# STUDENT RESPONSE CHARACTERISTICS TO OPEN-ENDED TASKS IN MATHEMATICAL AND OTHER ACADEMIC CONTEXTS

**DAVID CLARKE, PETER SULLIVAN AND URSULA SPANDEL**
**Mathematics Teaching and Learning Centre**
**Australian Catholic University (Victoria) - Christ Campus**

## INTRODUCTION

The research reported in this paper represents one component of a larger project addressing the consequences and implications of the use of a new class of open-ended mathematical tasks called "Good Questions", devised by Sullivan and Clarke (1991b). The particular characteristics of this type of task are the content-specific focus, and the opportunity for answers at different levels of sophistication. An extended research program is underway to investigate the use of this form of open-ended mathematics task for instruction and assessment in mathematics. The characteristics of these Good Questions have been discussed elsewhere (Sullivan and Clarke, 1988; Clarke and Sullivan, 1990, Sullivan and Clarke, 1991a and b, Sullivan, Clarke and Wallbridge, 1991). Each of the three postulated characteristics of Good Questions derives from a specific educational stance.

1. The task should require more than the recall or replication of a fact or procedure.
2. The task should be educative.
3. The task should be open-ended.

The rationale is that if mathematical power is to be associated with the ability to solve non-routine problems (NCTM, 1989), then our assessment and instructional tasks should have some of the characteristics of problem solving tasks. The Good Questions approach has been proposed as a step in this direction.

Perhaps the most consistent and significant finding to emerge from previous research (see Clarke & Sullivan (1991) for an overview of these findings) was the reluctance of pupils from grades 6 to 10 to provide more than a single answer to such open-ended questions as:

- The average of five numbers is 17.2, what might the numbers be?
- A number is rounded off to 5.8, what might the number be?
- A rectangle has a perimeter of 30 units, what might be its area?

The term "reluctance" is used here to indicate the disinclination of the pupils tested to give either multiple answers or a general statement. As was demonstrated in the earlier studies, this disinclination should not be confused with an inability to provide multiple answers. If open-ended tasks like those above were reworded so as to require pupils to list all possible multiple answers, then the number of multiple responses increased substantially. This demonstration of a capacity to provide multiple answers was also evident in the responses of pupils when the tasks were administered in interview situations. It is clear from these earlier studies (Clarke and Sullivan, 1991) that, while few students are equipped to produce general statements as solutions to open-ended tasks, many pupils have the capability to produce multiple answers but do not choose to do so. Key results from these previous studies are summarised below.

## PREVIOUS STUDIES

### Mathematics Only

In a typical administration of the investigation of open-ended mathematics tasks, a set of four questions was given to participant classes of schoolchildren. The criterion for selection of classes was the willingness of their teachers to participate. Even though no teacher declined the invitation no claims are made about representativeness of the results for other schools.

In the first administration of tasks, the questions asked, the procedure for administration and the response coding system were as follows:

**Subtraction**

Last night I did a subtraction task. I can remember some of the numbers.

```
   1 4 __
 -   7 __
--------
   __ 4
```

What might the missing numbers have been?

**Rounding**

A number has been rounded off to 5.8 What might the number have been?

**Area**

A rectangle has a perimeter of 30 m. What might be the area?

**Fraction as operator**

$\frac{2}{5}$ of the pupils in a school play basketball. How many pupils might there be in the school and how many might play basketball?

The format for administering the questions was the same in each class:

i) The question: "___ + ___ = 10 What might the missing number be?" was posed, and the responses suggested by the pupils were written on the chalkboard. The pupils in the class were asked to comment on what was different about this task from common mathematics questions. The response sought was that there are many possible answers.
ii) The first two questions were distributed (subtraction, rounding).
iii) The papers were collected and the answers reviewed. Again the possibility of multiple answers was discussed.
iv) The second two questions were distributed (area, fraction as operator)

The responses of the pupils to the tasks were coded. The coding was as follows:

0 meant no correct answers
1 meant only one correct answer
2 meant two or three correct answers
3 meant all or many correct answers
4 meant that a general statement was given

To illustrate the way that this code was applied to the rounding question, the following is the meaning of the codes. Individual correct answers were numbers such as 5.82 or 5.78. A code of "3" was given to a response like "5.75 5.76 5.77 5.78 5.79 5.80 5.81 5.82 5.83 5.84". Examples of responses which were considered to represent a general statement, "4", were "5.75 ... right up to 5.84999..." or "between 5.75 and 5.849".

