

**Developing Students' Understandings and
Representations of Statistical Covariation**

by

Jonathan Moritz

(B.A., B.Sc., Dip. Ed.)

**Submitted in fulfilment of the
requirements for the Degree of**

Doctor of Education

University of Tasmania

(November, 2006)

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution.

No material in this thesis has been previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright. Appendix 2 presents material previously published with content related to this thesis.

Signed: _____

Date: _____

Statement of authority of access

This thesis may be made available for loan and limited copying in accordance with the *Copyright Act 1968*.

Signed: _____

Date: _____

Abstract

Statistical covariation refers to the correspondence of variation of two statistical measures that vary along numerical scales. Reasoning about covariation commonly involves translation processes among three representations: (1) numerical data, (2) graphical representations, and (3) verbal statements such as “taller people tend to be heavier.” Two well-known translations are *graph production* and *graph interpretation*. Less well known is the process of *speculative data generation*, involving translating a verbal statement into a possible graph or other data representation. This study explored school students’ reasoning involving these three translation skills through various tasks in surveys and interviews. Evidence is presented concerning methods to assess these skills, and concerning how students as young as third-grade can engage covariation tasks involving familiar contexts. Interviews involved prompting for cognitive conflict using responses from other students, and provided evidence of limited engagement of ideas that were slightly more sophisticated than their own responses.

Responses for each of the three translation skills were described within assessment frameworks involving four levels – Nonstatistical, Single Statistical Aspect, Inadequate Covariation, and Appropriate Covariation – distinguished by the structure of combining correspondence and variation. Distinguishing features of the levels suggested stages of development that may inform instruction. For development from prior beliefs to data-based judgements, tasks involving counterintuitive covariation were designed to prompt students to engage data. For development from single variables to bivariate data, time was observed as a natural covariate, implicit in statements such as “it’s getting hotter,” with a connotation of order that supported pattern recognition of passing time being associated with

corresponding change in a measured variable. For development from single cases to global trends, many students represented correspondence in a single pair of values, at the expense of representing variation. Tasks involving discrete data with few cases, and the use of case labels in responses, were observed to support the view of two data values each linked to the same corresponding case label. This consolidated view of correspondence supported consideration of additional bivariate cases involving variation. Students tended to articulate covariation using the language of comparison and change.

Findings were related to issues in the historical development of coordinate graphing, to findings from educational research in statistics, algebra, science and psychology, and to recommendations within curriculum documents. Student representations of statistical covariation were observed to provide a window into statistical reasoning, and are advocated as a valuable basis for classroom discussions to help develop statistical literacy.

Acknowledgments

I would like to acknowledge my supervisor, Dr Jane Watson, University of Tasmania, for her support over many years. Her guidance in my research career, and particularly during my candidature, encouraged inquiry, rigour, and commitment. Her balance of nudging and patience were amazing. I would also like to acknowledge the examiners who provided helpful suggestions for amendments.

I would like to acknowledge my family, for their nudging and patience. They made personal sacrifices throughout my candidature, including listening to lengthy ramblings, accommodating volumes of paperwork, and permitting the tapping of keystrokes during all hours.

TABLE OF CONTENTS

	Page
CHAPTER 1. INTRODUCTION	1.1
1.01 OVERVIEW OF CHAPTER 1	1.1
1.02 STATISTICAL COVARIATION	1.1
1.03 UNDERSTANDINGS AND REPRESENTATIONS OF STATISTICAL COVARIATION	1.3
1.04 THE SIGNIFICANCE OF THE TOPIC	1.6
1.05 THE STRUCTURE OF THE STUDY	1.7
1.06 THE EVOLUTION OF THE STUDY	1.9
CHAPTER 2. LITERATURE REVIEW	2.1
2.01 HISTORY OF GRAPHING AND COVARIATION	2.1
2.01.01 Development of Coordinates and Functions	2.3
2.01.02 Empirical Data of the 17th and 18th Centuries	2.4
2.01.03 Playfair's Works	2.5
2.01.04 Increasing Recognition of the Power of Graphics	2.7
2.01.05 Contemporary Uses of Graphs	2.8
2.01.06 Historical Implications	2.12
2.02 SCHOOL CURRICULUM	2.13
2.02.01 The Broad Context for Statistics Education	2.13
2.02.02 Covariation in School Curricula	2.15
2.02.03 Algebra in School Curricula	2.16
2.02.04 Data Handling and Graphing in School Curricula	2.17
2.03 RESEARCH LITERATURE	2.21
2.03.01 Previous Research Reviews and Perceived Gaps	2.21
2.03.02 The Structure of the Review of Research Literature	2.24
2.03.03 Graph Interpretation	2.24
2.03.04 Covariation Interpretation	2.29
2.03.05 Graph Production	2.37
2.03.06 Speculative Data Generation	2.43
2.03.07 Summary of Research and Future Directions	2.47
2.03.08 Prior Research Overlapping this Thesis	2.48
2.03.09 Summary of Key Terms	2.49

	Page
CHAPTER 3. RESEARCH DESIGN AND METHODOLOGY	3.1
3.01 SUMMARY	3.1
3.02 RESEARCH AIMS AND EXPECTATIONS	3.1
3.02.01 Aim 1 – Assessment Frameworks of Student Responses	3.2
3.02.02 Aim 2 – Evidence of Student Performance	3.3
3.02.03 Aim 3 – Reasoning and Reactions to Conflicting Responses	3.4
3.02.04 Aim 4 – Conceptual Development across Skills	3.4
3.03 THEORETICAL ORIENTATION	3.5
3.03.01 Theoretical Developmental Model for Assessment	3.5
3.03.02 Misconceptions and Cognitive Conflict	3.7
3.04 DIMENSIONS OF RESEARCH METHODS OF STUDENT ASSESSMENT	3.12
3.04.01 Issues in Task Design for Paper-based Questions	3.13
3.04.02 Issues in Task Design for Interviews	3.17
3.05 DEVELOPING TASKS	3.20
3.05.01 Task Development – Data Collection 1	3.20
3.05.02 Task Development – Data Collection 2	3.23
3.05.03 Task Development – Data Collection 3	3.24
3.05.04 Survey Tasks – Data Collection 3	3.27
3.05.05 Interview Tasks – Data Collection 4	3.35
3.06 DESIGN AND PROCEDURES	3.35
3.06.01 Participants and Procedures Common Across Data Collections	3.35
3.06.02 Data Collection 1	3.37
3.06.03 Data Collection 2	3.40
3.06.04 Data Collection 3	3.40
3.06.05 Data Collection 4 – Interviews	3.42
3.06.06 Presentation of Results of Investigations	3.44
3.06.07 Summary of Research Design	3.45

CHAPTER 4. SPECULATIVE DATA GENERATION	4.1
4.01 SUMMARY	4.1
4.02 INTRODUCTION	4.1
4.03 INVESTIGATION 1: HEART DEATHS VERSUS USE OF MOTOR VEHICLES	4.4
4.03.01 Introduction and Aims	4.4
4.03.02 Task	4.4
4.03.03 Participants and Method	4.5
4.03.04 Quantitative Results	4.8
4.03.05 Level 0 – Nonstatistical	4.10
4.03.06 Level 1 – Single Statistical Aspect	4.12
4.03.07 Level 2 – Inadequate Covariation	4.14
4.03.08 Level 3 – Appropriate Covariation	4.17
4.03.09 Discussion of Investigation 1	4.19
4.04 INVESTIGATION 2: HEIGHT VERSUS AGE	4.22
4.04.01 Introduction and Aims	4.22
4.04.02 Task	4.22
4.04.03 Participants and Method	4.23
4.04.04 Quantitative Results	4.27
4.04.05 Results – Q1&2 – Height Versus Age	4.29
4.04.06 Results – Q3 – Height Versus Age By Sex	4.37
4.04.07 Discussion of Investigation 2	4.43
4.05 DISCUSSION OF SPECULATIVE DATA GENERATION AS EVIDENT IN INVESTIGATIONS 1 AND 2	4.45
4.06 INVESTIGATION 3A: HEIGHT VERSUS AGE (SURVEYS)	4.48
4.06.01 Introduction and Aims	4.48
4.06.02 Task	4.48
4.06.03 Participants and Method	4.48
4.06.04 Quantitative Results	4.49
4.06.05 Results – Q1&2 – Height Versus Age	4.51
4.06.06 Results – Q3 – Height Versus Age By Sex	4.55
4.06.07 Discussion of Investigation 3A	4.57

	Page
4.07 INVESTIGATION 3B: HEIGHT VERSUS AGE (INTERVIEWS).....	4.59
4.07.01 Introduction and Aims.....	4.59
4.07.02 Task.....	4.59
4.07.03 Participants and Method.....	4.61
4.07.04 Results – Overview.....	4.62
4.07.05 Results – Interview Dialogue.....	4.65
4.07.06 Discussion of Investigation 3B.....	4.69
4.08 INVESTIGATION 4A: TEST SCORES VERSUS STUDY TIME (SURVEYS)	4.72
4.08.01 Introduction and Aims.....	4.72
4.08.02 Task.....	4.72
4.08.03 Participants and Method.....	4.73
4.08.04 Quantitative Results.....	4.75
4.08.05 Results – Negative Covariation (Q1).....	4.78
4.08.06 Results – Positive Covariation (Q1*).....	4.88
4.08.07 Results – No Covariation (Q2).....	4.90
4.08.08 Discussion of Investigation 4A.....	4.93
4.09 INVESTIGATION 4B: TEST SCORES VERSUS STUDY TIME (INTERVIEWS)	4.94
4.09.01 Introduction and Aims.....	4.94
4.09.02 Task.....	4.94
4.09.03 Participants and Method.....	4.97
4.09.04 Results – Overview.....	4.97
4.09.05 Results – Interview Dialogue.....	4.100
4.09.06 Discussion of Investigation 4B.....	4.105
4.10 INVESTIGATION 5A: HEART DEATHS VERSUS USE OF MOTOR VEHICLES (SURVEYS)	4.107
4.10.01 Introduction and Aims.....	4.107
4.10.02 Task.....	4.108
4.10.03 Participants and Method.....	4.109
4.10.04 Quantitative Results.....	4.110
4.10.05 Results – Survey Responses.....	4.111
4.10.06 Discussion of Investigation 5A.....	4.114

4.11	INVESTIGATION 5B: HEART DEATHS VERSUS USE OF MOTOR VEHICLES (INTERVIEWS)	4.116
4.11.01	Introduction and Aims	4.116
4.11.02	Task	4.116
4.11.03	Participants and Method	4.117
4.11.04	Results – Overview	4.118
4.11.05	Results – Interview Dialogue	4.119
4.11.06	Discussion of Investigation 5B	4.125
4.12	GENERAL DISCUSSION – SPECULATIVE DATA GENERATION	4.126
	CHAPTER 5. COORDINATE GRAPH PRODUCTION	5.1
5.01	SUMMARY	5.1
5.02	INTRODUCTION	5.2
5.03	INVESTIGATION 6A: TEMPERATURE CHANGE OVER TIME (SURVEYS)	5.5
5.03.01	Introduction and Aims	5.5
5.03.02	Task	5.5
5.03.03	Participants and Method	5.7
5.03.04	Quantitative Results	5.8
5.03.05	Level 0 – Nonstatistical	5.9
5.03.06	Level 1 – Single Statistical Aspect	5.11
5.03.07	Level 2 – Inadequate Covariation	5.13
5.03.08	Level 3 – Appropriate Covariation	5.15
5.03.09	Discussion of Investigation 6A	5.17
5.04	INVESTIGATION 6B: TEMPERATURE CHANGE OVER TIME (INTERVIEWS)	5.20
5.04.01	Introduction and Aims	5.20
5.04.02	Task	5.21
5.04.03	Participants and Method	5.23
5.04.04	Results – Overview	5.23
5.04.05	Initial Response Level 0 – Nonstatistical	5.25
5.04.06	Initial Response Level 1 – Single Statistical Aspect	5.29
5.04.07	Initial Response Level 2 – Inadequate Covariation	5.33
5.04.08	Initial Response Level 3 – Appropriate Covariation	5.35

	Page
5.04.09 Discussion of Investigation 6B.....	5.37
5.05 GENERAL DISCUSSION – COORDINATE GRAPH PRODUCTION.....	5.38
CHAPTER 6. GRAPH INTERPRETATION.....	6.1
6.01 SUMMARY.....	6.1
6.02 INTRODUCTION.....	6.2
6.03 INVESTIGATION 7: TELEPHONE RATE VERSUS CALL DURATION.....	6.5
6.03.01 Introduction and Aims.....	6.5
6.03.02 Task.....	6.5
6.03.03 Participants and Method.....	6.7
6.03.04 Results – Verbal Graph Interpretation (Q1).....	6.7
6.03.05 Results – Numerical Graph Interpretation (Q2).....	6.11
6.03.06 Results – Numerical Graph Interpretation (Q3 and Q4).....	6.13
6.03.07 Discussion of Investigation 7.....	6.16
6.04 INVESTIGATION 8A: NOISE VERSUS NUMBER OF PEOPLE (SURVEYS).....	6.18
6.04.01 Introduction and Aims.....	6.18
6.04.02 Task.....	6.18
6.04.03 Participants and Method.....	6.20
6.04.04 Results – Verbal Graph Interpretation (Q1 and Q4).....	6.20
6.04.05 Results – Numerical Graph Interpretation (Q2 and Q3).....	6.23
6.04.06 Discussion of Investigation 8A.....	6.27
6.05 INVESTIGATION 8B: NOISE VERSUS NUMBER OF PEOPLE (INTERVIEWS).....	6.29
6.05.01 Introduction and Aims.....	6.29
6.05.02 Participants and Method.....	6.29
6.05.03 Results.....	6.30
6.05.04 Discussion of Investigation 8B.....	6.37
6.06 GENERAL DISCUSSION – GRAPH INTERPRETATION.....	6.38

	Page
CHAPTER 7. ASSOCIATIONS AMONG SKILLS	7.1
7.01 INTRODUCTION	7.1
7.02 SIMILARITIES AND DIFFERENCES AMONG ASSESSMENT FRAMEWORKS	7.1
7.03 ASSOCIATION OF TASK 1 (HEART DEATHS) AND TASK 2 (TELEPHONE CALL RATES)	7.3
7.04 RESPONSE COUNTS AMONG TASKS FOR SPECULATIVE DATA GENERATION	7.6
7.05 RESPONSE COUNTS AMONG SKILLS	7.10
7.06 EXAMPLES OF ASSOCIATIONS AMONG RESPONSES	7.15
7.07 DISCUSSION	7.25
CHAPTER 8. GENERAL DISCUSSION	8.1
8.01 ASSESSMENT FRAMEWORKS	8.1
8.02 TASKS TO ASSESS THE SKILLS	8.3
8.03 PERFORMANCE BY GRADE LEVEL AND ACROSS TASKS	8.4
8.04 COGNITIVE CONFLICT	8.7
8.05 EDUCATIONAL IMPLICATIONS	8.8
8.05.01 From Single Data Points to Global Trends	8.9
8.05.02 Ordering Variation to Establish Correspondence	8.11
8.05.03 From Single Variables to Bivariate Data	8.13
8.05.04 From Prior Beliefs to Data-Based Judgments	8.15
8.05.05 Selecting Variables and Structuring Representations	8.16
8.05.06 Graphing Conventions	8.18
8.05.07 Broader Statistical Literacy	8.21
8.06 LIMITATIONS OF THE CURRENT STUDY	8.22
8.07 FUTURE RESEARCH	8.24
8.08 CONCLUSION	8.26
REFERENCES	R.1
LIST OF APPENDICES ON CD-ROM	A.1

