69	0.05
Physical Sciences	0.69	0.21	0.89	0.02
Physics	1.43	0.19	0.86	0.05
Psychology	0.57	0.19	0.87	0.05
Religion and Philosophy	0.78	0.25	0.80	0.15
Science of Natural Resources	0.38	0.41	0.87	0.12
Sociology	0.45	0.27	0.87	0.04
Sport Science	0.39	0.19	0.90	0.03
Sports Science	0.33	0.23	0.91	0.05
Student Directed Inquiry	0.79	0.34	0.66	0.40
Studies of Religion	0.80	0.05	0.82	0.07
Theatre Performance	0.53	0.26	0.66	0.28

Table 8.3: Variations in cohort data for subjects 2005-2013: average overall achievement, correlation between awards and overall achievement

Subject	Study.count.Average	Study.count.Range	Prop.Female.Average	Prop.Female.Range	Age.Average	Age.Range
Art, Craft & Design - Appreciation	34.50	7	0.80	0.02	17.99	0.09
Australian Studies	37.38	76	0.59	0.54	18.81	4.61
Design Graphics Stage 6	40.50	36	0.19	0.07	17.73	0.50
German	43.11	39	0.68	0.32	18.79	8.88
Science of Natural Resources	45.00	34	0.38	0.10	17.64	0.50
Art Appreciation	56.57	44	0.80	0.09	18.01	0.86
Student Directed Inquiry	61.75	72	0.50	0.33	17.89	0.13
Music (Solo Performance)	63.00	28	0.61	0.12	17.76	0.05
Chinese (Specialist Level)	69.11	60	0.39	0.19	18.12	0.79
Dance Stage 5	74.50	13	0.89	0.07	17.60	0.26
Choreography and Dance Performance	80.33	44	0.90	0.01	17.45	0.19
Dance Choreography & Performance	83.75	31	0.83	0.06	17.50	0.12
Music Performance	84.67	16	0.47	0.03	17.81	0.14
Food Studies	88.33	52	0.81	0.18	17.80	0.16
Information Technology and Systems	103.60	41	0.14	0.03	17.66	0.21
Japanese	111.67	62	0.66	0.11	17.62	1.46
French	118.89	58	0.66	0.11	20.69	11.95
Outdoor Leadership	126.83	124	0.34	0.15	17.91	0.04
Audio Design	132.22	66	0.14	0.13	17.76	0.43
Art Studio Practice	132.80	137	0.70	0.07	17.98	0.04
Music	135.56	56	0.49	0.23	17.48	0.62
Media Production	151.57	124	0.44	0.18	17.62	0.28
Computer Science	158.78	35	0.05	0.03	17.54	0.42
English as a Second Language	171.89	70	0.45	0.14	18.50	1.05
Theatre Performance	185.71	37	0.63	0.09	17.96	0.31
Performance Stage 5	188.50	7	0.68	0.01	17.98	0.02
Housing and Design	210.00	229	0.49	0.23	17.67	0.50
Information Systems	218.33	137	0.16	0.06	17.69	0.43
Mathematics Specialised	234.78	61	0.29	0.08	17.98	0.27
Environmental Science	235.78	81	0.54	0.09	17.69	0.41
Geography	236.00	116	0.56	0.12	17.84	0.26
Drama	252.57	44	0.65	0.11	17.07	0.15
Economics	254.56	113	0.38	0.17	17.84	0.24
Computer Graphics and Design Stage 6	257.67	44	0.11	0.01	17.78	0.57
Computer Graphics and Design	277.20	42	0.16	0.06	17.69	0.30
Food and Nutrition	279.40	336	0.78	0.11	17.77	0.19
Drama Stage 5	307.50	3	0.69	0.01	17.77	1.42
Ancient Civilisations	311.00	126	0.60	0.13	17.75	0.49
Business Studies	313.33	311	0.50	0.06	17.68	0.19
Accounting	332.89	153	0.45	0.09	17.84	0.51
Studies of Religion	335.67	97	0.67	0.02	17.66	0.09
Physics	360.00	73	0.21	0.09	17.95	0.18
Australia in Asia and the Pacific	368.67	359	0.62	0.10	17.71	0.51
Religion and Philosophy	433.67	470	0.63	0.16	17.83	0.19
Modern World History	443.67	185	0.56	0.06	17.80	0.59
Legal Studies	479.44	217	0.60	0.11	17.83	0.27
Sport Science	496.75	114	0.50	0.06	17.63	0.51
Art, Craft & Design - Production	534.50	91	0.68	0.02	18.05	0.23
Sociology	542.44	169	0.78	0.04	17.70	0.44
Sports Science	548.00	85	0.46	0.08	17.56	0.07
Chemistry	616.67	157	0.47	0.07	17.97	0.16
English Studies	662.78	138	0.70	0.11	17.39	0.89
Art Production	672.14	169	0.71	0.06	17.64	0.42
English Writing	708.11	256	0.64	0.05	17.78	0.62
Biology	784.56	154	0.67	0.04	17.86	0.30
Mathematics Methods	793.44	126	0.41	0.07	17.43	0.88
Psychology	826.78	138	0.74	0.06	17.80	0.14
Health Studies	996.44	193	0.71	0.07	17.63	0.58
Physical Sciences	1329.56	154	0.44	0.05	17.26	1.35
Mathematics Applied	1599.33	476	0.50	0.03	17.49	0.65
English Communications	2007.00	309	0.55	0.04	17.35	0.74