### Key results

Key findings from the "Mathematics Only" phase of the research (from Sullivan, Clarke and Wallbridge, 1991; Clarke and Sullivan, 1991) included:

1. The capability to give multiple and general responses increased with age.
2. Revision of the "Good Question" format to specifically request multiple answers resulted in a significant increase in the proportion of grade 6 students providing multiple responses.
3. Results suggested a task-specific dependence in student levels of response.
4. Students were more likely to provide multiple responses in an interview context than in a test context.
5. The learning outcomes of a seven-lesson teaching program for grade 6 pupils based solely on the use of such open-ended mathematics tasks were not significantly different from the outcomes of a control class instructed to follow the program presented in the most commonly used text. These results render the hypothesised benefits problematic.

It is possible that the period of instruction for the experimental group was insufficient to overcome the many years of schooling during which the provision of a single correct answer has been the accepted student goal. It seemed to the authors that the distinction between the capability to provide multiple answers to an open-ended task and the inclination to do so required further investigation. In particular, it was not known whether the reluctance to provide multiple answers was specific to mathematics or common to other academic disciplines. This is the question which this study sought to address.

## THIS STUDY

### Mathematics and Other Academic Contexts

This paper is concerned primarily with the question:

> Is the reluctance of pupils to give multiple or general responses to open-ended mathematics tasks replicated with open-ended tasks in other disciplines?

The sample employed also made possible the investigation of the effects of grade level, gender, task-specificity, the presence or absence of explicit cueing, and the interaction of these variables with academic domain. Central to the study was the distinction between student response to open-ended tasks which did not explicitly request multiple answers ("uncued") and student response to open-ended tasks which explicitly requested multiple answers ("explicit cueing"). It was assumed that student responses to the uncued tasks would provide an indication of student "Inclination" to provide multiple or general responses, while student responses to tasks employing explicit cueing would indicate student "Capability" to provide multiple responses.

For this study, students at years 7 and 10 from four secondary schools (one single-sex boys' school, two single-sex girls' schools, and one co-educational school) were asked to respond to open-ended items from a variety of academic contexts. The communities served by the four schools had comparable socio-economic characteristics. Four separate research protocols were employed in the administration of the tasks. The following structural concerns were explicitly addressed in the protocol design: order of task administration; explicit or non-explicit cueing; and, task variation within a domain.

In particular, since the results arising from a single open-ended task in a given discipline may be distorted by task idiosyncracies such as the task context or the cognitive demands of the task, the investigation was repeated with a different task in each academic context. Over all four protocols, two tasks were used in each academic domain:

***Social Studies***
In a Victorian country town, the population fell by 50% over a period of 5 years.
Why might this have happened?

A 19 year old girl moved out of home.
Why might she have done this?

***English***
A fragment of an early sonnet by Shakespeare has been found. It includes the incomplete line:
"And ________________________ did mask the sun."
What might be missing?

A game of scrabble had only just started when the board was tipped over. The letters on the floor were: T, S, P, O.
How might the letters have been arranged before the board was tipped over?

***Science***
A plastic ice-cream container half full of water is left on the front lawn to keep cats and dogs away. One week later the container is empty. How might this have happened?

A farmer lost his entire crop.
Why might this have happened?

***Mathematics***
A rectangle has a perimeter of 30 units.
What might the area be?

The average of five numbers is 8.
What might the numbers be?

It was intended that the tasks be well within the cognitive capability of most students. The concern of the study was the distinction between student inclination (uncued) and capability (with explicit cueing) to give single or multiple answers, and it was not intended that task difficulty be a factor for the domains of Social Studies, English and Science. It is acknowledged, however, that student lack of familiarity with the context of a specific task might increase the task's cognitive demand and lead to task-specific differences. As will be seen in the following results section, student responses across the two tasks within a domain were sufficiently different as to provide an endorsement of the use of more than one task in each domain. The analysis of student response levels will combine both tasks as a representative sampling of responses in each domain. Both mathematics tasks had been used in earlier studies and provided a basis for comparison with earlier findings arising from the use of only mathematical tasks.

The protocols guiding the administration took a basic form derived from that for the earlier administration of mathematical tasks (see Clarke and Sullivan, 1991), and were administered in academically-neutral classrooms (with respect to the academic domains which formed the focus of this study) by a trained research assistant.

