

Random sampling project

2010 Report on random sampling of assessment in Authority subjects

Contents

1 Summary	1
1.1 Findings	1
1.2 Conclusions	2
1.3 Recommendations	2
2 Detailed report	2
2.1 Background	2
2.1.1 Purpose	2
2.1.2 Contribution to senior moderation	3
2.2 Project design	4
2.2.1 Sampling procedure	4
2.2.2 Random sampling review panel procedures	6
2.2.3 Analysis of results	6
2.2.4 Review by state review panellists/Senior Education Officers	6
2.3 Trial for managing the random sampling process	7
2.3.1 Background	7
2.3.2 Arrangements for the Modern History trial	7
2.3.3 Result of the trial	7
2.3.4 Recommendations	8
2.3.5 Conclusion	8
2.4 Findings	8
2.4.1 Overall differences	8
2.4.2 Subject analysis	10
2.4.3 District analysis	11
2.4.4 Serious disagreement	12
2.4.5 Reviews forwarded to the office of the QSA	13
2.4.6 Feedback on elements of school submissions	14
Appendix A: Form RS1	17
Appendix B: Form RS2	18
Appendix C: All subjects reviewed by year	20

List of tables

Table 1: Requested and received submissions and folios for the selected subjects	5
Table 2: Number of submissions requested from schools	5
Table 3: Comparison of absolute mean rung differences by subject	11
Table 4: Comparison of random sampling and home district mean differences	12
Table 5: Number of cases of rung differences where there has been a difference in level of achievement (LOA) awarded by random sampling panels	13
Table 6: Summary of SRP/SEO review (number of folios	13
Table 7: Subject submissions selected for review by SRPs/SEOs (Number of folios)	14
Table 8: Responses to statements about the submissions	14

List of figures

Figure 1: Number of folios sampled for review by random sampling review panels by year	5
Figure 2: Distribution of rung differences for folios	8
Figure 3: Distribution of level of achievement differences for folios	9
Figure 4: Comparison of percentage placed in same level of achievement	9
Figure 5: Comparison of rung differences across years	10
Figure 6: Absolute mean rung differences by district of the random sampling review panel	11
Figure 7: Absolute mean rung differences by district from which schools originate	12

1. Summary

Random sampling of school judgments of student achievement in Authority subjects (the random sampling project) is one of the Queensland Studies Authority's (QSA) quality-assurance procedures for senior certification. It has been conducted annually since 1994.

The principal purpose is to evaluate the quality of school-based assessment programs and the comparability of teacher judgments of student achievement in Authority subjects across the state after Senior Education Profiles (SEPs), including Queensland Certificates of Education (QCEs) and Senior Statements, have been issued. The key question for the random sampling project is therefore:

How consistently do teachers apply statewide standards in determining students' levels of achievement in Authority subjects?

The focus for this 2010 report was students completing Year 12 in 2009. The approach was similar to that of previous years; however, for the first time in three years, small and intermediate groups (fewer than 14 students) were included in the 2010 sample.

For selected Authority subjects, a random sample of schools submitted the exit folios of a stratified random sample of seven students. Where the subjects selected had either 13 districts or a number of combined districts, review panellists for that subject from a panel in another district reviewed each school submission of folios.

Panellists were asked to allocate a Form R6 rung placement to each sample folio and to comment on each submission.

1.1 Findings

- 2774 folios were reviewed from 416 school submissions involving a total of 230 schools across 22 subjects.
- Overall, there was substantial agreement between panels and schools: 84% of the folios were placed in the same level of achievement (LOA) by both the random sampling panel and the school; 83% differed by no more than one-third of a level of achievement (3 rungs or fewer).
- At 84%, the percentage agreement within a level of achievement was slightly below the range recorded for other years.
- The greatest variances were recorded for Drama, Graphics and Science21. In 2010, a smaller percentage of subjects had a higher than average variance than in 2009. Home Economics, Information Processing and Technology, Information Technology Systems and Modern History were above the average variance for agreement for the second consecutive year.
- There were no districts or schools within districts where only small differences between school and panel judgments were evident. For more than half of the districts, large differences were found across the subjects sampled.
- Serious disagreement (defined as eight or more rung differences, with a level of achievement difference) was recorded for 5.2% of folios, a figure that is consistent with previous years, though slightly less than in 2009.
- Based on the level of disagreement recorded by random sampling panels, 20 submissions were requested for further review by state review panellists (SRPs) and Senior Education Officers (SEOs) from the QSA. Subjects where a further review was requested were Agricultural Science, Chemistry, Dance, Drama, Geography, Graphics, Information Processing and Technology, Information Technology Systems, Mathematics C, Modern History, Physics, Science21, Study of Religion, and Technology Studies.
- Following the review by SRPs and SEOs, the number of folios with rung differences of three or more fell most for Graphics, Information Technology Systems, Mathematics C and Physics, with reviewers more likely to have agreed with the schools' placements. Conversely, there was an increase in the number of folios with a rung difference of three or more for Chemistry, Dance, Drama and Geography

after the additional review of these submissions. Further follow-up was recommended for some subject areas.

- Random sampling review panels generally found that the assessment packages provided broad course coverage (87%) but they were less likely to agree with grading. In particular, a number of submissions for Dance, Drama, Graphics and Study of Religion were identified as having high percentages of disagreement about the compatibility of syllabus standards with the grades awarded.

1.2 Conclusions

- The random sampling project supports the view that the school-based assessment and moderation process for Authority subjects continues to be an effective quality-assurance process, valued by schools and panels.
- The sampling methodology this year resulted in both an increase in the total sample size and in the number of subjects sampled, which, at 22 subjects, is considerably larger than the number of subjects sampled in the previous six years.
- The analysis of panel responses to the five statements about significant aspects of submissions identified a higher level of disagreement with the statements about compatibility of grading with syllabus standards as well as the grading awarded — in particular for Dance, Graphics and Study of Religion.

1.3 Recommendations

- The subjects Drama and Graphics are to be part of the random sampling process for 2010–11.
- Workshops are to be organised for the subject Modern History in 2011.
- Panel training for review panellists involved in random sampling will be trialled for 2010–11.

2. Detailed report

2.1 Background

2.1.1 Purpose

Random sampling of school judgments of student achievement in Authority subjects (the random sampling project) contributes to the processes of moderation for the levels of achievement awarded on the Senior Statement. The random sampling project has been conducted each year since 1994. Its purposes are to:

- evaluate the quality of school-based assessment and the comparability of teacher judgments of student achievement in Authority subjects across the state, that is, to assess the strength of school decision-making in the system of school-based assessment for senior certification
- provide information on the quality of assessment procedures and assessment judgments in various subjects and identify schools that need further assistance
- identify, at a systemic level, any issues concerning assessment and moderation that need further investigation.