Table 8.4: Variations in cohort data for subjects 2005-2013: counts, proportion female, age (sorted by number of students)

Subject	Rel.Prop.ATAR.Average	Rel.Prop.ATAR.Range	Prop.One.result.Average	Prop.One.result.Range
Art, Craft & Design - Appreciation	2.28	0.39	0.00	0.00
Australian Studies	1.70	0.87	0.08	0.22
Design Graphics Stage 6	1.18	0.39	0.08	0.17
German	0.91	0.87	0.06	0.14
Science of Natural Resources	1.31	0.44	0.09	0.07
Art Appreciation	2.37	0.67	0.01	0.04
Student Directed Inquiry	2.39	0.77	0.02	0.04
Music (Solo Performance)	2.25	0.39	0.01	0.03
Chinese (Specialist Level)	0.87	0.56	0.11	0.11
Dance Stage 5	0.78	0.11	0.15	0.00
Choreography and Dance Performance	1.32	0.18	0.11	0.03
Dance Choreography & Performance	1.25	0.19	0.09	0.06
Music Performance	2.37	0.28	0.04	0.05
Food Studies	1.74	0.78	0.09	0.14
Information Technology and Systems	1.23	0.41	0.09	0.08
Japanese	0.84	0.49	0.08	0.10
French	0.76	0.55	0.12	0.17
Outdoor Leadership	1.93	0.42	0.07	0.04
Audio Design	1.16	0.41	0.15	0.10
Art Studio Practice	2.30	0.30	0.05	0.05
Music	0.83	0.58	0.04	0.07
Media Production	1.54	0.61	0.04	0.04
Computer Science	1.12	0.49	0.09	0.05
English as a Second Language	0.95	0.38	0.14	0.11
Theatre Performance	2.27	0.29	0.07	0.08
Performance Stage 5	2.37	0.03	0.07	0.02
Housing and Design	1.40	0.75	0.09	0.07
Information Systems	1.29	0.43	0.13	0.09
Mathematics Specialised	2.82	0.30	0.00	0.01
Environmental Science	1.70	0.59	0.04	0.04
Geography	2.03	0.25	0.04	0.03
Drama	0.25	0.17	0.11	0.05
Economics	2.27	0.41	0.01	0.03
Computer Graphics and Design Stage 6	1.22	0.25	0.14	0.04
Computer Graphics and Design	1.21	0.19	0.10	0.04
Food and Nutrition	1.82	0.28	0.05	0.03
Drama Stage 5	0.19	0.03	0.11	0.02
Ancient Civilisations	1.48	0.24	0.07	0.05
Business Studies	1.57	0.45	0.08	0.07
Accounting	1.61	0.37	0.05	0.04
Studies of Religion	1.82	0.08	0.01	0.01
Physics	2.60	0.20	0.00	0.01
Australia in Asia and the Pacific	1.47	0.77	0.08	0.10
Religion and Philosophy	2.31	0.56	0.02	0.02
Modern World History	1.70	0.29	0.05	0.04
Legal Studies	1.94	0.30	0.04	0.04
Sport Science	1.43	0.16	0.06	0.03
Art, Craft & Design - Production	1.50	0.04	0.10	0.02
Sociology	1.70	0.34	0.03	0.03
Sports Science	1.38	0.05	0.07	0.03
Chemistry	2.72	0.30	0.00	0.01
English Studies	0.83	0.16	0.04	0.04
Art Production	1.22	0.47	0.11	0.06
English Writing	1.58	0.18	0.08	0.04
Biology	2.12	0.28	0.01	0.01
Mathematics Methods	0.73	0.23	0.01	0.01
Psychology	2.02	0.32	0.03	0.02
Health Studies	1.30	0.15	0.10	0.02
Physical Sciences	0.24	0.07	0.04	0.03
Mathematics Applied	0.92	0.27	0.05	0.02
English Communications	0.65	0.31	0.07	0.03

Table 8.5: Variations in cohort data for subjects 2005-2013: relative proportion with an ATAR, proportion with only one result (sorted by number of students)

