

**STUDENT RESPONSES TO CRITICAL  
THINKING TASKS IN THE NEMP  
ASSESSMENTS**

**A National Education Monitoring Project  
Probe Study Report**

**Gordon Knight**

**July 2006**

**RESEARCHER:**

**Dr Gordon Knight  
8 Centennial Drive  
Palmerston North**

## **CONTENTS**

<b>1.</b>	<b>INTRODUCTION</b>	<b>1</b>
1.1	Background	1
1.2	Research Questions	2
1.3	Intended Outcome	2
1.4	Use expected to be made of the findings	2
<b>2.</b>	<b>METHODOLOGY</b>	<b>2</b>
2.1	The choice of tasks for analysis	2
2.2	The student sample	3
2.3	Analysis	3
	2.3.1 The tasks	3
	2.3.2 The videotapes	3
<b>3.</b>	<b>THE SCIENCE TASK – ENVIRONMENTAL ISSUES</b>	<b>4</b>
3.1	The task	4
3.2	The marking criteria	5
3.3	The reporting of student performance	5
3.4	The nature of the thinking required in the task	5
3.5	Examples of student responses	6
3.6	Identifying critical thinking in this task	9
	3.6.1 Arranging the pictures	10
	3.6.2 Thinking time	11
	3.6.3 The number of pictures in each row	12
	3.6.4 The justifications	12
3.7	COMMENTS	14
<b>4.</b>	<b>THE ASPECTS OF TECHNOLOGY TASK – NUT CRACKER</b>	<b>16</b>

4.1	<b>The task</b>	<b>16</b>
4.2	<b>The marking criteria</b>	<b>17</b>
4.3	<b>The reporting of student performance</b>	<b>17</b>
4.4	<b>The nature of the thinking required in the task</b>	<b>18</b>
4.5	<b>Identifying critical thinking in this task</b>	<b>18</b>
4.6	<b>Comments</b>	<b>19</b>
5.	<b>LISTENING AND VIEWING TASK – LOOKING AROUND</b>	<b>20</b>
5.1	<b>The task</b>	<b>20</b>
5.2	<b>The marking criteria</b>	<b>21</b>
5.3	<b>The reporting of student performance</b>	<b>21</b>
5.4	<b>The nature of the thinking required in the task</b>	<b>21</b>
5.5	<b>Identifying critical thinking in this task</b>	<b>22</b>
5.6	<b>Examples of student responses</b>	<b>22</b>
	5.6.1 Choosing a very good picture	22
	5.6.2 Choosing a not very good picture and making suggestions for improvement	23
5.7	<b>COMMENTS</b>	<b>24</b>
6.	<b>THE ART TASK – TWO SCULPTURES</b>	<b>25</b>
6.1	<b>The task</b>	<b>25</b>
6.2	<b>The marking criteria</b>	<b>27</b>
6.3	<b>The reporting of student performance</b>	<b>27</b>
6.4	<b>The nature of the thinking required in the task</b>	<b>27</b>
6.5	<b>Identifying critical thinking in this task</b>	<b>28</b>
6.6	<b>Examples of student responses</b>	<b>28</b>
	6.6.1 Evaluating individual sculptures	28
	6.6.2 Comparing the two sculptures	31



6.7	<b>Comments</b>	<b>32</b>
7.	<b>DISCUSSION</b>	<b>32</b>
7.1	<b>Identifying critical thinking in the interviews</b>	<b>32</b>
7.1.1	Is it possible?	32
7.1.2	Visual clues	32
7.1.3	Explanations and justifications	33
7.2	<b>The tasks</b>	<b>33</b>
7.2.1	The Science task	33
7.2.2	The Aspects of Technology task	34
7.2.3	The Listening and Viewing task	34
7.3.4	The Art task	34
7.3	<b>The nature and quality of the critical thinking in each task</b>	<b>35</b>
7.3.1	The Science task	35
7.3.2	The Aspects of Technology task	35
7.3.3	The Listening and Viewing task	35
7.3.4	The Art task	36
7.4	<b>Thinking in context</b>	<b>36</b>
7.4.1	Fact or opinion	36
7.4.2	Subject area differences	37
7.5	<b>Age group differences</b>	<b>37</b>
8.	<b>CONCLUSIONS</b>	<b>38</b>
9.	<b>REFERENCES</b>	<b>40</b>
9.1	<b>NEMP reports</b>	<b>40</b>
9.2	<b>Other references</b>	<b>40</b>

# 1. INTRODUCTION

## 1.1 BACKGROUND

In 2005, the researcher conducted a probe study (Knight, 2005) in which the NEMP assessment tasks from 1996 to 2003 were considered to identify those which had significant potential for assessing different kinds of thinking. One of the objectives of the research was to identify particular tasks, presented in a one-to-one interview task approach with videotape, which might be used in subsequent studies to explore the nature of the thinking which was actually used by a sample of students in completing the tasks.

The classification of thinking used in the study was that of the New Zealand Curriculum framework document (Ministry of Education, 1993, p17) which states, among other attributes, that:

*Students will:*

- *Think critically, creatively, reflectively and logically.*

For the purpose of the research, the following working definitions of these kinds of thinking were:

**Critical thinking** is thinking which involves evaluation and, perhaps, challenge.

**Creative thinking** is directed towards solving a problem in one's own way. It often involves imagination and initiative.

**Reflective thinking** involves looking back on one's previous thinking, knowledge and understanding.

**Logical thinking** is directed towards making deductions or presenting arguments.

(Knight, 2005, p3)

The terms are used in the same way in this study.

In identifying suitable tasks for subsequent studies a test format suggested by Halpern (2003, p 361) was used. She writes:

*“My own preference for test format, when the goal is to assess critical thinking, is to use an ecologically valid example with an open-ended response format, followed by specific questions that probe the reasoning behind an answer.”*

The NEMP tasks which have the potential required are those which:

- are in a one-to-one interview format
- are open-ended
- ask for explanations or justifications

The research identified 22 such tasks. 11 of these were critical thinking tasks, 10 reflective thinking tasks and one logical thinking task.

In this study the responses of students to some of the critical thinking tasks is examined.

The 11 critical thinking tasks identified came from four different assessment areas: Science, Art, Technology, and Listening and Viewing.

The best critical thinking task from each of these areas was selected and the videotapes of random samples of about 30 year 4 students and 30 year 8 students responding to these tasks were analysed.

## **1.2 RESEARCH QUESTIONS**

1. What is the nature of the thinking actually used by students in completing the critical thinking tasks?
2. Are there differences between year 4 and year 8 students?
3. Is it possible to identify differences between subject areas in this regard?

## **1.3 INTENDED OUTCOME**

A better understanding of the nature of the thinking of students in completing the NEMP assessment tasks.

## **1.4 USE EXPECTED TO BE MADE OF THE FINDINGS**

The findings may be useful to:

- NEMP in designing further tasks to assess critical thinking skills.
- Teachers who wish to encourage and assess critical thinking skills in their classrooms.

## **2. METHODOLOGY**

### **2.1 THE CHOICE OF TASKS FOR ANALYSIS**

The four tasks chosen for this study were selected on the basis of the work done in the 2005 Probe Study (Knight, 2005).

The tasks, and the reports from which they were taken, are:

Report 13	Science 1999	Page 53	<i>Environmental Issues</i>
Report 18	Aspects of Technology 2000	Page 22	<i>Nut Cracker</i>
Report 10	Listening and Viewing 1998	Page 26	<i>Looking Around</i>

## 2.2 THE STUDENT SAMPLE

The Educational Assessment Research Unit at the University of Otago were asked to supply 30 videotapes, selected at random, at both year 4 and year 8 for each of the four selected tasks, a total of 240 tapes.

At EARU 31 ID numbers for each task were selected expecting that there might be occasional tapes missing. There were also a few tapes without recordings of the task required or which were inaudible. In the end 231 tapes analysed. The numbers of tapes analysed for each task are given below.

	Year 4	Year 8
Science	26	29
Technology	28	29
Listening and Viewing	30	31
Art	<u>29</u>	<u>29</u>
Total	<u>113</u>	<u>118</u>

## 2.3 ANALYSIS

### 2.3.1 The tasks

Each of the tasks was closely examined looking at:

- the marking criteria for the task
- the reporting of student performance
- the involvement of the interviewer in administering the task
- the nature of any probes suggested
- the nature of the thinking which it seemed likely would be required to complete each part of the task.

### 2.3.2 The videotapes

Initially all the videotapes were viewed, focussing particularly on those parts of the task which required critical thinking. Brief notes were written for each student and those tapes containing responses which the researcher thought might be worthy of further consideration were identified.

Each subject area was then considered in turn. The tapes which had been identified as potentially useful were revisited and some of the responses were transcribed.

Comments were written concerning the responses in each subject area

The results of this consideration are presented in the next four sections of this report.

### 3. THE SCIENCE TASK – ENVIRONMENTAL ISSUES

#### 3.1 THE TASK

### Environmental Issues

**Approach:** One to one

**Level:** Year 4 and year 8

**Focus:** Judge the relative importance of seven environmental threats and justify these decisions.

**Resources:** 7 pictures showing environmental disasters.

**Questions/instructions:**

I am going to give you 7 pictures of things happening on the earth that may be of concern to people. You are to put these pictures in 3 rows on the table.