LIST OF TABLES

	Page
Table 2.01. <i>Selected Developments in the History of Graphing, Covariation and Functions</i>	2.2
Table 2.02. <i>Summary of Key Terms</i>	2.50
Table 3.01. <i>Numbers of Students Surveyed by Grade and Gender</i>	3.41
Table 3.02. <i>Numbers of Students Interviewed by Gender and Grade</i>	3.44
Table 3.03. <i>Summary of Research Design</i>	3.45
Table 4.01. <i>Characteristics of Four Levels of Speculative Data Generation</i>	4.3
Table 4.02. <i>Response Counts and Percentages of Responses by Grade at each Level and Category of Speculative Data Generation</i>	4.9
Table 4.03. <i>Response Counts and Percentages of Responses by Sex and Grade at each Level and Category of Speculative Data Generation for Q1&2</i>	4.28
Table 4.04. <i>Response Counts and Percentages of Responses by Sex and Grade at each Level and Category of Speculative Data Generation for Q3</i>	4.28
Table 4.05. <i>Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 3 Q1&2</i>	4.50
Table 4.06. <i>Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 3 Q3</i>	4.51
Table 4.07. <i>Summary Descriptions of Interview Dialogue for Q1&2</i>	4.64
Table 4.08. <i>Summary Descriptions of Interview Dialogue for Q3</i>	4.65
Table 4.09. <i>Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 4</i>	4.77
Table 4.10. <i>Summary Descriptions of Interview Dialogue (Negative Covariation format)</i>	4.99
Table 4.11. <i>Summary Descriptions of Interview Dialogue (Positive Covariation format)</i>	4.100

Table 4.12.	<i>Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 1 (revised)</i>	4.111
Table 4.13.	<i>Summary Descriptions of Interview Dialogue</i>	4.119
Table 5.01.	<i>Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Coordinate Graph Production</i>	5.9
Table 5.02.	<i>Summary Descriptions of Interview Dialogue</i>	5.24
Table 6.01.	<i>Characteristics of Four Levels of Verbal and Numerical Graph Interpretation</i>	6.4
Table 6.02.	<i>Response Counts and Percentages of Responses by Grade at each Level of Verbal Graph Interpretation to Q1</i>	6.8
Table 6.03.	<i>Response Counts and Percentages of Responses by Grade at each Level of Numerical Graph Interpretation to Q2</i>	6.11
Table 6.04.	<i>Response Counts and Percentages of Responses by Grade at each Level of Numerical Graph Interpretation to Q3 and Q4</i>	6.14
Table 6.05.	<i>Response Counts and Percentages of Responses by Gender and Grade at each Level of Verbal Graph Interpretation to Q1 and Q4</i>	6.21
Table 6.06.	<i>Response Counts and Percentages of Responses by Gender and Grade at each Level of Numerical Graph Interpretation to Q2 and Q3</i>	6.24
Table 6.07.	<i>Students' Level of Verbal Graph Interpretation by Level of Numerical Graph Interpretation</i>	6.31
Table 7.01.	<i>Frequency of Student Responses at each Level of Speculative Data Generation for Task 1 by Level of Verbal Graph Interpretation and Numerical Graph Interpretation for Task 2</i>	7.4
Table 7.02.	<i>Response Counts and Correlations Among Pairs of Response Levels</i>	7.6
Table 7.03.	<i>Frequency of Student Responses at each Level of Speculative Data Generation for Tasks 3, 4 and 1 (revised) by Level of Speculative Data Generation for Task 3</i>	7.8

	Page
Table 7.04. <i>Frequency of Student Responses at each Level of Speculative Data Generation for Task 1 (revised) by Level of Speculative Data Generation for Task 4</i>	7.9
Table 7.05. <i>Frequency of Comparison of Response Levels of Speculative Data Generation for Four Tasks</i>	7.9
Table 7.06. <i>Response Counts and Correlations Among Pairs of Response Levels</i>	7.11
Table 7.07. <i>Frequency of Student Responses at each Response Level for Tasks 5 and 6 by Response Level for Tasks 4 and 5</i>	7.12
Table 7.08. <i>Frequency of Student Responses at each Level of Numerical Graph Interpretation for Task 6 by Level of Verbal Graph Interpretation for Task 6</i>	7.13
Table 7.09. <i>Frequency of Comparison of Response Levels of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation</i>	7.14

LIST OF FIGURES

	Page
<i>Figure 1.01.</i> Forms of representing statistical covariation and skills to translate them.....	1.4
<i>Figure 2.01.</i> Graph from Humphrey Bear, Nine Network, 2001.....	2.10
<i>Figure 2.02.</i> Weather graph from WIN television news (19 January, 2002).....	2.11
<i>Figure 2.03.</i> Percentiles of height versus age for females (Department of Health and Human Services, Tasmania, 2000).....	2.12
<i>Figure 2.04.</i> Scatterplot multiple-choice test item.....	2.21
<i>Figure 3.01.</i> Five possible response sequences.....	3.10
<i>Figure 3.02.</i> A sample task and coding scheme from TIMSS (1995).....	3.16
<i>Figure 3.03.</i> Task 1: Heart deaths versus use of motor vehicles.....	3.22
<i>Figure 3.04.</i> Task 2: Telephone call rate versus call time, based on a newspaper extract.....	3.23
<i>Figure 3.05.</i> Task 3: Height versus age.....	3.24
<i>Figure 3.06.</i> Task 1 (revised): Heart deaths versus use of motor vehicles.....	3.28
<i>Figure 3.07.</i> Task 4: Test scores versus study times.....	3.30
<i>Figure 3.08.</i> Graph of average mathematics scores by students' report on time spent daily on mathematics homework at grades 4, 8, and 12 from the National Assessment of Educational Progress, 2000 Mathematics Assessment.....	3.31
<i>Figure 3.09.</i> Task 5: Temperature change over time.....	3.33
<i>Figure 3.10.</i> Task 6: Noise level versus number of people.....	3.34
<i>Figure 4.01.</i> Task 1 to assess Speculative Data Generation based on a newspaper extract (The Mercury, June 11, 1991, p.2).....	4.5
<i>Figure 4.02.</i> Levels and categories of Speculative Data Generation for Task 1.....	4.7
<i>Figure 4.03.</i> Student responses – Level 0 – Nonstatistical.....	4.11
<i>Figure 4.04.</i> Student responses – Level 1 – Single Statistical Aspect.....	4.13
<i>Figure 4.05.</i> Student responses – Level 2 – Inadequate Covariation.....	4.16
<i>Figure 4.06.</i> Student responses – Level 3 – Appropriate Covariation.....	4.18
<i>Figure 4.07.</i> Task 3 to assess Speculative Data Generation.....	4.23

	Page
<i>Figure 4.08.</i> Coding scheme for levels and categories of response to Q1&2.....	4.25
<i>Figure 4.09.</i> Coding scheme for levels and categories of response to Q3.....	4.26
<i>Figure 4.10.</i> Student responses to Q1&2 – Level 1 – Single Statistical Aspect.....	4.30
<i>Figure 4.11.</i> Student responses to Q1&2 – Level 2 – Inadequate Covariation.....	4.32
<i>Figure 4.12.</i> Student responses to Q1&2 – Level 2 – Inadequate Covariation.....	4.34
<i>Figure 4.13.</i> Student responses to Q1&2 – Level 3 – Appropriate Covariation.....	4.36
<i>Figure 4.14.</i> Student responses to Q3 – Level 1 – Single Statistical Aspect.....	4.38
<i>Figure 4.15.</i> Student responses to Q3 – Level 2 – Inadequate Covariation.....	4.40
<i>Figure 4.16.</i> Student responses to Q3 – Level 3 – Appropriate Covariation.....	4.42
<i>Figure 4.17.</i> Student responses to Q1&2 – Level 1 – Single Statistical Aspect.....	4.52
<i>Figure 4.18.</i> Student responses to Q1&2 – Level 2 – Inadequate Covariation.....	4.53
<i>Figure 4.19.</i> Student responses to Q1&2 – Level 3 – Appropriate Covariation.....	4.55
<i>Figure 4.20.</i> Student response to Q3 – Level 1 –Single Statistical Aspect.....	4.56
<i>Figure 4.21.</i> Student response to Q3 – Level 2 – Inadequate Covariation.....	4.56
<i>Figure 4.22.</i> Student responses to Q3 – Level 3 – Appropriate Covariation.....	4.57
<i>Figure 4.23.</i> Prompts displayed to students to promote cognitive conflict for Q1.....	4.60
<i>Figure 4.24.</i> Prompts displayed to students to promote cognitive conflict for Q2.....	4.60
<i>Figure 4.25.</i> Prompts displayed to students to promote cognitive conflict for Q3.....	4.61
<i>Figure 4.26.</i> Response to Q2 with difficulty showing non-growth (G9f1).....	4.68
<i>Figure 4.27.</i> Task 4 to assess Speculative Data Generation.....	4.73
<i>Figure 4.28.</i> Levels and categories of Speculative Data Generation for Task 4.....	4.75
<i>Figure 4.29.</i> Student responses – Level 0 – Nonstatistical.....	4.79
<i>Figure 4.30.</i> Student responses – Level 1 – Single Statistical Aspect.....	4.80
<i>Figure 4.31.</i> Student responses – Category 2A. Double and Group Comparison.....	4.81

	Page
<i>Figure 4.32.</i> Student responses – Level 2 – Inadequate Coordinate.....	4.82
<i>Figure 4.33.</i> Student responses – Category 3A. Series Comparison.....	4.84
<i>Figure 4.34.</i> Student responses – Category 3B. Table.....	4.86
<i>Figure 4.35.</i> Student responses – Category 3C. Coordinate Variable.....	4.87
<i>Figure 4.36.</i> Student responses – Level 1 – Single Statistical Aspect.....	4.88
<i>Figure 4.37.</i> Student responses – Level 2 – Inadequate Covariation.....	4.89
<i>Figure 4.38.</i> Student responses – Level 3 – Appropriate Covariation.....	4.91
<i>Figure 4.39.</i> Student responses to No Covariation task format (Q2).....	4.92
<i>Figure 4.40.</i> Graphs used as prompts (Negative covariation).....	4.95
<i>Figure 4.41.</i> Graphs used as prompts (Positive covariation).....	4.96
<i>Figure 4.42.</i> Student response – Double Variable (G3f3).....	4.101
<i>Figure 4.43.</i> Student response – Double Comparison (G7m3).....	4.103
<i>Figure 4.44.</i> Task 1 (revised), to assess Speculative Data Generation (The Mercury, June 11, 1991, p. 2).....	4.109
<i>Figure 4.45.</i> Student responses – Level 0 – Nonstatistical.....	4.112
<i>Figure 4.46.</i> Student responses – Level 1 – Single Statistical Aspect.....	4.112
<i>Figure 4.47.</i> Student responses – Level 2 – Inadequate Covariation.....	4.113
<i>Figure 4.48.</i> Student responses – Level 3 – Appropriate Covariation.....	4.114
<i>Figure 4.49.</i> Responses used as prompts to create cognitive conflict.....	4.117
<i>Figure 4.50.</i> Survey response of an interviewee at Level 1.....	4.120
<i>Figure 4.51.</i> Survey response of an interviewee at Level 2.....	4.121
<i>Figure 4.52.</i> Survey responses of interviewees at Level 3.....	4.122
<i>Figure 5.01.</i> Levels and categories of Coordinate Graph Production.....	5.4
<i>Figure 5.02.</i> Task 5 to assess Coordinate Graph Production.....	5.6
<i>Figure 5.03.</i> Student responses – Level 0 – Nonstatistical.....	5.10
<i>Figure 5.04.</i> Student responses – Level 1 – Single Statistical Aspect.....	5.12
<i>Figure 5.05.</i> Student responses – Level 2 – Inadequate Covariation.....	5.14
<i>Figure 5.06.</i> Student responses – Level 3 – Appropriate Covariation.....	5.16
<i>Figure 5.07.</i> Graphs used as prompt in interviews.....	5.22
<i>Figure 5.08.</i> Survey responses – Level 0 – Nonstatistical.....	5.26
<i>Figure 5.09.</i> Survey responses – Level 1 – Single Statistical Aspect.....	5.30
<i>Figure 5.10.</i> Survey responses – Level 2 – Inadequate Covariation.....	5.33

	Page
<i>Figure 5.11.</i> Survey responses – Level 3 – Appropriate Covariation.....	5.36
<i>Figure 6.01.</i> Task 2 to assess Verbal Graph Interpretation (Q1) and Numerical Graph Interpretation (Q3 and Q4; Q2) based on a newspaper extract (The Mercury, 22 July, 1993, p. 17)	6.6
<i>Figure 6.02.</i> Task 6 to assess Verbal Graph Interpretation (Q1 and Q4) and Numerical Graph Interpretation (Q2 and Q3)	6.19
<i>Figure 6.03.</i> Student response – Level 2 – Inadequate Covariation.....	6.25
<i>Figure 6.04.</i> Student response – Level 3 – Appropriate Covariation.....	6.26
<i>Figure 7.01.</i> Proposed development order of four skills for understanding and representing statistical covariation.....	7.15
<i>Figure 7.02.</i> Responses from a Grade 3 female.....	7.16
<i>Figure 7.03.</i> Responses from a Grade 3 female.....	7.18
<i>Figure 7.04.</i> Responses from a Grade 5 male.....	7.19
<i>Figure 7.05.</i> Responses from a Grade 7 female.....	7.21
<i>Figure 7.06.</i> Responses from a Grade 7 male.....	7.22
<i>Figure 7.07.</i> Responses from a Grade 7 male.....	7.23
<i>Figure 7.08.</i> Responses from a Grade 9 male.....	7.24
<i>Figure 8.01.</i> Two graphs of the same data for X and Y.....	8.18

CHAPTER 1. INTRODUCTION

1.01 OVERVIEW OF CHAPTER 1

An overview of aspects of understanding and representing of statistical covariation is provided. The significance of the topic is considered in relation to the general importance of understanding basic statistical covariation for statistical literacy. The significance of the study is considered in relation to the existing research literature for statistical literacy, graphical understanding and aspects of algebraic understanding, and the lack of research linking these basic concepts for students at primary- and middle-school levels. The structure of the thesis is then outlined chapter by chapter, including discussion of what each chapter contributes to the exploration of the topic.

1.02 STATISTICAL COVARIATION

Covariation, in broad terms, concerns correspondence of variation of two variables. The nature of the covariation may be categorized according to the variation possible in the measure of each variable involved. The values of the variables may be said to involve some form of relationship, association, function, dependency, or correspondence. For logical variables, which can be either True or False, the logical statement $A = \text{NOT}(B)$ expresses *logical covariation* between A and B , since varying the value of A from True to False entails a corresponding variation in the value of B from False to True to maintain the equation as true. The equation $y = 2x$ expresses *numerical covariation* between real-number variables x and y , since a variation in the value of either x or y entails a corresponding variation in the value of the other variable. Other polynomial and piecewise functions also express numerical covariation.

Statistical covariation refers to the correspondence of variation of two statistical variables that vary along numerical scales. An example of statistical covariation is the relationship between the height and weight of people: taller people tend to be heavier. Statistical covariation may be represented in scatterplots using a Cartesian coordinate system that shows the correspondence of the ordination of each variable. Coordinate graphs employ a general feature that “position denotes value” and apply this feature in two-dimensional space to represent values of two variables. They represent both (a) the data points, emphasizing the *correspondence* of values of two variables, and (b) general trends, emphasizing *variation* of the two variables due to the ordination of the values along each axis.

The more general term *statistical association* may refer also to associations between two categorical variables, commonly represented in two-way frequency tables, and between one categorical and one interval variable, often formulated as the comparison of groups. Statistical association involves more than just a relation of values; it is a relation of measured quantities of distinct characteristics because data are “not merely numbers, but *numbers with a context*” (Moore, 1990, p. 96). Much work in the social and physical sciences concerns attempts to use statistical association as evidence of causal association between two characteristics, which may be used to enhance the prediction or control of one variable by knowledge or manipulation of the other variable. Many statistical associations do not fit perfectly the deterministic models of logical or numerical covariation just described; that is, there is variation from the model. Tests of statistical significance are required to measure the degree to which data fit or vary from one of these models. Formal measures of statistical covariation depend on the type of variation of the measures of each variable involved: χ^2 tests may be used to judge the significance of the

association between categorical variables, and *t*-tests or analyses of variance may be used to judge the significance of differences of mean values of an interval variable across groupings of a categorical variable. For statistical covariation, which involves two numerical variables, Pearson correlation coefficients are commonly used to measure the degree of association and may be tested for the significance of the linear fit of covariation between the variables. Much of the discussion in this study focuses on statistical covariation in the restricted sense of being considered in relation to degree of fit to a linear function, as opposed to polynomial or piecewise models.

