

Students' grades in Rich Tasks 2004

This report provides statewide data about the grades students received in Rich Tasks in 2004. It points to some notable features of these data, and discusses some possible interpretations of them.

This report includes only aggregated statewide data. Each school with students who received a New Basics report in 2004 has received summary data about its own students' grades.

The importance of the data for the system

The Rich Tasks offer the Queensland system its only information in the compulsory years about student achievement in authentic assessment tasks, graded against statewide standards. "How are students performing in these tasks?", "How are schools making the tasks available to students?" and "How well are the tasks themselves working?" are questions that may provide both gratifying and uncomfortable answers. A system that both values its achievements and seeks to build on them needs to hear both kinds of answers.

The importance of the data for individual schools

The grades achieved in Rich Tasks by students at a particular school have meaning regardless of the grades achieved by students at other schools. This is one function of statewide standards. An *A* is an outstanding result with a specific meaning, however many other students also received one; likewise a *C* is a commendable performance in its own right, signifying that the student has successfully completed a challenging, large-scale, transdisciplinary task.

Nevertheless, schools will want to know how their students performed in comparison with other students across the state. This is another function of statewide standards; and the better the standards are being applied across the state, the greater the comparability of grades, and the greater the value of the data. Comparing one school's results with results from across the state, however, needs to be approached with caution, as there are many factors that influence grades which are not apparent in the numbers alone.

These data also provide schools with a context in which to reflect on their own experiences in implementing (or deciding not to implement) particular tasks. Questions these data can help schools answer might include:

- We ensured that all students undertook all tasks in their suite — did everyone do this?
- None of our students achieved more than a *C* in this task — did everyone find it that difficult?

The Research Premise of the New Basics — *Reconstruction of curriculum, pedagogy and assessment needs to be explicitly guided by documented analysis and rigorous discussion of current school practices* — refers to school-level, as well as system-level, analysis and discussion.

The importance of the data for the Rich Task Team

The Research Premise applies also to the Rich Task Team. The grades awarded to students are the visible end-product of the moderation (or standards validation) process, and are essential elements in any evaluation of that process or in considering future developments.

Where the results point to tasks that students seem to be finding more difficult than others, or ones that schools are more unwilling to offer their students than others, this may suggest areas where teachers could benefit from additional support, or new kinds of support. On the other hand, it may also suggest aspects of Rich Tasks or Grading Masters that should be considered for revision or clarification.

The scope of this report

This report:

- points out some features of the statewide data that may be of particular interest to schools;
- suggests directions in which discussion of these features of the data may proceed;
- seeks possible explanations of some of these features in the context of the 2004 moderation reports on individual Rich Tasks.

The results, 2004

Aggregated statewide results in each Rich Task in 2004 are presented in Table 1. A shaded cell represents a grade that was not available in a particular Rich Task (for example, the grading scheme for 3#1 does not include a *B* grade).

In the Year 9 results, it is important to recognise that Rich Tasks 9#7 and 9#7B are alternative, mutually exclusive tasks; the percentages given here for these tasks are based on the *whole cohort* of Year 9 students.

Table 1. For students, grades awarded statewide, 2004, by Rich Task, as a percentage of the cohort, and for schools, percentage of schools submitting all *Ns*

Task	Students							Schools	
	Grade						Code		All <i>Ns</i> (%)
	A	B/A	B	C/B	C	C-	U	N	
3#1	1.9	20.9		34.2	24.8	1.2	4.7	12.3	9
3#2	10.6		47.3		20.3	1.1	2.4	18.2	13
3#3	0.3		30.6		39.3	2.1	4.0	23.8	15
3#4	1.8		46.6		30.5	0.5	1.8	18.8	20
3#5	8.3		45.3		26.5	0.2	5.0	14.6	7
6#1	1.1	7.3		24.2	32.5	1.2	11.4	21.9	22
6#2	0.9	13.9		29.0	33.5	1.1	6.8	14.4	9
6#3	0.0	0.7		9.0	27.0	6.9	13.4	42.7	24
6#4	1.9	12.9		24.4	30.2	0.5	3.6	26.3	20
6#5	0.9		17.8		27.4	0.7	3.9	48.9	43
6#6	2.4	9.7		23.5	26.7	5.7	8.0	23.7	26
6#7	1.2	7.4	11.4	14.2	14.5	0.6	3.4	46.9	43
9#1	3.3	10.7		17.0	31.4	2.0	26.2	9.3	6
9#2	1.8		6.5		10.9	29.4	26.0	25.3	22
9#3	2.1	5.5	9.5	16.2	18.1	2.4	17.6	28.7	11
9#4	1.5	5.1	11.6	9.9	26.2	4.2	18.5	23.1	11
9#5	0.6	6.4		22.3	46.5	1.0	11.8	11.5	6
9#6	5.0		28.8		38.8	1.2	16.1	10.1	11
9#7	0.7	1.3		2.3	4.5	0.0	6.4	84.9	67
9#7B	1.9	6.3		12.3	28.3	2.4	21.4	27.4	28
9#8	0.7	2.7	9.4	15.0	20.1	1.5	28.7	21.8	11