Subject	MeanPoly.Average	MeanPoly.Range	Spearman.cor.Average	Spearman.cor.Range
Art, Craft & Design - Appreciation	0.90	0.00	0.78	0.12
Australian Studies	0.25	0.55	0.88	0.14
Design Graphics Stage 6	0.56	0.17	0.75	0.27
German	1.06	0.54	0.47	0.72
Science of Natural Resources	0.38	0.41	0.87	0.12
Art Appreciation	0.80	0.30	0.67	0.30
Student Directed Inquiry	0.79	0.34	0.66	0.40
Music (Solo Performance)	0.96	0.02	0.61	0.20
Chinese (Specialist Level)	0.18	0.24	0.68	0.21
Dance Stage 5	0.14	0.01	0.83	0.03
Choreography and Dance Performance	0.30	0.12	0.43	0.29
Dance Choreography & Performance	0.35	0.27	0.60	0.34
Music Performance	0.90	0.18	0.53	0.12
Food Studies	0.24	0.45	0.86	0.10
Information Technology and Systems	0.15	0.24	0.78	0.14
Japanese	0.78	0.27	0.61	0.45
French	1.13	0.30	0.71	0.21
Outdoor Leadership	0.36	0.23	0.83	0.18
Audio Design	0.27	0.25	0.79	0.23
Art Studio Practice	0.57	0.14	0.68	0.09
Music	0.67	0.23	0.75	0.15
Media Production	0.41	0.32	0.71	0.18
Computer Science	0.51	0.33	0.88	0.07
English as a Second Language	0.27	0.34	0.77	0.15
Theatre Performance	0.53	0.26	0.66	0.28
Performance Stage 5	0.63	0.04	0.69	0.05
Housing and Design	0.31	0.22	0.80	0.08
Information Systems	0.12	0.35	0.80	0.08
Mathematics Specialised	1.72	0.19	0.66	0.20
Environmental Science	0.57	0.22	0.88	0.05
Geography	0.55	0.23	0.86	0.06
Drama	0.33	0.17	0.81	0.12
Economics	0.87	0.39	0.88	0.05
Computer Graphics and Design Stage 6	0.22	0.13	0.78	0.11
Computer Graphics and Design	0.30	0.14	0.81	0.13
Food and Nutrition	0.33	0.25	0.87	0.05
Drama Stage 5	0.35	0.25	0.81	0.02
Ancient Civilisations	0.43	0.25	0.86	0.07
Business Studies	0.39	0.09	0.86	0.05
Accounting	0.51	0.21	0.86	0.10
Studies of Religion	0.80	0.05	0.82	0.07
Physics	1.43	0.19	0.86	0.05
Australia in Asia and the Pacific	0.27	0.26	0.88	0.05
Religion and Philosophy	0.78	0.25	0.80	0.15
Modern World History	0.63	0.17	0.87	0.04
Legal Studies	0.58	0.26	0.89	0.03
Sport Science	0.39	0.19	0.90	0.03
Art, Craft & Design - Production	0.49	0.06	0.57	0.08
Sociology	0.45	0.27	0.87	0.04
Sports Science	0.33	0.23	0.91	0.05
Chemistry	1.40	0.09	0.85	0.06
English Studies	0.74	0.19	0.82	0.06
Art Production	0.46	0.16	0.57	0.20
English Writing	0.52	0.13	0.78	0.06
Biology	0.93	0.17	0.89	0.02
Mathematics Methods	1.04	0.15	0.76	0.09
Psychology	0.57	0.19	0.87	0.05
Health Studies	0.20	0.17	0.89	0.05
Physical Sciences	0.69	0.21	0.89	0.02
Mathematics Applied	0.46	0.15	0.84	0.05
English Communications	0.43	0.16	0.82	0.02

Table 8.6: Variations in cohort data for subjects 2005-2013: average overall achievement, correlation between awards and overall achievement (sorted by number of students)