1.03 UNDERSTANDINGS AND REPRESENTATIONS OF STATISTICAL COVARIATION

The three representations of statistical covariation considered in this study include (a) raw numerical data or tabular summary data, (b) graphical representations such as scatterplots, line graphs, or bar graphs, and (c) verbal statements about statistical covariation such as “taller people tend to be heavier.” More formal representations of statistical covariation appropriate to higher levels of mathematical competence include symbolic expressions of algebraic functions indicating the best fit of numerical covariation.

Reasoning about statistical covariation, to demonstrate understanding, commonly involves translation processes among raw numerical data, graphical representations, and verbal statements about statistical covariation and causal association. A comprehensive taxonomy of translations among words, graphs, tables of data, and algebraic formulae was described in seminal work by Janvier (1978; Bell & Janvier, 1981) and others who subsequently considered these translation skills (e.g., Coulombe, 1997; Coulombe & Berenson, 2001; Gagatsis & Shiakalli, 2004; Kieran, 1993; Swan, 1985, 1988). Modelling skills included plotting and sketching

graphs and curve fitting, and interpreting skills included reading and interpreting values and patterns. Similar processes of re-presenting data, with the intent to obtain or convey meaning about the measures in the contextual situation, have been described as “transnumeration” by Pfannkuch and Wild (2004).

The translation processes for statistical covariation investigated in this study are shown in Figure 1.01. The representations exclude algebraic equations and include causal statements for two reasons: the investigations employed in this study focus on covariation in statistical contexts often with causal connotations and focus on assessing common translations that are relevant to most school students, rather than with algebraic expressions that may be expected of only secondary mathematics students.

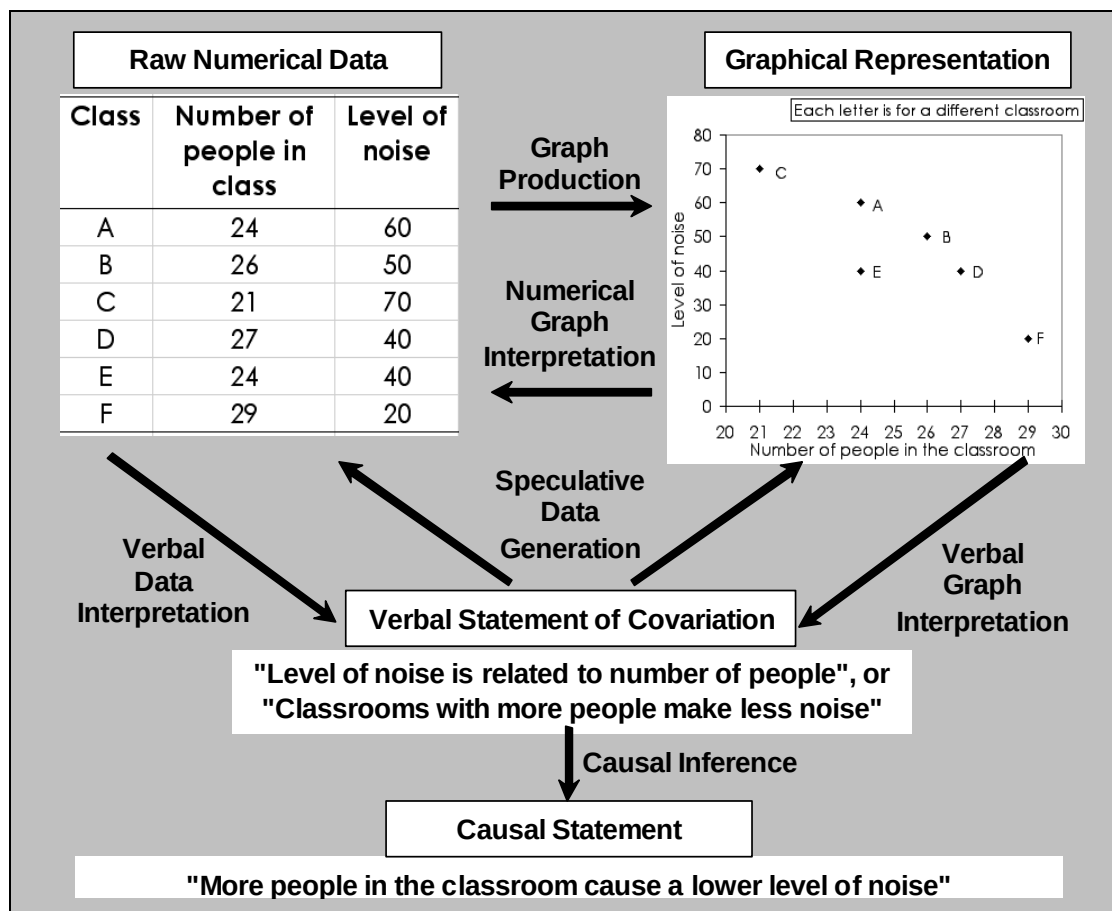


Figure 1.01. Forms of representing statistical covariation and skills to translate them.

Graph production and graph interpretation are frequently recommended as topics for students in schools. In daily life such as reading the newspaper, however, adults rarely engage in the data analysis sequence of Graph Production, Verbal Graph Interpretation, followed by Causal Inference. Many newspaper reports and advertisements make verbal statements that involve causal claims, but only some use graphs to illustrate the statistical data that lie behind the claims. More commonly, adults read a causal statement based on a statistical association, and in order to understand and evaluate it critically, they must imagine what statistical data lie behind it, that is, they must engage in *Speculative Data Generation*. For verbal statements of covariation, Speculative Data Generation requires an understanding of numerical covariation, an understanding of statistical variation, and a contextual understanding of how the data might have been collected and measured. Tasks of Speculative Data Generation have some degree of freedom in the speculation of what was lost in the forward process of data interpretation to arrive at the verbal statement. For assessment purposes, this reverse type of task may be more informative of student understanding than interpretation tasks, as students are required to supply more detail in their responses. Speculative Data Generation, Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation were investigated in this study. Verbal Data Interpretation and Causal Inference were not explored in this study, however appendices provide preliminary evidence of these skills.

Developing students' understandings and skills is at the heart of education. An assessment orientation involves determining valued aspects of understanding upon which to evaluate, and criteria for evidence of this understanding, including intermediate steps towards developing understanding. A teaching orientation to development of understanding concerns interventions to promote learning in students

along these developmental steps. The current study approached the topic with an assessment orientation to create and explore frameworks that detailed steps in students' understandings and representations of statistical covariation. Evidence for these frameworks was provided by observation of features of students' responses that were evaluated according to levels. The discussion of characteristics of responses aimed to provide a clear detailing of possibilities and issues for educational assessment, as well as implications of these that may form the basis for teaching approaches.

1.04 THE SIGNIFICANCE OF THE TOPIC

Statistics is becoming increasingly significant in daily life. Newspaper reports and advertisements, for example, make verbal statements based on research studies that involve statistical data and use graphs to illustrate statistical data. Interpreting these verbal statements and graphs is important in order for students to develop an appreciation of what is presented and to evaluate critically the conclusions that are drawn from data. Statistics educators recently have described characteristics of statistical literacy that encompass the skills and dispositions necessary to apply statistical reasoning in social contexts (e.g., Gal, 2002, 2004). Watson (1997), for example, proposed a statistical literacy hierarchy including three tiers: a basic understanding of statistical terminology, an understanding of statistical concepts when embedded in applied contexts, and a questioning attitude to challenge claims made without proper statistical foundation. Gal also suggested that general literacy and contextual knowledge are necessary to engage with authentic settings in everyday living.

Statistical covariation, or more broadly statistical association, is a core concept in the social and physical sciences. These sciences commonly aim to infer

causation by collecting bivariate data that involve statistical association and by controlling for other variables. Some examples of causal inferences include “People’s sex influences their income,” or “Increased caffeine intake increases the risk of heart problems.” In order to understand such statements, it is necessary to be aware of the translation processes among raw numerical data, graphical representations and verbal summaries, and to have an understanding of what constitutes a statistical association when presented in one of these forms. It is also important that students are sensitive to the possibility of bias or error in these translations, and to the need to consider other potentially confounding variables when attempting to draw inferences of causal association.

Most research into developing understanding of statistical covariation has come from tasks involving graphs. The broader research literature on graphing has often reported on pointwise tasks of construction and interpretation, such as plotting points or locating values (Leinhardt, Zaslavsky, & Stein, 1990). Tasks involving variation and qualitative graphs—that is, without specific data values—have been considered by some researchers (Leinhardt et al., 1990) to be an underutilized avenue for exploring understanding of general features of graphs, including covariation. Drawing a graph to illustrate a verbal statement of covariation requires Speculative Data Generation; such tasks are rarely found in curricula or research.

1.05 THE STRUCTURE OF THE STUDY

The current chapter forms an introduction to the topic of the study defining the problem in broad terms and its significance.

Chapter 2 is a literature review providing background about the history of approaches to covariation and graphing, about school curricula and the place of covariation, and about previous research relevant to the current study. *The history*

illustrates some of the sources, difficulties, milestones and the protracted period involved in the evolution of understanding and representing covariation, particularly of coordinate graphs. These historical developments are noted to have educational implications as students engage similar issues in learning within a few years what took academics centuries to develop. *School curricula* from various English speaking countries and Australian states are surveyed in relation to the expectations for school students' understandings and representations of statistical covariation. *Previous research* literature is considered in relation to interpreting graphs and covariation, and then producing graphs and generating speculative data. The chapter concludes with a summary of gaps in the literature that indicate potential for investigating and informing the thesis topic, and preliminary local research overlapping this thesis.

Chapter 3 presents the research design used to investigate the topic. Research aims are formulated and plans to address these questions are outlined. Tasks are presented to assess students' understanding and details are provided of the procedures and samples of student responses data. Coding methods are described.

Chapters 4, 5, and 6 report analyses of students' responses to tasks for each of the skills of understanding and representing covariation: Chapter 4 concerns Speculative Data Generation, Chapter 5 concerns Coordinate Graph Production, and Chapter 6 concerns Verbal Graph Interpretation and Numerical Graph Interpretation. These skills are further sub-divided and related to individual tasks that were investigated by written survey and/or interview methods. Each analysis describes student responses within frameworks of up to four levels of response, and details further categories and features of developing each skill.

In Chapter 7, relationships among various skills and responses to various tasks are discussed and indications of the ordering of conceptual development are

considered. Chapter 8 discusses the findings and educational implications.

Limitations of the study and possibilities for future research and teaching initiatives are outlined.

A comprehensive set of references cited is provided. Appendices on CD-ROM include papers published as part of this doctoral study, administrative aspects of the studies, samples of student survey data, video extracts and transcripts from interviews, analysis artefacts, and a list of additional resources consulted but not cited in this study.

1.06 THE EVOLUTION OF THE STUDY

The study was based upon four phases of data collection undertaken (see Chapter 3). In the first instance, two survey tasks, both from newspaper articles, were used as part of a wider research project, and were the basis for formulating the remainder of the study. Of interest were the difficulties encountered by students. Were these difficulties due to the nature of the variables and how they were embedded in the newspaper context, or were they due to the nature of thinking about statistical covariation between two variables? This prompted the use of an additional survey task of height versus age, which obtained graphical responses showing more evidence of appropriate understanding. It was only after these preliminary stages that the final data collection occurred with a view of comparing and contrasting the skills of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation. This historical nature to the study is detailed further in Chapter 3, and is signalled by the summary of findings from Investigations 1 and 2, which informed later investigations in relation to task design of surveys and interviews.

CHAPTER 2. LITERATURE REVIEW

This chapter provides a history of approaches to covariation and graphing, and a review of current school curricula and the place of the covariation within it. The review of previous research literature relevant to the current study concludes with a discussion of the gaps in the literature and preliminary research overlapping the current study.

2.01 HISTORY OF GRAPHING AND COVARIATION

Various accounts of the historical development of graphing (Beniger & Robyn, 1978; Biderman, 1990; Friendly & Denis, 2002; Funkhouser, 1937b; Hankins, 1999; Reidhaar, 1984; Tilling, 1975; Wainer & Velleman, 2001) indicate some of the sources, stages, difficulties, and the long time involved in the evolution of graphs, particularly in relation to understanding functional and statistical covariation. Table 2.01 shows a selection of milestones prior to 1800, in particular developments in the 17th and 18th centuries. The following subsections chronicle some of this historical development. The significance of this historical development for the current study is that students may engage similar issues in their own development of graphical understanding.

Table 2.01.

Selected Developments in the History of Graphing, Covariation and Functions

Year	Person	Development
c3200BC	Egyptians	Coordinate systems for surveying land
c1000AD	Unknown	Curves of planetary orbits on a time grid
c1350	Nicole Oresme	Proto-bar graph of theoretic functions
c1600	Galileo Galilei	Functional relation between period and length of pendulum
1637	René Descartes	Coordinate system for analytical geometry
c1660	Christopher Wren	Automatic recording of temperature over time
1669	Christiaan Huygens	Coordinate plot of data for proportion survival vs age; curve-fitting to interpolate life expectancies
1685	Robert Plot	Broken line graph of daily barometric readings
1686	Edmund Halley	Bivariate plots of barometric reading vs altitude
1755	Leonhard Euler	Function defined as quantity depending on another variable quantity
1760-5	Johann Heinrich Lambert	Curve fitting and interpolation; measurement error
1765	Joseph Priestley	Time lines to compare life-spans
1786	William Playfair	Bar chart
1796	James Watt	Automatic recording of pressure vs volume in a steam engine
1800s	Francis Galton, Charles Edward Spearman and Karl Pearson	Measures of rank correlation and product-moment correlation
1830s-1930s	Lejeune Dirichlet and Nicolas Bourbaki	Gradual acceptance of functions defined as arbitrary correspondence between real numbers, rather than based upon the notion of dependency

2.01.01 Development of Coordinates and Functions

Statistical graphs, including time-series graphs and scatterplots, were infrequent before the development of graphic design by Playfair in the late 1700s (Tuft, 1983, p. 9). As early as 3200 BC, Egyptian surveyors used coordinate systems as features of maps in which the space of the map corresponds to physical space (Beniger & Robyn, 1978; Funkhouser, 1937b). The oldest known attempt to show graphically values changing over time is a tenth-century graph illustrating planetary orbits (Funkhouser, 1937a), and by the thirteenth-century musical notation was standardized as a time series (Beniger & Robyn). About 1350, Oresme had “a brilliant thought—why not draw a picture or graph of the way in which things vary?” (Boyer, 1991, p. 264). Oresme drew pipes for the graphic representation of idealized curve functions (Biderman, 1990) with the insight that “everything measurable can be represented by a line” (Clement, 1989, p. 84). Using language of longitudes and latitudes that we would apply today in geographic mapping, he illustrated how the horizontal axis could be segmented into equal spacing representing equal time units, and vertical lines could represent velocity, that is the distance travelled in that time spacing. In 1637 Descartes developed the formal coordinate system, although this graphing was aligned with a rational philosophy that was not readily adopted for recording empirical data and in fact may have hindered the development of statistical graphing until Playfair’s creative developments in 1786 (Biderman; Wainer & Velleman, 2001).

Kieran (1993) contrasted two views of functions: the formal set-theoretic notion of a static correspondence of sets of numbers, and the dependency notion of change in a variable and related change in another variable. Her brief history of the development of the concept of function refers to Oresme’s graph about 1350 (see

Table 2.01), and the developments of (a) the concept of real numbers, (b) creation of symbolic algebra, (c) the study of motion, and (d) the wedding of algebra to geometry. Descartes is noted for contribution to (b) and (d), and Galileo to (c).

Kieran suggests the influence of Descartes in the mid 1600s permitted equations to express relationships between variables, however it was a century later, in 1755, that Euler clearly defined functions with the notion of one variable quantity depending upon another. The concept of function of an arbitrary correspondence of two sets emerged in the early 1800s and was generalized in the early 1900s. Kieran notes that the use graphs to view functional relationships of physical quantities, especially those as a function of time, predated the use of algebraic language of these relationships. This historical development supports advocating the use of the notion of dependency to build an appreciation of covariation. In particular measuring the effect of manipulations of one variable upon another, may be a natural starting point for considering covariation (see also Coulombe, 1997). It should be noted, however, that the view of covariation as a dependent relationship carries with it causal implications that may make it difficult for students to distinguish covariation from causation.