The percentage of schools submitting all *Ns* is provided because while students receive grades (and the code *U*) on the basis of their work, they may receive the code *N* as a result of a school decision (not to offer the task to those students).

The names of the tasks are provided in Table 2. The names and IDs used are those that were current in 2004.

Table 2. Names of the Rich Tasks, 2004

<i>Task ID</i>	<i>Task name</i>
3#1	Webpage Design
3#2	Multimedia Presentation on an Endangered Plant or Animal
3#3	Physical Fitness
3#4	Read and Talk about Stories
3#5	Historical and Social Aspects of a Craft
6#1	Travel Itineraries
6#2	Narrative Text: Away with Words
6#3	Personal Health Plan
6#4	A Celebratory, Festive or Artistic Event or Performance
6#5	Oral Histories and Diverse and Changing Lifestyles
6#6	Design, Make and Display a Product
6#7	Space Futures
9#1	Science and Ethics Confer
9#2	Improving Wellbeing in the Community
9#3	The Built Environment: Designing a Structure
9#4	Australian National Identity: Influences and Perspectives
9#5	Personal Career Development Plan
9#6	Opinion-making Oracy
9#7	Pi in the Sky
9#7B	The Shape We're In
9#8	International Trade

The generic descriptor for each grade and code used is given in Table 3. (Note that the grade of C- was introduced in 2003, and was not intended to be a permanent feature of the system.)

Table 3. Grades and codes, and their descriptors, 2004

<i>Grade</i>	<i>Standards Descriptor</i>
A	Excellent work, displaying the features of the top grade available for reporting.
B/A	Work surpassing the standard for the B-grade, but not satisfying the requirements for the award of the A-grade.
B	High-quality work, surpassing the requirements for task completion in a number of respects.
C/B	Work surpassing the standard for the C-grade, but not satisfying the requirements for the award of the B-grade.
C	Commendable work. Task was completed within specifications.
C-	Work with aspects at C standard or above, but not meeting the task specifications in full.
Code	Descriptor
U	The student did not comply with major specifications of the task.
N	The student had no opportunity to undertake the particular task.

Throughout this report, “grade” is used as in this table. *U* and *N* are not considered grades, but codes that describe the way a student did *not* receive a grade.

The percentages in Table 1 are represented graphically, for each suite, in Figures 1–3.

Figure 1. Year 3 grades awarded statewide, 2004, by Rich Task, as a percentage of the cohort