In the **top row** put the picture or pictures which you think people should be most concerned about; the things that you think have the worst effect on the environment of the world.

In the **middle row**, put the picture or pictures that show problems that are not quite so bad.

In the **bottom row** put the picture or pictures that you feel are not as important as the rest.

**Give the student the pictures.**

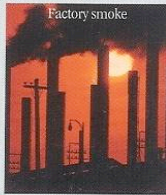
Here are the pictures. Put them into the 3 rows I have described.

**Read instructions again if the student needs it.**

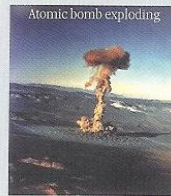
1. Let's look at the top row. You have put [read titles of pictures] in the top row. I want you to tell me why you think these are the most important concerns.

**Encourage use of picture titles in spoken explanations.**

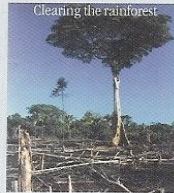
2. Let's look at your bottom row of pictures. You have put [read titles of pictures] in the bottom row. I want you to tell me why you think these are not as important as the other concerns.



**Plastic rubbish on sea animals**



**Factory smoke**



**Oil spill**

**Dumping rubbish at sea**



**Forest fires**

**% responses**  
1999 ('95) 1999 ('95)  
**year 4 year 8**

top row 54 (55) 41 (46)  
middle row 34 (30) 40 (37)  
bottom row 12 (15) 19 (17)

**Atomic bomb exploding**  
top row 65 (71) 69 (79)  
middle row 25 (19) 19 (14)  
bottom row 10 (10) 12 (7)

top row 40 (44) 43 (44)  
middle row 33 (34) 33 (35)  
bottom row 27 (22) 24 (21)

**Clearing the rainforest**  
top row 29 (38) 34 (46)  
middle row 36 (32) 35 (28)  
bottom row 35 (30) 31 (26)

top row 47 (45) 56 (56)  
middle row 37 (36) 37 (37)  
bottom row 16 (19) 7 (7)

top row 52 (49) 54 (49)  
middle row 36 (37) 40 (40)  
bottom row 12 (14) 6 (11)

top row 73 (66) 48 (41)  
middle row 19 (23) 35 (31)  
bottom row 8 (11) 17 (28)

TOP CONCERNS 1999	
YEAR 4	YEAR 8
Forest fires 73%	Atomic bombs 69%
Atomic bombs 65%	Oil spill 56%
Plastic on sea animals 54%	Dumping at sea 54%
Dumping at sea 52%	Forest fires 48%
Oil spill 47%	Factory smoke 43%
Factory smoke 40%	Plastic on sea animals 41%
Clearing rainforest 29%	Clearing rainforest 34%

**Commentary:**

The environmental concerns of year 4 and year 8 students were quite similar, except that year 8 students placed substantially less importance on forest fires (the top concern in 1999 of year 4 students, but only fourth concern of year 8 students). Year 8 students showed greater change from 1995 to 1999, with reduced concern about atomic explosions and clearing the rainforest, but more concern about forest fires. In their reasons both year 4 and year 8 students gave greatest weight to the danger to animal life, then to the danger to human life. Year 8 students gave much greater weight than year 4 students to global (as opposed to local) effects

### **3.2 THE MARKING CRITERIA**

The interviewers were asked to record:

- In which row each picture was placed
- The number of rows created

The student justifications were classified according to whether they included:

- heavy focus
- medium focus
- light focus
- not mentioned or implied

in relation to each of the following possible justifications:

- Geographic spread across the planet
- Danger to human life (in general)
- Danger to plant life
- Lack of reversibility / can not be undone / fixed up (do not include can be fixed or will regrow)
- Direct consequences for personal life
- Danger to animal life

There was then a global rating of judgements as:

- very strong
- strong
- moderate
- weak
- very weak

### **3.3 THE REPORTING OF STUDENT PERFORMANCE**

For each environmental threat the percentage of students placing the picture in the top, middle, or bottom row, was recorded for both year 4 and year 8 students.

The threats were also ranked to show the top concerns at each level.

The only reporting of reasons for choice was in a statement in the *Commentary* that:

*In their reasons both year 4 and year 8 students gave greatest weight to the danger to animal life, then to danger to human life. Year 8 students gave much greater weight than year 4 students to global (as opposed to local) effects.*

### **3.4 THE NATURE OF THE THINKING REQUIRED IN THE TASK**

This was almost certainly the most challenging of the four thinking tasks used in this study.

Initially reflective thinking is required. The students must reflect on their knowledge of the effects of each of the environmental threats. The critical thinking is required when they are asked to perform the ranking of these effects.

This is difficult because:

- all the effects are bad
- some criterion for the ranking is required and, as indicated in the marking schedule, there are at least six possible ranking criteria

This produces a situation which, it seems, is a feature, of many tasks involving critical thinking. If the students has been relatively unsuccessful in the reflective identification of the effects of the environmental threats then the ranking is relatively straightforward. For example if a student only identifies the danger to animal life in each of the threats, the ranking and subsequent justification, is not too difficult. However, if the student were to consider each of the six possible factors mentioned in the marking criteria in relation to each of the seven threats, ranking them would be extremely difficult.

This does not, of course, make it an inappropriate task. On the contrary, in a less structured interview situation the potential to explore the students' critical thinking would be much greater in a question of this kind than in a simpler one.

After they had put the pictures into three rows, the students were asked to justify their ranking. There may be an element of logical thinking in this, but it seemed that for most students it was a matter of thinking reflectively on their previous critical thinking, or lack of it.

### **3.5 EXAMPLES OF STUDENT RESPONSES**

#### **Student A Year 8**

##### **Arranging pictures**

Interviewer reads the instructions and hands all the pictures to the student in a pile.

The student looks through the pile considering each picture in turn.

*Picks out Forest Fire and puts it in top row.*

*Looks through all pictures again and puts Atomic Bomb in top row.*

*Looks through all pictures again and puts Factory Smoke in top row.*

Pause

*Puts Dumping Rubbish in middle row.*

Long pause looking at cards on table and those in hand. Says "I don't know" and then, after a further pause, asks "Can I swap?" Is told that he can.

*Replaces Atomic Bomb in top row with Rainforest and puts Atomic Bomb in middle row.*

*Adds Oil Spill to middle row.*

*Puts Plastic Rubbish in bottom row.*

Final arrangement:

Forest Fire	Rainforest	Factory Smoke
Oil Spill	Atomic Bomb	Dumping Rubbish
Plastic Rubbish		

### Justification

#### Top row:

Forest fire: Need trees for oxygen, and that, and so birds and insects can live in it.

Rainforest: Same for rainforest.

Factory smoke: Is like pollution, you know.

#### Bottom row:

Plastic rubbish: Because they're just fish.

### Student B Year 8

#### Arranging pictures

Interviewer reads out instructions and hands all the pictures to the student in a pile.

Interviewer tells the student to "spread them out".

Student spreads the pictures out seemingly at random.

Initial arrangement:

Plastic Rubbish		
Atomic Bomb	Forest Fire	Dumping Rubbish
Rainforest	Oil Spill	Factory Smoke

Student says "OK" indicating that she is about to start. Her eyes move over all the pictures quite quickly.

*Moves Rainforest to the top row.*

Short pause, looks around all the pictures again.

*Moves Atomic Bomb to top row and Oil Spill to middle.*

*Picks up Dumping Rubbish and Factory Smoke and considers them quite carefully.*

*Puts Dumping Rubbish in top row and Factory Smoke in the bottom.*

*Puts Forest Fire in middle row and Plastic Rubbish in the bottom.*

Final arrangement:

Rainforest	Atomic Bomb	Dumping Rubbish
Oil Spill	Forest Fire	
Factory smoke	Plastic Rubbish	



## Justification

### Top row:

- Rainforest: Well because when you clear the rainforest, like the rainforest affects the lives of everyone. Trees take in carbon dioxide and give out fresh oxygen and if they are gone there won't be much oxygen.
- Atomic Bomb: Atomic bombs exploding is dangerous (*pause*). I don't have exact reasons for that.
- Dumping Rubbish: Because it's pollution and can affect the sea life.

### Bottom row:

- Factory smoke: It is pollution but it can't kill somebody. If they breathed in too much they could die.
- Plastic Rubbish: It's bad but it's like dumping rubbish at sea, if you don't dump rubbish at sea it wouldn't happen to the animals.

## Student C Year 4

### Arranging pictures

Interviewer reads out instructions and hands all the pictures to the student in a pile.

*Student spreads out cards at random.*

Student asks for clarification "What does 'concerned' mean?" Interviewer explains and repeats instructions for top row.

*Student quite quickly selects Forest Fire, Atomic Bomb and Dumping Rubbish for top row.*

Student asks for criteria for middle row to be repeated.