2.01.02 Empirical Data of the 17th and 18th Centuries

From 1663, ideas for time-series graphs were developed involving mechanical devices. One invention could record temperature change over time “on a moving chart by means of pen attached to a float on the surface of a thermometer” (Tilling, 1975, p. 195). It is interesting to note, however, that “such automatic graphs were considered useless for analysis and were routinely translated into tabular logs” (Beniger & Robyn, 1978, p. 2). Also in the 17th century, a number of graphs were produced for analysis of measures from the physical or social sciences, however each

graph was apparently without immediate successors. In 1669, Huygens produced a statistical graph showing ages on the horizontal axis and numbers of survivors on the vertical axis, to calculate life expectancy by analysis of the age at which 50% of people survived (Boyer, 1947). In 1685, Plot produced a line graph of barometric readings and in 1686, Halley graphed barometric readings against altitude (Biderman, 1990).

For almost a century, there is little evidence of graphs of empirical data. Then, in 1764, James Watt constructed an instrument to record pressure versus volume in a steam engine automatically, but he was concerned by “irregularities in the results” (Tilling, 1975, p. 198), deviations from an ideal curve which today we would acknowledge as experimental errors. In 1765, Johann Lambert described using lines of best-fit for data plotted in Cartesian coordinates, and thus “graphics could describe and characterize relations between variables—thus displaying the essential evidence necessary for establishing cause and effect” (Tufte, 1997, p. 16). Also in 1765 Joseph Priestley used individual bars to compare values on a time-line, acknowledged by Playfair as a forerunner to his development of bar graphs in 1786 (Beniger & Robyn, 1978), but taking a further 15 years until 1801 to gain acceptance by Playfair himself as a useful graph form for comparison of data values (Beniger & Robyn, p. 3).

2.01.03 Playfair’s Works

Around 1780, William Playfair was employed as a draughtsman under Watt (Biderman, 1990), which may have influenced his work for graphing data rather than idealized curves. Playfair’s 1786 publication, *The Commercial and Political Atlas and Statistical Breviary*, made popular the use of graphs of empirical data. The atlas included 44 charts, most of which were line graphs of monetary amounts over time, although one lacked time-series data and resulted in his apology for a chart that is

credited as the first bar graph. Playfair's inventive draftmanship allowed him to use graphic space to display data available without scholastic assumptions of fitting to theoretical forms (Biderman, 1990). The significance of Playfair's work was based in the following insight.

[A]nything that could be expressed in numbers could be represented as well by lines [...] these uses of 'lineal arithmetic' he learned in his boyhood [...] keeping a register over time of the readings of the thermometer by drawing lines on a divided scale. Here, lengths of the thermometer column were literally what was observed and these lengths, indeed, were the only scale values available for expressing temperature quantitatively. (Biderman, 1990, p. 9)

This insight meant graphs could represent empirical data, including social data, which need not be assumed to fit a theoretical relationship. Graphs hence gained a key purpose as a representation to support exploratory data analysis. It was revolutionary that spatial position was used to denote a quantity that did not have a physical correlate such as in a map or in a direct reading of a scale from a physical measuring device. Playfair argued that tables of values recorded precise numerical values, whereas graphs had an alternate purpose for conveying a global perspective.

The advantage proposed, by this method, is not that of giving a more accurate statement than by figures, but it is to give a more simple and permanent idea of the gradual progress and comparative amounts, at different periods, by presenting to the eye a figure, the proportions of which correspond with the amount of the sums intended to be expressed [...] as much information may be obtained in five minutes as would require whole days to imprint on the

memory [...] by a table of figures. (Playfair 1801, pp. xi-xiii, cited by Funkhouser, 1937b, p. 281)

It is a sign of those times, and the pioneering of Playfair, that such a defence of graphing was felt necessary.

2.01.04 Increasing Recognition of the Power of Graphics

In the decades following Playfair, graphing became widely accepted in scientific and government publications (for a review, see Friel, Curcio, & Bright, 2001). The power of pictorial representation is commonly acknowledged in the proverb “A picture is worth a thousand words,” quoted in introductions to two editions of a book on graph comprehension for mathematics educators (Curcio, 1989, 2001). English variants of the proverb can be traced back to two advertisements that appeared in the advertising trade journal *Printers’ Ink* in the 1920s (Hepting, 1999). The first, in 1921, stated, “One look is worth a thousand words,” and attributed it to a famous Japanese philosopher, whereas the second in 1927, stated, “One picture is worth ten thousand words,” and displayed Chinese characters and claimed it was a Chinese proverb. The closest Confucian proverb found may be translated, “One look is worth being told a thousand times” (Pei, 2002). Whereas the Confucian version suggests doubt in hearing second hand what may be better observed first hand, the versions of the 1920s are advertisements for the power of pictures in advertising. Pictures are eye-catching because humans have well-developed visual processing to recognise quickly objects and spatial relationships between them, whereas verbal descriptions must be decoded from their visual (or auditory) form and processed sequentially. Various early studies in the 20th century concerned information design to improve comprehension, often as advice for graphing in scientific, government, or advertising publications (Culbertson & Powers, 1959; Feliciano, Powers, & Kearn,

1963; Malter, 1952; Peterson & Schramm, 1954; Vernon, 1946, 1950; Washburne, 1927).

2.01.05 Contemporary Uses of Graphs

Today, graphs are commonly accepted as valid representations for conveying data, perhaps more so than tables of data. Twenty years ago, it was estimated that over 900 billion statistical graphs are printed every year (Tufte, 1983). Daily newspapers print a variety of graphs, from simple trend lines for financial time-series data, to unusual picture-graphs developed by advertising graphic artists to catch readers' attention. One reason graphs are so common is the potential quickly to convey an overall sense of a large set of data. Tufte (1983) commented,

At their best, graphics are instruments for reasoning about quantitative information. Often the most effective way to describe, explore, and summarize a set of numbers—even a very large set—is to look at pictures of those numbers. Furthermore, of all methods for analyzing and communicating statistical information, well-designed data graphics are usually the simplest and at the same time the most powerful. (Tufte, 1983, p. 9)

These strengths are described for graphs “at their best” that are “well-designed.” The strength of graphs for quick communication is also a dangerous feature: graphs may be well-designed to mislead, or they may simply be poorly presented. Both problems need to be addressed in school curricula. Misleading graphs exploit the initial impression of the graph gained by the casual reader. The graph designer, for example, may change the scale, so that although the data are portrayed accurately for the given scale, the impression given to the reader is quite different. As a result, it is important to assist students to read and interpret, that is, make sense of, graphs, including those that might be misleading. Poorly presented

graphs are also common. Wainer (1992) commented, “though there is ample evidence that the ability to understand graphically presented material is hard-wired in, there is even more evidence that the ability to draw graphs well is not” (p. 18).

Bar graphs are today probably the most universally recognized graphs, despite the late historical development of this form. Line graphs are frequently used to represent time-based data, for example in newspaper and television reports of financial trends or temperature change. Time-based data are the most common form of coordinate graph. Coordinates were used to represent neither position nor time in less than 10% of newspaper and magazine graphs surveyed by Tufte (1983), although this figure was 20-70% in the case of science magazines, science texts and standardized educational tests. School algebra, on the other hand, emphasizes coordinate graphs, often known as Cartesian graphs from Descartes’ work.

Three examples of publicly presented graphs are shown in Figures 2.01, 2.02, and 2.03, which have been selected to illustrate issues closely aligned with some tasks presented in this thesis. The graph in Figure 2.01 appeared on the Australian young children’s television show *Humphrey Bear*. During the show, Humphrey rotated the graph 90°, and the narration stated, “That’s great Humphrey, now the graph is going up and up and up instead of down and down and down.” Perhaps the most interesting aspect of this graph is that one axis is explicitly about money, denoted by the \$ symbol, whereas the other axis is labelled with an arrow “^,” indicating that time is assumed implicitly. That such a times-series graph, with implicit use of time, has been used for a young audience indicates something of the early exposure and emphasis on time-based line graphs.



Figure 2.01. Graph from *Humphrey Bear*, Nine Network, 2001.

Many weather reports in Australia include a line graph showing temperature change over times of the day. This is another example of a time-based line graph being common in public use. Figure 2.02, from a television news report, displays forecast temperature change for a period 5 am to 7 pm, with key points noted every 3 hours. An interesting aspect of temperature reporting is the dominance of reporting the maximum and minimum, which appear near the top of the screen, but also appear within the graph despite the fact that 11°C at 6 am does not appear to be the minimum with the range shown from 5 am.

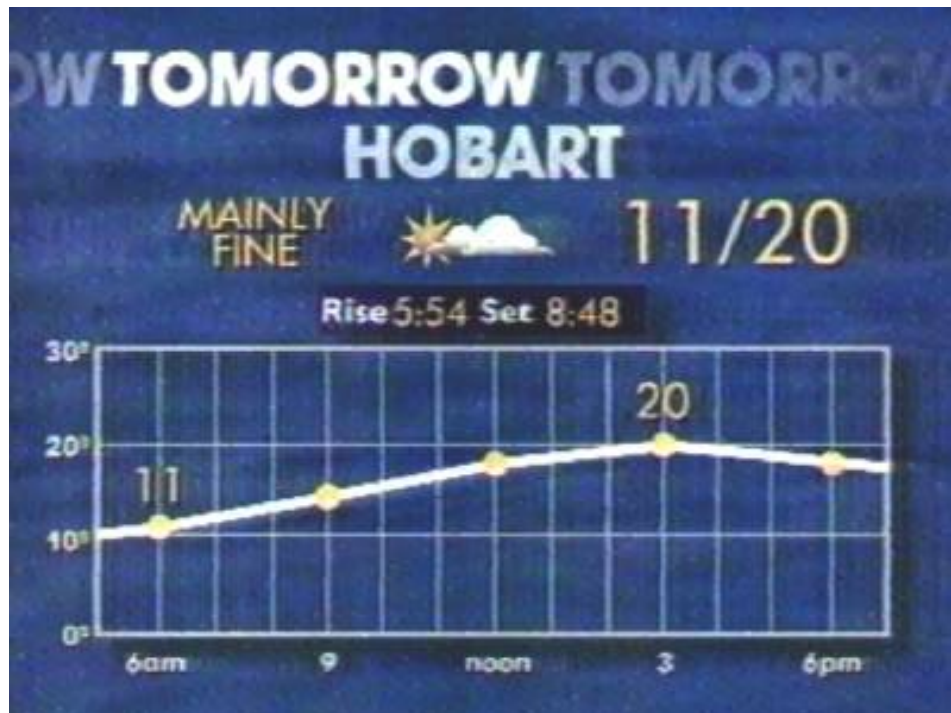


Figure 2.02. Weather graph from WIN television news (19 January, 2002).

A third example, shown in Figure 2.03, shows covariation of increasing height versus age, and includes the statistical distribution of the population by representing various percentile lines. Such a graph comes from health booklets supplied to new parents by government health services. It should not be assumed that a high proportion of the public can accurately read or interpret this graph, however it should be noted that there is public exposure to a graph showing non-linear covariation, and incorporating a third data complexity, namely variability from the norm shown in percentiles.

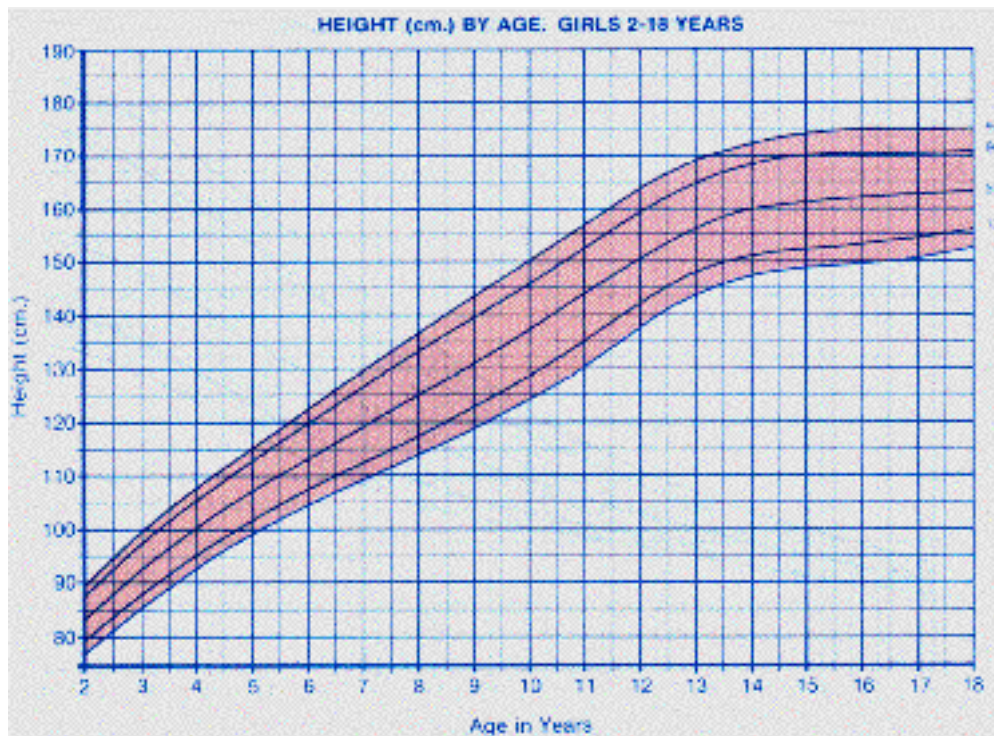


Figure 2.03. Percentiles of height versus age for females
(Department of Health and Human Services, Tasmania, 2000).

2.01.06 Historical Implications

These historical developments may indicate a possible ordering of difficulty of graphing. Firstly, there may be questionable purpose to graph data values if there is no anticipation of a pattern being shown, since the data values alone could just as simply be read from a list or tables. Second, maps are likely to be the most intuitive use of coordinates, particularly if the map is cross-sectional and the vertical dimension corresponds to a height measure, and hence the representation is a stylized picture. Third, time is a natural covariate, often used implicitly in English use of tense such as “it’s going up (over time).” Times-series graphs can be viewed as trends of continuous data or as comparison of values of discrete data cases over time, providing a conceptual bridging of two different schools of thought: graphing idealized functions using Cartesian coordinates, and comparison of data values in bar graphs.

2.02 SCHOOL CURRICULUM

2.02.01 The Broad Context for Statistics Education

During the last two decades, data handling has received increased attention in school level mathematics curriculum documents from a number of countries (e.g., Australian Education Council [AEC], 1991, 1994; Department for Education and Employment [DEE], 1999; Ministry of Education [ME], 1992; National Council of Teachers of Mathematics [NCTM], 1989, 2000). Statistics educators recently have described characteristics of statistical literacy that encompass the skills and dispositions necessary to apply statistical reasoning in social contexts (e.g., Gal, 1998, 2002, 2004; Watson, 1997, 2000). Watson (1997), for example, proposed a statistical literacy hierarchy including three tiers: a basic understanding of statistical terminology, an understanding of statistical concepts when embedded in applied contexts, and a questioning attitude to critique the statistical basis for inferential claims. Gal (1998) commented, “the development of students’ ability to generate sensible and justifiable opinions (e.g., [...] about the validity of arguments that rely on or make reference to statistical data) should thus become a target area for instruction in statistics education” (p. 278).

Gal (2002) claimed that adults are consumers of statistical messages more often than they are producers of them. Skills of interpreting verbal statements of covariation in social contexts, for example, are likely to be required in everyday living more often than skills of graph production from raw data and also more often than graph interpretation, because verbal statements are not always accompanied by graphs. Interpreting and representing verbal statements, without data provided, require appreciating the statistical data that might lie behind such statements, the

skill of Speculative Data Generation as shown in Figure 1.01. This skill may be further decomposed into two understandings based on the observation that data are “not merely numbers, but *numbers with a context*” (Moore, 1990, p. 96):

- (a) a *numerical* understanding of covariation as the correspondence of variation of two sets of numbers, for example, that higher values in one set tend to correspond with higher values of the other set, and
- (b) a *contextual* understanding of data elements concerning how the data might have been collected and measured.