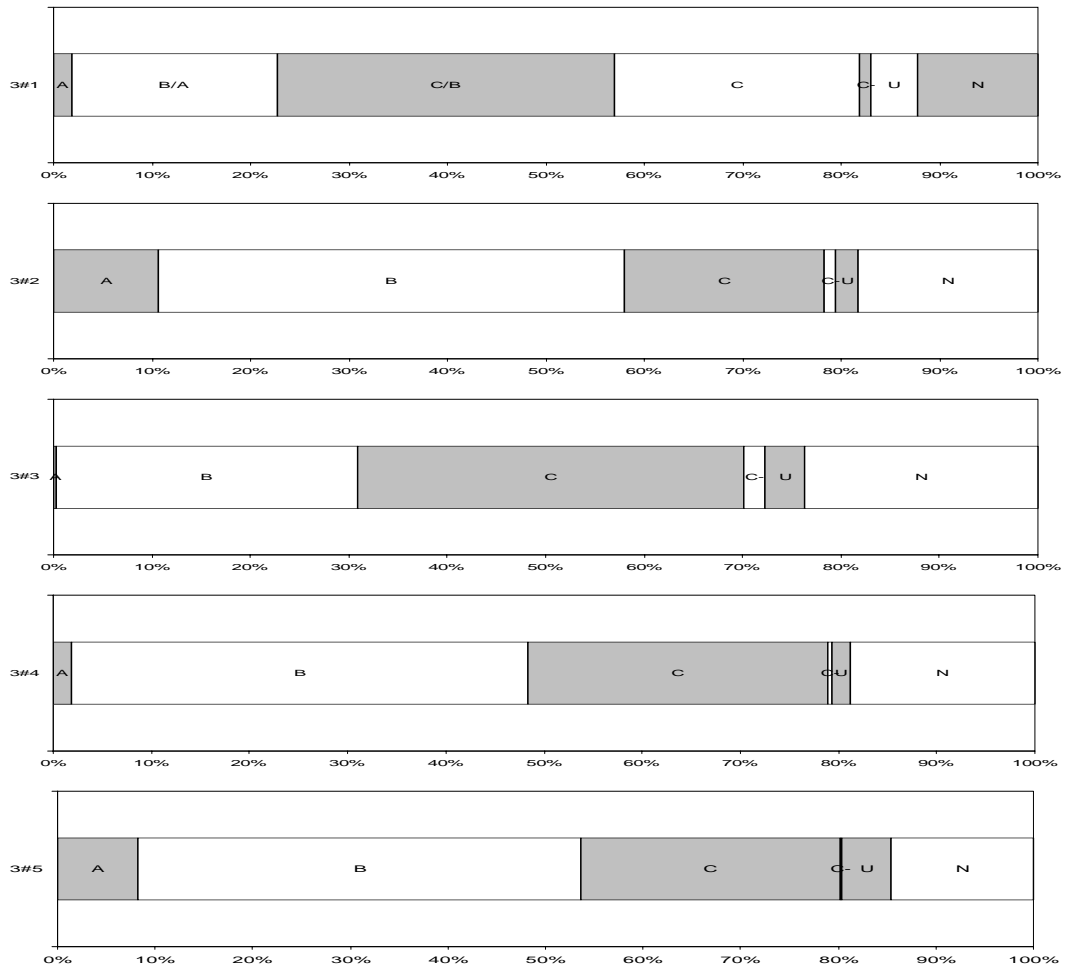


Figure 2. Year 6 grades awarded statewide, 2004, by Rich Task, as a percentage of the cohort