Subject	Study.count.Average	Study.count.Range	Prop.Female.Average	Prop.Female.Range	Age.Average	Age.Range
Computer Science	158.78	35	0.05	0.03	17.54	0.42
Computer Graphics and Design Stage 6	257.67	44	0.11	0.01	17.78	0.57
Audio Design	132.22	66	0.14	0.13	17.76	0.43
Information Technology and Systems	103.60	41	0.14	0.03	17.66	0.21
Computer Graphics and Design	277.20	42	0.16	0.06	17.69	0.30
Information Systems	218.33	137	0.16	0.06	17.69	0.43
Design Graphics Stage 6	40.50	36	0.19	0.07	17.73	0.50
Physics	360.00	73	0.21	0.09	17.95	0.18
Mathematics Specialised	234.78	61	0.29	0.08	17.98	0.27
Outdoor Leadership	126.83	124	0.34	0.15	17.91	0.04
Economics	254.56	113	0.38	0.17	17.84	0.24
Science of Natural Resources	45.00	34	0.38	0.10	17.64	0.50
Chinese (Specialist Level)	69.11	60	0.39	0.19	18.12	0.79
Mathematics Methods	793.44	126	0.41	0.07	17.43	0.88
Physical Sciences	1329.56	154	0.44	0.05	17.26	1.35
Media Production	151.57	124	0.44	0.18	17.62	0.28
Accounting	332.89	153	0.45	0.09	17.84	0.51
English as a Second Language	171.89	70	0.45	0.14	18.50	1.05
Sports Science	548.00	85	0.46	0.08	17.56	0.07
Music Performance	84.67	16	0.47	0.03	17.81	0.14
Chemistry	616.67	157	0.47	0.07	17.97	0.16
Music	135.56	56	0.49	0.23	17.48	0.62
Housing and Design	210.00	229	0.49	0.23	17.67	0.50
Student Directed Inquiry	61.75	72	0.50	0.33	17.89	0.13
Business Studies	313.33	311	0.50	0.06	17.68	0.19
Sport Science	496.75	114	0.50	0.06	17.63	0.51
Mathematics Applied	1599.33	476	0.50	0.03	17.49	0.65
Environmental Science	235.78	81	0.54	0.09	17.69	0.41
English Communications	2007.00	309	0.55	0.04	17.35	0.74
Modern World History	443.67	185	0.56	0.06	17.80	0.59
Geography	236.00	116	0.56	0.12	17.84	0.26
Australian Studies	37.38	76	0.59	0.54	18.81	4.61
Ancient Civilisations	311.00	126	0.60	0.13	17.75	0.49
Legal Studies	479.44	217	0.60	0.11	17.83	0.27
Music (Solo Performance)	63.00	28	0.61	0.12	17.76	0.05
Australia in Asia and the Pacific	368.67	359	0.62	0.10	17.71	0.51
Religion and Philosophy	433.67	470	0.63	0.16	17.83	0.19
Theatre Performance	185.71	37	0.63	0.09	17.96	0.31
English Writing	708.11	256	0.64	0.05	17.78	0.62
Drama	252.57	44	0.65	0.11	17.07	0.15
Japanese	111.67	62	0.66	0.11	17.62	1.46
French	118.89	58	0.66	0.11	20.69	11.95
Biology	784.56	154	0.67	0.04	17.86	0.30
Studies of Religion	335.67	97	0.67	0.02	17.66	0.09
Art, Craft & Design - Production	534.50	91	0.68	0.02	18.05	0.23
German	43.11	39	0.68	0.32	18.79	8.88
Performance Stage 5	188.50	7	0.68	0.01	17.98	0.02
Drama Stage 5	307.50	3	0.69	0.01	17.77	1.42
English Studies	662.78	138	0.70	0.11	17.39	0.89
Art Studio Practice	132.80	137	0.70	0.07	17.98	0.04
Art Production	672.14	169	0.71	0.06	17.64	0.42
Health Studies	996.44	193	0.71	0.07	17.63	0.58
Psychology	826.78	138	0.74	0.06	17.80	0.14
Sociology	542.44	169	0.78	0.04	17.70	0.44
Food and Nutrition	279.40	336	0.78	0.11	17.77	0.19
Art Appreciation	56.57	44	0.80	0.09	18.01	0.86
Art, Craft & Design - Appreciation	34.50	7	0.80	0.02	17.99	0.09
Food Studies	88.33	52	0.81	0.18	17.80	0.16
Dance Choreography & Performance	83.75	31	0.83	0.06	17.50	0.12
Dance Stage 5	74.50	13	0.89	0.07	17.60	0.26
Choreography and Dance Performance	80.33	44	0.90	0.01	17.45	0.19

Table 8.7: Variations in cohort data for subjects 2005-2013: counts, proportion female, age (sorted by proportion female)