These two aspects reflect two of the tiers of statistical literacy in Watson (1997), namely (a) the statistical terminology of “covariation,” and (b) how this terminology is applied in the context of the specific variables.

Extending Moore’s (1990) description, Tufte (1997) noted that in social contexts, “Not a great many substantive problems, however, are exclusively two-dimensional. Indeed, the world is generally multivariate” (p. 17). Ross and Cousins (1993b) suggested that to be authentic to social settings, tasks should involve contextual elements in multivariate situations, and should require identifying which variables are of interest for the statistical association and deciding whether other variables are ignored or used as covariates to control their affect on the association of interest.

Consumers of information need not only the skills to interpret statistical messages in social contexts, but also the skills to question claims based on statistics (Watson, 1997). Gal (2002, 2004) listed ten sorts of “worry questions,” involving the appropriateness of the study source, the sample, the measurement instruments, the numerical analyses, the graphical representations, the statistical significance, and the causal inference when other variables may be intervening. In providing questions for

consumers, Gal acknowledged a further challenge for research and assessment of students' statistical literacy. A task to critique a media report may assess students' *ability* to question claims, but not their *disposition* to do so of their own initiative. Aware of this issue, Watson and Moritz (2000b) noted that it is important to distinguish students' responses when prompted to question specific aspects of a media story, such as sample bias, from their unprompted responses when simply asked to comment on any unusual features of the story.

The aspects of statistical literacy described above—statistical terminology, numerical understanding in context, awareness of relevant variables, awareness of issues to be questioned, and the disposition to question—are important to consider in developing a variety of tasks that encompass the statistical literacy expected of consumers. These have been reflected in the large scale international assessment projects of the Trends in International Mathematics and Science Study (TIMSS: see www.timss.org/), and particularly Programme for International Student Assessment (PISA: see www.pisa.oecd.org/), where terms such as “numeracy” and “quantitative literacy” reflect the engagement of mathematical concepts in real contexts.

2.02.02 Covariation in School Curricula

Constructing and interpreting graphs and understanding covariation appear as parts of statistics and algebra in school mathematics curricula in Australia (AEC, 1991, 1994), England (DEE, 1999), New Zealand (ME, 1992), and the United States (NCTM, 2000). These curriculum documents place statistical investigation as the central focus of teaching statistics, with syntactic skills of representation considered as tools to aid the investigation. Students are asked to engage steps in a multistep process (a) to hypothesize a relationship between two variables, (b) to collect data, (c) to represent the data graphically or analyse them numerically, and (d) to draw

conclusions about the relationship in verbal statements. This multi-step process reflects professional use in the social and physical sciences, in which covariation is often observed within bivariate data sets, and causal inferences are made. Curriculum recommendations for algebraic understanding and for data handling and graphing are considered in the following sections.

2.02.03 Algebra in School Curricula

Apart from statistical contexts, curricula (e.g., AEC, 1991; NCTM, 2000) for early algebra courses include covariation relating familiar variables. Time-series data are also considered as an important avenue to introduction of algebra. In the United States (NCTM, 2000), students Pre-K-2 should “describe qualitative change, such as a student’s growing taller” (p. 90). By grades 3-5, as part of the algebra standard of “analyze change,” students should “represent and analyze patterns and functions, using words, tables, and graphs” (NCTM, p. 158) and “investigate how a change in one variable relates to a change in a second variable” (p. 158), such as for the growth of a plant, “describe how the rate of growth varies over time” (p. 163). Australian primary students should have experiences with functions to “represent (verbally, graphically, in writing and physically) and interpret relationships between quantities” (AEC, 1991, p. 193). Possible activities include “Sketch informal graphs to model familiar events such as variations in hunger through the day” and “Given a sketch graph (e.g. of the depth of water in the farm water tank), write a story about it” (AEC, 1991, p. 193). In New Zealand, suggested learning activities for upper-primary students include “sketching, interpreting, and writing stories about graphs of familiar situations” (ME, 1992, p. 139), with example graphs illustrating change in a variable over times in the day.

A perusal of selected textbooks for Australian students illustrates agenda for student reasoning in accord with these curriculum documents. An introductory algebra text for secondary students (Lowe, Johnston, Kissane, & Willis, 1993), based on the package *The Language of Functions and Graphs* (Swan, 1985), introduces Cartesian graphs with a variety of variables. In *Book 1*, some examples show axes with two data points, and ask students to compare the values of the points on each variable, and other examples use continuous measures over time. In *Book 2*, a chapter on “graphing relationships” is introduced using change in one variable over time. Similar emphases are presented by Barnes (1991). A series of books from the United States addresses continuous functions for primary schools students (Tierney, Nemirovsky, & Shulman-Weinberg, 1995; Tierney, Weinberg, & Noble, 1996; Wright, Nemirovsky, & Tierney, 1997).

2.02.04 Data Handling and Graphing in School Curricula

Representing data using graphs and interpreting graphs using verbal descriptions are included as part of the Chance and Data strand in Australia (AEC, 1991, 1994), and within the statistics standards in the United States (NCTM, 1989). A number of curriculum statements express conceptual outcomes as part of broader skills and dispositions with contexts. In the *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), graphing is considered as a tool for investigations using data for students of grades 5-8.

[C]onstructing simple graphs, and reading data points as answers to specific questions are important activities, but they reflect only a very narrow aspect of statistics. Instead, instruction in statistics should focus on the active involvement of students in the entire process: formulating key questions; collecting and organizing data; representing the data using graphs, tables,

frequency distributions, and summary statistics; analyzing the data; making conjectures; and communicating information in a convincing way. (NCTM, 1989, p. 105)

Similarly, *A National Statement on Mathematics for Australian Schools* (AEC, 1991), advises that upper primary students should “represent, interpret and report on data in order to answer questions posed by themselves and others.”

Children should learn that data can be displayed in a variety of ways and that the choice of display depends upon the question being asked of the data.

Graphs should not be regarded as an end in themselves; rather they should serve purposes which are clear to children. As the children perceive the need for increasingly sophisticated forms of data representation, the teacher can assist them by introducing new methods of representation. (AEC, 1991, p. 168)

Represent data in tables and graphs and compare different representations of the same data, considering how well they communicate the information (e.g. correct, clear, misleading). [...] Discuss and interpret information presented in graphs and tables found in newspapers, magazines and text materials. (AEC, 1991, p. 172)

It is recommended primary students represent univariate data in pictographs and bar graphs, employing baselines and consistent scales. Understanding scales and number lines also appear within measurement and number strands, and coordinates appear as part of the space strand, for example in maps (AEC, 1991). Bivariate data are recommended for secondary school students, although time-based variation is often introduced earlier. In Australia, representation tasks are suggested for lower

secondary students, such as “represent two-variable data in scatter plots and make informal statements about relationships” (AEC, 1994, p. 93), and “represent bivariate time series data in line graphs” (p. 109). In England (DEE, 1999), secondary students are expected to draw scatter graphs and line graphs for time-series data, to “look for cause and effect when analysing data” (p. 40), and to “draw lines of best fit by eye, understanding what these represent” (p. 41). In the United States (NCTM, 2000), it is recommended that sixth- to eighth-grade students use scatterplots as an important tool in data analysis, and students are encouraged to interpret lines of fit. Causal inference is also considered in curricula, for example, secondary students in Australia should “investigate and interpret relationships, distinguishing association from cause and effect” (AEC, 1991, p. 178).

The New Zealand curriculum (ME, 1992) gives particular attention to representation of and interpretation of time-series data at many levels. In upper primary grades, students should be “collecting and graphing simple time-series data such as the height of a classroom-grown bean plant at midday each day” (p. 179). Secondary students should be “devising ways to display data showing variations of variables over time and using conventional time-series displays” (p. 187) in order to “make statements about time-related variation” (p. 188). This attention to time-series graphs is prior to the mention of graphing of scatterplots to assess bivariate association, which is suggested for senior secondary school years.

Some Australian documents provide lists of specific learning outcomes (AEC, 1994). These are intended to be read in conjunction with the broader purposes of curriculum statements (AEC, 1991) to avoid the possible connotations of checklists of isolated skills, however they do provide detail concerning particular issues, such as engaging various data types and graph forms. *Mathematics - A*

curriculum profile for Australian schools (AEC, 1994), for example, refers to graphing under three strand organisers, “Organising data,” “Displaying and summarising data,” and “Interpreting data” (p. 13). For the latter two, the curriculum outcomes for early secondary students include “Displays one-variable and two-variable data in plots” (p. 93) and “Reads and describes information in histograms, plots” (p. 93), with pointers to evidence the outcomes including “Represent two-variable data in scatter plots and make informal statements about relationships [...] Informally interpret relationships and reach conclusions from scatter plots [...] Write or present an accurate summary of the information displayed in a range of tables and graphs” (p. 93).

A curriculum developed in the state of Victoria (Victorian Curriculum and Assessment Authority, 2002) follows a similar model of listing specific learning outcomes at various levels of schooling. The learning outcome expected of secondary students to “Informally interpret relationships in bivariate data” (p. 209) is described as evident when students “make informal statements about the association between bivariate variables suggested by a scatter plot [...] make predictions from a line of good fit” (p. 209). A corresponding assessment item, administered to Year 7 students, is shown in Figure 2.04. A Victorian text for senior secondary students titled *Reasoning and Data*, includes a chapter on related variables (Fitzpatrick, Galbraith, & Henry, 1991) that addresses representation in scatterplots, methods of fitting data, description and measurement of correlation, distinguishing correlation and causation, and a section devoted to handling time-series data.

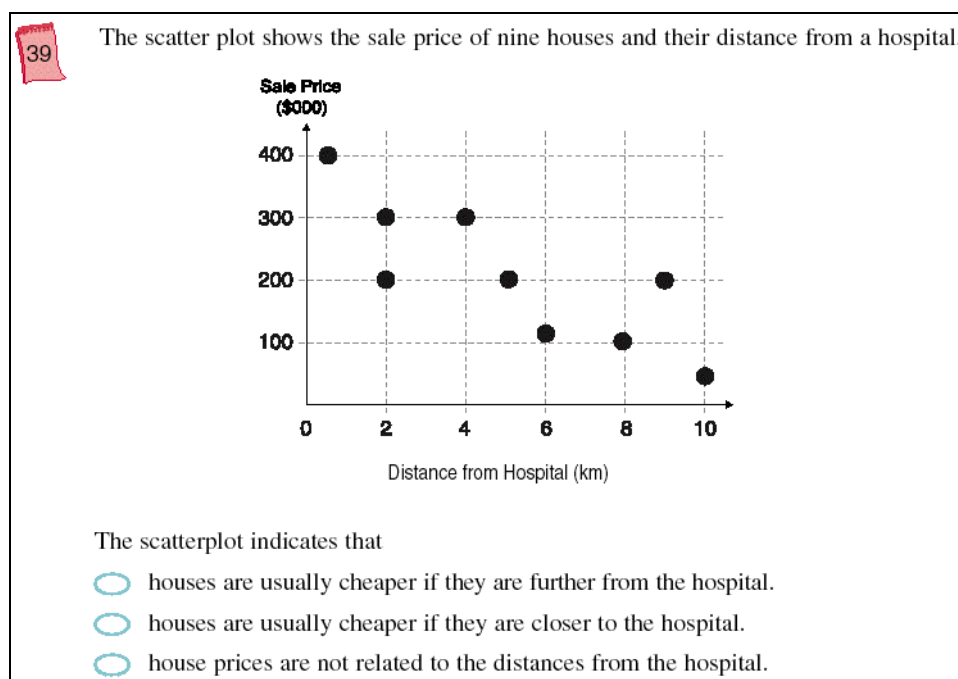


Figure 2.04. Scatterplot multiple-choice test item.

Source: <http://www.vcaa.vic.edu.au/prep10/aim/testing/tests/7M03.PDF>

In the state of Tasmania where this study was conducted, significant reform to the documented curriculum appeared in 2002, as part of the major curriculum reform titled “Essential Learnings” (see www.parentinfo.education.tas.gov.au/parentinfo/teaching/els). Eighteen key elements were described, one titled “Being Numerate,” and another titled “Inquiry.” Both included strong reference to data handling and data interpretation skills.

2.03 RESEARCH LITERATURE

2.03.01 Previous Research Reviews and Perceived Gaps

The need for research into school students’ statistical concepts, and into how these concepts develop over time or change as a result of instruction, was reiterated in a number of reviews over a decade ago (Garfield & Ahlgren, 1988; Green, 1993; Shaughnessy, 1992; Shaughnessy, Garfield, & Greer, 1996). Much of the research related to graphing, graph interpretation and detecting covariation prior to this time

involved adults and university age students. The need to conduct research with school age students was identified around the same time that school curriculum reform was placing greater emphasis on data handling and statistical interpretation.

The broad research literature on graphing has often reported on pointwise tasks of construction and interpretation, such as plotting points or locating values (Leinhardt et al., 1990). Reviews of the research on graph comprehension, with emphases on both data with contexts (e.g., Friel, Curcio, & Bright, 2001; Shah & Hoeffner, 2002) and mathematical functions (e.g., Dugdale, 1993; Kieran, 1993; Leinhardt et al., 1990), indicate some of the breadth of disciplines in which graph interpretation studies have taken place. In their review of research on graph comprehension, Shah and Hoeffner (2002) commented:

The scope of this review is limited, in two ways. First, it focuses on the *interpretation* of graphs depicting meaningful data [...], generally in the context of science and social science. Thus, this article does not review the large body of research examining the understanding of mathematical functions and the role of graphs in mathematics education. Second, research on graph interpretation has taken place in a diverse variety of disciplines [...], so it is impossible to cover all the research available. The research presented here focuses on a representative sample of those studies for which there are clear educational implications. (Shah & Hoeffner, 2002, pp. 49-50)

They cited many studies from the disciplines of psychology and cognitive science (e.g., Cleveland & McGill, 1984, 1985; Pinker, 1990), information and graphic design (e.g., Kosslyn, 1985, 1989; Tufte, 1983), mathematics education (e.g., Kaput, 1997; Leinhardt et al., 1990), and science education (e.g., McKenzie & Padilla, 1986).

Research into the development of understanding of statistical association has often involved frequency tables of categorical variables (e.g., Batanero, Estepa, Godino, & Green, 1996). Considerable psychological research (e.g., Alloy & Tabachnik, 1984) has explored aspects such as influence of prior expectations, direction of association, and complexity of stimulus. Research more specifically on statistical covariation has often come from tasks involving graphs, and there have been calls for further research (Kaput, 1993). Tasks involving variation and qualitative graphs—that is, without specific data values—have been considered by some researchers (Leinhardt et al., 1990) to be an underutilized avenue for exploring understanding of general features of graphs, including covariation.

Ross and Cousins (1993b) identified several shortcomings of previous research of correlational reasoning.

1. Tasks were often bivariate and involved dichotomous variables, whereas out-of-school contexts are more likely to be multivariate and involve continuous variables.
2. Students were often asked to assess association, but rarely asked to represent it, and never asked to engage issues of data collection, such as sampling or data measurement. As a result, levels of responses were restricted to a single dimension of correlational reasoning, whereas Ross and Cousins distinguished a number of skills such as representing, reading the graph, and drawing conclusions about associations.
3. Some researchers predetermined aspects of student growth based on task demands, rather than observing in student responses the ways that students engaged the tasks, possibly including ways not anticipated by the researchers.

2.03.02 The Structure of the Review of Research

Literature

The following review of research literature is arranged in four sections: graph interpretation, covariation interpretation, graph production, and speculative data generation. This order to some degree follows the historical development of the research reviewed. The research of interpretative skills is considered first as it preceded later studies of production skills, and because conceptual issues identified in research of interpretation of graphs and covariation was foundational for research of issues of production of graphs and data showing covariation. This ordering also corresponds to first reviewing issues and themes from among the vast research on graph interpretation, then drawing on these issues and themes via reviews of research on covariation interpretation and graph production, and finally considering in more detail the limited research on speculative data generation and the gaps in the literature that this thesis aims to fill.