*After a little more consideration picks Oil Spill and Rainforest for middle row.*

*Puts Plastic Rubbish and Factory Smoke in bottom row.*

When asked to justify the bottom row *she moved Plastic Rubbish to the top row*

Final arrangement:

Forest Fire	Atomic Bomb	Dumping Rubbish	Plastic Rubbish
Oil Spill	Rainforest		
Factory Smoke			

## Justification

### Top row

- Forest Fire: Because people could get killed in the fire and badly hurt.
- Atomic Bomb: Because it's bad.
- Dumping Rubbish: Because the animals could get killed and be extinct.
- Plastic Rubbish: Because animals could get killed and couldn't breathe and become extinct

**Bottom row**

Factory Smoke: Because people don't usually get near it and they can't easily get killed from it.

**Student D Year 4****Arranging pictures**

Interviewer reads out instructions and hands all the pictures to the student in a pile.

Student considered each picture in turn and put them directly into the three rows.

*Dumping Rubbish in middle row*

*Atomic Bomb in top row*

*Oil spill in middle row*

*Plastic Rubbish in middle row*

*Rainforest in bottom row*

*Forest Fire in top row*

*Factory Smoke in middle row*

Student looked at last two rows briefly and *moved Plastic Rubbish to bottom row.*

Final arrangement:

Atomic Bomb	Forest Fire	
Dumping Rubbish	Oil Spill	Factory Smoke
Rainforest	Plastic Rubbish	

**Justification****Top row**

Atomic Bomb: Could harm people, like more than one or two.

Forest Fire: Could burn widely and start getting into towns and villages, sort of.

**Bottom row**

Plastic Rubbish: Because we can always ask people to stop. If we see anybody throwing rubbish in the water, plastic, we can tell them to stop doing it because they could harm the animals and they could die.

Rainforest: It's a matter of asking council, or whoever is in charge of the forest, to stop doing it because you can harm some animals and all the trees. There might be some native trees in there.

**3.6 IDENTIFYING CRITICAL THINKING IN THIS TASK**

It seemed that there were four aspects of the video recordings which might give some indication of the student's thinking:

- the physical movements in arranging, and rearranging the pictures in rows
- the time taken to do this
- the number of pictures in each row

- the justifications

### **3.6.1 Arranging the pictures**

The Assessment Result Report indicates that interviewers were asked to read the complete instructions to students before giving them the pictures and then to read the instructions again if necessary. The majority of interviewers in the sample did this, but a few gave the students the pictures first. There were also variations in the way that interviewers gave the pictures to the students. Some handed them to the student in a pile, some handed them one at a time reading the title, and a number spread the cards out in front of the student or suggested to the student that the pictures should be spread out. It was interesting, and perhaps important, to see how these differences might have influenced the responses.

One year 8 student was handed the pictures before the instructions were read and as soon as he heard that he was to put them into 3 rows he did just that and had finished before the interviewer had completed the instructions. When challenged by the teacher he did make some changes, but they were very quick and did not seem to the researcher to be the result of careful consideration. It is unlikely, of course, that this student would have produced a very thoughtful response even if he had waited to be told that this was required.

When the pictures were handed to the students in a pile some just spread them out anyway. Others considered each picture individually placing each on the table in a position relative to the ones before. They seemed to be combining the reflective thinking concerning what were the effects on the environment together with the critical thinking of evaluation against those which they had previously considered. When all the pictures were out most students considered the arrangement as a whole, usually making changes. Another group of students looked right through the pile in their hand one at a time, then went through again picking out any likely candidates for the top row, then for the other rows with more consideration.

If the pictures were initially spread out at random, either by the interviewer or the student, a different strategy seemed to emerge. Rather than considering each picture individually and reflecting on the importance of the environmental threat involved, many of the students seemed to try to deal with all the information in front of them at once. Their eyes flicked from one picture to another looking for a particularly 'bad' one to put in the top row. While this decision was evaluative, and hence involved critical thinking as we have defined it, it seldom seemed to be based on careful reflection. The same, almost intuitive, procedure followed for the rest of the top row and the middle row. Those that were left became the bottom row by default. In many cases it did not seem that the bottom row pictures had really been considered at all which led to considerable difficulty when the students were asked to justify putting them in this position. After completing

the three rows a number of students reviewed their arrangement and made changes, sometimes after obvious reflection and sometimes not.

### 3.6.2 Thinking time

Thinking takes time and it did seem likely that the pauses between actions or responses might be indicators that thinking was taking place.

This task had three sections:

- the arranging of the pictures in rows
- the justification for the top row
- the justification for the bottom row

Interestingly, pauses were a significant feature of the first and third of these sections, but not the second. This seems to reflect the different nature of the sections.

In arranging the pictures the students, as mentioned earlier, needed to combine reflective thinking on the environmental effects of the situations portrayed with the critical thinking required for the ranking these effects. The researcher felt that pauses in this section nearly always indicated careful consideration.

When they came to justifying the top row the students were, in general, just reporting the results of the thinking in the first section and little further reflective or critical thinking was required. There were few pauses.

This was not the case in the justification of the bottom row. For many students the pictures in this row were there by default, they had just not been chosen for the top or middle row. Consequently, pauses here often seemed to indicate a lack of thinking in the original arrangement rather than any positive thinking at the time. This is reinforced by the nature of the responses in which many students, after significant pauses, were unable to provide a justification.

Consequently the researcher felt that, while pauses in the arrangement of the pictures were generally positive indications of thinking, those in the justification sections were more likely to be negative.

The range of times taken to complete the arrangement of the pictures was:

Year 8	38 – 121 seconds	mean: 67 seconds
Year 4	39 – 275 seconds	mean: 113 seconds

The range of times taken to justify the bottom row was:

Year 8	5 – 48 seconds	mean: 22 seconds
Year 4	8 – 240 seconds	mean: 53 seconds

There is a clear difference between the year groups in this regard with year 4 students taking much longer, on average, than year 8 students. The task was very challenging for

most of the younger students. They had limited prior knowledge and understanding of the environmental issues which made both reflective and critical thinking much more difficult. The more successful among them took longer because they were required to think through the consequences for the environment from scratch. The less successful took longer because they simply could not think of any reason to rank one threat above another. They were all bad.

### 3.6.3 The number of pictures in each row

The researcher thought that the number of pictures in each row, or perhaps the number of rows used, might indicate something about the thinking of the students.

The table below indicates the number of pictures in each row in the final arrangement at both year levels.

	Number of pictures										
Top	6	5	5	4	4	4	3	3	3	2	2
Middle	1	2	1	3	2	1	4	3	2	3	2
Bottom	0	0	1	0	1	2	0	1	2	2	3
Year 8	3	1	1	1	6	1	0	10	4	2	0
Year 4	2	1	0	0	4	0	1	10	4	1	1

The instructions specifically asked for three rows and the interviewers varied in how they approached responses which did not conform to this.

One year 8 student only wanted one row “because they are all bad” but was pushed by the interviewer and eventually produced a 322 arrangement. A year 4 student initially produced a 016 arrangement with nothing in the top row. She was asked if that was what she intended and she said yes. The interviewer asked if the student could remember what the top row was supposed to be and she could not. The interviewer then repeated the instructions and the student produced a 421 arrangement in what seemed to be a fairly random manner. The interviewers seemed reasonably happy if the students did not want to put any pictures in the bottom row.

The distribution is quite similar at year 4 and year 8 which perhaps tells us more about the way that the question was interpreted than about the thinking involved.

### 3.6.4 The justifications

After arranging the pictures the students were asked to justify their choice of pictures for the top row and then the bottom row.

#### Issues of most concern

The students were asked to look at their first row and the interviewer said “*I want you to tell me why you think these are the most important concerns*”.

In fact no students in the sample really did this. Nearly all just described what they believed to be the effects of the environmental threats in the top row without indicating why these effects were more important than those in the other two rows.

The following are some examples in addition to those given above:

Year 8	Forest fire:	“Might spread all over the world and burn all the houses and everything.”
	Oil spill:	“Might kill all the animals in the sea and be stink.”
	Atomic bomb:	“Might land on something like radio.”
	Plastic rubbish:	“Kill nearly everything in the sea if they go through it like choke them or squeeze them.”

Year 8	Oil spill:	“...like affects most of the animal life in the sea.”
	Factory smoke:	“...that pollutes a lot of the ozone layer.”
	Atomic bomb:	“...its like all around it, wherever they test it, it wrecks all around it”.

Year 4 “Well, forest fire could burn down all the trees and atomic bomb could blow up the sea and oil spill could catch on fire.”

Two students in the sample did indicate some general principle behind their top row choice, but did not relate that principle to the other rows.

Year 8 Atomic bomb and rainforest:  
“Because they both have a long term effect.”

Year 4 Plastic rubbish, atomic bomb and rainforest:  
“Because they kill a lot of animals and it’s bad for the environment.”

Both went on to describe the effects in more detail.

### **Issues of least concern**

The students were asked to look at their bottom row of pictures and the interviewer said “*I want you to tell me why you think these are not as important as the other concerns*”.

Surprisingly, nearly all the students seemed to realise that, in this case, it would not be sufficient to just describe the effects of the concerns. This presented some difficulty because in many cases the pictures were there by default, they had just not been chosen for the other two rows. Many students gave up. The following are responses of some of these students.

Year 8 “I don’t know.”  
“I don’t like them.”  
“They are all important.”  
“They are just not important.”

“They are important but not as important.”  
“Can’t think of one.”

Year 4 “Not sure.”  
“Don’t know.”  
“Doesn’t look important.”  
“Quite bad but not as bad.”  
“I just put it there.”