The process of reviewing student folios for the random sampling project occurs in the year after the students have left school and after they have been issued with their SEP. Therefore, the outcome does not influence the levels of achievement awarded to that cohort of students. Rather, the random sampling project checks the quality of school-based judgments after they have been made. However, feedback provided to the sampled schools can contribute to the calibration of their future judgments. As well, the findings can contribute to further improvements in moderation processes.

2.1.2 Contribution to senior moderation

Moderation is the set of processes designed to:

- support the integrity of school-based assessment in Authority subjects
- strengthen the quality of teacher judgments of student achievement in Authority subjects
- ensure a high degree of comparability in certified levels of achievement in Authority subjects
- maintain the credibility and acceptability of the SEP.

Moderation begins with the approval of work programs for Authority syllabuses. Other key processes of moderation are **monitoring**, **verification**, **confirmation** and **random sampling**.

Monitoring of Year 11 folios occurs at the end of the first half of a course, that is, at the end of February. Review panels consider evidence of the school's delivery of their courses of study and of their programs of assessment. They also consider school judgments of student achievement in Authority subjects, based on a sample of student folios from each school. Advice is given to schools early in Year 12 so that schools can be reassured about, or helped with, their delivery of approved courses of study and their standards judgments.

Verification occurs towards the end of Year 12. Schools submit sample student folios in September each year for the verification meeting in October. School submissions of a sample of student folios in each Authority subject offered by the school are sent to the relevant (usually district) review panel. These submissions consist of a sample of folios of work for students about to complete the course of study, together with the school's judgments of interim levels of achievement for those students. Panellists survey the folios for evidence to confirm the school's judgments, confer with other panellists (and in the case of different opinions, the Chair), and formulate advice to the school. If the panel cannot confirm a school's proposal, consultation between the school and the District Review Panel Chair (DRPC) takes place. Where agreement cannot be reached between the school and the DRPC on all sample folios in a submission, the complete submission is sent to the relevant state review panel for further consideration.

The role of the state review panel is to check that comparable standards are maintained in their subject across all districts. They do this by examining sample submissions from each district and validating the judgments of the district panels.

Confirmation occurs following completion of Year 12. Schools forward their exit proposals for levels of achievement to the office of the QSA immediately after the finishing day for Year 12 in November. The period between receipt of schools' proposals for exit levels of achievement and the printing of SEPs is referred to as the confirmation period. SEOs review any changes to the levels of achievement that had been agreed to at verification. Legitimate changes can occur as a result of assessment in the final term of Year 12. The confirmation phase concludes when the QSA reaches agreement with the school on its proposed results for recording on students' Senior Statements.

Random sampling focuses on student *exit* folios. This means that it occurs after the issue of SEPs. No changes in the recorded results in SEPs occur as a consequence of random sampling.

Random sampling refers to the process of sampling schools and students. However, subjects are not randomly selected, and some (smaller) subjects were not previously randomly sampled at all. Subjects typically have been selected on the basis of their size (total number of students), stage of implementation or implementation issues.

Schools are chosen randomly within each subject. In the past, to be included, the school needed to have a large group (14 or more students) in that subject. Small (nine students or fewer) and intermediate (10–13 students) groups were generally not included because most of these students' folios of work were assumed to have already been reviewed by their district or state panel. However, this discounted the value of including such groups to allow these groups to be reviewed by other districts, which is a key aspect of the random sampling project. From 2005, some small and intermediate groups were included and the number of folios per school (submission) was reduced from nine (as in previous years) to seven to reduce the load on panels.

For each chosen group, a random sample of students is selected, stratified by levels of achievement awarded to the students. The school is asked to provide the exit folios for these students (known as the

random sampling submission) including each student's level of achievement and rung placement (recorded on the Form RS1, see Appendix A).

Random sampling submissions are allocated randomly to other districts. The other district panel is referred to as the random sampling review panel when it is reviewing random sampling submissions.

District review panels (DRPs), acting as random sampling review panels, review random sampling submissions in February at the same time that panels meet for Year 11 monitoring.

2.2 Project design

2.2.1 Sampling procedure

This random sampling project focused on the Year 12 cohort of 2009.

Subjects were selected deliberately to include those with large statewide enrolments as well as other subjects of interest, such as those that had not previously been sampled or had not been sampled in recent years. Strategic interests included:

- subjects not selected for two or more years (e.g. Business Organisation Management, Graphics, Health Education, Technology Studies)
- subjects with Year 12 for the second time on a new or revised syllabus (e.g. Information Technology Systems)
- subjects that have had high "inter-rater" variation in previous random sampling reviews (e.g. Information Processing and Technology, and Modern History).

For subjects with 13 QSA district panels, schools were selected randomly within each of the districts across the state under the following constraints (where possible):

- no more than three subject groups from one school
- a maximum of 25 school subject groups for any one subject.

In 2005 the number of folios per school (submission) was reduced from nine (as in previous years) to seven to differentiate the random sampling process from monitoring and verification. This is the sixth year in which seven submissions have been sampled.

A stratified random sample of student folios was selected within each school subject group (submission) with the following specifications:

- folios are selected by the QSA, not the school
- if there are fewer than the required number of folios at any given level of achievement, folios are selected from the next level of achievement (moving towards the centre)
- if there are fewer than two SA folios, folios are selected, in turn, from HA, VHA, LA, or VLA.

The outcome of this selection process is shown in Table 1 overleaf. The final number of submissions was 416. The number of folios received was 2839 (versus the targeted 2912). Some of the requested folios were unavailable because they were required for other purposes, such as requests for verification of Senior Statements or review of Tertiary Entrance Statement results. Of the 2839 folios received, a further 65 were not reviewed by a random sampling review panel as there was insufficient evidence or information to make a judgment.

Subjects were distributed across 214 panels.

Table 1: Requested and received submissions and folios for the selected subjects

SUBJECT	Number of schools	Number of folios requested	Number of folios recieved	Number of folios reviewed
Agricultural Science	8	56	56	55
Ancient History	25	175	168	161
Business Organisation and Management	12	84	84	83
Chemistry	25	175	173	172
Dance	14	98	98	98
Drama	25	175	166	148
Geography	25	175	175	175
Graphics	25	175	174	166
Health Education	12	84	83	81
Home Economics	25	175	168	168
Hospitality Studies	8	56	56	56
Information Processing and Technology	24	168	159	150
Information Technology Systems	11	77	77	77
Legal Studies	25	175	174	174
Marine Studies	12	84	84	82
Mathematics C	25	175	175	173
Modern History	25	175	154	146
Physical Education	25	175	174	173
Physics	25	175	175	175
Science21	8	56	56	54
Study Of Religion	12	84	84	84
Technology Studies	20	140	126	123
Total	416	2912	2839	2774