2.03.03 Graph Interpretation

2.03.03.01 Aspects of Graph Interpretation

Previous reviews of graph comprehension and interpretation (e.g., Friel et al., 2001; Shah & Hoeffner, 2002) provide a sense of the wealth of previous research, and an indication of the aspects commonly considered as factors in graph comprehension. Friel et al. (2001) structured their review by first defining graphs, in general, as using spatial characteristics to represent quantity. They listed structural components of graphs as including the *framework* (e.g., axes, scales), *specifiers* used to represent data values (e.g., lines, bars), and *labels*. They defined graph comprehension in relation to various tasks, suggesting three levels of tasks that have

been described by a number of studies (Carswell, 1992; Curcio, 1987; McKnight, 1990; Wainer, 1992):

1. an elementary extraction of information,
2. an intermediate level of reading data to find relationships, and
3. an overall interpretive level that moves beyond the data to include, for example, topic knowledge or inference.

Friel et al. (2001) suggested that critical factors affecting comprehension include (a) purposes for using graphs, (b) characteristics of graph decoding, judgment tasks, and context, (c) characteristics of the discipline, such as variation, data types and graph complexity, and (d) characteristics of graph readers. Shah and Carpenter (2002) described the three broad factors involved in graph comprehension as (a) visual characteristics of the graph, (b) knowledge about graphs, and (c) the topic content of the graph. Visual characteristics considered have generally concerned the form of the representation (e.g., line graph, bar graph, or tabular forms) and the structuring of complex data, particularly involving more than two variables.

Pinker (1990) suggested that graph comprehension divides at the most fundamental level into (a) comprehension of the axis framework and scale, and (b) comprehension of the data elements. The scale is necessary for reading numerical values, whereas the data cases without the scale permit trend identification and qualitative comparison of cases. This is the basis for the distinction between skills of Numerical Graph Interpretation and Verbal Graph Interpretation sustained in this study, and parallels the balance between graphs locally representing specific data values and globally representing general trends or relationships in the data, as observed historically by Playfair and in student responses (e.g., Bell & Janvier, 1981;

Ben-Zvi & Arcavi, 2001; Gerber, Boulton-Lewis, & Bruce, 1995; Leinhardt et al., 1990).

Graph interpretation has been described (Curcio, 2001) as having various purposes including reading data values, reading beyond the data by prediction based on global trends, and an intermediate level of reading between the data, such as comparing values. When constructing graphs, students' responses may be influenced by their beliefs about these purposes for the representation more than by any lack of ability to represent a given graph form (Roth & McGinn, 1997). Similarly, when interpreting graphs, the purpose of the tasks and the graph provided may influence students' responses in relation to local or global features more than their ability to engage these features. Curcio (1987) found predictors of graph comprehension included reading achievement, mathematics achievement, and prior knowledge of the topic, mathematical content, and form of the graph. McKnight (1990) considered different levels of data-based tasks including (a) observation of facts in a graph such as extracting a numerical value, (b) observation of relationships in graphs such as "the curves tend to increase," (c) interpretation of relationships in the real-world context, and (d) critical evaluation of inferential claims. For open-ended tasks, Gerber et al. (1995) interviewed students aged 8, 11, 14, and 16 years using multiple maps and graphs of data about fictitious countries, and identified stages in which students (a) interpreted graphs as drawings, (b) interpreted features of the individual graphs in a global or local sense, (c) identified patterns in single graphs, and (d) related patterns in graphs to other data or prior beliefs about relationships. Preece (1983b) asked 122 students, aged 14-15 years, to explain what a graph is, and found that 59% believed a graph was a useful way of displaying information, 16% said it was like a picture, 11% thought it was like a table, and only 11% indicated a graph

shows the relationship between two variables. Across a number of tasks, she identified four levels applicable to both graph interpretation and construction: (a) a graph is viewed as a picture, (b) points are interpreted, (c) comparisons across intervals are considered, and (d) continuous changes in gradient are interpreted. She found that quantitative questions were more often answered correctly than qualitative questions involving the context and the data.

2.03.03.02 Local Data Values and Global Trends

Graph interpretation tasks, and students' abilities on tasks, differ according to the degree to which they consider graphs either locally to represent values or globally to represent general trends or relationships in the data (e.g., Bell & Janvier, 1981; Ben-Zvi & Arcavi, 2001; Gerber et al., 1995; Guthrie, Weber, & Kimmerly, 1993; Leinhardt et al., 1990). Many early studies of graph interpretation concerned aspects of information design to improve the readability of graphs (Anscombe, 1973; Culbertson & Powers, 1959; Feliciano et al., 1963; Hermann, 1973; Kirk, Eggen, & Kauchak, 1978; MacDonald-Ross, 1977; Malter, 1952; Thomas, 1933; Washburne, 1927; see also Meyer, 1997). Studies by Vernon (1946, 1950) suggested that a logical system or statement concerning relationships between variables was important in underpinning graph interpretation.

Various studies have concluded that students construct and read graphs as individual numerical points rather than a global whole (e.g., Bell et al., 1987a; Brasell & Rowe, 1993). With respect to reading individual points using the scale, Bryant and Somerville (1986) found many students as young as 6 and 9 years can read coordinates from either axis with reasonable accuracy. The tendency to consider graphs as a global whole, however, may depend on appropriate task design. When a variety of tasks were compared, Meyer, Shinar, and Leiser (1997) found trend

judgments from line graphs and bar graphs were performed faster and more accurately than tasks (a) to read values, (b) to compare values from the same data series for different X values (X comparisons), (c) to compare values from different data series with the same X value (series comparisons), or (d) to identify the maximum. These findings may be because verbal trend interpretations generally involve reading the data series, whereas numerical interpretations also require reading the scales (Pinker, 1990). Principles of perceptual grouping support the reading of values in bar graphs, and the reading of trends in line graphs (Pinker, 1990; Zacks & Tversky, 1999).

2.03.03.03 Confusing Features of the Graph

At a basic level, some students interpret graphs as if they were pictures (Bell, Brekke, & Swan, 1987a; Janvier, 1978; Swan 1985). A graph involving changing speed over time, for example, might be perceived as a picture of a roller-coaster, that is of changing height over horizontal space. A number of studies, particularly involving kinematics graphs, have documented the confusion of value, denoted by position, and rate, denoted by slope of a continuous graph (Beichner, 1994; Brasell & Rowe, 1993; Hermann, 1973; Hitch, Beveridge, Avons, & Hickman, 1982; McDermott, Rosenquist, & van Zee, 1987; Padilla, McKenzie, & Shaw, 1986; Thompson, 1994b; Tobin & Capie, 1981; van Zee & McDermott, 1987; Wainer, 1980; Wavering, 1984, 1985). Exposure to teaching, particularly using computers to produce the graphs in real-time with the situational context, supported improvements in graph interpretations and construction skills (Adams & Shrum, 1988, 1990; Avons, Beveridge, Hickman, & Hitch, 1983; Brasell, 1987; Jackson, Edwards, & Berger, 1993; Linn, Layman, & Nachmais, 1987; Mokros & Tinker, 1987; Testa, Monroy, & Sassi, 2002).

The scale of graphs is well known to affect interpretations (Ben-Zvi, 1999, 2000). Other features, such as depth or area, can interfere with reading accuracy compared to simple features such as position, length, or angle (Carswell, 1992; Zacks, Levy, Tversky, & Schiano, 1998), and even with strong skills in reading two-dimension graphs, reading data involving a third variable proves challenging conceptually (Shah & Carpenter, 1995).

2.03.04 Covariation Interpretation

2.03.04.01 Aspects of Covariation

Coordinate graphs employ a general graphic feature that position denotes value and apply this feature in two-dimensional space to represent values of two variables. They represent both data points emphasizing the *correspondence* of values of two variables and general trends emphasizing variation of the two variables due to the *ordination* of the values along each axis.

Piaget's theory of cognitive development (e.g., Piaget, 1983) highlights some of the key concepts of students' development of reasoning about covariation. *Correspondence* (to confirm identity or a one-one mapping), *classification* (to identify as one of a class or group), and *seriation* (to order a series) were among the logical operations Piaget observed across many studies and considered to be universally fundamental to cognitive development. *Conservation* is perhaps the most renowned indication of the developmental stage called concrete operations. When pouring a given quantity of liquid from a thin glass to a wide glass, for example, most young children attend to only one aspect, such as the height, and proclaim the thin glass has more. The coordination (correspondence of seriations) of height and width is what encourages the learner to rely not on the configurations but rather the

transformation or operations (Piaget, 1983, p. 122). Some researchers have attributed difficulties with correlational reasoning to limited thinking at the formal stage in Piagetian terms (e.g., Adi, Karplus, Lawson, & Pulos, 1978; Berg & Phillips, 1994; Lawson & Bealer, 1984).

Teaching and reasoning about covariation often focus on either correspondence of bivariate data points, or variation within variables, and aim to build one aspect upon the other. Nemirovsky (1996a) described these two approaches with reference to algebra teaching as (a) a *pointwise approach* of comparing bivariate pairs to identify the functional rule for translating one to the other, and (b) a *variational approach* that considers change in a single variable across a number of cases. The two approaches are similar to two competence models for Cartesian graphing of covariation suggested by Clement (1989): a static model involving translating bivariate data values to points in coordinate space, and a dynamic model involving concepts of variation. Clement noted that a basic qualitative form of the dynamic model involves simply the direction of change with no indication of how the variables are quantitatively measured (e.g., “the more I work, the more tired I’ll get,” p. 80). These correspond to local and global aspects of graphs (Ben-Zvi & Arcavi, 2001). These approaches are also similar to two Piagetian schema that Wavering (1989) suggested are developed in reasoning to create bivariate graphs: (a) one-to-one correspondence of bivariate data values as static points, and (b) seriation of values of a variable, necessary for linear scaling of graphs to produce a coordinate system showing continuous variation along each scale.

The variational approach has been advocated by researchers of early algebra learning (e.g., Nemirovsky, 1996a, 1996b; Yerushalmy, 1997) and those using continuous real-time graphs, often concerning kinematics (e.g., Beichner, 1994) and

generated by computer (e.g., Linn, Layman, & Nachmais, 1987; Mokros & Tinker, 1987). Consideration of variation has also been described as a key element of statistical thinking (Pfannkuch & Wild, 2004; Reading & Shaughnessy, 2004). Nemirovsky (1996b) discussed the importance of time-based mathematical narratives without specific data values, with verbal and graphical language both read left to right to express generalities of how a quantity varies over time. Yerushalmy (1997) used various graphic icons with computer software to provide a graphic language that corresponds to verbal terms *increasing*, *decreasing*, and *constant*, often with time as the implicit covariate. These studies indicate that verbal phrases and graphs are important forms for understanding covariation, echoing links emphasized by Bell and Janvier (1981). Carlson, Jacobs, Coe, Larsen, and Hsu (2002) have proposed a framework for how such qualitative understanding further develops to reasoning about rates of change.

2.03.04.02 Positive Covariation Bias

Subjects' ability to detect covariation in a variety of situations has been investigated by researchers in social psychology (e.g., Alloy & Tabachnik, 1984; Crocker, 1981), science education (e.g., Donnelly & Welford, 1989; Ross & Cousins, 1993a, 1993b; Swatton, 1994; Swatton & Taylor, 1994), and statistics education (e.g., Batanero et al., 1996; Batanero, Estepa, & Godino, 1997; Estepa & Batanero, 1996; Estepa, Batanero, & Sánchez, 1999; Konold, Pollatsek, Well, & Gagnon, 1997).

Many studies have followed Inhelder and Piaget (1958) in considering association of dichotomous variables in contingency tables, whereas few have considered covariation of two numerical variables (Ross & Cousins, 1993b). In judging covariation between variables X and Y in a 2x2 contingency table, the

normative approach for inference to the population is a Chi-squared test. A simpler descriptive comparison, indicative of this normative approach, is to consider the conditional probability $P(Y|X)$, that is involving only two of the four cells of the contingency table. Early studies (e.g., Smedslund, 1963; Jenkinds & Ward, 1965) indicated that adults often consider only the frequency of confirmatory cases, that is involving the single cell of the table where both X and Y are true, without reference to discrepant evidence. Some researchers (e.g., Seggie & Endersby, 1972) suggested that in relevant contexts and task conditions, adults were capable of appropriate covariation judgements. Shultz and Mendelson (1975) observed students younger than 7 often considered only facilitory causes, that is confirmatory evidence, and considered that causes need not precede effects, whereas students by age 11 also considered inhibitory causes, that is dis-confirmatory evidence, and that causes should precede effects. A review by Sedlak and Kurtz (1981) suggested that although some students had a bias to confirmatory evidence, some of this was attributable to task effects rather than an inability to do so. Crocker (1982) identified that the biased question of a positive association affected responses. Others using various questioning method re-affirmed the positive covariation bias (Beyth-Marom, 1982; Kuhn, Phelps, & Walters, 1985). A series of studies by Shaklee and colleagues (Shaklee, 1983; Shaklee & Elek, 1988; Shaklee & Hall, 1983; Shaklee, Hall, & Paszek, 1982; Shaklee, Holt, Elek, & Hall, 1988; Shaklee & Mims, 1981, 1982; Shaklee & Paszek, 1985) indicated that students from grade 2 can use simple covariation judgement rules based on confirmatory cases (Shaklee & Paszek, 1985), and that although normative rule use increases with age (Shaklee & Mims, 1981), even with teaching, the proportional reasoning needed to support the normative strategy is difficult to teach to middle-school students (Shaklee et al., 1988). The

complexity of proportional reasoning in middle school for interpreting contingency tables (Shaklee et al., 1988; Thompson, 1994b) echoes the findings about the difficulties of slope as a rate in kinematics context (e.g., McDermott et al., 1987).

Positive covariation bias has also been observed in studies of correlation between two continuous variables, for example in scatterplots. Erlick and Mills (1967) found university students' estimates were more accurate for positive correlations, and generally had a positive bias. Strahan and Hansen (1978) and Bobko and Karren (1979) found graduate-level students tended to under-estimate positive correlation coefficients from scatterplots, particularly in scatterplots involving extreme slopes or attenuating outliers.

2.03.04.03 Causal Reasoning From Prior Topic Knowledge

Crocker (1981) outlined six steps for statistically correct judgments of covariation in social settings, as well as some common errors at each step. The six steps included deciding what data are relevant, sampling cases, classifying instances, recalling evidence, integrating the evidence, and using the covariation for predictions. A frequent finding when deciding what data are relevant and integrating the evidence, has been that people judge association inadequately due to misuse of topic knowledge or numerical data (Ahn, Kalish, Medin, & Gelman, 1995; Alloy & Tabachnik, 1984; Arcuri & Forzi, 1988; Broniarczyk & Alba, 1994; Jennings, Amabile, & Ross, 1982). In using topic knowledge, people often hold prior beliefs about causal associations between the real-world variables that may influence judgments (e.g., Jennings et al., 1982). Topic knowledge may result in ignoring the available data (Alloy & Tabachnik, 1984; Batanero et al., 1996), or dismissing statistical covariation in the data because there is no apparent causal relationship or

because other variables are more plausible causes (Batanero et al., 1997; Crocker, 1981; Estepa & Batanero, 1996).

2.03.04.04 Data Types and Data Presentation

Students' judgements of statistical association appear to be influenced by the data representation provided and by how students handle data complexity. Students tend to assign higher estimates of association for data presented in scatterplots than for the same data presented in two numerical lists (Lane, Anderson, & Kellam, 1985). It has also been found that judgments comparing differences between numerical scores of two samples (Batanero et al., 1997; Estepa et al., 1999) are easier than either judgements of association of frequencies of two categorical measures in two-way tables (Batanero et al., 1996) or judgements of association of two numerical measures in scatterplots (Estepa & Batanero, 1996). This finding may be due to the data complexities of the different types of statistical association. If one variable is categorical, then each category may be summarized to a single value and a simple comparison can be made of the values; the values might be the means or rough modal values based on where "most" of the data are (Watson & Moritz, 1999), or proportions of frequency data in different categories (Konold et al., 1997). Reducing data complexity to a comparison of representative values or proportions is not, however, a ready option for representing and judging association between two continuous variables, unless one splices the data (Cobb, McClain, & Gravemeijer 2003). These associations depend on interpreting covariation across all data in a single judgement, that is, looking globally at the data rather than locally at individual values. Though representing associations and judging associations may be distinct skills (Ross & Cousins, 1993b), each involves students' conceptions of what constitutes a statistical association.