Subject	Rel.Prop.ATAR.Average	Rel.Prop.ATAR.Range	Prop.One.result.Average	Prop.One.result.Range
Computer Science	1.12	0.49	0.09	0.05
Computer Graphics and Design Stage 6	1.22	0.25	0.14	0.04
Audio Design	1.16	0.41	0.15	0.10
Information Technology and Systems	1.23	0.41	0.09	0.08
Computer Graphics and Design	1.21	0.19	0.10	0.04
Information Systems	1.29	0.43	0.13	0.09
Design Graphics Stage 6	1.18	0.39	0.08	0.17
Physics	2.60	0.20	0.00	0.01
Mathematics Specialised	2.82	0.30	0.00	0.01
Outdoor Leadership	1.93	0.42	0.07	0.04
Economics	2.27	0.41	0.01	0.03
Science of Natural Resources	1.31	0.44	0.09	0.07
Chinese (Specialist Level)	0.87	0.56	0.11	0.11
Mathematics Methods	0.73	0.23	0.01	0.01
Physical Sciences	0.24	0.07	0.04	0.03
Media Production	1.54	0.61	0.04	0.04
Accounting	1.61	0.37	0.05	0.04
English as a Second Language	0.95	0.38	0.14	0.11
Sports Science	1.38	0.05	0.07	0.03
Music Performance	2.37	0.28	0.04	0.05
Chemistry	2.72	0.30	0.00	0.01
Music	0.83	0.58	0.04	0.07
Housing and Design	1.40	0.75	0.09	0.07
Student Directed Inquiry	2.39	0.77	0.02	0.04
Business Studies	1.57	0.45	0.08	0.07
Sport Science	1.43	0.16	0.06	0.03
Mathematics Applied	0.92	0.27	0.05	0.02
Environmental Science	1.70	0.59	0.04	0.04
English Communications	0.65	0.31	0.07	0.03
Modern World History	1.70	0.29	0.05	0.04
Geography	2.03	0.25	0.04	0.03
Australian Studies	1.70	0.87	0.08	0.22
Ancient Civilisations	1.48	0.24	0.07	0.05
Legal Studies	1.94	0.30	0.04	0.04
Music (Solo Performance)	2.25	0.39	0.01	0.03
Australia in Asia and the Pacific	1.47	0.77	0.08	0.10
Religion and Philosophy	2.31	0.56	0.02	0.02
Theatre Performance	2.27	0.29	0.07	0.08
English Writing	1.58	0.18	0.08	0.04
Drama	0.25	0.17	0.11	0.05
Japanese	0.84	0.49	0.08	0.10
French	0.76	0.55	0.12	0.17
Biology	2.12	0.28	0.01	0.01
Studies of Religion	1.82	0.08	0.01	0.01
Art, Craft & Design - Production	1.50	0.04	0.10	0.02
German	0.91	0.87	0.06	0.14
Performance Stage 5	2.37	0.03	0.07	0.02
Drama Stage 5	0.19	0.03	0.11	0.02
English Studies	0.83	0.16	0.04	0.04
Art Studio Practice	2.30	0.30	0.05	0.05
Art Production	1.22	0.47	0.11	0.06
Health Studies	1.30	0.15	0.10	0.02
Psychology	2.02	0.32	0.03	0.02
Sociology	1.70	0.34	0.03	0.03
Food and Nutrition	1.82	0.28	0.05	0.03
Art Appreciation	2.37	0.67	0.01	0.04
Art, Craft & Design - Appreciation	2.28	0.39	0.00	0.00
Food Studies	1.74	0.78	0.09	0.14
Dance Choreography & Performance	1.25	0.19	0.09	0.06
Dance Stage 5	0.78	0.11	0.15	0.00
Choreography and Dance Performance	1.32	0.18	0.11	0.03

Table 8.8: Variations in cohort data for subjects 2005-2013: relative proportion with an ATAR, proportion with only one result (sorted by proportion female)

Subject	MeanPoly.Average	MeanPoly.Range	Spearman.cor.Average	Spearman.cor.Range
Computer Science	0.51	0.33	0.88	0.07
Computer Graphics and Design Stage 6	0.22	0.13	0.78	0.11
Audio Design	0.27	0.25	0.79	0.23
Information Technology and Systems	0.15	0.24	0.78	0.14
Computer Graphics and Design	0.30	0.14	0.81	0.13
Information Systems	0.12	0.35	0.80	0.08
Design Graphics Stage 6	0.56	0.17	0.75	0.27
Physics	1.43	0.19	0.86	0.05
Mathematics Specialised	1.72	0.19	0.66	0.20
Outdoor Leadership	0.36	0.23	0.83	0.18
Economics	0.87	0.39	0.88	0.05
Science of Natural Resources	0.38	0.41	0.87	0.12
Chinese (Specialist Level)	0.18	0.24	0.68	0.21
Mathematics Methods	1.04	0.15	0.76	0.09
Physical Sciences	0.69	0.21	0.89	0.02
Media Production	0.41	0.32	0.71	0.18
Accounting	0.51	0.21	0.86	0.10
English as a Second Language	0.27	0.34	0.77	0.15
Sports Science	0.33	0.23	0.91	0.05
Music Performance	0.90	0.18	0.53	0.12
Chemistry	1.40	0.09	0.85	0.06
Music	0.67	0.23	0.75	0.15
Housing and Design	0.31	0.22	0.80	0.08
Student Directed Inquiry	0.79	0.34	0.66	0.40
Business Studies	0.39	0.09	0.86	0.05
Sport Science	0.39	0.19	0.90	0.03
Mathematics Applied	0.46	0.15	0.84	0.05
Environmental Science	0.57	0.22	0.88	0.05
English Communications	0.43	0.16	0.82	0.02
Modern World History	0.63	0.17	0.87	0.04
Geography	0.55	0.23	0.86	0.06
Australian Studies	0.25	0.55	0.88	0.14
Ancient Civilisations	0.43	0.25	0.86	0.07
Legal Studies	0.58	0.26	0.89	0.03
Music (Solo Performance)	0.96	0.02	0.61	0.20
Australia in Asia and the Pacific	0.27	0.26	0.88	0.05
Religion and Philosophy	0.78	0.25	0.80	0.15
Theatre Performance	0.53	0.26	0.66	0.28
English Writing	0.52	0.13	0.78	0.06
Drama	0.33	0.17	0.81	0.12
Japanese	0.78	0.27	0.61	0.45
French	1.13	0.30	0.71	0.21
Biology	0.93	0.17	0.89	0.02
Studies of Religion	0.80	0.05	0.82	0.07
Art, Craft & Design - Production	0.49	0.06	0.57	0.08
German	1.06	0.54	0.47	0.72
Performance Stage 5	0.63	0.04	0.69	0.05
Drama Stage 5	0.35	0.25	0.81	0.02
English Studies	0.74	0.19	0.82	0.06
Art Studio Practice	0.57	0.14	0.68	0.09
Art Production	0.46	0.16	0.57	0.20
Health Studies	0.20	0.17	0.89	0.05
Psychology	0.57	0.19	0.87	0.05
Sociology	0.45	0.27	0.87	0.04
Food and Nutrition	0.33	0.25	0.87	0.05
Art Appreciation	0.80	0.30	0.67	0.30
Art, Craft & Design - Appreciation	0.90	0.00	0.78	0.12
Food Studies	0.24	0.45	0.86	0.10
Dance Choreography & Performance	0.35	0.27	0.60	0.34
Dance Stage 5	0.14	0.01	0.83	0.03
Choreography and Dance Performance	0.30	0.12	0.43	0.29