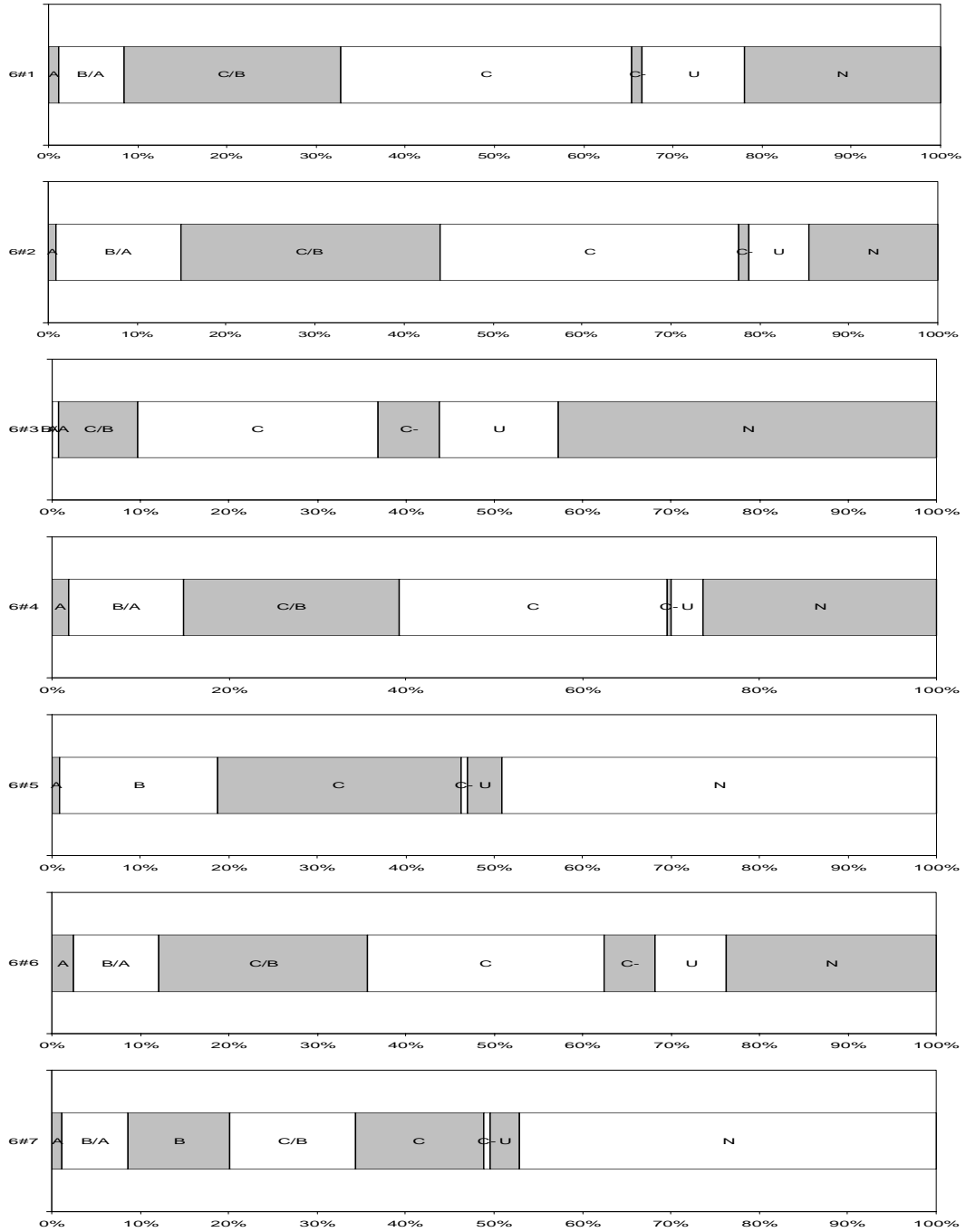
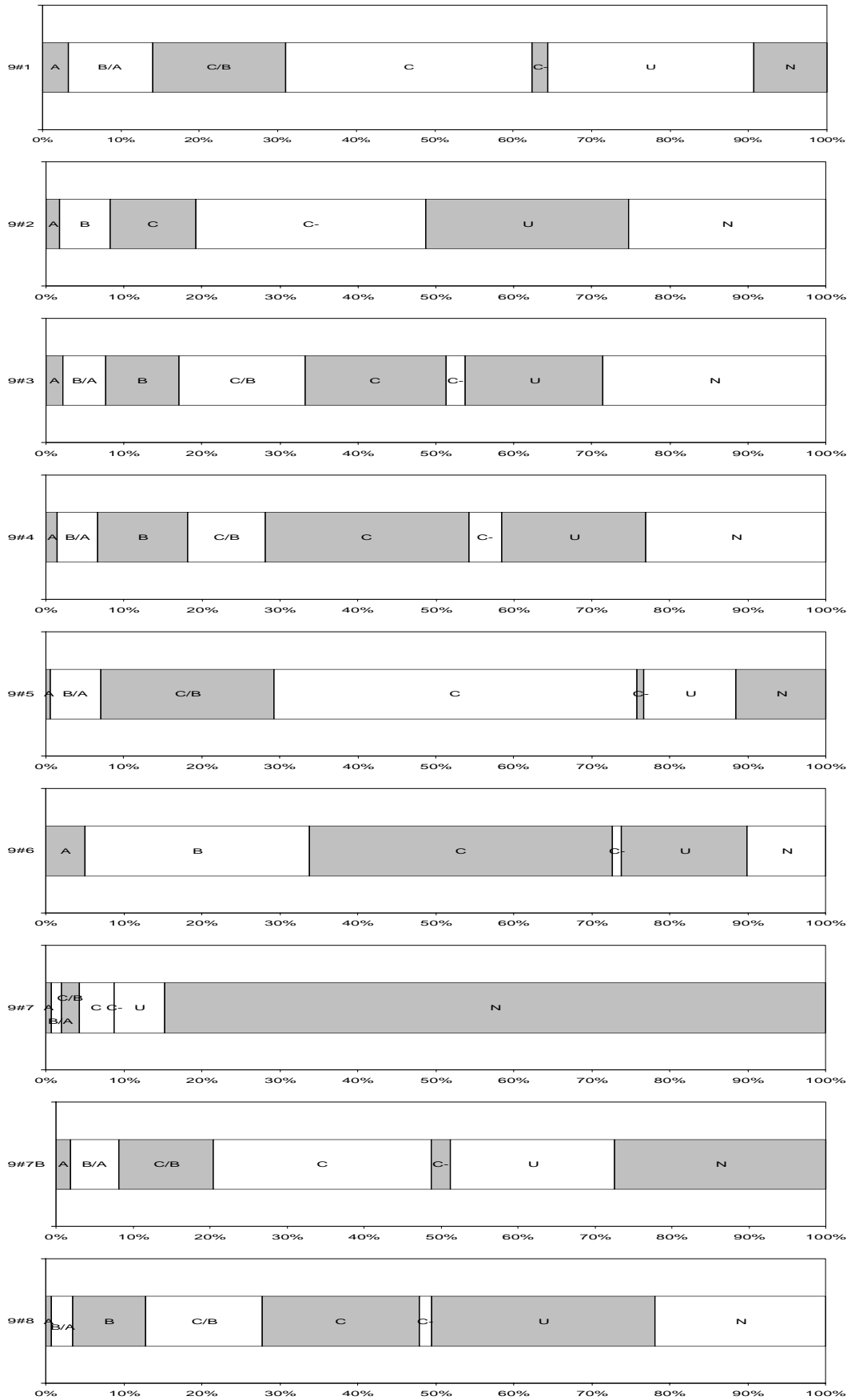