2.03.04.05 Reducing Data complexity

Some students attend to selected data or selected variables as a means of reducing the complexity of the data (Bell et al., 1987a). Attention to selected data points may involve only the extreme points in a scatterplot (Batanero et al., 1997) or the cells with confirming cases in contingency tables (e.g., Batanero et al., 1996; Crocker, 1981; Inhelder & Piaget, 1958). Attention to selected variables has been observed in some studies that have identified levels of response based on the number of variables students have referred to when asked to provide a general verbal statement from given data (e.g., Donnelly & Welford, 1989; Ross & Cousins, 1993b; Swatton, 1994; Swatton & Taylor, 1994). Ross and Cousins (1993b) asked students from grades 5 to 13 to “find out if there was a relationship” between two continuous variables in situations where a third, categorical, variable was involved. Their analysis concerned the numbers of variables students appropriately ordered or described, including 0, 1, 2, or 2 with acknowledgment of the differential effect of the third, categorical, variable. Swatton (1994) showed sixth-grade students scatter graphs and line graphs and asked, “what do you notice about [X] and [Y]?” Level 0 responses involved only the context of the data or syntactic/visual patterns in graphs, Level 1 responses described univariate data patterns, Level 2 involved both variables, and Level 3 responses involved both variables with appropriate directionality. Proportions of responses at each level differed across three tasks, from 16% to 51 % at Level 3, 5% to 23% Level 1, and 35% to 52% non-response.

Swatton and Taylor (1994) employed a variety of graphs, often time-series data with points connected by lines, and found only 2% to 16% of 11-13 year olds provided an adequate pattern statement, but up to 50% of students referred to the independent or dependent variable, and about 50% of the students referred to the

direction of the graph, often as simply as “it goes up/down.” Such comments have reduced the bivariate situation to a univariate one, or one involving no variables at all, similar to the observations of implicit reference to time noted previously (e.g., Nemirovsky, 1996b; Yerushalmy, 1997). In contrast, Swatton and Taylor found 31% to 71% of the same students interpolated values, and they commented, “It suggests that the ability to describe the relationships between variables is of a very different nature to that required to read and manipulate data in both these co-ordinate forms” (p. 235). Donnelly and Welford (1989) gave 15-year-old students bivariate and multivariate data tables. Successful generalizations were provided by 76% of students for a positive association task with 5 bivariate integer values, and 45% for a negative association task with 6 bivariate decimal values and with labels (country) for each. For a negative association task with 8 multivariate integer data points, 18% generalized and a further 11% did so but failed to exclude a variable with no association. These differences may be attributed to task differences including direction of association, number of data points, context, and numerical value complexity. The authors also noted that asking about generalizations might prevent students detecting non-associations.

2.03.04.06 Interpreting Variation and Analysing Data

Slices

In using numerical data, some people hold a deterministic concept of association (Batanero et al., 1996, 1997; Crocker, 1981). They may not consider a relationship if there is variation from a monotonic trend, that is, if there is evidence from subsets of data that is counter to the more general trend.

Noss, Pozzi, and Hoyles (1999) asked nurses to consider scatterplots of a positive relationship of age and blood pressure. Despite knowing a tendency for a

positive association, they often failed to see the relationship in the scatterplot. The researchers taught the nurses to consider the scatterplot in vertical *data slices*, and by noting the distribution and average of each slice, they could then note the differences between these groupings. Konold and Higgins (2003) reviewed this study and commented that failure to see the relationship was likely due to “the trend apparently [being] masked by the variability in the data” (p. 210). They commented that the slicing technique may not be trivial to engage in scatterplots. “Emerging evidence suggests that it may be easier for students to perceive relationships in two-column tables or case-value plots in which the values of one variable have been ordered and displayed next to the corresponding values of the other variable” (p. 211). Further consideration of this form of graph is provided in Section 2.03.05.05.

2.03.05 Graph Production

2.03.05.01 Levels of Graph Production and Other Skills

A number of studies have proposed frameworks of levels of graph construction and interpretation. Some have involved univariate graphs for primary school students (Watson & Moritz, 2001), and some have involved generic levels that may be applied to univariate or bivariate graphs (Jones, Thornton, Langrall, Mooney, Perry, & Putt, 2000). Preece (1983b) identified four levels applicable to both graph interpretation and construction: (a) a graph is viewed a picture, (b) points are interpreted, (c) comparisons across intervals are considered, and (d) continuous changes in gradient are interpreted.

Wavering (1989) asked students in grades 6 to 12 to construct graphs and then to identify the relationship in the data sets provided for each of three items, involving a positive slope, a negative slope, and an exponential curve. Nine

hierarchical categories of response were distinguished. Categories 1 to 3 ranged from no response to pre-Cartesian graphs in which one-to-one correspondence was evident but the data were not ordered. Successive improvements in the use of scale were observed in responses in categories 4 to 7, so that each data point was represented not only by its numerical value, but also by its position along the axis. The highest two categories involved students interpreting their graphs to state the relationship between the variables involved; that is, the structure of the task and the response hierarchy assumed students could represent an association prior to making a judgement of association.

Ross and Cousins (1993b) asked school students to represent data that involved comparing bivariate associations of continuous variables on each of two levels of a dichotomous variable. They asked students from grades 5 to 13 to “find out if there was a relationship” between two continuous variables, and concluded that “organizing information” in a graph was correlated with skills of “locating or selecting data” and “drawing a conclusion.” Each of these skills was assessed on a three to five point marking scheme. One limitation of this simple coding was the loss of the richness of student responses, which was evident in the few examples presented. Another limitation of their research was their self-stated restriction of not exploring students’ questioning of the collection of the data provided.

Watson and colleagues (Chick & Watson, 1998, 2001; Watson & Callingham, 1997; Watson, Collis, Callingham, & Moritz, 1995) provided upper primary level students with an open-ended multivariate exploration task and identified three levels of interpretation and three levels of representation. Representation levels included (1) lists of data values with no attempt at aggregation, (2) univariate bar graphs, and (3) bivariate scattergraphs. Notably, levels of

representation concerned the ways students engaged the data itself, and as the open nature of the task allowed students to graph the complexity of data they chose, most students drew bar graphs. Often students could hypothesize associations between two variables using topic knowledge (e.g., “eating fast foods increases a person’s weight”), sometimes focussing on individual data cards and observing a one-to-one correspondence (e.g., “the heaviest person also ate the most fast foods”). Many of these students, however, reverted to univariate bar graphs when asked to represent these associations. This reversion may be an instance where students selected only one variable to reduce the complexity of the task. It may also be that although students could readily employ topic knowledge of causal associations, they had not had sufficient experiences with techniques for processing and representing bivariate data, such as in scatterplots.

2.03.05.02 Conventional Graphing

Brasell and Rowe (1993) asked 84 twelfth-grade physics students to construct a graph of five pairs of data values indicating heights from which a ball was dropped and to which it rebounded. Students were instructed to pay attention to labeling and scaling axes, to plotting points, and to using a line of best fit. Some students drew pictures of bouncing balls. The majority of students drew line graphs that were connected point-to-point, and reversal of axes and inadequate labeling were each a problem for approximately 50% of students. The researchers concluded that most students in their study constructed graphs by plotting data points of ordered pairs, without appreciating the purpose of the graph to show an association between the two variables. Of the 84 students, for example, 11 drew a graph as a picture for which one axis represented both the height of the drop and the height of the bounce, and the other axis represented the horizontal motion of the ball. Four other students

represented two data series in a double line graph, with the horizontal axis indicating the ball number. It is important to note, however, that though unconventional, the data provided could be read from such representations. Rather than placing each height on a separate axis in a Cartesian form, these 15 students placed both height measures on the same axis, indicating not an attempt at Cartesian *association* expected by the researchers, but rather an attempt at *comparison* of values, a reminder that the choice of representation is dependent on its perceived purpose.

2.03.05.03 Intuitive Graph Forms

Some researchers of students' skills for graph production have exploited contexts in which students have prior beliefs about covariation and there is a natural mapping of height on the vertical axis and time on the horizontal axis. Ainley (1995; Ainley, Pratt, & Nardi, 2001) found that intuitions about the context of height growth allowed primary school students to construct bivariate graphs and correct plotting errors that did not fit the trend of growth. Compared to the graphs observed by other researchers, these tasks resulted in remarkable success by students of a young age in representing covariation trends in data, possibly because of familiarity with the covariation and with the measurement of the variables.

2.03.05.04 Inventing Alternative Graph Forms

More recent studies of students' graphing have explored with younger students the ways they choose to construct graphs for their own purposes. diSessa, Hammer, Sherin and Kolpakowski (1991; diSessa, 2001) described children as *inventing graphing*. diSessa and Sherin (2000) suggested students have criteria for judging what is adequate for the graphing purpose. A number of researchers have commented about the importance of a sense of the context of the data prior to, or to support, representation production or interpretation (Lehrer & Romberg, 1996; Pratt,

1995). Roth and colleagues (Roth & Bowen, 1994, 2001; Roth & McGinn, 1997) emphasized the social dimensions of the practice of graphing, and that graphing can be highly dependent on the context or data being represented. Roth and McGinn observed that when graphs are drawn for the individual to interpret rather than for another audience, then conventions may be of little importance.

Sherin (2000) observed that students' basic representations of motion often involved visual-spatial features of height and position, like a drawing, and these students did not clearly separate time, position, and speed until situations were described that required this separation of variables, such as zero speed. These were similar to findings of various studies in kinematics and other physics concepts (Adams & Shrum, 1990; Beichner, 1994; McDermott et al., 1987; Brasell, 1987; Padilla et al., 1986; Testa et al., 2002).

Nemirovsky and Tierney (2001) observed that students are often concerned to represent all features they know about the narrative context, such as the events or people involved. They also found some students omit elements of a conventional scaled graph, for example periods of constant value, for purposes of efficiency in telling the narrative. For these reasons students may not choose to create "homogeneous spaces," that is using two-dimensional space for which scaled position denotes interval value.

2.03.05.05 Case-Value Graphs

Konold (2002) suggested that a variety of graph forms are valid alternatives to scatterplots for representing and interpreting covariation. One such example is ordered *case-value bars*, which involve a bar graph ordered by cases of one variable to examine any pattern in the other variable, similar to those observed by Brasell and Rowe (1993). For ordered case-value bars, ordering was considered important to

assist scanning values to offer a global summary. Two bar graphs of teeth-brushing times and of plaque levels, both across cases of student names, were ordered by brushing times. As each univariate bar graph had corresponding positions for student names, scanning the plaque graph for evidence of a trend corresponded to detecting covariation with brushing time. Pfannkuch and Rubick (2002) also observed middle school students drawing case-value graphs comparing data series when analyzing multivariate data cards.

Cobb et al. (2003) conducted classroom investigations of the way a class of grade 8 students learned about statistical covariation. They first allowed students to graph a bivariate data set, anticipating students would draw double bar graphs as well as scatter plots, with subsequent discussion of how well the graphs “enabled them to assess how one of the measured quantities varied as the other increased” (p. 16). An aspect of their planning, significant for the current study, concerned how the classroom discourse would need to establish certain foundations prior to working effectively with scatterplots, including the following.

[...] the importance of it becoming normative that bivariate data consist of the measures of two attributes of each of a number of cases. [...] We conjectured that both here and in the subsequent discussions of the students’ analyses, it would be important to develop ways of talking that referred explicitly to cases whose attributes had been measured rather than to speak solely in terms of the measures. This, we reasoned, might support the view that each dot on a scatter plot signifies a single case whose measures are indicated by its location with respect to the axes. (Cobb et al. 2003, p. 16)

Perhaps most interesting about this extract was that they did not advocate the power of double case-value bar graphs for assisting the view of bivariate data cases. They

observed that by ensuring that “the interpretation of bivariate data as consisting of the measures of two attributes of each of a number of cases did become normative,” that “the convention of inscribing such data as dots in a scatter plot emerged relatively unproblematically” (p. 73). They also noted that students “typically reduced scatter plots to lines that signified fixed relationships of covariation rather than conjectured relationships about which the data were distributed” (p. 75). They suggested that further development of student reasoning using scatterplots might be built by shifting consideration from individual bivariate data points to *vertical splices of the data*—that is those points that had common values of the variable shown on the horizontal axis—and then comparing the distributions of various data slices, as described by others (Noss et al., 1999; Konold & Higgins, 2003; see Section 2.03.04.06).

2.03.06 Speculative Data Generation

Historically, few research studies directly addressed how students translate verbal descriptions or numerical forms into graphical representations (Leinhardt et al., 1990), and only recently have student-constructed representations begun to receive attention (Brasell & Rowe, 1993; Mevarech & Kramarsky, 1997; Moritz, 1999, 2000; Ross & Cousins, 1993a, 1993b; Watson & Moritz, 1997; Wavering, 1989). Students have often been asked to plot points (e.g., Kerslake, 1977) and to read and compare data values (e.g., Curcio, 1987); thus it is not surprising that researchers have often found that students construct and read graphs pointwise (e.g., Kerslake, 1977; Bell et al., 1987a). To encourage student attention to representing general trends rather than specific points, some researchers have asked students to graph relationships between variables based not upon given data values, but upon

verbal descriptions (e.g., Bell et al., 1987a, 1987b; Coulombe & Berenson, 2001; Krabbendam, 1982; Mevarech & Kramarsky, 1997; Swan, 1985, 1988).

2.03.06.01 Early Studies of Translations and Functions

Janvier's seminal work and following works (Bell & Janvier, 1981; Janvier, 1978, 1981, 1982, 1987, 1989; Janvier, Girardon, & Morand, 1993; Preece & Janvier, 1992) formulated the conceptualisation of translations between words, graphs, situations, and tables, which influenced the skills as described in Figure 1.01. His emphasis on qualitative graphs and functions gained momentum in a conference in 1982 (Bergeron & Herscovics, 1982; Freudenthal, 1982; Herscovics, 1982; Janvier, 1982; Krabbendam, 1982; Swan, 1982; Van Weering, 1982). Krabbendam gave 12- to 13-year-olds various graphing tasks, such as a newspaper text about the gathering and dispersion of a crowd of people. He concluded, "it appears to be rather difficult for children to keep an eye on two variables" (p. 142), but that "time could play an important part in recording a relation" (p. 142) provided it is seen to pass gradually, that is continuously, thus supporting a view of continuous variation rather than a pointwise approach. Swan (1985, 1988; Bell et al., 1987a, 1987b, 1987c) developed a range of materials for junior secondary students on the language of functions and graphs, including a section on sketching graphs from words. The materials involved various worksheets addressing student misconceptions. One involved how graphs are not pictures but representations that follow conventions, because some students interpret graphs as pictures, particularly when height is involved in the situation. Another concerned how students can coordinate bivariate data represented in a Cartesian coordinate system, because it was observed that in bivariate situations some students tend to fixate on one variable (Bell et al., 1987a). For a task to represent "how the price of each ticket will vary with the size of the

party” on a bus with a fixed total cost, Swan (1988) found that 37% of 192 students aged 13 to 14 years drew a graph that was decreasing.

2.03.06.02 Recent Studies and Recurring Findings

Chazan and Bethell (1994) briefly described a range of dilemmas students encounter in graphing verbal statements of relationships, including identifying the variables, specifying the units of measurements, deciding which variables are independent and dependent, and deciding whether to represent a continuous line or discrete points.