Some students did explain their choice but without explicit comparison with the other rows. Some examples follow:

Year 8	Factory smoke:	“It just goes up.”
	Rainforest:	“The wind might have done it.”
	Forest fires:	“We need forest fires.”

Year 4	Factory smoke:	“It won’t do anything.”
	Forest fires:	“Because it happens.”

A few students did relate their reasons to the previous rows.

Year 8	Forest fires:	“More natural and more easily fixed.”
	Atomic bomb:	“Not killing as many creatures because they usually do it in the desert.”

Year 4	Rainforest:	“Not kill the animals as much as all these.”
	Factory smoke:	“Not as dangerous as the other ones, like there’s a lot of factory smoke, it just blows up.”
	Factory smoke:	“Because not as many things die from factory things as the top row.”

### 3.7 COMMENTS

The following are initial comments based on this task. They will be discussed later in conjunction with the comments on the other tasks.

- This is a very difficult task. It involves critical thinking concerning factual knowledge which is not necessarily available to all students, particularly those at year 4. There are also a large number of possible criteria on which the environmental effects may be judged.
- This does not invalidate the task as one in which to evaluate critical thinking. On the contrary, the challenge which the task presented ensured that almost all students were really trying to provide a justifiable arrangement of the pictures. The very obvious ‘thinking pauses’ in the video tapes of this task are evidence of this.
- The most successful strategy in arranging the pictures seemed to be a reflection – action – reflection cycle, with the action being the result of critical thinking. The cycle was repeated until the student was satisfied with the resulting arrangement.

- It seemed that the initial reflection which began the thinking about the task was quite important. Those students who considered all the pictures together initially, because they were spread out by the interviewer or the students themselves, were unlikely to spend time reflecting on a single picture for long with their eyes moving over all the pictures looking for a bad one. Many of them finished up shuffling the pictures around, constantly reviewing the complete arrangement. The students who began by considering each picture individually were much more likely to combine a reflection on each picture with a critical evaluation in relation to those that went before. The difference between these two approaches was particularly evident in the bottom rows of the arrangements. For those considering the group as a whole the bottom row tended to be a default group, they had never really been considered. For those considering each card individually the bottom row had been chosen.
- All the students justified their choices in the top row by explaining what they thought the effects of the environmental threats they had chosen were. Hardly any students explained why they thought these effects were worse than those of the other rows. A further probe would have been valuable. None of the interviewers provided one.
- In the justification of the bottom row choices more students recognised that they would have to compare their choices to those in the other rows. They found this very difficult.



## 4. THE ASPECTS OF TECHNOLOGY TASK – NUT CRACKER

### 4.1 THE TASK

NEMP Report 18: Technology 2000

#### Nut Cracker

#### Trend task

Approach: One to one

Level: Year 4 and year 8

Focus: Describe the features of two different nutcrackers and evaluating them.

Resources: 2 nut crackers, bag of walnuts.

#### Questions/instructions:

In this activity you will be investigating how a nut cracker works and finding out how useful it is.

1. Try to crack open the shell of this nut using only your hands. How did your hands feel trying to crack the nut shell open?



Give student the metal nut cracker.

	% responses	
	2000 ('96) year 4	2000 ('96) year 8
2. Have you ever used one of these?		
yes	32 (39)	45 (49)

3. Put the nut in the nut cracker and try to crack the shell open. How did your hands feel using this nut cracker?

4. How useful do you think the nut cracker is? Why?

Responses to 3 & 4	it cracks nut	53 (51)	54 (42)
it's easy to use ( <i>doesn't hurt hand</i> )		31 (34)	46 (37)
it's easy to control amount of cracking		1 (0)	2 (3)

Give student the wooden nut cracker.

5. Have you ever used one of these?		
yes	6 (8)	9 (5)

6. Now try to crack the shell with this nut cracker. How do your hands feel using this nut cracker?

7. How useful do you think this nut cracker is? Why?

Responses to 6 & 7	it cracks nut	45 (41)	35 (35)
it's easy to use ( <i>doesn't hurt hand</i> )		25 (31)	27 (36)
it's easy to control amount of cracking		4 (0)	6 (11)
fragments are contained		4 (0)	10 (15)

#### Commentary:

Slightly higher percentages of year 8 than year 4 students gave better descriptions of the key differences between the two nutcrackers, and better justifications for their choice of nutcracker. Overall there was very little change



	% responses	
	2000 ('96) year 4	2000 ('96) year 8
8. Can you describe the differences between the two nut crackers and how they work?		
different material ( <i>metal vs. wood</i> )	23 (27)	21 (21)
different visual design	14 (19)	12 (14)
operating principle ( <i>screw vs. lever</i> )	73 (69)	81 (71)
9. Do you think one of these nut crackers is better than the other? Which one?		
Choice: wooden	58 (73)	51 (67)
metal	29 (18)	35 (17)
none	13 (9)	14 (15)
Why?		
Explanation: strong justification	2 (2)	7 (5)
moderate justification	39 (33)	51 (51)
limited justification	53 (59)	37 (38)
no justification	6 (6)	5 (6)
10. Why do you think we have different types of nut crackers?		
different consumer preferences, resources	71 (64)	82 (73)
inventors come up with new ideas	4 (2)	10 (10)
companies, individuals competing for market share	1 (0)	4 (3)
11. What do you think people in the future might use to crack open nuts?		
[not marked]		

## **4.2 THE MARKING CRITERIA**

For both types of nut cracker, the markers were asked to record:

If the student had used one before

If, in discussing the usefulness of the nut cracker, the student mentioned:

It cracks nuts

It's easy to use

It's easy to control the amount of cracking

For the wooden nut cracker, the markers were also asked to record if the student mentioned that the fragments were retained.

In describing the differences markers recorded:

Different materials

Different visual design

Different operating principle

In the explanations for choice between the two types the following possible explanations were suggested to markers:

- how it works
- how durable it is
- how nicely / cleanly it cracks the nut
- how easy it is to use
- how it looks
- what happens to the broken pieces

Markers ranked the explanations as being:

strong, well justified choice

moderately justified choice

limited justification

no justification

In recording the responses to why we have different types, markers recorded:

Different consumer preferences, needs, resources

Inventors come up with new ideas

Companies competing for market share

Monetary gain/profit for person

The last question concerning the future was not marked.

## **4.3 THE REPORTING OF STUDENT PERFORMANCE**

The percentage of students at year 4 and year 8 responding in each of the above categories was reported.

#### 4.4 THE NATURE OF THE THINKING REQUIRED IN THE TASK

This was a much more straightforward thinking task than the previous Environmental Issues one. There were only two things to compare instead of seven and the criteria for comparison were relatively simple. The principal criterion, ease of use, was suggested in the interviewer's questions.

Again the task required reflective thinking first, reflecting on the experience of using each nut cracker and describing the differences between them. Critical thinking was required in evaluating and comparing these experiences.

The question: "why do we have different kinds of nut crackers?" requires reflective rather than critical thinking.

#### 4.5 IDENTIFYING CRITICAL THINKING IN THIS TASK

There did not seem to be many visual clues in the video tapes concerning the nature of the thinking being used. The only physical activity was the use of the nut crackers and the nature of this varied very little between students. While there were some 'thinking pauses' they were much less of a feature here than in the previous task. Students seemed to be able to make their choice and give some sort of justification quite quickly. Consequently it is the verbal justification of the choice between the nut crackers which might indicate the nature of the critical thinking used.

##### The justifications

The initial attempt to crack the nut in their hands, and asking the students how their hands felt after using each cracker, would almost certainly influence the way the students evaluated the two implements. The numbers in the categories of explanation suggested in the marking criteria were not reported, but in the sample reviewed the majority of students only considered ease of use in the justification of their choice. The percentages of students who:

- gave ease of use as their only justification
- said that there was no difference between the two
- gave other reasons

is given below:

	Ease of use	No difference	Other reasons
Year 4	50%	18%	32%
Year 8	68%	7%	25%

Some examples of choice and justification in each category are given below:

### **Ease of use**

Year 4	Wood	You just turn it.
	Metal	It's just easier.
	Wood	Just easier to turn, the other needs a lot of pressure on it.
Year 8	Wood	Better because you don't have to use much muscle.
	Metal	Just have to squeeze it, in this one you have to turn.
	Wood	Not a lot of strain on your hands.

### **No difference**

Few students were asked to justify this response

Year 4          Both good, they both crack nuts.

### **Other reasons:**

Year 4	Wood	The bits stay inside.
	Metal	More stronger, you have more grip on it.
	Metal	Because this one (wood) takes more time twisting it.
	Wood	I just like turning things – bolts and stuff. Probably just as good but some prefer different sorts.
Year 8	Metal	Its easy to use and fits in the drawer.
	Metal	The other just crushes it.
	Wood	It cracks it more efficiently even though this (metal) is faster
	Metal	Might be better for someone with a disability like could only use one hand.
	Wood	Don't have to do much and it keeps the bits inside.
	Iron	It's easier to store and wash. Could use for other things like pliers.