Protocol 1 Introductory discussion employing the example "The sum of two numbers is ten; what might the numbers be?"
Distribution of two questions: One mathematics, and one science.
Response sheets collected and differences between answers discussed.
Distribution of two questions: One language and one social studies, with the direct request that multiple solutions be sought.

Protocol 2 As for 1 with the Mathematics/Science and Language/Social Studies pairs exchanged (that is, Language and Social Studies questions given first).

Protocol 3 As for 1 with different tasks.

Protocol 4 As for 2 with different tasks.

Since the focus of this study is on student inclination to provide multiple answers to open-ended tasks, analysis disregarded the correctness of student response in applying the earlier coding of student responses. The responses of the pupils to the tasks were coded as follows:

0 meant no answer provided
1 meant only one answer provided
2 meant two or three answers provided
3 meant all or many distinct answers provided
4 meant that a general statement was attempted.

With respect to order of task administration, a preliminary administration of the protocol 1 and 2 tasks to a smaller sample (n = 87) without any explicit cueing of multiple responses, established that a different order of administration of the domain-specific tasks did not lead to significant differences in response level. As a consequence, the protocol structures and sampling design anticipated an analysis with respect to year level, gender, degree of cueing, and task specificity, but not task order. The sample construction is shown in Table 1.

**Table 1:** Sample construction

| | Protocol 1 | | Protocol 2 | | Protocol 3 | | Protocol 4 | | Total Sample | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Yr 7 | Yr 10 | Yr 7 | Yr 10 | Yr 7 | Yr 10 | Yr 7 | Yr 10 | Yr 7 | Yr 10 |
| Male | 70 | 69 | 72 | 67 | 33 | 30 | 27 | 32 | 202 | 198 |
| Female | 76 | 71 | 52 | 65 | 32 | 25 | 32 | 23 | 192 | 184 |

## RESULTS

To summarize the data collection from which the following results were derived: two tasks were administered for each domain (Task One and Task Two). Each task was given in both cued and uncued modes. No one student attempted more than one task from a domain, and no student attempted both cued and uncued forms of a given task. Student level of response is reported in two forms: the mean level of response for a given group, and the distribution of student responses among the five response categories employed (0, 1, 2, 3, 4).

A more extensive discussion of results can be found in Clarke, Sullivan and Spandel (1992).

### Task-specificity

There appeared to be some dependence of response level on the specific task employed. A two-factor analysis of variance for uncued and cued response level, by task and domain revealed statistically significicant differences attributable to both task, domain, and their combination.

The differences evident in student response level to different tasks within the one academic domain made clear the task-specificity of responses to open-ended tasks. For the remainder of the analyses employed in this paper, responses to both tasks within a specific domain were combined.

### Inclination or Capability - the significance of explicit cueing

The results of previous mathematics-only studies had revealed significant differences in student response level dependent on whether or not an explicit request was made for multiple answers. It was postulated as a consequence of these results that the differences between responses to "uncued" and "cued" open-ended tasks reflected a fundamental distinction between student Inclination and Capability to provide multiple and general responses. This study set out to investigate whether the significance of this distinction pertained in other academic contexts. Figure 1 sets out the differences between uncued and cued mean response levels for the four academic domains addressed in this study.