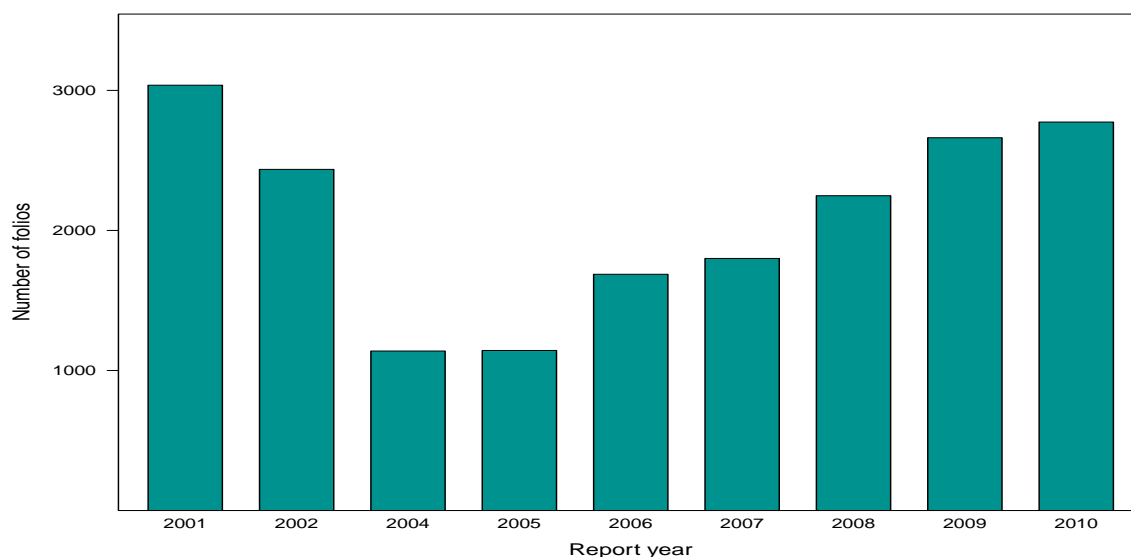
A full list of all subjects sampled for the past 10 years is contained in Appendix C. Most schools were required to provide only one submission. No school was requested to provide more than three submissions (see Table 2).

Table 2: Number of submissions requested from schools

Number of submissions	Number of schools
1	102
2	70
3	58
Total schools	230

Following the recommendation in the 2005 random sampling report, the sample size was further increased this year (see Figure 1).

Figure 1: Number of folios sampled for review by random sampling review panels by year



Note: 2003 data is incomplete due to work bans in certain panel districts and therefore has not been included.

2.2.2 Random sampling review panel procedures

Members of the district review panels (acting as random sampling review panels) examined each of the folios in the school submissions allocated to their panel and decided a specific rung placement (ten rungs within each level of achievement). Two panellists, selected by the DRPC, reviewed each submission independently. Following discussion between the panellists to reach consensus, and usually in consultation with the DRPC, the panel judgments were recorded.

District review panellists were provided with advice about how to ensure that two independent reviews of the two submissions allocated to their district took place (see below).

Advice to district panellists

Panellists will need to exchange submissions so that both panellists can consider each submission. There are two options available for this exchange. The DRPC should discuss the alternatives with the two chosen panellists, and inform the district coordinator of the method to be used by the panel.

Option 1

Panellists can arrange to meet briefly and exchange submissions. The second submission is then pre-reviewed in the panellist's own time. After the second pre-reviewing, panellists meet again for approximately one hour to reach consensus on the two submissions. This meeting could be held before, after or on the day of the monitoring meeting. If the meeting is on the day of monitoring, then it could be at the time set aside during the meeting, or after all monitoring submissions have been completed, or before the monitoring meeting.

This option is appropriate if the panellists live or work near each other. Submissions are not to be posted between panellists.

Option 2

After undertaking independent pre-reviewing of one submission, panellists meet, and at this meeting exchange and independently pre-review the second submission and reach a consensus on both submissions. This meeting could be held before, after or on the day of the monitoring meeting. If the meeting is on the day of monitoring, then it could be at a time set aside during the meeting, or after all monitoring submissions have been completed, or before the monitoring meeting.

If the panellists do not live or work near each other it would be most appropriate for them to meet to pre-review the second submission and reach consensus on the day of the monitoring meeting.

Where it is proposed that the consensus meeting (at which the second random sampling review takes place) be held on the day of monitoring, the DRPC should determine the viability of such a meeting after considering the monitoring workload of the panel as a whole.

Panellists were asked to complete a summary form rating each submission (of seven folios) on five characteristics of assessment and application of standards.

2.2.3 Analysis of results

Rung-achievement placements allocated by schools and random sampling review panels were converted to a numerical scale of 1–50. The rung or level difference was calculated by subtracting the school's exit rung (or level) placement from that of the panel. Negative differences therefore mean that the panel judged the schools' placement to be lower.

2.2.4 Review by state review panellists/Senior Education Officers

Folios with a significant difference (defined as eight or more rungs difference) between school judgment and panel judgment were identified and the following criteria were used to select submissions for further review by SRPs and SEOs:

- highest proportion of subjects with eight or more rungs difference
- three or more students identified as having a change to rung level and/or level achievement
- the Random Sampling Review Panel identified serious concerns with:
 - the assessment package

- meeting syllabus requirements
- insufficient evidence available to support overall standards awarded.

State review panellists independently reviewed these submissions and, after comparing the school judgment with the panel judgment, determined the appropriate action to be taken. Follow-up involved one or more of the following:

- schools contacted and provided with support to assist in areas where difficulties were identified
- home DRPCs contacted to discuss identified difficulties
- random sampling DRPCs contacted to discuss aspects of decisions.

In addition, the information gained from this review provided input to the format of future professional development sessions in specific subject areas.

2.3 Trial for managing the random sampling process

2.3.1 Background

For the 2010 random sampling project, the office of the QSA trialled a new approach for managing the arrangements as a way of clarifying the purpose and process of random sampling. The trial was undertaken for the subject Modern History and was to separate, to some extent, the random sampling and monitoring processes. Twenty-five submissions, 13 DRPCs and review panellists for each district were involved.

2.3.2 Arrangements for the Modern History trial

- The DRPC and one other member of the panel were to be the reviewing panellists. In regional districts, the panellist needed to be from a local school.
- The consensus meeting was to be held after the monitoring meeting day, but before 3 March 2010.
- Two hours was allocated for the consensus meeting.
- The meeting could be held at either the district office or at the DRPC's school.
- The Modern History random sampling submissions were sent prior to the monitoring meeting.
- The relevant random sampling package was sent prior to the beginning of the process.

2.3.3 Result of the trial

Eight of the 13 DRPCs provided feedback on the trial.

Positive feedback included:

- ability to concentrate on submissions for random sampling, rather than both random sampling and monitoring submissions
- good for networking with another panellist outside panel time
- allowed for substantive conversation between a panellist and DRPC around two submissions from a different district
- freed up time during monitoring meeting and as a result provided appropriate time to complete random sampling at second meeting.

Negative feedback included:

- difficult to find time to meet for random sampling outside of normal time allocated for monitoring
- two hours was insufficient time to review the second submission and have two separate conferences.