Table 8.9: Variations in cohort data for subjects 2005-2013: average overall achievement, correlation between awards and overall achievement (sorted by proportion female)

Subject	Study.count.Average	Study.count.Range	Prop.Female.Average	Prop.Female.Range	Age.Average	Age.Range
Information Systems	218.33	137	0.16	0.06	17.69	0.43
Dance Stage 5	74.50	13	0.89	0.07	17.60	0.26
Information Technology and Systems	103.60	41	0.14	0.03	17.66	0.21
Chinese (Specialist Level)	69.11	60	0.39	0.19	18.12	0.79
Health Studies	996.44	193	0.71	0.07	17.63	0.58
Computer Graphics and Design Stage 6	257.67	44	0.11	0.01	17.78	0.57
Food Studies	88.33	52	0.81	0.18	17.80	0.16
Australian Studies	37.38	76	0.59	0.54	18.81	4.61
Audio Design	132.22	66	0.14	0.13	17.76	0.43
English as a Second Language	171.89	70	0.45	0.14	18.50	1.05
Australia in Asia and the Pacific	368.67	359	0.62	0.10	17.71	0.51
Computer Graphics and Design	277.20	42	0.16	0.06	17.69	0.30
Choreography and Dance Performance	80.33	44	0.90	0.01	17.45	0.19
Housing and Design	210.00	229	0.49	0.23	17.67	0.50
Food and Nutrition	279.40	336	0.78	0.11	17.77	0.19
Drama	252.57	44	0.65	0.11	17.07	0.15
Sports Science	548.00	85	0.46	0.08	17.56	0.07
Drama Stage 5	307.50	3	0.69	0.01	17.77	1.42
Dance Choreography & Performance	83.75	31	0.83	0.06	17.50	0.12
Outdoor Leadership	126.83	124	0.34	0.15	17.91	0.04
Science of Natural Resources	45.00	34	0.38	0.10	17.64	0.50
Sport Science	496.75	114	0.50	0.06	17.63	0.51
Business Studies	313.33	311	0.50	0.06	17.68	0.19
Media Production	151.57	124	0.44	0.18	17.62	0.28
English Communications	2007.00	309	0.55	0.04	17.35	0.74
Ancient Civilisations	311.00	126	0.60	0.13	17.75	0.49
Sociology	542.44	169	0.78	0.04	17.70	0.44
Mathematics Applied	1599.33	476	0.50	0.03	17.49	0.65
Art Production	672.14	169	0.71	0.06	17.64	0.42
Art, Craft & Design - Production	534.50	91	0.68	0.02	18.05	0.23
Computer Science	158.78	35	0.05	0.03	17.54	0.42
Accounting	332.89	153	0.45	0.09	17.84	0.51
English Writing	708.11	256	0.64	0.05	17.78	0.62
Theatre Performance	185.71	37	0.63	0.09	17.96	0.31
Geography	236.00	116	0.56	0.12	17.84	0.26
Design Graphics Stage 6	40.50	36	0.19	0.07	17.73	0.50
Environmental Science	235.78	81	0.54	0.09	17.69	0.41
Art Studio Practice	132.80	137	0.70	0.07	17.98	0.04
Psychology	826.78	138	0.74	0.06	17.80	0.14
Legal Studies	479.44	217	0.60	0.11	17.83	0.27
Modern World History	443.67	185	0.56	0.06	17.80	0.59
Performance Stage 5	188.50	7	0.68	0.01	17.98	0.02
Music	135.56	56	0.49	0.23	17.48	0.62
Physical Sciences	1329.56	154	0.44	0.05	17.26	1.35
English Studies	662.78	138	0.70	0.11	17.39	0.89
Religion and Philosophy	433.67	470	0.63	0.16	17.83	0.19
Japanese	111.67	62	0.66	0.11	17.62	1.46
Student Directed Inquiry	61.75	72	0.50	0.33	17.89	0.13
Art Appreciation	56.57	44	0.80	0.09	18.01	0.86
Studies of Religion	335.67	97	0.67	0.02	17.66	0.09
Economics	254.56	113	0.38	0.17	17.84	0.24
Music Performance	84.67	16	0.47	0.03	17.81	0.14
Art, Craft & Design - Appreciation	34.50	7	0.80	0.02	17.99	0.09
Biology	784.56	154	0.67	0.04	17.86	0.30
Music (Solo Performance)	63.00	28	0.61	0.12	17.76	0.05
Mathematics Methods	793.44	126	0.41	0.07	17.43	0.88
German	43.11	39	0.68	0.32	18.79	8.88
French	118.89	58	0.66	0.11	20.69	11.95
Chemistry	616.67	157	0.47	0.07	17.97	0.16
Physics	360.00	73	0.21	0.09	17.95	0.18
Mathematics Specialised	234.78	61	0.29	0.08	17.98	0.27