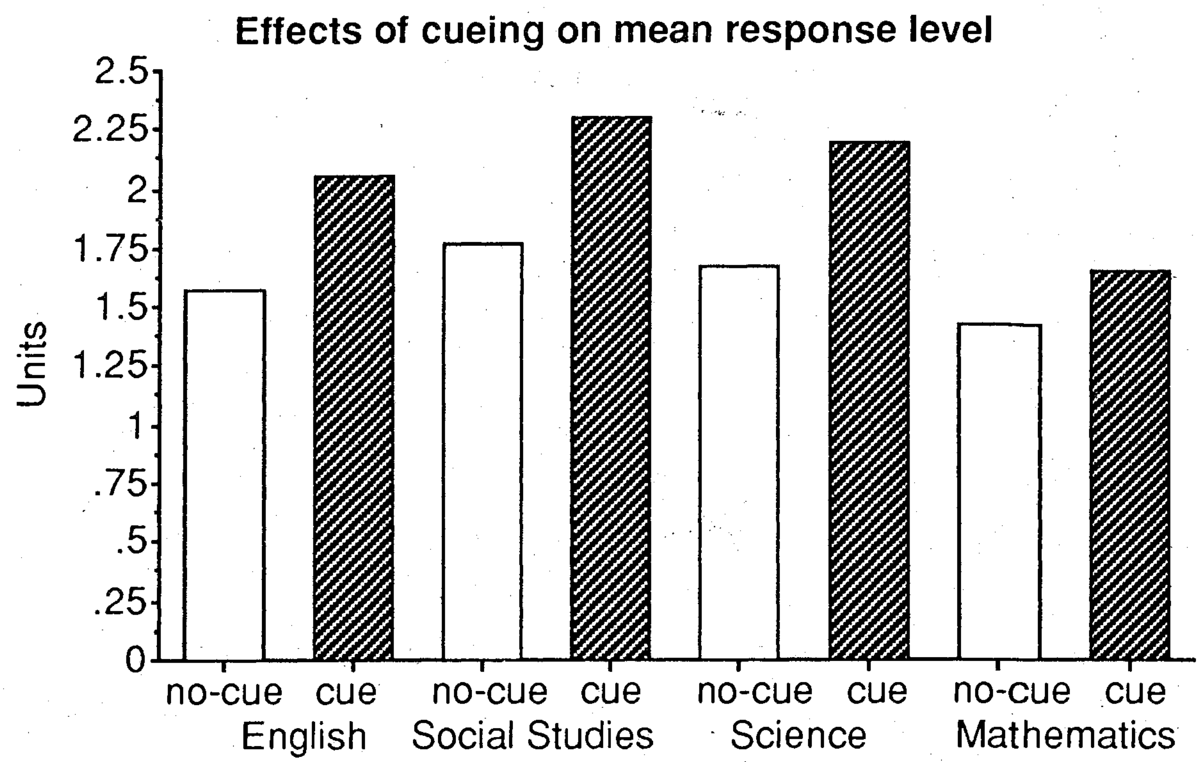


**Figure 1:** Comparison of the uncued and cued mean response levels for English, Social Studies, Science and Mathematics.

It is evident from Figure 1 that the difference in response level arising from the inclusion or omission of the specific request for multiple responses identified in previous studies was present to an even greater extent in non-mathematical contexts. Effect sizes for each of the non-mathematics domains exceeded 0.5, and separate t-tests comparing uncued and cued response levels for each academic context confirmed the significance of the differences displayed in Figure 1. The detail of the differences in response levels is provided in the following section.

**Domain-specificity**

The key question addressed by this paper concerned whether the response patterns observed with open-ended mathematics tasks would be replicated with open-ended tasks in other academic domains. In the results which follow the distinction is maintained between Inclination (uncued responses) and Capability (cued responses).

If either inclination or ability to provide sophisticated (predominantly multiple) responses to open-ended tasks is a reflection of some general intellectual capacity, then one would expect a high correlation of response levels across all academic domains. Alternatively, if either the capacity or the inclination to give more sophisticated responses is a consequence of expertise in the specific domain, then one would not expect high correlations for the response levels across different domains. The sample design employed in this study did not permit correlations to be calculated for both uncued and cued responses for all possible paired combinations of academic domain. For the purposes of this study it was felt more important to ensure that all participating students supplied both cued and uncued responses. This maximized the diversity of the sample which formed the basis of the investigation of

Inclination and Capability. No individual was required to respond to both cued and uncued tasks in the one domain in case completion of one task type affected performance in the other task type. As a consequence, this study could not address the two hypotheses outlined above directly.

It is certainly not possible to distinguish between the "Intellectual Capacity" and "Domain-specific Expertise" hypotheses on the basis of this study. Further research is required to adequately investigate these two hypotheses.

**Year level**

Previous mathematics-only studies employed exclusively uncued task formats and included consideration of the mathematical correctness of the response in the coding, as set out earlier in this report. These mathematics-only studies identified an association between (uncued) response level and year level. The study reported here found that the association between response level and year level differed in character for uncued and cued task administrations. These results are shown in Figures 2 and 3.

Continuing to interpret student response level to uncued open-ended tasks as indicative of the level of inclincation, this study made it clear that the difference between year levels in the inclination to provide multiple answers is only significant for Social Studies and Science. That is, pupils in years 7 and 10 were similarly disinclined to give multiple responses to open-ended tasks in English and Mathematics (when not explicitly requested to do so).