Figure 3. Year 9 grades awarded statewide, 2004, by Rich Task, as a percentage of the cohort



Observations

In no task did all schools offer all students the opportunity to undertake the task.

The highest percentage of students to undertake a particular task was 90.7% (in 9#1).

The lowest percentage of students to undertake a particular task (apart from 9#7, which is one of a mutually exclusive pair of tasks) was 51.1% (in 6#5).

The highest percentages of schools to submit all *Ns* for a particular task (apart from in 9#7 and 9#7B) were 43% (in 6#5 and 6#7), and 24% (in 6#3).

In every task but one (6#3), some students received an *A*.

The highest percentage of *As* awarded in a task was 10.6% (in 3#2).

In every task but two, every available grade and code was utilised (the exceptions being *A* in 6#3, and *C-* in 9#7).

In every task, some students received a *U*.

In every task, more students received a grade (*A* to *C-*) than a *U*.

The proportion of students who received a grade (*A* to *C-*) to students who received a *U* in a task ranged from 42.9 (in 3#4) to 1.7 (in 9#7).

In every task but four, more students received a *C* than any single higher grade. All four tasks where this was *not* the case were from the first suite: 3#1, 3#2, 3#4 and 3#5.

In nine tasks, more students received a *C* than all higher grades combined. One of these tasks was from the first suite (3#3), two from the second suite (6#3 and 6#5), and six from the third suite (9#1, 9#2, 9#5, 9#6, 9#7 and 9#7B).

The percentage of students receiving a *C-* was generally low, but highest in 9#2 (29.4%), 6#3 (6.9%) and 6#6 (5.7%).

The range of participation and performance, by suite

The extent of variation within the context established by the above observations can be suggested by looking at the extreme values within each suite, in relation to four basic questions:

- How many students were given the opportunity by their school to undertake a particular task (that is, received a result other than an *N*)?
- How many students performed to at least the commendable, within-specification standard (that is, received at least a *C*)?
- How many students performed above the commendable standard (that is, received a grade higher than *C*)?
- How many students performed to the highest standard (that is, received an *A*)?

In the tasks in the first suite:

- between 75.7 and 87.7% of students were given the opportunity by their school to *undertake a particular task*;
- between 69.3 and 81.8% of students performed to *at least the commendable standard*;
- between 29 and 57.6% of students performed *above the commendable standard*;
- between 0.3 and 9.5% of students performed *to the highest standard*.

In the tasks in the second suite:

- between 49.2 and 85.2% of students were given the opportunity by their school to *undertake a particular task*;
- between 36.1 and 77.3% of students performed to *at least the commendable standard*;

- between 8.5 and 44.4% of students performed *above the commendable standard*;
- between 0 and 2.2% of students performed *to the highest standard*.