2.3.4 Recommendations

- There is inconclusive evidence from the trial for the subject Modern History to establish if it was effective.
- A form of panel training for random sampling is to be implemented.

2.3.5 Conclusion

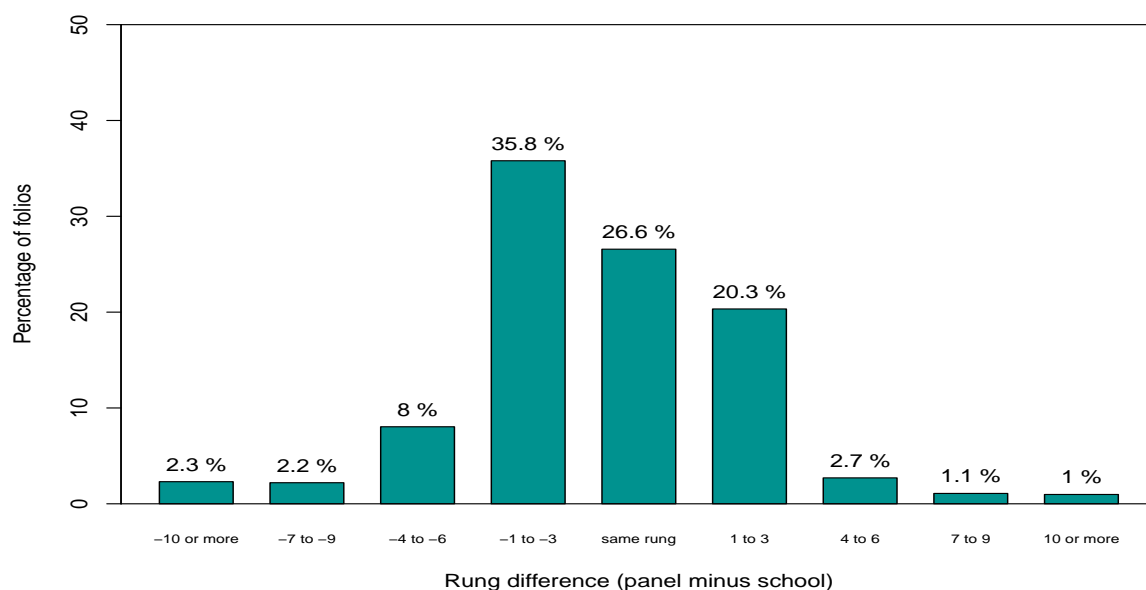
The recommendation to the Assessment and Moderation Committee on 24 May 2010 to separate random sampling from moderation was not endorsed.

2.4 Findings

2.4.1 Overall differences

The random sampling panels were asked to comment on the standards evident within each school submission as applied by schools. As shown in Figure 2, 26.6% of folios had no rung difference and 82.7% were found to have been appropriately placed within **three rungs** on the Form R6 by their schools. As noted in previous studies, there is a greater tendency for random sampling review panels to rate folios lower than the schools.

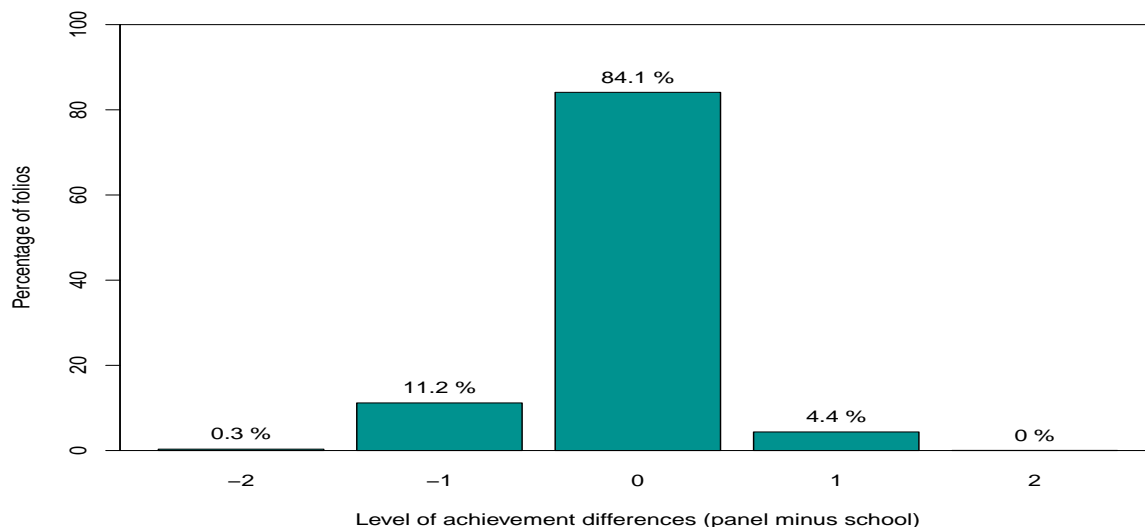
Figure 2: Distribution of rung differences for folios



Note: due to rounding the total of this graph may not equal 100%.

There was a high level of agreement between the random sampling review panels and the schools about **levels of achievement** awarded to folios. Figure 3 indicates that reviewers found that 84.1% of the levels of achievement awarded by the school were able to be supported. While a number of folios (11.5%) were judged to have been placed 1–2 levels of achievement too high at exit, 4.4% of folios were found to have been awarded 1–2 levels too low.

Figure 3: Distribution of level of achievement differences for folios



Note: due to rounding the total of this graph may not equal 100%.

Figure 4 shows that there has been some variation over time (between 84% and 93%) in the percentage of folios considered by random sampling review panels to have been placed appropriately in terms of level of achievement overall. While the current result is fairly consistent with previous years, 84% in 2010 (as it also was in 2009) is the lowest level of agreement since the first year of random sampling in 1994.