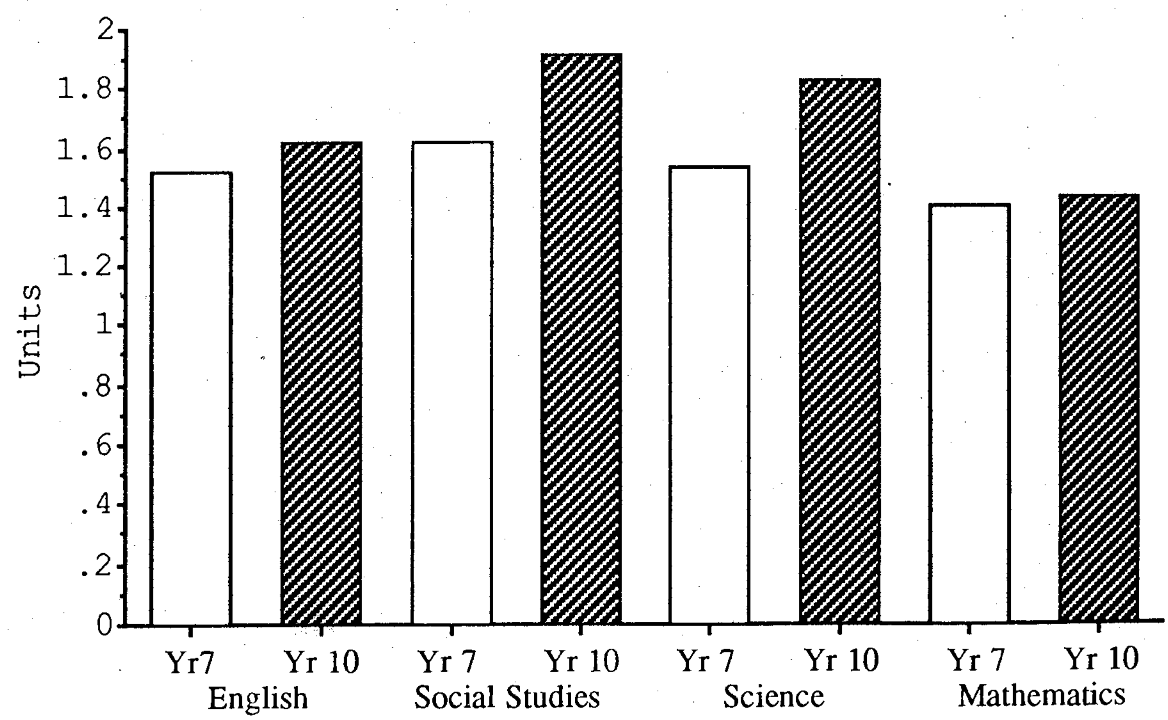


**Figure 2:** Mean uncued response levels (Inclination) by year level for each academic domain.

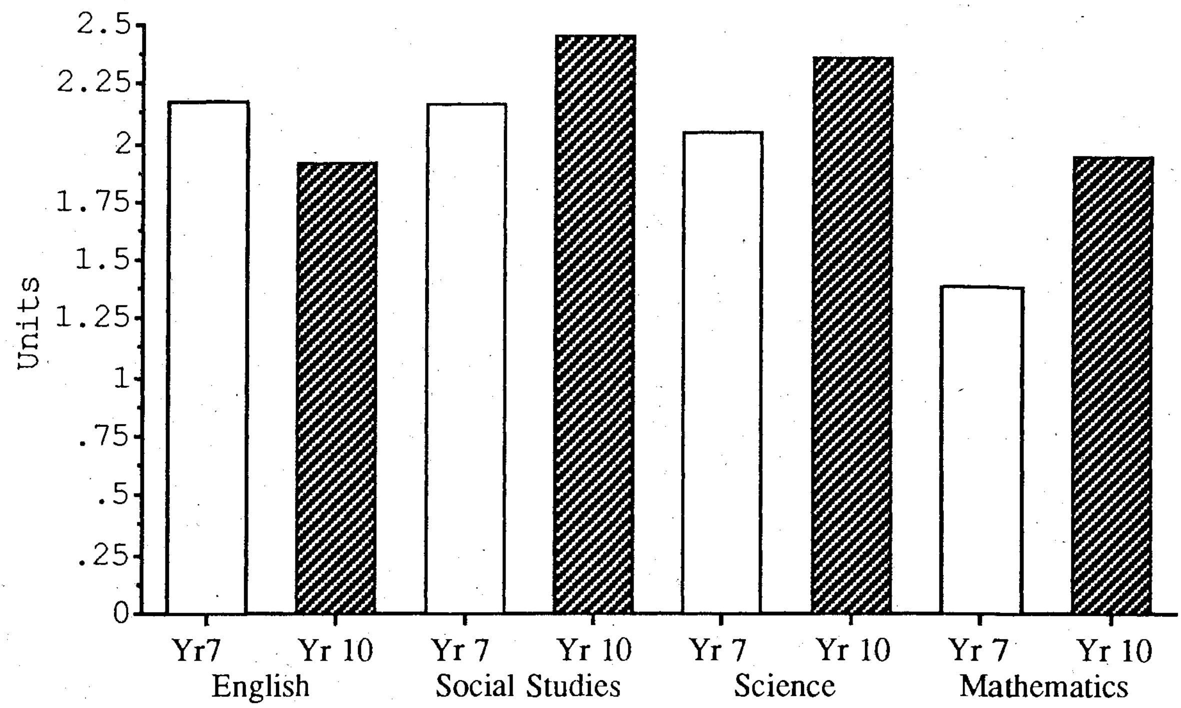


**Figure 3:** Mean cued response levels (Capability) by year level for each academic domain.

When an explicit request for multiple responses is included in the task, student response levels at year 10 are significantly higher than those for year 7 in each of the academic domains studied, with the notable exception of English. We interpret these results as suggesting that the Capability to respond to an explicit request for multiple answers to an open-ended task increases with year level, except in the case of English.

## Gender

In contrast with the findings of Thomas (1989), girls in this study were more likely than boys to provide multiple answers, when