## **4.6 COMMENTS**

- This is a much more straightforward evaluation task than the science one. All students at both year 4 and year 8 in the sample were able to make and justify a choice.
- It is likely that the emphasis on how the student's hand felt in the questions directed the thinking of the students to ease of use in their justifications. From the point of view of an assessment of critical thinking, a more neutral initial discussion of any strengths and weaknesses of the designs, together with some prompts, might have produced more thoughtful responses.



## 5. LISTENING AND VIEWING TASK – LOOKING AROUND

### 5.1 THE TASK

26

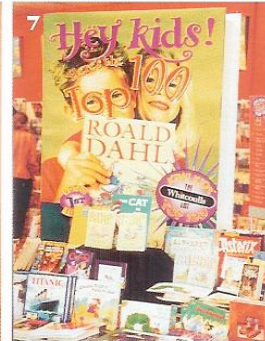
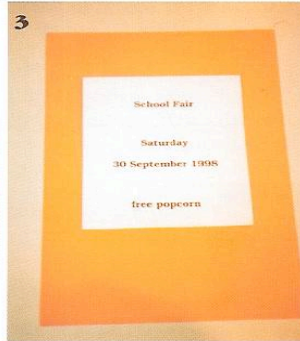
NEMP Report 10: Listening and Viewing 1998

#### Looking Around

Approach: One to one

Level: Year 4 and year 8

Resources: Photographs of seven signs or posters.



#### Questions/instructions

There are lots of interesting signs and posters that you can see in all sorts of places. Some examples are shown in these 7 photographs.

Arrange all 7 pictures in front of student.

- First, I want you to choose one sign or poster from these pictures that you think is very good because of the way it has been made.

Picture chosen: 1 43 32



2 7 10

3 1 0

4 9 17

5 6 6

6 3 3

7 31 32

Allow time. Write picture number on recording sheet.

- Now look carefully at the picture you have chosen. I want you to try to explain to me everything about the way this sign/poster has been made that makes it very good.

Allow time. Possible elements include attractive features, colour, symbolism.

clear with multiple ideas 37 58

relevant, not fully developed 29 30

on right track, but vague 26 10

very limited 8 2

- Now choose one sign or poster that you think is not very good because of the way it has been made.

Allow time.

Picture chosen: 1 1 0



2 8 5

3 63 86

4 6 1

5 9 4

6 10 3

7 3 1

- Now look carefully at the picture you have chosen as being not very good. Try to explain to me why it is not very good — because of the way it has been made.

Student responds. Quality of explanation:

clear with multiple ideas 29 54

relevant, not fully developed 34 35

on right track but vague 24 9

very limited 13 2

- What are some of the things that could be done to make this a really good sign?

Quality of explanation:

clear with multiple ideas 45 75

relevant, not fully developed 26 20

on right track but vague 19 4

very limited 10 1

#### Commentary

This task was designed to assess students' ability to explain why some signs or posters seem effective and others do not. Year 4 and year 8 students chose the same effective and ineffective examples, but year 8 students were substantially better at explaining the reasons for their choices.

## 5.2 THE MARKING CRITERIA

The markers recorded:

The number of the first picture chosen

An evaluation of the explanation of the choice on a scale:

Clear with multiple ideas

Relevant but not fully developed

On the right track but vague

Any other response

Markers were asked to look for: *Ability to pick out attractive features, colour, images and symbolism.*

The number of the second picture chosen.

An evaluation of the explanation of the choice on the same scale as above.

Markers were asked to look for: *Ability to pick out features for comment, and justify why picture is not very good.*

An evaluation of the ideas to improve the picture on the same scale as above.

## 5.3 THE REPORTING OF STUDENT PERFORMANCE

The percentage of students choosing each picture as being either very good or not very good was reported.

The quality of the three explanations was reported by giving the percentage of student responses placed in each of the categories:

Clear with multiple ideas

Relevant, not fully developed

On right track but vague

Very limited

## 5.4 THE NATURE OF THE THINKING REQUIRED IN THE TASK

The task has many similarities to the Science task – Environmental Issues. In both cases the students are presented with seven pictures to evaluate. However, in the Science task students have to place each of the environmental issues represented in the pictures in one of three categories, while in this task they are just asked to pick one sign or poster which is very good and one which is not. They do not necessarily have to be the best and the worst. Consequently less reflectional and critical thinking is involved in this task.

The students still need to review all the pictures and reflect on what might make some of them better than the others. They then need to think critically about the two choices they have made in order to be able to offer an explanation and make suggestions for improving the not very good one

This is consistent with the statement in the commentary which reports:

*This task was designed to assess students' ability to explain why some signs or posters seem effective and others do not.*

## **5.5 IDENTIFYING CRITICAL THINKING IN THIS TASK**

Again, there did not seem to be many visual clues in the video tapes concerning the nature of the thinking being used. The interviewers were asked to arrange all 7 pictures in front of the students and they did this in every case in the sample. There was little movement of the pictures by the students, just a visual scanning over the collection.

There were variations in the time taken both to make the choices and to come up with the explanations, but they were not nearly as significant as in the science task.

As in the nut cracker task, it was the verbal responses in the explanations which indicated the nature of the critical thinking involved.

## **5.6 EXAMPLES OF STUDENT RESPONSES**

### **5.6.1 Choosing a very good picture**

Below are a range of student choices together with the explanations offered for the choice:

#### **Year 8**

**Toy World** The way it's really bright and colourful captures attention. It's got, like, a mascot so everybody knows what it is. It's got a little Lego man standing outside the door and it's got lots of toys and they will catch more people's attention. It's a big building, it's not real small. You can see the words Toy World everywhere it's not, like, tiny and its open seven days, you don't have to worry about going there on Sunday.

**Book display** It's in the shape of a big book. It's got 'Hey Kids' in big letters so kids will see it and come down and read it and then it's got all the top 100 Ronald Dahl books and then it's got all the little books around it. It's colourful and the way they've got the two kids, they've put the 0s in the 100 round their eyes.

**Flames** It's a bit, like, colourful pictures. It's set up real nice, it's a really good design, everything's really done properly, it's really nice.

**Clothes** It's got pictures of clothes you can buy.

**Flames** The lettering, it makes it look like fire.

Clothes Shows, like, how much different kind of clothes there are for kids.

#### **Year 4**

Toy World The writing on it, the people will notice it. They've got their own mascot. The writing is fancy and different colours, it's bright. It's on different angles like the bear is facing that way and the signs facing this way.

Flames Flames, because flames help you cook your food and this sign here 'Flames' is designed like a flame.

Books Says Top 100 and I like Roald Dahl. The Top 100 looks funny because it's like glasses. It looks flash, the colours and the glasses make it look cool.

Flames Because they could eat, and if they don't eat they could die.

Books It's got lots of books.

#### **5.6.2 Choosing a not very good picture and making suggestions for improvement**

Below are a range of responses giving the choice of a not very good picture, the explanation for the choice and the suggestions for improvement.

#### **Year 8**

##### **School Fair**

Explanation It's been made very, very regular and it's normal and many things are normal and people just won't notice it because everything is the same. The typing, it doesn't us like big words, like big words and small words so that if people were at a distance they probably wouldn't see the words. And it's not funny.

Suggestions I would cut these edges and make different shapes and I would type this one with fully lettering and make it bigger. I would put more colour in it.

##### **School Fair**

Explanation Because it's just an orange piece of cardboard and a white sheet of paper with some writing on it, like it doesn't catch anyone's eye. It's not that colourful and it's all plain, nothing too exciting about it. The only thing exciting about it is that that it has got free popcorn.

Suggestions Make it bigger, like a huge sign and real colourful. More details of the school fair like what rides there are going to be and stuff. Make it real big and have popcorn like cups and that with popcorn in it. Yes just give it more colour.



#### School Fair

- Explanation It's got no pictures on it. Stink background.  
Suggestions Add some pictures round here of what's going on.

#### Lotto

- Explanation It doesn't give us any information, it doesn't tell us what Lotto is.  
Suggestions Put more information on, make it look interesting.

### Year 4

#### School Fair

- Explanation It doesn't have very much information, which school? Looks very boring and doesn't have any bright colour or design. It's been done on a computer and they haven't put much work into it.  
Suggestions Put little pictures in the corners here. Put some more information, say which school or whatever and lighten it up with more colours.

#### School Fair

- Explanation Because it's boring, like it's only got three colours, it could have heaps and heaps and heaps of colours. It's not in big letters.  
Suggestions Put borders around it, funny faces, funny pictures around it.

#### School Fair

- Explanation Because they haven't put enough stuff on it.  
Suggestions No ideas.

#### Lotto

- Explanation It's just one word between two posts.  
Suggestions Put it out in the main street, make it bigger.