Figure 4: Comparison of percentage placed in same level of achievement

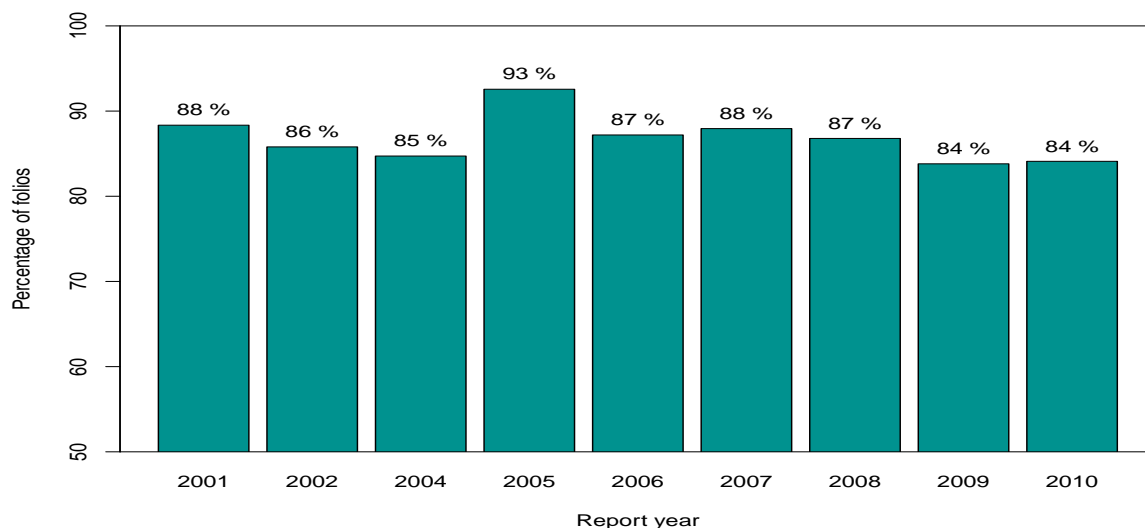
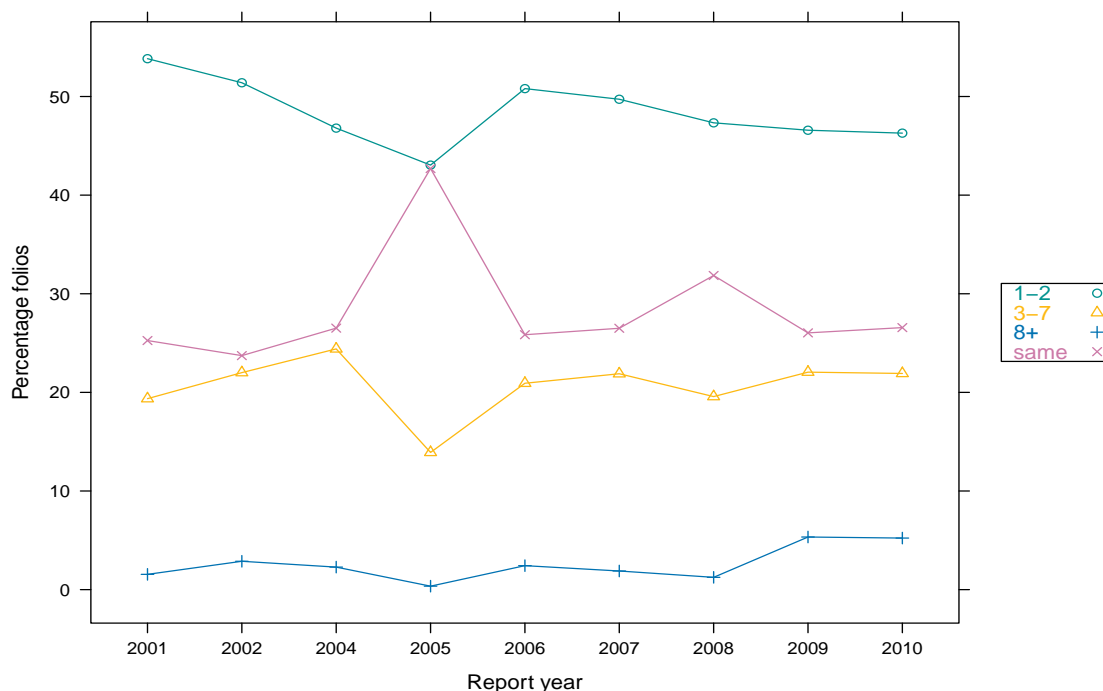


Figure 5 shows the historical comparisons for rung differences from 2001 to the present. While the 2010 results are generally consistent with past results, the percentage of 8+ rung placements in this sample remains higher than all other years, except 2009.

Figure 5: Comparison of rung differences across years



Note: 2003 data is incomplete due to work bans in certain panel districts and therefore has not been included.

2.4.2 Subject analysis

Table 3 summarises the absolute mean rung differences by subject, over time. The absolute mean does not take into consideration the direction of difference and therefore provides an overall indication of the degree of consistency in judgment based on rungs. The overall mean and standard deviation is for all subjects sampled in that year (not just those appearing in the table).

The table reinforces the trend noted in Figure 5, in which there has been a lower level of agreement of standards as reflected by a higher absolute mean value than for any year, except 2009. It also highlights the fact that there were eight common subjects in the 2009 and 2010 samples (Geography, Home Economics, Information Processing and Technology, Information Technology Systems, Legal Studies, Mathematics C, Modern History and Study of Religion).

The 2010 random sampling for Home Economics, Information Processing and Technology, Information Technology Systems and Modern History showed a higher than average absolute mean difference for the second consecutive year.

Drama, Graphics and Science21 have the largest average mean differences, while Dance, Home Economics, Information Processing and Technology, Information Technology Systems, Modern History and Technology Studies also have larger than average mean differences.

The absolute mean for Dance, Drama, Geography, Graphics and Hospitality is higher than for any previous year, whereas the results for Modern History and Technology Studies have been above the absolute mean in each year of sampling.

Business Organisation Management, Legal Studies and Physical Education have the smallest absolute mean rung differences.

Table 3: Comparison of absolute mean rung differences by subject

Subject name	2001	2002	2004	2005	2006	2007	2008	2009	2010
Agricultural Science							1.93		1.98
Ancient History		1.87				2.16	1.45		1.76
Business Organisation and Management	1.87	2.71			1.43				1.13
Chemistry	1.42						1.2		1.83
Dance				.46			.89		2.52
Drama	1.35	1.3	2.09		1.19		1.45		2.78
Geography		1.97	1.36		1.51			1.52	2.04
Graphics	2.18	2.13	1.75		2.23	1.98			3.08
Health Education	2	2.37		1.18					1.86
Home Economics	2.12			1.28			1.52	2.39	2.31
Hospitality Studies				1.29			.98		1.84
Information Processing and Technology	1.98	1.77				1.69		2.69	2.17
Information Technology Systems				2.53	2.87	1.37		2.3	2.71
Legal Studies	1.7		1.74			1.71		2.1	1.52
Marine Studies		1.75					1.45		1.68
Mathematics C	1.58		1.96			1.62		2.42	1.88
Modern History	1.75					1.75	2.17	2.71	2.34
Physical Education	1.87	1.99	1.75			1.18	1.11		1.43
Physics	1.03						2.16		1.97
Science21									3.02
Study of Religion		1.75	2.11		1.57			2.43	2.3
Technology Studies	2.81	2.67			2.49				2.31
Standard deviation	1.74	2.32	2.09	1.44	2.01	1.89	1.79	2.72	2.71
Overall mean(abs)	1.62	1.9	1.88	1.15	1.78	1.7	1.55	2.11	2.09

Note: 2003 data is incomplete due to work bans in certain panel districts and therefore has not been included.

2.4.3 District analysis

Figure 6 compares absolute mean rung differences for random sampling panels in each district. Panels in the Rockhampton and Townsville districts had the highest level of disagreement with the decisions made by schools about student placement across all subjects reviewed by the district.