Table 8.10: Variations in cohort data for subjects 2005-2013: counts, proportion female, age (sorted by average overall achievement)

Subject	Rel.Prop.ATAR.Average	Rel.Prop.ATAR.Range	Prop.One.result.Average	Prop.One.result.Range
Information Systems	1.29	0.43	0.13	0.09
Dance Stage 5	0.78	0.11	0.15	0.00
Information Technology and Systems	1.23	0.41	0.09	0.08
Chinese (Specialist Level)	0.87	0.56	0.11	0.11
Health Studies	1.30	0.15	0.10	0.02
Computer Graphics and Design Stage 6	1.22	0.25	0.14	0.04
Food Studies	1.74	0.78	0.09	0.14
Australian Studies	1.70	0.87	0.08	0.22
Audio Design	1.16	0.41	0.15	0.10
English as a Second Language	0.95	0.38	0.14	0.11
Australia in Asia and the Pacific	1.47	0.77	0.08	0.10
Computer Graphics and Design	1.21	0.19	0.10	0.04
Choreography and Dance Performance	1.32	0.18	0.11	0.03
Housing and Design	1.40	0.75	0.09	0.07
Food and Nutrition	1.82	0.28	0.05	0.03
Drama	0.25	0.17	0.11	0.05
Sports Science	1.38	0.05	0.07	0.03
Drama Stage 5	0.19	0.03	0.11	0.02
Dance Choreography & Performance	1.25	0.19	0.09	0.06
Outdoor Leadership	1.93	0.42	0.07	0.04
Science of Natural Resources	1.31	0.44	0.09	0.07
Sport Science	1.43	0.16	0.06	0.03
Business Studies	1.57	0.45	0.08	0.07
Media Production	1.54	0.61	0.04	0.04
English Communications	0.65	0.31	0.07	0.03
Ancient Civilisations	1.48	0.24	0.07	0.05
Sociology	1.70	0.34	0.03	0.03
Mathematics Applied	0.92	0.27	0.05	0.02
Art Production	1.22	0.47	0.11	0.06
Art, Craft & Design - Production	1.50	0.04	0.10	0.02
Computer Science	1.12	0.49	0.09	0.05
Accounting	1.61	0.37	0.05	0.04
English Writing	1.58	0.18	0.08	0.04
Theatre Performance	2.27	0.29	0.07	0.08
Geography	2.03	0.25	0.04	0.03
Design Graphics Stage 6	1.18	0.39	0.08	0.17
Environmental Science	1.70	0.59	0.04	0.04
Art Studio Practice	2.30	0.30	0.05	0.05
Psychology	2.02	0.32	0.03	0.02
Legal Studies	1.94	0.30	0.04	0.04
Modern World History	1.70	0.29	0.05	0.04
Performance Stage 5	2.37	0.03	0.07	0.02
Music	0.83	0.58	0.04	0.07
Physical Sciences	0.24	0.07	0.04	0.03
English Studies	0.83	0.16	0.04	0.04
Religion and Philosophy	2.31	0.56	0.02	0.02
Japanese	0.84	0.49	0.08	0.10
Student Directed Inquiry	2.39	0.77	0.02	0.04
Art Appreciation	2.37	0.67	0.01	0.04
Studies of Religion	1.82	0.08	0.01	0.01
Economics	2.27	0.41	0.01	0.03
Music Performance	2.37	0.28	0.04	0.05
Art, Craft & Design - Appreciation	2.28	0.39	0.00	0.00
Biology	2.12	0.28	0.01	0.01
Music (Solo Performance)	2.25	0.39	0.01	0.03
Mathematics Methods	0.73	0.23	0.01	0.01
German	0.91	0.87	0.06	0.14
French	0.76	0.55	0.12	0.17
Chemistry	2.72	0.30	0.00	0.01
Physics	2.60	0.20	0.00	0.01
Mathematics Specialised	2.82	0.30	0.00	0.01