In the tasks in the third suite:

- between 71.3 and 90.7% of students were given the opportunity by their school to *undertake a particular task* (or in the case of 9#7 and 9#7B, to undertake *either* of these mutually exclusive tasks);
- between 19.3 and 75.7% of students performed to *at least the commendable standard* in a particular task (or in the case of 9#7 and 9#7B, in *either* of these mutually exclusive tasks);¹
- between 8.3 and 33.7% of students performed *above the commendable standard* (or in the case of 9#7 and 9#7B, in *either* of these mutually exclusive tasks);
- between 0.6 and 5% of students performed *to the highest standard* (or in the case of 9#7 and 9#7B, in *either* of these mutually exclusive tasks).

Discussion

Schools offering tasks to students

In 2004 it was an expectation that all students would undertake all tasks in the appropriate suite. This did not occur in any task. Even in 9#1, the task with the highest participation rate, 9.3% of students did not undertake the task.

Although in 2005 schools are no longer required to do all tasks, investigating the pattern of offered and withheld tasks in 2004 may still reveal important information, or suggest valuable lines of enquiry, about how schools regard the suites of tasks.

A notable feature of the data about schools giving (or not giving) students the opportunity to undertake a task is that it is not just particular tasks that have been withheld. *Each* task has been withheld from a notable proportion of students. If it was the same few tasks that were not being offered across a range of schools, the clear implication would be that those tasks were being judged to be flawed (too difficult, perhaps, or in some other way inappropriate or unpalatable). It still may be that schools consider some tasks to be flawed, but they do not seem to think the *same* tasks are flawed.

Although all tasks were withheld to some extent, however, the highest instances of low participation are clustered in the second suite. In three of the second suite's seven tasks, over 40% of students did not undertake the task. Not taking 9#7 into account (again, because it is one of a mutually exclusive pair of tasks), the next highest percentage of students not undertaking a task is 28.7% (for 9#3). There is not a striking difference between participation rates in the first and the third suites, so this pattern for the second suite deserves consideration.

It may be that the seven tasks in the second suite (as opposed to five in the first) are too much for many teachers and their classes, at a year level where one teacher may be responsible for teaching all tasks. If so, an issue that could be pursued is whether it is the required *time* that is excessive, or the required depth of teachers' *knowledge* across a range of disciplines. It could be argued that the depth and breadth of knowledge required in the first suite provide less of a challenge to Year 3 teachers, while for Year 9 teachers — operating often within subject

¹ For 9#7 and 9#7B, it is straightforward to combine the number of students who did *not* receive an N in either of these tasks to find the number of students who were given the opportunity by their school to undertake one of these tasks. Combining the number of students who received particular grades in either of these tasks, however (for example, an A, or "more than a C"), is simply a convenient way of positioning the split results within the whole cohort. It is justifiable in relation to the generic definitions of grades in Table 3; it does not imply that a particular grade in one of these tasks is "the same" as the corresponding grade in the other.

specialties — the depth is less of a problem and the breadth is dealt with by having a range of teachers involved, perhaps quite independently of each other. Even allowing for a transdisciplinary approach to the tasks, it would not generally be up to one teacher to see that all tasks in the third suite are offered.

Students' performance in tasks

The pattern of results varies between tasks and between suites, but within these variations it can be observed that in all but two tasks at least *some* students performed to a standard that matched each of the available grades (the exceptions, as noted above, being the *A* grade in 6#3 and the *C-* grade in 9#7). The highest grades, then — despite their explicitly aspirational nature — were generally achievable by a generally small number of students (as many as 10.6% in 3#2, and as few as 0.3% in 3#3). The grade that predominates in the results, however, is *C*. In all tasks in the second and third suites, more students received a *C* than any single higher grade. In six of the nine tasks in the third suite, more students received a *C* than all the higher grades combined. Some students in each task did not perform to a level that matched a grade (as many as 28.7% in 9#8, and as few as 1.8% in 3#4).

A reasonable question to ask is, “What do these distributions of grades tell us about how well the students actually performed?” Each Rich Task is intended to be complex and demanding in its own right. A *C* grade signifies successful completion of a task. If the tasks were widely regarded as being undemanding, then successful completion could be considered a minor achievement. This, however, does not seem to be a widely held view of the tasks.

Accordingly, the predominance of the *C* grade in most tasks cannot be regarded as a sign of poor performance. An issue worth considering, however, is whether the heavily populated *C* grade, together with (often) the relatively thinly populated higher grades, provides sufficient discrimination for a detailed and informative reporting system.