Figure 6: Absolute mean rung differences by district of the random sampling review panel

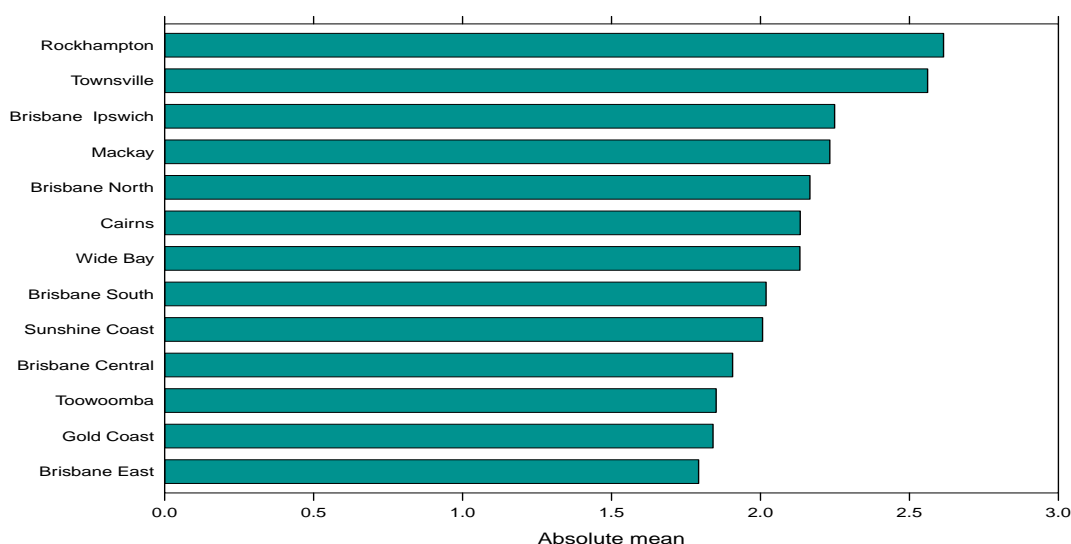


Figure 7 compares absolute mean rung differences for schools in each district. Mean rung differences across all subjects were largest for schools in the Mackay district, followed by Sunshine Coast, Cairns and Wide Bay.

Figure 7: Absolute mean rung differences by district from which schools originate

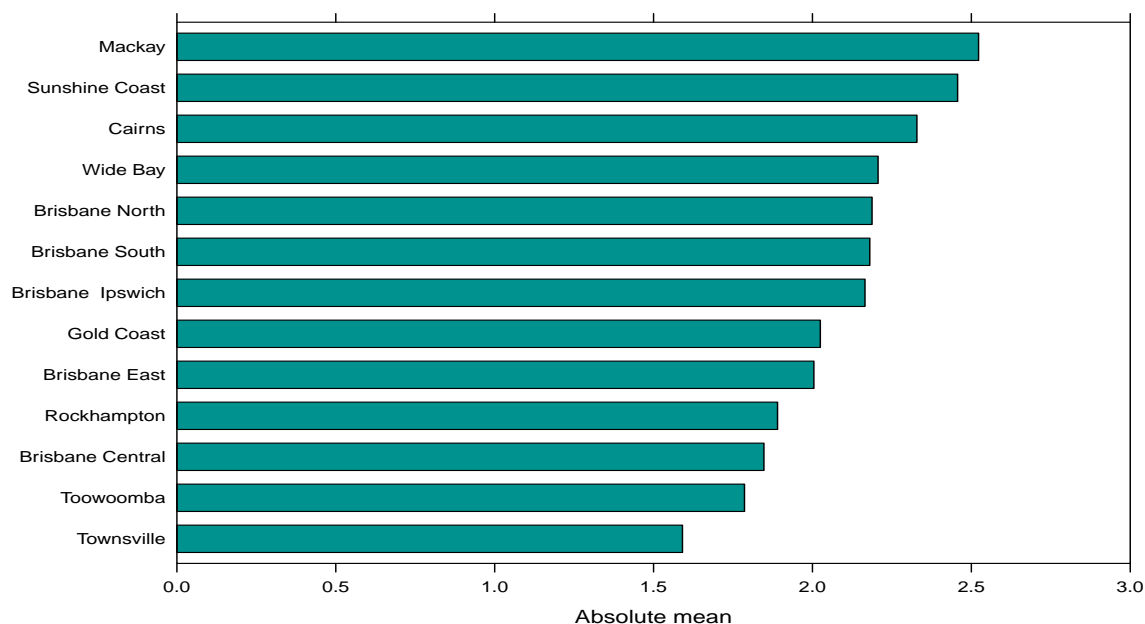


Table 4 compares absolute mean rung differences for random sampling panels in each district with those of the schools' districts. Differences have been classified as large, medium or small to facilitate this analysis (where large is equivalent to greater than 2 rungs absolute mean difference and small is equivalent to less than 1.5 rungs absolute mean difference). It should be noted that sample sizes received by district panels ranged from 168 to 292 folios and each district reviewed a different range of subjects. The most evident patterns to emerge from this table were:

- no small differences were found by random sample district panels or in schools within districts, as was found in the 2009 report
- Brisbane East and Gold Coast districts found medium differences in the folios they reviewed while other panels found large differences for schools in their districts
- while Townsville district panels had one of the highest levels of disagreement with folios they reviewed, schools in the Townsville district had the lowest level of disagreement when reviewed by panels from other districts
- more than half of the districts showed large differences for panels and schools.

Table 4: Comparison of random sampling and home district mean differences

School's district	Random sampling district panels		
	Large	Medium	Small
Large	Brisbane Ipswich Brisbane North Brisbane South Cairns Mackay Sunshine Coast Wide Bay	Brisbane East Gold Coast	
Medium	Rockhampton Townsville	Brisbane Central Toowoomba	
Small			

2.4.4 Serious disagreement

Earlier random sample reports quote figures for the level of serious disagreement over the exit level of achievement awarded to folios. Table 5 summarises the rung differences where there has been a level

of achievement difference. Over the years, the percentage of folios considered to have serious disagreement has ranged from 1% to 3%. This remains consistent in 2010.

Table 5: Number of cases of rung differences where there has been a difference in level of achievement (LOA) awarded by random sampling panels (refer to Appendix C for all years)

	2001	2002	2004	2005	2006	2007	2008	2009	2010
Sample size	3037	2436	1139	1143	1687	1800	2248	2662	2774
Different LOA with 1–2 rungs	176	136	55	46	79	80	114	150	146
	6%	6%	5%	4%	5%	4%	5%	6%	5%
3–7 rungs	141	158	94	36	99	105	160	191	209
	5%	6%	8%	3%	6%	6%	7%	7%	8%
8+ rungs	37	42	24	3	36	32	23	68	64
	1%	2%	2%	0%	2%	2%	1%	3%	2%
Total with different LoA	354	346	174	42	216	217	297	431	441

Note: 2003 data is incomplete due to work bans in certain panel districts and therefore has not been included.