## 5.7 COMMENTS

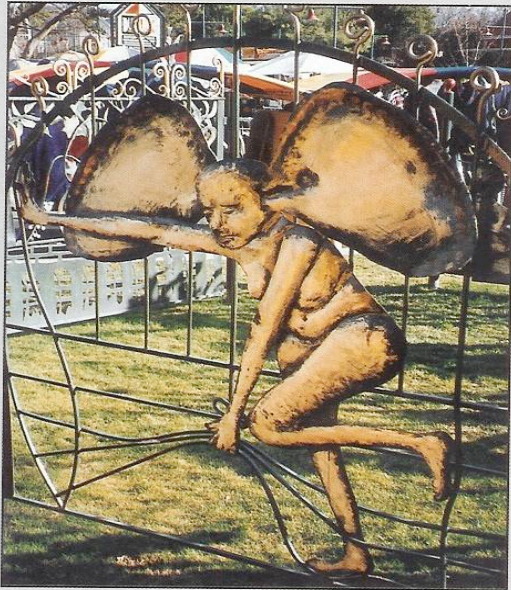
- The quality of the explanations offered by students in this task seemed to be much higher than in the two previous tasks. Multiple ideas were much more common.
- The poster for the school fair was a clear favourite for the not very good poster. The assessment results report indicates that 86% of students at year 8, and 63% at year 4, chose it as not being very good. Perhaps the question would have been better without quite such an obvious choice.
- A number of students, particularly at year four, seemed to be influenced by whether or not they liked the product being advertised as much as by the effectiveness of the advertising. This may account for the difference in the proportion of students choosing the school fair as a not very good poster.
- The combination of being asked to pick out a picture of a sign or poster which was not very good and then to suggest ways of improving it seemed to give a better indication of the student's thinking than the two considered individually.

## 6. THE ART TASK – TWO SCULPTURES

### 6.1 THE TASK



Year 8 sculpture



Year 8 and year 4 sculpture



Year 4 sculpture



## Two Sculptures

**Approach:** One to One Interview

**Level:** Year 4 and year 8

**Resources:** Video recording of two sculptures; colour photographs of each sculpture

### Description

The student viewed a short video recording of two sculptures in public places. The student viewed the first sculpture and responded to interview questions on that sculpture before seeing the second sculpture and responding to associated questions. Colour photographs of the sculptures were used to support discussion after viewing the video (see opposite page).

The teacher administrator asked a sequence of questions to investigate the student's knowledge, understandings and ideas about each sculpture in turn, then the two sculptures in relation to each other. Prompts (in italics) were given only when it was considered necessary to seek or encourage clarification from the student.

The following questions were asked:

- 1 I would like you to describe what you saw in the video. This picture of the sculpture will help you.  
(Give the student the photo)
- 2 What do you think this sculpture is about?  
*What is the purpose of the sculpture?*  
*Why do you think the artist made the sculpture in this way?*  
*What is the artist trying to say with this sculpture?*
- 3 Why do you think this sculpture has been put in a public place?  
*Why do you think the place was chosen for the sculpture?*
- 4 How does the sculpture make you feel?  
*Do you like it? Can you explain why?*

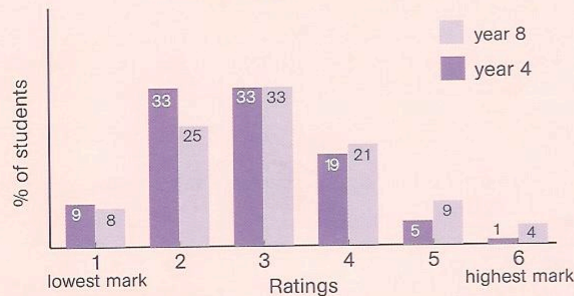
Questions relating the two sculptures.

- 5 I would like you to explain any similarities between the two sculptures — are there any ways in which they are similar? Are there any other ways in which they are similar?
- 6 Could you now explain the important differences between the two sculptures. *Are there any other important differences that you notice?*
- 7 We haven't got a title for the sculptures. I would like you to think of a title or a name for each one.  
*What name or title would you give to each sculpture?*
- 8 Which of the two sculptures do you like most?  
*Why do you like that one most?*  
*Are there any other reasons why you like that one most?*

Although the assessment procedures, interview questions and marking scheme were the same for years 4 and 8, the stimulus pictures of sculptures differed. Accordingly, the results should be read separately.

94 not  
7/8

**Global ratings: Two sculptures**



There is a very similar distribution of ratings across the 6 point scale for both year 4 and 8 students. Although one sculpture was changed between year 4 and 8, the task procedure was the same.

Key attributes (1 low – 4 high)	Qualities	Mean score	
		year 4	year 8
description	identification of images; detail of description; art making information (use of colour, dimensionality, etc.)	2.2	2.4
responsiveness	sense of engagement; curiosity; confidence; feelings, empathy.	2.2	2.3
interpretation	accounting for detail; narrative - tells a story; relevant use of art vocabulary; fluency of ideas; backing opinions with reasons; use of metaphor.	2.0	2.4
<b>Global Rating (1 low – 6 high)</b>		<b>2.8</b>	<b>3.1</b>

## 6.2 THE MARKING CRITERIA

Rather than evaluating each response separately the markers gave an overall ranking of the student's performance in relation to three key attributes:

Description (describes exactly...)

- identification of image
- detail of description
- art making information

Responsiveness (how it makes you feel...)

- sense of engagement
- curiosity
- confidence
- feelings/empathy

Interpretation (naming, why painted this way, similarities/differences...)

- accounting for detail
- narrative – tells a story
- relevant use of art vocabulary
- fluency of ideas
- backing of opinion with reasons
- use of metaphor

In each case the student response was graded on a four point scale:

<i>Undeveloped</i>	<i>Slightly Developed</i>	<i>Moderately Developed</i>	<i>Highly Developed</i>
--------------------	-------------------------------	---------------------------------	-----------------------------

There was then a global assessment of the student's performance on the task as a whole on a six point scale:

*Very Poor*    *Poor*    *Fair*    *Good*    *Very Good*    *Excellent*

## 6.3 THE REPORTING OF STUDENT PERFORMANCE

The mean scores for both years on each of the key attributes was reported together with a graph representing the distribution of the global ratings of students.

## 6.4 THE NATURE OF THE THINKING REQUIRED IN THE TASK

It seems that the majority of thinking in this task is reflective.

Reflecting on:

- what the student saw in the video
- what the sculpture might be about
- why the sculpture is in a public place
- how the sculpture makes the student feel
- the similarities and differences between the sculptures

Suggesting a title for the pictures involves creative as well as reflective thinking.

Critical thinking is required:

- in the prompt to the question which asks the students how the sculpture makes them feel when they are asked if they like the sculpture and why.
- in the last question in which the students are asked which sculpture they like the most and, at least in the prompts, are asked to justify their choice.

## **6.5 IDENTIFYING CRITICAL THINKING IN THIS TASK**

Again, there did not seem to be many visual clues in the video tapes concerning the nature of the thinking being used.

There were variations in the time taken to complete the critical thinking tasks, but they did not seem to be significant.

It was the verbal responses in the explanations which indicated the nature of the critical thinking involved.

## **6.6 EXAMPLES OF STUDENT RESPONSES**

### **6.6.1 Evaluating individual sculptures**

There was much more variation in the input by interviewers in the Art tasks than in the other subject areas. Many interviewers introduced their own prompts and some gave the task question and the prompts at the same time. The interviewers' input is given in italics in the examples of responses below.

#### **Year 8**

##### **Sculpture 1**

##### **Wooden Archway**

##### **Student 1**

*How does the sculpture make you feel?* That it's quite clever.

*Do you like it?* Yes it's OK, it's quite artistic and the shapes they've done are quite good.

##### **Student 2**

*How does it make you feel, this sculpture? Do you like it? Can you explain why?* Because it looks really attractive and has all these different pictures.

**Student 3**

*How does the sculpture make you feel?* Um, nothing, I don't feel anything really, it's just a piece of art.

*And you don't feel anything about works of art?* No not really, it's just a sculpture, it's quite good.

*What do you mean it's quite good?* Well, I don't really like it.

*What is it you don't like?* It's a bit surreal, all you see is bits hanging out and I don't like it.

**Student 4**

*How does the sculpture make you feel?*

It makes me feel quite happy because it's got like lots of different designs.

*Do you like it?* Yes.

*Any other reasons apart from the different designs?* Um, all the ideas sort of show through and you want to know if it means anything.

**Sculpture 2****Grotesque figure on fence/gate****Student 6**

*How does the sculpture make you feel?* It makes me think strangely, like as if you are in another world looking at something strange happening, or a legend.

*Do you like it?* Yes, I think it's quite interesting and nobody would really choose it, they would just choose the same old statue or something, but this sort of - it suits the place.

**Student 7**

*How does the sculpture make you feel?* Stink.

*Do you like it?* No, it just doesn't look any good.

**Student 8**

*How does the sculpture make you feel?* Um, interesting, it reminds me of fairy tales because fairy tales are just made up things just like this and these people being imaginary people.

*What do you mean imaginary people?* I reckon that dog with two, like a normal person and this one with big ears like Dumbo, making a Dumbo and their patterns, I like their patterns.

**Student 9**

*How does the sculpture make you feel?* It makes me feel different about myself. Don't know. Well, I don't think Dick, my brother, has got big ears any more because they stick out but these ones are bigger.

*Do you like it?* Yes, don't know, it's just different, the things you see.

**Student 10**

*How does the sculpture make you feel? Do you like it?* Yes, it's really sort of interesting and it's different the way they've used different metals and stuff to make it.

**Year 4**

**Sculpture 1**

**Boy on tree stump**

**Student 11**

*How does the sculpture make you feel?* Happy.

*Do you like it?* Yes.

*Can you explain why?* Um, because it's a boy instead of a girl.