specifically requested to do so, in all four domains. However, only in English and Science were these differences statistically significant. Figure 4 sets out the respective Inclinations of girls and boys to provide multiple answers. It is evident that boys and girls do not differ significantly in their Inclination (or reluctance) to provide multiple or general answers to open-ended tasks, except in Social Studies, for which girls display a significantly higher mean response level than do boys.

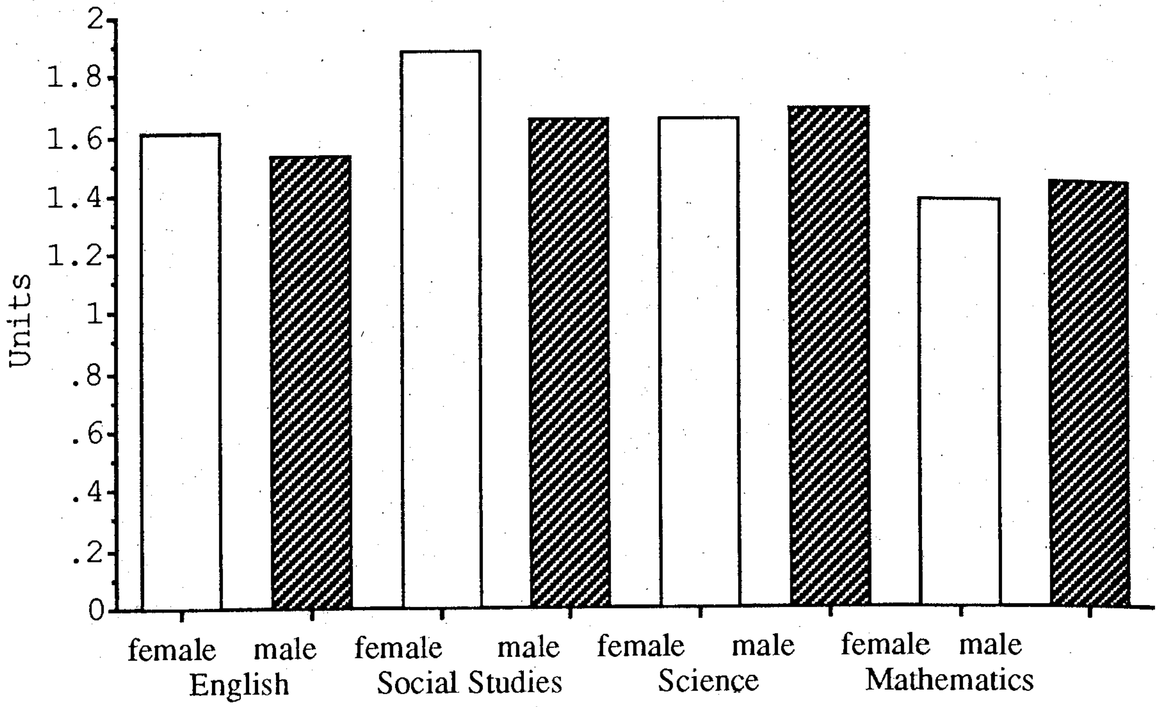


**Figure 4:** Gender-specific mean uncued response levels (Inclination) for each academic domain.

Figure 5 examines whether a significant difference exists for any of the academic domains studied between the Capability of girls and boys to provide multiple answers when explicitly requested to do so.