2.4.5 Reviews forwarded to the office of the QSA

Submissions were requested for further, independent review by SRPs and SEOs when two or more of the following criteria were met:

- at least one folio had eight or more rungs difference
- three or more students were identified as having a change of three or more rungs or a level of achievement difference
- the random sampling review panel identified serious concerns with:
 - the assessment package
 - meeting syllabus requirements
 - insufficient evidence available to support overall standards awarded.

Based on the above criteria, 20 submissions (containing 140 folios) were recalled for additional review. SRPs and SEOs were unaware of the placements given by either the school or random sampling review panel and were asked to provide an independent assessment of the selected folios. The results of this review are summarised in Table 6.

Table 6: Summary of SRP/SEO review (number of folios)

	After random sampling	After second review*
Rung differences (8+)	22	23
Rung differences (3–7)	84	73
Rung differences (1–2)	24	36
Different LoA	71	65
Mean rung difference	1.52	.52
Absolute mean rung difference	4.99	4.55

* Calculations are based on the differences between school and review judgments.

Following the additional review, the mean rung difference has declined. Despite an increase in the number of folios with small differences to rung placement, there is a smaller number with a level of achievement difference. This is despite the fact that there was a slight increase in the number of folios with eight or more rungs difference. Further review of additional folios is more likely to reduce the amount of disparity between the schools' and the random sampling review panels' judgments.

State review panellists provided feedback about each of the submissions they reviewed and SEOs will follow up with schools and district panel chairs as appropriate. The following broad categories of feedback were provided:

- Schools will be contacted and provided with support to assist in areas where difficulties have been identified.
- Home district review panel chairs will be contacted to discuss identified difficulties in general.
- District review panel chairs of random sampling panels will be contacted to discuss aspects of review.

Table 7 summarises the subjects reviewed by SRPs and SEOs. Information Technology Systems had the largest number of folios requiring additional review, followed by Drama, Geography, Graphics and Technology Studies. Following the additional review there were fewer folios with disagreements, indicating that the state panellists were more likely to have been in agreement with the schools' placement of students.

After the review by SRPs and SEOs, the greatest number of discrepancies remained for Drama, Geography and Information Processing and Technology.

Table 7: Subject submissions selected for review by SRPs/SEOs (Number of folios)

Subject	Submissions	Total folios	Folios with 3+ rung difference	
			Before SRP review	After SRP review
Agricultural Science	1	7	3	3
Chemistry	1	7	2	5
Dance	1	7	3	5
Drama	2	14	8	9
Geography	2	14	8	10
Graphics	2	14	7	3
Information Processing and Technology	1	7	6	6
Information Technology Systems	3	21	11	5
Mathematics C	1	7	3	1
Modern History	1	7	4	3
Physics	1	7	5	3
Science21	1	7	2	1
Study of Religion	1	7	3	3
Technology Studies	2	14	6	5
Total	20	140	71	62

2.4.6 Feedback on elements of school submissions

The random sampling panels were asked to respond, using a five-point scale, to five statements about each school's assessment packages and application of standards (see Appendix B).

Table 8 shows responses to these statements as provided on a consensus form for each submission (416 submissions). Some responses were missing from the data and have been excluded from the calculations.

Table 8: Responses to statements about the submissions (in percentages)

Elements of the school's submission	Strongly disagree	Disagree	Unsure	Agree	Strongly agree
The assessment package provides broad course coverage	*	8	5	67	20
The assessment package provides opportunities for a range of achievement	1	13	7	60	20
The assesment package allows discrimination between students' achievements	*	11	9	63	17
The grading/marking of student work is compatible with syllabus standards	3	23	13	51	9
Sufficient evidence is available to support the overall standards awarded	3	21	10	54	12

* denotes less than 1%.

When considering schools' submissions, panels were most strongly in agreement with the statement that the assessment package provides broad course coverage (87%). In addition, relatively high levels of agreement were registered for the following two statements:

- The assessment package provides opportunities for a range of achievement (78%).
- The assessment package allows discrimination between students' responses (82%).

There was less agreement with the way in which standards had been applied than with other aspects of submissions:

- The grading/marking of student work is compatible with syllabus standards (25% disagreed).
- Sufficient evidence is available to support the overall standards applied (21% disagreed).

Further analysis indicates that more than 70% of those subjects with student work which was sent for review by SRPs and SEOs were also considered (by district random sampling panellists) to have insufficient evidence to support the overall standards awarded. Of these, 100% were deemed to demonstrate grading/marking that is incompatible with syllabus standards.

The compatibility of the grading of student work with syllabus standards and concerns about the amount of evidence to support the overall standards awarded have traditionally been areas of greater disagreement and variation. While this has continued to be the case in the current random sampling process, there has been a decrease in the number of submissions for which there were concerns about the amount of supporting evidence.

While additional analysis of the data by subject area has been conducted, sample sizes of responses should be considered. Engineering Technology and Drama had fewer than six responses. In addition, there were a number of missing forms that further reduced the sample size for some subjects. Comments therefore focus predominantly on the subjects with ten or more responses.

The assessment package provides broad course coverage

As noted in Table 8, panels agreed that assessment packages provided broad course coverage in 87% of submissions. The level of agreement was even higher for the subjects Ancient History, Dance, Home Economics, Legal Studies, Physics and Study of Religion (95% to 100%). However, 33% of Agricultural Science and Drama submissions did not provide sufficiently broad course coverage. There were also some concerns about 29% of the Science21 submissions.

The assessment package provides opportunities for a range of achievement

For all subjects randomly sampled, panels agreed with this statement in 78% of cases. There was particularly high agreement for Hospitality Studies (100%), Geography (95%), Technology Studies (93%) and Dance (91%). There were relatively higher levels of disagreement for Drama (47%), Science21 (43%) and Agricultural Science (33%).

The assessment package allows discrimination between students' responses

Some 82% of random sampling panels agreed that the assessment packages allowed for discrimination between students' responses. Geography and Home Economics (both 100%) and Legal Studies (96%) rated most highly in this regard. However, panels were less able to agree on submissions for Science21 (43%) and Drama (33%).

The grading/marking of student work is compatible with syllabus standards

As noted above, this aspect had lower agreement compared to the other statements and has been the issue of greatest concern over a number of years.

For all randomly sampled subjects, 61% agreed, 25% disagreed and 14% were unsure. However there were differences by subject areas with a higher level of **agreement** for this statement in Science21 (86%) and Geography (84%). A higher level of **disagreement** was noted, in particular, for Dance (60%), Drama (54%), Study of Religion (45%) and Graphics (42%).

Sufficient evidence is available to support the overall standards awarded

Across all subjects randomly sampled, panels generally agreed that 61% of schools had provided sufficient evidence to support the overall standards awarded. Levels of agreement with this statement were generally high, especially for Science21 (100%), Health Education (91%) and Geography (89%). Some subjects had a high level of disagreement. These included Information Technology Systems (57%), Hospitality (56%), Study of Religion (45%), Graphics (42%) and Dance (40%).