*You like sculptures about boys?* (student nods)

*Can you explain why?* Because I'm a boy.

**Student 12**

*How does the sculpture make you feel? Do you like it?* Yes.

*How does it make you feel?* If I walked past it I would probably want to climb on it.

**Student 13**

*How does the sculpture make you feel?* Pretty good.

*Do you like it?* Yes.

*Can you explain why?* I like the fairy and the people under the log.

**Student 14**

*How does the sculpture make you feel? Do you like it?* Um, no not really, too dark, should have more colour. I mean it says to me all dark like and something.

**Sculpture 2**

**Grotesque figure on fence/gate**

**Student 15**

*How does the sculpture make you feel?* Funny. It's a cool one I reckon.

*Why does it make you feel like that?* Because they've put in all different kinds of characters together, that one's got big ears.

**Student 16**

*How does the sculpture make you feel?* It makes me want to have a look round it.

*Do you like it?* Yes.

*Can you explain why?* Because some sculptures are funny, they look kind of neat.

**Student 17**

*How does the sculpture make you feel? Do you like it?* No.

*Why don't you like it?* Looks like it's rude.

*OK, it looks like it's rude, why do you say that?* Because people have no clothes on.

**Student 18**

*How does the sculpture make you feel? Do you like it?* Yes.

*Why do you like it?* Because it's about something and because it's interesting.

## Student 19

*How does the sculpture make you feel?* It makes me feel – it makes me laugh quite a lot.  
*Do you like it?* Yes. I like it because it's really funny, it's got funny pictures and they're are like making – and he's got big ears.

### 6.6.2 Comparing the two sculptures

*Which of the sculptures do you like the most and why?*

#### Year 8

- |                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Figure in fence | Well, I think it just attracts me more than this one here. Um, I don't know, it probably wouldn't to adults, it may do but I think this is more of a kid's thing with like the gold and black, mainly the colours, I think, and just the way it's made.                                                                                                                                                                                                  |
| Wooden arch     | Because of my religion. <i>Any other reason?</i> Yeah, because it's a carving.                                                                                                                                                                                                                                                                                                                                                                           |
| Wooden arch     | I like that one because it doesn't take up much room. Like that one's a big circle. It's natural, it's wood, it comes from the trees and stuff and it's got more detail in it. More colours, like that one is just two colours and that one is (counts) three (laughs) but more colourful than that one and it's lighter and darker wood.                                                                                                                |
| Wooden arch     | Because it's more welcoming, it sort of invites you in and the wood gives it a nice tone. <i>Any other reasons?</i> Well this one seems to make more sense, I don't really know what this one is about.                                                                                                                                                                                                                                                  |
| Figure in fence | I'd choose that one because, I wouldn't say more detail, but it may have taken longer and it sort of shows the way someone feels about maybe a dog or something. <i>Are there any other reasons why you like that the most?</i> Because it would probably be something that I couldn't do. I couldn't do one of those but I probably could do one of those because it's wood and you can chip pieces out of wood but you can't chip pieces out of metal. |

#### Year 4

- |                   |                                                                                                                                                                                                                                  |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Boy on tree stump | Because you could go and play on it. <i>Any other reason?</i> You could pretend, you could go and find a log and do what he is doing.                                                                                            |
| Boy on tree stump | Because you understand what it's about more. It's more like real life. Like that's like things that are made out of different animals like half human and half of something else so it's hard to understand what it's all about. |



- Figure in fence      Because it's got lots of background, lots of gates, more than one statue there, and it shows that the statues are actually doing something.
- Boy on a tree stump      Because it's really good done and it's like a little bird and it's done really good. It's not like next to a fence, it's like in nature.
- Boy on a tree stump      It's done properly and there aren't any mistakes and it looks real.

## 6.7 COMMENTS

- A number of interviewers included the first prompt with the initial question and asked: *How does the sculpture make you feel? Do you like it?* In every case where this happened the student ignored the first question and answered the second.
- On the whole, the weaker students at both year 4 and year 8 seemed much more comfortable with this task than with any of the others. This was, perhaps, because in the previous tasks they may have felt that there was a 'correct' answer which they were struggling to find. In this task it was made clear that the interviewer was just interested in their feelings about the sculptures without any indication that the student should be feeling, or thinking, in a particular way. Consequently, there were very few students in the sample who were unable to give an explanation of why they liked or disliked the sculptures.
- The questions and prompts in this task guide the students thinking much more than in the other tasks considered. Before they are asked to evaluate the sculptures they have been asked to consider a number of different factors which might influence their evaluation.
- The additional prompts by interviewers were nearly always productive in exploring the students' thinking.

## 7. DISCUSSION

### 7.1 IDENTIFYING CRITICAL THINKING IN THE INTERVIEWS

#### 7.1.1 Is it possible?

The crucial question in this research is whether or not it has been possible to identify and evaluate the students' critical thinking from the videotapes of their interviews.

If the question is applied to the videotape of an individual student working, the answer is almost certainly no. Neither the tasks nor the interviews were designed for this purpose. They were designed to provide an overview of student performance on tasks which involve critical thinking. To explore in depth the critical thinking of an individual student would require a much less structured interview where each response was followed by further carefully tailored questions that probe the reasoning behind each response. This is clearly not necessary, or even possible, within the framework of the NEMP assessments.

However, if the question is applied to the whole sample of students in his study then the answer is a qualified yes. The researcher feels that, while he cannot be sure what any of the individual students were thinking when they made their responses, there do seem to be some patterns in the responses from which tentative conclusions can be drawn which might be useful.

### **7.1.2 Visual clues**

The only task of the four which involved any physical activity, other than the mechanical nut cracking of the Technology task, was the Science task. Students were to arrange the pictures into three rows. There was undoubtedly a connection between this activity and the students' thinking. Students would pick up a picture study it carefully and, looking at those already on the table, put it in its place. There was also a significant amount of swapping around when all the pictures were down. However there were too many possible reasons for these movements for the researcher to draw any conclusions about the nature of the thinking involved.

Similarly in all the tasks, but particularly in the Science task, there were variations in the time students took to respond to, or complete, sections of the tasks. Some of this time was almost certainly 'thinking time' but, on the evidence of the videotape, it was seldom possible to distinguish between that and 'I'm stuck' time.

### **7.1.3 Explanations and justifications**

The researcher has no doubt that the only reliable evidence of the nature and quality of the critical thinking of the students is in the explanations and justifications which they gave in their responses.

The features which seemed to best indicate effective critical thinking in these responses were the inclusion within the explanation or justification of a choice:

- Multiple ideas
- Comparisons with the object not chosen
- Innovative criteria for comparisons

## **7.2 THE TASKS**

The four tasks used in this study, although similar in that they depended on both reflective and critical thinking for their successful completion, presented different challenges to the students. The major characteristics of each task are discussed below.

### **7.2.1 The Science task**

This was the most difficult task of the four. Some of the reasons for this difficulty are:

- The context of the task is factual and depends significantly on prior knowledge. Each of the issues raised does have a particular effect on the environment.
- The student is required to use reflective thinking to determine this effect for each of the seven environmental threats.

- The student is then required to use critical thinking to compare the effect of each threat against all the others.
- There are a large number of possible criteria which might be used in this comparison.
- The interview protocol contains no prompts to assist the students. For example, none of the students in the sample really explained why the effects of the threats they had put in the top row were more important than those in the other two rows. None of the interviewers probed this lack of response.

### **7.2.2 The Aspects of Technology task**

This task was much more straightforward.

- Although the context still has a factual basis, what the nut crackers actually do, it does not depend on prior knowledge. The student conducts an experiment to determine what the tools do.
- There are only two situations to compare.
- The criteria for comparison are relatively simple and the major one, ease of use, is strongly suggested in the task.
- Although there were no prompts in the protocol the students were prepared for the evaluation by questions concerning the usefulness of each tool and were asked to describe the differences between them.

### **7.2.2 The Listening and Viewing task**

Although it has similarities to the Science task in that there were seven pictures to compare, this task was more straightforward.

- The context was based on opinion more than fact.
- The student only had to select one picture which was very good and one which was not. They did not need to be the best and the worst.
- The criteria for selection are relatively straightforward.
- The students were asked to explain everything about the way the sign/poster was made which made it very good.
- There were no prompts in the protocol but the researcher found few places where he felt they would have been especially useful.

### **7.2.3 The Art task**

This was a challenging task, but in a different way from the Science task.

- The context is based much more on opinion than fact. The students are asked what they think the sculptures are about and how the sculptures make them feel.
- There were only two objects to compare, but unlike the nut crackers, the objects were very different in character.
- The criteria for comparison are less obvious and more subjective than in the other tasks.
- There were a number of prompts in the protocol and these were supplemented by some interviewers.
- As in the Technology task, the interview questions led the students towards the final comparison question.

## **7.3 THE NATURE AND QUALITY OF THE CRITICAL THINKING IN EACH TASK**

### **7.3.1 The Science task**

In the Science task, because it was complex and little help was available in the form of prompts or intermediate questions, students had to devise their own thinking strategies. This seemed to be beyond the ability of many students, particularly those in year 4. A reflection-action-reflection cycle seemed to be successful.

The visual evidence of thinking in the videotapes, both in the physical movement of the cards and in the 'thinking pauses', was much more significant in this task than in the others. However, the verbal responses were still the key to the nature of the thinking.