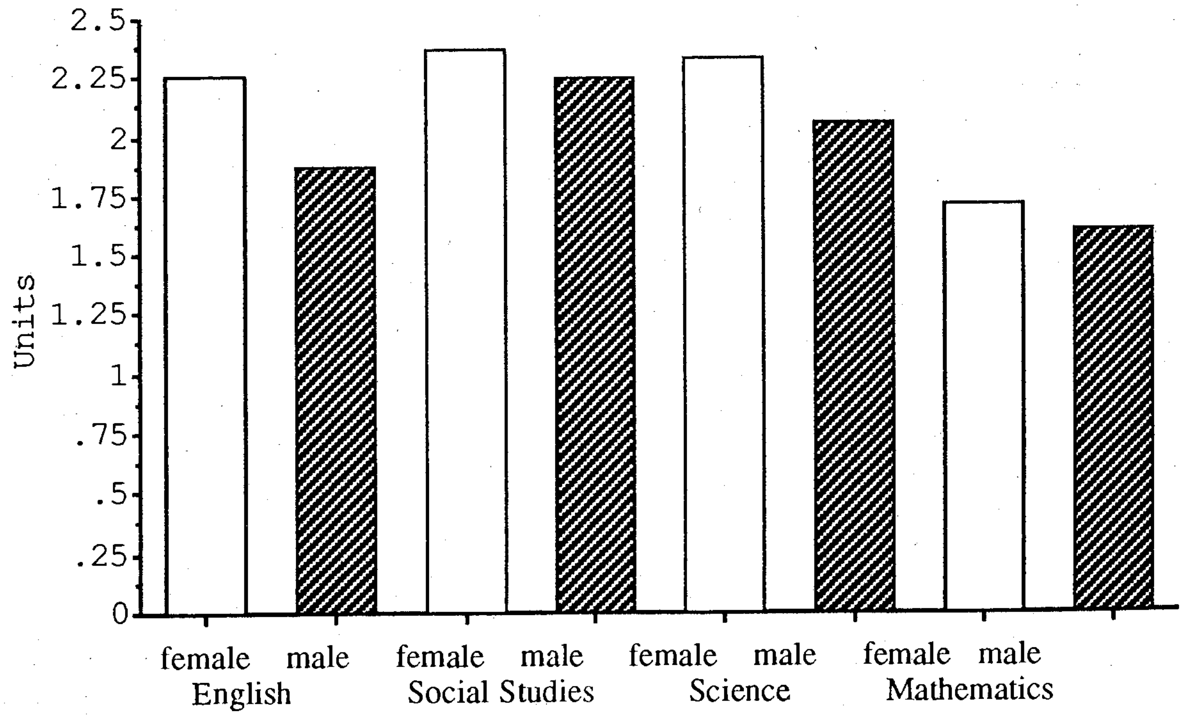


**Figure 5:** Gender-specific mean cued responses (Capability) for each academic domain.

For the sample of pupils used in this study, girls displayed a greater Capability to provide multiple answers than did boys. This difference was significant in the academic domains of English and Science.

## CONCLUSIONS

Conclusions which follow from these results include:

- that the inclination to give single responses (or the reluctance to give multiple responses) is a product of schooling, and not peculiar to mathematics;
- that the explicit request of multiple responses produces a significant increase in the response level in all academic contexts;
- that the capability to give multiple responses increases significantly with year level, except in the context of English;
- that gender-related differences in response level are evident, and where these exist they favour girls.

Substantial additional research is required if we are to understand the significance of the meanings constructed by students in responding to open-ended tasks. Such research must address those student conceptions of legitimate mathematical activity on which their response inclinations are predicated (Clarke, Wallbridge and Fraser, 1991, and this research), and issues of cognitive load or working memory capacity and related developmental theories of learning outcomes which determine student response capability (for example, Collis, 1991; Sweller, 1989). The associated matters of inclination and capability must be understood if we are to employ such tasks with success in mathematics classrooms for the purposes of either instruction or assessment.