A small percentage of submissions were missing some items:

- a completed work program (1%)
- a set of all assessment instruments (6%)
- all seven folios complete with responses (6%).

Missing items included assessment items from folios (24 submissions), work programs, stimulus materials and student profiles (either missing or incomplete). A small number of panellists commented that additional evidence/tasks had been supplied when they were not required. There was also acknowledgment that, in some instances, schools included letters of explanation with their submissions.

In commenting about significant positive and/or negative aspects of submissions, random sampling review panels mentioned the following positive aspects:

- *good/appropriate range of opportunities for students to achieve*
- *well presented submission*
- *appropriate judgments*
- *positive teacher feedback.*

Negative comments included:

- *concern regarding appropriateness/design/clarity/complexity/range of assessment tasks*
- *marking inconsistent with evidence/inflated/lenient/hard to justify*
- *syllabus requirements not met*
- *concerns re criteria sheets (incomplete/missing/not task-specific)*
- *insufficient evidence to support judgments/standards.*

Appendix A: Form RS1



School Recording Form

School relative achievement awarded

Form RS1

School: «schoolName»
 Subject: «subjectName1»
 District:

School code: «schoolCode»
 Subject code: 0«subject»
 Panel code:

Exit Rung	Level of Achievement
10	Very High Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	High Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Sound Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Limited Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Very Limited Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	

Instructions for preparation of Random Sampling submission:

- The submission should include:
 - The approved work program
 - A set of summative assessment instruments
- Include in the submission, all student responses used to make teacher judgments about the *Relative Achievement* of the student at *Exit*.
- If the subject requires additional information for *Verification*, such as a video or audio tape of student performance standards or the transcripts of the listening tests, these should also be included.
- Indicate on this form (Form RS1), the rung placement of the student at *Exit* by student code below. Add any comments that may assist in substantiating the *Exit Achievement*.
- Place the Form RS1 **on top** of the random sampling submission.
- Dispatch the random sampling submission directly to the QSA local district office.

Selected students:

Student A: «studentA»
 Student B: «studentB»
 Student C: «studentC»
 Student D: «studentD»
 Student E: «studentE»
 Student F: «studentF»
 Student G: «studentG»

Appendix B: Form RS2



Panellist Recording Form

Form RS2

Panel achievement awarded

School: «SchoolName» Subject: «SubjectName» District:	School code: «School» Subject: «Subject» Panel code:
---	--

Instructions:

- After making judgments about the student work in the folio, indicate the rung placement by student code that reflects the relative achievement of the student at Exit.

Selected students:

- Student A: «studentA»
 Student B: «studentB»
 Student C: «studentC»
 Student D: «studentD»
 Student E: «studentE»
 Student F: «studentF»
 Student G: «studentG»

Panel Comments:

Panelist Name: _____

Exit Rung	Level of Achievement
10	Very High Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	High Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Sound Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Limited Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	
10	Very Limited Achievement
9	
8	
7	
6	
5	
4	
3	
2	
1	

This information is collected so that the legislated functions of the QSA concerning moderation can be carried out. Personal information is not usually disclosed to anyone other than relevant QSA staff unless required or authorised by law, permitted under the Queensland Government privacy policy, or so that the legislated function can be completed.
 Queensland Studies Authority, Ground floor, 296 Ann Street Brisbane, PO Box 307 Spring Hill Qld 4004.
 Phone: (07) 3864 0296; fax: (07) 3221 2653; email: office@qsa.qld.edu.au; website: www.qsa.qld.edu.au

Please ensure only ONE choice bubble is filled in for each question. Fill each bubble DARKLY and FULLY using BLACK or BLUE pen.

Office Use			
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
School Code		Subject Code	

Please check whether the school submission is complete and it contains:

	Yes	No	
1	<input type="checkbox"/>	<input type="checkbox"/>	the work program
2	<input type="checkbox"/>	<input type="checkbox"/>	the set of all assessment instruments used for making decisions about exit LoAs
3	<input type="checkbox"/>	<input type="checkbox"/>	seven student folios complete with responses

Comment on any missing or additional material.

Office Use	
3a	<input type="checkbox"/> <input type="checkbox"/>
3b	<input type="checkbox"/> <input type="checkbox"/>

Reflect on the elements of the school submission and select the most appropriate response to the following statements.

		Strongly disagree	Disagree	Unsure	Agree	Strongly agree
4	The assessment package provides broad course coverage.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	The assessment package provides opportunities for a range of achievement.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	The assessment package allows discrimination between students' responses.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	The grading/marking of student work is compatible with syllabus standards.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	Sufficient evidence is available to support the overall standards awarded.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Comment on significant positive and/or negative aspects of the submission.

Office Use	
8a	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
8b	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
8c	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

For the questions below, reflect on the task sheets and standards schemas presented to support teacher judgments about individual students' assigned grades or standards.

		Not at all	Appropriately applied	Cannot be determined
9	Was there any evidence of non-compliance with QSA policy on late submission of student work?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10	Please comment on your observation about how the school has applied the policy on late submissions.	 		
11	Was there any evidence of non-compliance with QSA policy on non-submission of student work?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12	Please comment on your observation about how the school has applied the policy on non-submission.	 		

83549



Appendix C: All subjects reviewed by year

Subject name	Subject ID	2002	2004	2005	2006	2007	2008	2009	2010
English	1	*		*	*		*	*	
French	5						*		
Italian	8			*					
Japanese	9				*		*		
Chinese	11			*					
Ancient History	20	*				*	*		*
Modern History	21					*	*	*	*
Aboriginal and Torres Strait Islander Studies	23			*					
Geography	24	*	*		*			*	*
Economics	27	*	*			*		*	
Study Of Society	28	*							
Legal Studies	29		*			*		*	*
Mathematics A	36		*			*		*	
Mathematics B	37		*			*		*	
Mathematics C	38		*			*		*	*
Chemistry	40						*		*
Physics	41						*		*
Biology	42	*		*			*	*	
Earth Science	43			*					
Marine Studies	45	*					*		*
Science21	46								*
Agricultural Science	51						*		*
Accounting	60				*			*	
Business Organisation and Management	62	*			*				*
Business Communication and Technologies	63	*			*			*	
Information Technology Systems	65			*	*	*		*	*
Health Education	67	*		*					*
Physical Education	68	*	*			*	*		*
Home Economics	71			*			*	*	*
Hospitality Studies	72			*			*		*
Engineering Technology	74			*				*	
Graphics	76	*	*		*	*			*
Technology Studies	78	*			*				*
Visual Art	80	*	*		*		*	*	
Dance	85			*			*		*
Study Of Religion	86	*	*		*			*	*
Information Processing and Technology	87	*				*		*	*
Drama	88	*	*		*		*		*
Music	91	*		*			*	*	
Film, Television and New Media	93						*	*	

Note: 2003 data is incomplete due to work bans in certain panel districts and therefore has not been included.