In addition to each Rich Task being complex and demanding in its own right, each suite of Rich Tasks is intended to contain a challenging diversity of tasks. What may seem acceptable performance in the context of a single task, may assume more impressive proportions in the context of a *range* of such performances across diverse disciplines.

The general pattern of grades awarded in the first suite has a marked difference from that in the other suites, especially the third. As noted above, in four of the five tasks in the first suite there was a higher grade than *C* which was awarded to more students than was the *C* grade, while in six of the nine tasks in the third suite, more students received a *C* than all the higher grades combined. Similarly, the percentage of *Us* in the first suite ranged from 1.8 to 5.0%, while in the third suite (again combining grades for 9#7 and 9#7B) it ranged from 11.8% to 28.7%. In two of the first suite tasks, students received markedly more *As* than in any tasks from the other suites: 10.6% in 3#2, and 8.3% in 3#5.

These differences are sufficient to suggest that different conditions may operate in the early years, compared with late primary or, more markedly, early secondary. It may be that the differences pertain to actual student performance, or to teacher *assessment* of student performance, in the early years, or to the nature of the tasks themselves. It would not be unexpected if the culture of teachers in the early years tended towards encouragement and away from stringency to a greater extent than that of teachers of Year 9. If this is indeed what is influencing the grades awarded, it raises questions about the nature of standards that are developed in the early years.

On the other hand, the nature of the first suite tasks themselves may be influencing these patterns of performance. Only one of the tasks in the first suite is entirely an individual task (3#1), and it is possible that the cooperative nature of these tasks allows some students to perform to a higher level than they would generally perform by themselves.

The relatively low incidence of *C-* grades in most tasks suggests that where students are producing work “with aspects at *C* standard or above”, they are also, in nearly all cases, successfully completing the task within specifications. If it were otherwise, it might suggest

that there were elements in the tasks that either place unreasonable hurdles in front of students who would otherwise successfully complete them, or are not being understood by teachers as being essential to the task. This appears to be not the case.

Looking for explanations

Despite patterns of results *across* tasks that suggest interpretations such as those canvassed above, there are particular tasks with highly individual patterns of results. These include:

- 6#3, which 42.7% of students did not undertake, and for which no student received an A and only 0.7% received the next highest grade, B/A;
- 9#2, in which the proportion of students who received a C- almost equalled the combined proportion of students who received a higher grade (A, B or C).

While interpretations of patterns across tasks or across suites tend to be speculative, there is a secondary source of data about particular tasks that should help explain such individual patterns: the moderation reports for 2004.

The moderation report for 6#3, for example, among a range of other comments, points to two features of this task's implementation that could help explain the small number of students who received a high grade. One of these relates mainly to what students did, the other mainly to what teachers did:

- If the science experiment had not been set up correctly, it was extremely difficult for students to present data on the overall aspect and both factors or actions, and thus demonstrate gradable performances. Students proceeding with experiments that yielded insufficient data also found that this restricted the depth of analysis that could be demonstrated.
- Some teachers developed templates for students that heavily scaffolded their learning experiences through the Task. Some templates were useful in guiding students to demonstrate what is highly valued in this Task, while other templates seemed to constrain student demonstrations to the extent that high quality performances were not often realised. (<http://education.qld.gov.au/corporate/newbasics/docs/mr6-3.doc>)

The report emphasises the need for enhanced scientific knowledge on the part of Year 6 teachers. It is possible that a contributing factor to schools' relative reluctance to offer this task is some teachers' own sense of an inadequate background in science.

The moderation report for 9#2 likewise presents information about the implementation of the task that could readily account for the very high proportion of students who received the grade of C-, which is reserved for work with aspects at C standard or above, but not meeting the task specifications in full. For example:

- Four products are required by this Task. Their presence is necessary evidence in supporting judgments on the quality of student work. The inclusion of student diaries is a "hidden" fifth product. Of these five products, the collaboratively developed community plan is the centrepiece of the entire Task, not only of the first part of the Task, which requires "students to work with a local community to develop a plan for improving an aspect of the wellbeing of this community". This product is frequently missing from student demonstrations. (<http://education.qld.gov.au/corporate/newbasics/docs/mr9-2.doc>)

Moderation reports, then, are an important source of information that schools may use in making sense of the data.

The aggregated statewide data that have been examined in this report will probably be of interest to schools, but their benefit to schools really occurs when they are used to illuminate school-specific data, within the context of direct experience of implementing the Rich Tasks.