Among the various concerns expressed about the use of open-ended tasks for assessment in mathematics, student unfamiliarity with the required mode of response is frequently cited. Student unfamiliarity with the demands of open-ended tasks was also raised as a concern in the context of reading assessment. It should be noted that, for the case of mathematics, even when multiple responses were given the explicit endorsement of an instructional intervention no corresponding increase in response level ensued (Sullivan, Clarke, and Wallbridge, 1991; Sullivan, Clarke, Spandel, and Wallbridge, 1992). This would suggest that student reluctance or incapacity to provide multiple responses to open-ended tasks was attributable to more than simply lack of familiarity with the task format. It is possible that while the brief instructional intervention effectively addressed the issue of familiarity, student inclination to give single responses to academic questions is a consequence of extensive training, and that a much more extended instructional intervention would be required to challenge this 'training' effect. Such a training effect could be attributable to the explicit training of students to provide single responses to academic tasks, or to a lack of explicit training in the provision of multiple or general responses. This study has demonstrated that this training effect is evident in other academic contexts besides mathematics, and may be more appropriately seen as a consequence of schooling rather than just mathematics instruction.

There are other issues besides those raised above, which arise if open-ended items are to be used to assess student learning. The research reported here has raised additional concerns. For instance:

- Task specific differences in pupil performance were evident throughout the study and warrant investigation.

- Inferences concerning student learning drawn from student performance on open-ended tasks must take into account student beliefs concerning appropriate answers in academic contexts or risk mistaking student inclination for student capability.

- It appears that young children find it substantially more difficult than older children to provide multiple answers to mathematics tasks.

The legitimacy of relating student responses to non-routine and open-ended tasks to curricular content currently being studied continues to be the subject of research. It may not be realistic to expect students of any age to access recently-acquired skills in open-ended or problem-solving situations.

Substantial additional research is required if we are to understand the meanings constructed by students in responding to open-ended tasks sufficiently well to employ such tasks with success in mathematics classrooms.

## REFERENCES


Clarke, D.J., & Sullivan, P. (1990). Is a question the best answer? *The Australian Mathematics Teacher,* 46 (3), 30 - 33.

Clarke, D.J., & Sullivan, P. (1991). The sssessment implications of open-ended tasks. Paper given at the "Assessment and the Mathematical Sciences" conference, held at the Institute of Educational Administration, Geelong, November 20 - 24.

Clarke, D.J., Sullivan, P., & Spandel, U. (1992). *Student response characteristics to open-ended tasks in mathematical and other academic contexts. Research Report 7.* Oakleigh: Mathematics Teaching and Learning Centre.

Clarke, D.J., Wallbridge, M., & Fraser, S. (1991). *The other consequences of a problem-based curriculum. Research Report No. 3.* Oakleigh, Vic.: MTLC.

Collis, K.F. (1991). Assessment of the learned structure in elementary mathematics and science. Paper presented to the Conference on Assessment in the Mathematical Sciences, held at the Institute of Educational Administration, Geelong, November 20-24.

National Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

Sullivan, P., & Clarke, D.J. (1988). Asking better questions. *Journal of Science and Mathematics Education in South East Asia,* June14-19.

Sullivan, P., & Clarke, D.J. (1991a). Catering to all abilities through "Good" questions. *Arithmetic Teacher*, 39 (2), pp. 14 - 21.

Sullivan, P., & Clarke, D.J. (1991b). *Communication in the classroom: The importance of good questioning*. Geelong: Deakin University Press.

Sullivan, P., Clarke, D.J., Spandel, U., & Wallbridge, M. (1992). *Using content-specific open questions as a basis of instruction: A classroom experiment. Research Report No. 4*. Oakleigh: Mathematics Teaching and Learning Centre.

Sullivan, P., Clarke, D.J., & Wallbridge, M. (1991). *Problem solving with conventional mathematics content: Responses of pupils to open mathematical tasks. Research Report 1*. Oakleigh: Mathematics Teaching and Learning Centre.

Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 81 (4), pp. 457-466.

Thomas, B. (1989). On their own - Student responses to open-ended tests in science. Massachusetts State Department of Education, Quincy Bureau of Research, Planning and Evaluation.