While students were quite good at explaining why they thought the environmental threats were bad, which depends principally on reflective thinking, they found the critical thinking task of explaining why some were worse than others very much more difficult. Although a global rating of judgements was in the marking criteria, the results of this assessment were not given in the assessment report. However, the researcher felt that, while there was plenty of evidence of good reflective thinking in the student responses in the sample, the critical thinking of students in this task was weaker than in any of the other tasks.

### **7.3.2 The Aspects of Technology task**

This was a much simpler task which did not seem to require the students to come up with their own thinking strategy. They were guided through the task by the questions. Having only two objects to compare and not having to reflect on prior knowledge also made this task much more straightforward.

The critical thinking was needed when the students were asked which tool was better than the other and were asked to explain their choice. The ratings of these explanations were given in the assessment report and 41% of year 4 students and 58% of year 8 students gave a moderate or strong justification. Not many students in the sample saw beyond the fairly obvious criterion for judgement of 'ease of use'. This criterion was also suggested in the task questions. Consequently, the critical thinking of the sample of students in this relatively simple task was still quite weak.

### **7.3.3 The Listening and Viewing task**

The students seemed much more comfortable with this task, perhaps because they did not feel that they were looking for a 'correct' answer. They were being asked for their opinions. The ratings of the quality of the three explanations that the students were asked to make were given in the assessment report. The average percentage of students who were in the top category – clear with multiple ideas – was 37% at year 4 and 62% at year 8. If we include the

second category, for comparison with the Technology task, the average percentage of students who were rated in one or other of these was 63% at year 4 and 74% at year 8.

The task itself did not seem to the researcher to be as simple as the Technology task. The students had to choose two items from seven and they were not led by the interview questions, or prompts, in the way that they should be thinking about this part of the task. In spite of this, or perhaps because of it, the quality of the critical thinking seemed significantly higher. Multiple ideas in the explanations were much more common than in either the Science or the Technology tasks. This might have been because of the instruction to tell the interviewer everything about the picture which made it very good.

### **7.3.4 The Art task**

Like the Listening and Viewing task this task is much more about opinions than facts. However it does have some difficult questions to answer: What do you think the sculpture is about? How does it make you feel? To help with these questions there are a number of prompts for the interviewers to use and a number of interviewers introduced prompts of their own. The quality of the explanations given, which is the major indicator of critical thinking, was not assessed independently in the marking criteria, but was included with other attributes in an 'interpretation' category.

The researcher felt that the quality of critical thinking in the sample of students was highest in this task. It seemed that the context of art appreciation, in which the questions leading to the critical thinking asked the students how they felt and what they liked, gave the students more freedom and led to many more responses containing multiple ideas.

## **7.4 THINKING IN CONTEXT**

### **7.4.1 Fact or opinion**

One difference in the context of the four questions which seemed to influence the quality of the student responses was whether the principal focus of the students' critical thinking was their knowledge of the facts of the situation presented or their opinions concerning the situations. All of the tasks involved elements of both of these but the emphasis did differ significantly between the questions.

- The Science task requires students to identify "...the things that have the worst effect on the environment of the world."

The context is essentially factual. The students are required to think critically about their knowledge of the environmental threats. This should have been followed by critical thinking about their opinion that some situations were worse than others but very few students completed this part of the task successfully.

- The Technology task requires them to say which nut cracker they think "...is better than the other".

In this task the context is almost entirely factual. The students are required to think critically about their experience of using the tools.

- The Listening and Viewing Task requires them to choose something which is very good "...because of the way it is made."

The instruction for students to focus on the way in which the sign or poster is made also introduces a factual context into this question. However, it is clear on reviewing the tapes that there was a strong opinion element in the responses with the students looking something which they liked rather than something that was well made.

- The Art task does not require the student to choose the best, or even a good, sculpture but asks them "Which of the two sculptures do you like the best?"

Here the students are asked to think critically about their opinions and feelings rather than the facts of the situation.

There seems little doubt that this does make a difference. The majority of students seemed much more confident and at ease with the opinion questions, and, in general, their critical thinking responses were much stronger in the opinion based questions than in the factual ones.

The critical thinking responses did, to the researcher, seem to be strongest in art and weakest in science

#### **7.4.2 Subject area differences**

The differences noted above between subject area questions are not surprising. Factual knowledge has a very significant place in science and technology and value judgements are a major part of both listening and viewing and art. Consequently, students may come to a science question with a mind set based on dealing with the facts and to an art question expecting to discuss opinions.

It also seems possible that students are more likely to have been expected to think critically about their opinions in an art class than they are to have been asked to think critically about the facts that they are learning in a science class. This might account for the fact that the researcher felt that the students were much more at ease answering the Listening and Viewing and Art tasks than those in Science and Technology.

#### **7.5 AGE GROUP DIFFERENCES**

Only in two subject areas, Technology and Listening and Viewing, the assessment of the quality of the explanations or justifications was reported and this will give some quantitative measure of the critical thinking at each level in these areas.

In the Technology task the assessment of the explanation given for the choice of nut cracker was:

percentage of responses

	year 4	year 8
strong justification	2	7
moderate justification	39	51
limited justification	53	37
no justification	6	5

In this relatively simple task, not relying very much on prior knowledge, the difference in performance between the two age groups is not very great.

In the Listening and Viewing task the average assessment for the quality of the three explanations required was:

	percentage of responses	
	year 4	year 8
clear with multiple ideas	37	62
relevant, not fully developed	30	28
on right track, but vague	23	8
very limited	10	2

The differences between year 4 and year 8 are more apparent here with the year 8 students better at producing multiple ideas.

The year 4 students certainly found the Science task very difficult. It relies on a considerable amount of prior knowledge and there are a large number of possible criteria for assessing the environmental threats. It was not really a good task to assess the critical thinking skills of year 4 students. It must, of course, be acknowledged that it was not designed to do this.

In the Art task there were certainly some very good responses from year 4 students. However it did seem that year 8 students were a little more likely to use multiple ideas in their justifications.

Overall, there is not enough evidence in this research to make major judgements about the differences in critical thinking skills between year 4 and year 8.

## 8. CONCLUSIONS

Each of the following conclusions should be read remembering that the NEMP assessment tasks used in this study were not principally designed to assess the critical thinking of the students.

1. It was not possible from the videotapes to reliably identify the nature of the critical thinking of individual students. This would require a much less structured interview in which each student response is probed by further questions to determine the reasoning behind it. This is clearly not possible in the context of the NEMP assessments.

2. There did, however, seem to be some patterns in the responses of the samples of students from which tentative conclusions can be drawn.
3. There were possible visual clues to the students thinking in the tapes such as the physical activity of moving pictures around and the periods of inactivity which might have been 'thinking pauses'. However, the researcher became convinced that the only potentially reliable evidence of the nature and quality of the critical thinking of students is in the explanations and justifications which they gave in their responses.
4. The features within these explanations and justifications which seemed to best indicate effective critical thinking were:
  - the presence of multiple ideas
  - comparisons between objects selected and those not selected
  - innovative criteria for comparisons.
5. The quality of the critical thinking used by students did seem to vary between the tasks. The thinking in the Art task seemed to be the best followed by the Listening and Viewing task and then the Aspects of Technology task. The weakest student critical thinking performance was in the Science task. This is discussed in sections 7.2 and 7.3 of this report.
6. A significant factor in this subject area difference may be that the students seemed to find it much more difficult to think critically about their factual knowledge than they did about their opinions. It seems possible that students are more likely to have been expected to think critically about their opinions in an art class than they are to have been asked to think critically about the facts they are learning in the science class.
7. This difference may be important and is worthy of further research.
8. While the critical thinking of year 8 students was, in general, better than that of year 4 students with year 8 students being more likely to use multiple ideas in their judgements, there was not enough evidence in this research to make major judgements about any differences in critical thinking between the groups.
9. Because the interviews are, of necessity for the purposes of NEMP, strongly structured, the influence of individual interviewers was minimal. However, in the Science task some interviewers varied from the instructions in handing out the pictures, and in the Art task some interviewers either gave the question and the probes together or introduced probes of their own. The way in which this influenced the responses is a strong reminder of how important the presentation of the task is in tasks which are used to assess student thinking.

## **9. REFERENCES**

### **9.1 NEMP REPORTS**



*National Education Monitoring Report: Assessment Results.* Dunedin: EARU, University of Otago

2.	<i>Art</i>	1995
10.	<i>Listening and Viewing</i>	1998
13.	<i>Science</i>	1999
18.	<i>Aspects of Technology</i>	2000

## **9.2 OTHER REFERENCES**

Fasko, Daniel, Jr. (ed). (2003). *Critical thinking and reasoning. Current research, theory, and practice.* Hampton Press Inc., Cresskill NJ.

Halpern, Diane F. (2003). *The “how” and “why” of critical thinking assessment.* In Fasko (2003) pp 355 – 366.

Knight,G. (2005). *Critical, creative, reflective and logical thinking in the NEMP assessments* A National Education Monitoring Project Probe Study Report. Dunedin: EARU, University of Otago.

Ministry of Education (1993). *The New Zealand Curriculum Framework.* Wellington: Ministry of Education.