Table 8.11: Variations in cohort data for subjects 2005-2013: relative proportion with an ATAR, proportion with only one result (sorted by average overall achievement)

Subject	MeanPoly.Average	MeanPoly.Range	Spearman.cor.Average	Spearman.cor.Range
Information Systems	0.12	0.35	0.80	0.08
Dance Stage 5	0.14	0.01	0.83	0.03
Information Technology and Systems	0.15	0.24	0.78	0.14
Chinese (Specialist Level)	0.18	0.24	0.68	0.21
Health Studies	0.20	0.17	0.89	0.05
Computer Graphics and Design Stage 6	0.22	0.13	0.78	0.11
Food Studies	0.24	0.45	0.86	0.10
Australian Studies	0.25	0.55	0.88	0.14
Audio Design	0.27	0.25	0.79	0.23
English as a Second Language	0.27	0.34	0.77	0.15
Australia in Asia and the Pacific	0.27	0.26	0.88	0.05
Computer Graphics and Design	0.30	0.14	0.81	0.13
Choreography and Dance Performance	0.30	0.12	0.43	0.29
Housing and Design	0.31	0.22	0.80	0.08
Food and Nutrition	0.33	0.25	0.87	0.05
Drama	0.33	0.17	0.81	0.12
Sports Science	0.33	0.23	0.91	0.05
Drama Stage 5	0.35	0.25	0.81	0.02
Dance Choreography & Performance	0.35	0.27	0.60	0.34
Outdoor Leadership	0.36	0.23	0.83	0.18
Science of Natural Resources	0.38	0.41	0.87	0.12
Sport Science	0.39	0.19	0.90	0.03
Business Studies	0.39	0.09	0.86	0.05
Media Production	0.41	0.32	0.71	0.18
English Communications	0.43	0.16	0.82	0.02
Ancient Civilisations	0.43	0.25	0.86	0.07
Sociology	0.45	0.27	0.87	0.04
Mathematics Applied	0.46	0.15	0.84	0.05
Art Production	0.46	0.16	0.57	0.20
Art, Craft & Design - Production	0.49	0.06	0.57	0.08
Computer Science	0.51	0.33	0.88	0.07
Accounting	0.51	0.21	0.86	0.10
English Writing	0.52	0.13	0.78	0.06
Theatre Performance	0.53	0.26	0.66	0.28
Geography	0.55	0.23	0.86	0.06
Design Graphics Stage 6	0.56	0.17	0.75	0.27
Environmental Science	0.57	0.22	0.88	0.05
Art Studio Practice	0.57	0.14	0.68	0.09
Psychology	0.57	0.19	0.87	0.05
Legal Studies	0.58	0.26	0.89	0.03
Modern World History	0.63	0.17	0.87	0.04
Performance Stage 5	0.63	0.04	0.69	0.05
Music	0.67	0.23	0.75	0.15
Physical Sciences	0.69	0.21	0.89	0.02
English Studies	0.74	0.19	0.82	0.06
Religion and Philosophy	0.78	0.25	0.80	0.15
Japanese	0.78	0.27	0.61	0.45
Student Directed Inquiry	0.79	0.34	0.66	0.40
Art Appreciation	0.80	0.30	0.67	0.30
Studies of Religion	0.80	0.05	0.82	0.07
Economics	0.87	0.39	0.88	0.05
Music Performance	0.90	0.18	0.53	0.12
Art, Craft & Design - Appreciation	0.90	0.00	0.78	0.12
Biology	0.93	0.17	0.89	0.02
Music (Solo Performance)	0.96	0.02	0.61	0.20
Mathematics Methods	1.04	0.15	0.76	0.09
German	1.06	0.54	0.47	0.72
French	1.13	0.30	0.71	0.21
Chemistry	1.40	0.09	0.85	0.06
Physics	1.43	0.19	0.86	0.05
Mathematics Specialised	1.72	0.19	0.66	0.20

Table 8.12: Variations in cohort data for subjects 2005-2013: average overall achievement, correlation between awards and overall achievement (sorted by average overall achievement)