Coulombe (1997) asked 121 eighth- and ninth-grade students six survey items involving covariation, to address all translations among verbal statements, graphs, and tables of data. Four items involved multiple segments of covariation (e.g., increasing then constant). Responses were assessed in relation to four covariation themes: dependency, multiple patterns, linear patterns, and generalizability. For example, for a task of translating a verbal statement about the height of water in a bath from beginning to run the water to the completion of draining, 78% of students responded showing the multiple patterns of covariation, but only 51% demonstrated all aspects of the dependency, with examples indicating difficulties with representing the constant function. Findings of differences in difficulty across the tasks might be explained by differences in contexts and complexities of the data sets/situations. Generalizability was assessed using tasks of translating tables of discrete data to words or graphs, with evidence, for example, of graphing of line segments to demonstrate awareness of possible interpolation and extrapolation. As the task description did not specify this expectation, however, it is arguable that lack of representing continuity might reflect not a limited ability to generalise, but rather a conscious decision to faithfully present discrete data. In the current study involving

statistical covariation, discrete data cases were often expected based on the context of the data collected.

Mevarech and Kramarsky (1997) asked 92 eighth-grade students to construct graphs to represent the following four girls' claims concerning an association between time studying and achievement grades in school:

1. the more she studies, the better her grades;
2. no matter how long she studies, she always gets the same grade;
3. up to three hours, the longer she studies the better her grades, but beyond three hours, she becomes tired and her grades become lower; and
4. when she studies more, her grades decrease.

The task to graph these verbal descriptions of association, without providing any data, was used to encourage students to focus on the association rather than to plot individual data points. Approximately 10% of responses were excluded from analysis, from students who either represented a general graph schema with no correspondence to the description given or represented relations between the two variables by idiosyncratic methods that were difficult to interpret. Approximately 55% of students appropriately represented claims 1, 2, and 4 using a labeled two-axis graph, whereas only 38% of students correctly represented claim 3 involving a curvilinear function. For incorrect student responses, three common alternative conceptions were identified: (a) only a single point was represented in a graph, (b) only one factor was represented in each of a series of graphs, and (c) an increasing function was represented irrespective of task requirements. The first two conceptions may reflect students' attempts to reduce the complexity of bivariate data sets. These three conceptions were represented in approximately 25%, 30%, and 5% of the 92 students' responses respectively (students could exhibit more than one alternative

conception). Students were then given 12 textbook-based lessons concerning conventions of Cartesian graphing, various data types and corresponding graph forms, and interpretations of a variety of graph forms including distance-time graphs and graphs found in newspaper articles. A subsequent post-test on the same items indicated that students improved at including relevant features such as labels and scales, and that idiosyncratic and generalized graphs were eliminated. There was also a reduction in incidence of the three alternative conceptions to approximately 15%, 15%, and 5% respectively, however for some students these conceptions were robust to instruction, and certain students exhibited these conceptions on the post-test but not on the pre-test. The researchers concluded that further research was required to identify students' alternative conceptions in a variety of contexts before appropriate instruction could be recommended that would directly address these conceptions.

2.03.07 Summary of Research and Future Directions

The research foundations concerning students' understandings and representations of statistical covariation comes from diverse bodies of research, including researchers in psychology, science education, algebra and statistics education. A number of researchers have identified key themes, such as local versus global reading of graphs, confusing rate and value, unpacking verbal statements, positive covariation bias, influence of prior topic knowledge, issues in reducing data complexity, conventions of graphing and alternative forms to suit various purposes. Some research has established frameworks of levels, in some cases across various skills.

A future direction for research, addressed by this study, is a focus on Speculative Data Generation, involving covariation in statistical contexts, and how this skill links to graph production and graph interpretation. The few studies that

have focussed on Speculative Data Generation have tended to involve continuous change only, without the issues of underlying statistical data assumed from the verbal statement. These previous studies also have not assessed responses for Speculative Data Generation in levels similar to assessment frameworks for graph interpretation and graph production.

2.03.08 Prior Research Overlapping this Thesis

A number of studies using tasks, and in some cases response sets, identical to those used in this study have been previously published. In some cases, these studies are reproduced in Appendix 2.

Watson (2000; Watson & Moritz, 1997) asked students to represent “an almost perfect relationship between the increase in heart deaths and the increase in the use of motor vehicles” (p. 55) as reported in a newspaper article. Some students’ graphs were pictures of the context or basic graphs with no context. Some compared single values of each measure without variation, whereas others showed variation but just for one measure. Successful responses were those that displayed the relationship in a Cartesian coordinate system, or by displaying two data series compared over time on the horizontal axis. This analysis was revised and extended in Investigation 1 of this study.

Moritz and Watson (1997) explored students’ interpretations of a graph showing decreasing telephone call rates with increasing call duration. This analysis was revised and extended in Investigation 7.

Moritz (2000) asked students to represent height versus age graphs, a context used by Janvier (1978). It was found that about 90% of upper primary students drew a graph to show that people grow taller as they get older, albeit unconventionally, and most adopted the natural mapping of height as represented vertically. More than

30% of students, however, had difficulties graphing the constant function “when you are 20 years old, you stop growing.” The constant function was identified as a significant task feature requiring students to display two independent axes, one for each variable. A multivariate task incorporating differences between males and females proved more difficult: Some students represented a single comparison of one male and one female to reduce complexity, some a double comparison of heights for two specific ages, and some a series comparison of two trend lines over a series of ages. This analysis was revised in Investigation 2.

Moritz (2004a) asked students to produce a graph of data values for temperature with corresponding times, as reproduced in Investigation 6A. Moritz (2004b) explored two tasks. One task involved Speculative Data Generation in the content of test scores and study times, and the analysis is reproduced in Investigation 4A. The other task involved Verbal Graph Interpretation and Numerical Graph Interpretation of a scattergraph involving noise levels, and the analysis is reproduced in Investigation 8A. Moritz (2003) further interviewed students for the task of noise levels, and the analysis is reproduced in Investigation 8B.

2.03.09 Summary of Key Terms

A summary of key terms is presented in Table 2.02. These terms were introduced in Chapters 1 and 2. They are reproduced here to support the reading of the remaining chapters, and to link to the operational codes used in analysing the data presented.

Table 2.02.

Summary of Key Terms

Term	Section	Description
Covariation terms		
Covariation	1.02	Correspondence of variation of two variables. Types (logical, numerical, and statistical) depend on data types of the variables
Statistical covariation	1.02	Correspondence of variation of two statistical variables that vary along numerical scales
Correspondence	1.02	Relationship, association, function, or dependency between two measures (values or set of values)
Variation within variables	1.02	Distinctly different values for a distinct characteristic
Statistical association	1.02	A relation of values representing measured quantities of distinct characteristics
Causal association	1.02	A covariation relation in which one characteristic causes another
Translation skills		
Speculative Data Generation	1.03	Producing a representation of data (tabular, graphical) by speculating values based on a verbal statement
Graph production	1.03	Producing a graph form, based on values (real or imagined)
Graph interpretation	1.03	Producing responses involving numerical values or verbal statements, based on a graph
Representational forms		
Coordinate graphs	2.01.01	Graphs of bivariate data involving Cartesian axes, in which horizontal and vertical position represent the values of the two variables
Times-series data	2.01.01	A series of data values which have corresponding times
Local data values	2.03.03.01	Data cases with similar values on at least one measure
Global trends	2.03.03.01	Patterns of covariation across entire data sets
Data slices	2.03.04.06	Data points within a subset of the range of one measure
Double line graphs	2.03.05.02	Two line graphs, superimposed on the same framework, to compare the values represented in the graph lines across a common baseline
Case-value graphs	2.03.05.05	Graphs in which one axis or feature represents the data case, such as categorical label, and the other axis or feature represents the data value.

CHAPTER 3. RESEARCH DESIGN AND METHODOLOGY

3.01 SUMMARY

This chapter reports the research design and methodology used in this study. Four research aims are described in relation to assessment frameworks, student performance, development of reasoning, and development across skills. Methodological issues are discussed, in particular the advantage of including both quantitative analysis of survey responses, and qualitative evidence from individual interviews. Tasks are presented, as well as the procedures used and the characteristics of the samples of students surveyed and interviewed. Methods of data recording and qualitative analysis are discussed. Coding methods are described in generic terms and in principle; details are deferred to subsequent chapters as coding is related to the specific assessment framework for each skill.

3.02 RESEARCH AIMS AND EXPECTATIONS

In exploring the development of student understandings and representations of statistical covariation, the current research aimed to build on the previous research of students' graph productions and graph interpretations by also exploring further aspects of Speculative Data Generation involving statistical covariation, as shown in Figure 1.01. The previous research related to Speculative Data Generation, based on drawing graphs from verbal statements, was limited and generally involved contexts of continuous functions over time (e.g., Leinhardt et al., 1990; Mevarech & Kramarsky, 1997; Swan 1985). This study aimed to explore student understandings of statistical covariation, related to understandings of graphs documented by

researchers from statistics education and science education, and also related to understandings of functions as documented by researchers of early algebra development. Four research aims are described in relation to assessment frameworks, student performance, development of reasoning, and development across skills. Further details are provided in the introduction to each investigation to indicate how the investigation contributes to one or more of these aims.

3.02.01 Aim 1 – Assessment Frameworks of Student Responses

The primary aim of the current study was to identify and describe common student responses for representing statistical covariation within assessment frameworks of levels of success. Identification of these levels was based on characteristics fundamental for success on the task, that is representation of appropriate variables and of covariation between those variables (cf. Ross & Cousins, 1993a, 1993b). It was expected that three to six levels would likely be evident in responses to each task, and that these levels would be described in frameworks that would be similar across tasks for each skill.

In addition to response levels, a process of observing qualitative differences in responses without predetermining the groupings demanded allowing for the identification of response categories within each level. Students' conceptions exhibited in responses depend upon the tasks students are asked to undertake; in particular open-ended tasks allow students to exhibit alternative conceptions. Hence it was expected that response categories at each level might differ for various tasks, even for tasks assessing the same skill. For tasks concerning statistical covariation, it was expected that correspondence and variation were relevant components, although just how these components were involved was not pre-determined. The results that

follow show that emphasis on one or other of these often characterised two categories within a given level.

The clarification of the assessment frameworks was intended to identify progressions (levels) and aspects of progress (categories) contributing to the demonstration of competence for given skills in engaging with statistical covariation. These assessment frameworks were considered to provide support for teaching in assessing levels of student responses, and in informing aspects that might promote development of students' understandings and representations of statistical covariation. Jones, Langrall, Mooney, and Thornton (2004) discussed further using models of development in statistical reasoning not only to assess students, but also in the design and implementation of instruction.

3.02.02 Aim 2 – Evidence of Student Performance

A second aim of the current study was to quantify the proportions of students from grade 3 to grade 9 responding within each level and category as identified for Aim 1. Such results, in broad terms, may inform teachers of appropriate expectations for student responses in relation to the curriculum. More specifically, however, quantifying the proportions of students responding with demonstrated performance to tasks that teachers may not have asked of their students, such as tasks of Speculative Data Generation, aimed to inform teachers of ways students might be able to engage tasks above their teachers' expectations. It was expected that the tasks would be appropriately challenging for students in years 3 to 9 of schooling, and that students in higher grade levels would tend to respond at higher levels. In addition, although Aim 1 was to develop assessment frameworks of levels that could be consistently applied across tasks of the same skill, it was acknowledged that response categories may differ across tasks. Similarly, it was expected that student

performance would differ across tasks addressing the same skill, and that these differences would assist in clarifying the component aspects of tasks that affect performance, and hence aspects that would make tasks appropriately challenging to various year levels.

3.02.03 Aim 3 – Reasoning and Reactions to Conflicting Responses

A third aim of the current thesis was to explore students' reasons for responses, and their reactions to the responses of other students. Whereas Aim 1 concerned assessing students' understandings and representations of statistical covariation evident in written responses, Aim 3 concerned exploring thought processes of students when verbally explaining their reasoning, and when exposed to the ideas of other students, some of which were higher level, lower level, or simply alternative to their own responses. This aim concerned how receptive students might be to learning in relation to statistical covariation, and what aspects of students or the stimulus during the interaction might affect this receptiveness.

3.02.04 Aim 4 – Conceptual Development across Skills

A final aim was to explore how conceptual development in one skill is related to other skills, both (a) related conceptually based on the assessment frameworks, and (b) related empirically based on student response data. This aim built upon the outcomes of the first two aims. Conceptual links between each pair of skills were considered based on the assessment frameworks developed as research outcomes in relation to Aim 1. Empirical associations between student performances on different skills were considered in relation to these assessment frameworks. Case studies of student responses across tasks for different skills were also considered. The

outcomes of investigating Aim 4 were intended to provide evidence to address issues in ordering curriculum elements.

3.03 THEORETICAL ORIENTATION

3.03.01 Theoretical Developmental Model for Assessment

In this study, responses were assessed at one of four levels, consistent with the SOLO taxonomy (Biggs & Collis, 1982). The SOLO taxonomy refers to the Structure of the Observed Learning Outcome, a general framework for assessing observed responses, without direct implications for student understanding or ability. The framework gains its generality by referring not to the topic content of the response, but to the structure of the response with respect to whatever elements are relevant to the given task. Response levels differ in the number of relevant elements evident in responses and how these are structured to provide evidence of understandings aligned with valued educational outcomes. The response levels include:

- (1) *prestructural* responses, which do not involve relevant elements,
- (2) *unistructural* responses, which use one relevant element,
- (3) *multistructural* responses, which use multiple relevant elements without integration, and
- (4) *relational* responses, which appropriately relate elements relevant for the task.

This framework was developed into a package of assessment tasks of mathematical problem solving, each task successively harder to elicit various levels of response (Collis & Romberg, 1992). The taxonomy has since been used in a wide variety of

works including mathematics education research (Watson & Moritz, 2000) and large-scale science assessment programs (Ministerial Council on Education, Employment, Training and Youth Affairs, 2004). The generic nature of these descriptors allows them to be applied in the current study with respect to the relevant elements, particularly correspondence and variation, involved in representing statistical covariation.

An additional feature distinguishing these levels was the coherence or consistency among elements: relational responses involve a coherent integration of elements, whereas multistructural responses include multiple elements that may involve inconsistencies, which the student may be aware of but does not resolve to an integrated relational view. Students may offer a sequence of unistructural responses without acknowledging them as a sequence, and thus not be aware of inconsistencies among these responses. This recognition and resolution of inconsistencies is discussed in the next section in relation to the role of cognitive conflict in promoting conceptual development.

Beyond relational responses, further development of responses was termed *extended abstract* by Biggs and Collis (1982), in which a higher level of abstraction or symbolisation is used to address the task. This higher-level response may occur in another mode of functioning. The SOLO taxonomy was part of a more comprehensive neo-Piagetian developmental theory involving modes of functioning (sensori-motor, ikonic, concrete-symbolic, and abstract) that were similar to Piaget's stages both in their characteristics and the ages at which they were typically expected to emerge. Of particular interest for school-based tasks was the concrete-symbolic mode. Further research studies (e.g., Watson & Moritz, 1999) have observed two

cycles of Unistructural-Multistructural-Relational learning for responses in the concrete-symbolic mode.

For the purposes of this study, modes were not considered. It was assumed that most of the evidence observed in student responses would be in the concrete-symbolic mode, referring to the context of the data, rather than in an abstract mode. Responses not relevant to this mode were termed pre-structural rather than ikonic, indicating such responses were naïve-intuitive, even though they may include connotations of imagery. In addition, the four levels were considered as being related to previous research that described levels of graph comprehension (Curcio, 2001) and understanding associations (Swatton, 1994). Although the coding of responses was open to the possibility of fewer or more levels being identified, a four-level framework was found to result in sustainable and meaningful distinctions among responses.

3.03.02 Misconceptions and Cognitive Conflict

The Latin root of *education* implies leading out, widening horizons, and encountering new information or experiences. Piaget (1983) described the way sense is made of the world as equilibration, a balancing act of what is already believed and what new information is encountered (Woolfolk, 1993). When exposed to new information, Piaget suggested individuals may invoke one of two processes. *Assimilation* involves interpreting the new information according to some similarity to existing cognitive schemes, a process which may involve distorting or ignoring the new information. If there is awareness that existing schema do not fit, disequilibration or cognitive conflict occurs, and involves either further search for similar schema or the process of *accommodation*, that is building a new cognitive scheme to house the new information (Posner, Strike, Hewson & Gertzog, 1982). As

noted in relation to the SOLO taxonomy, a key feature of learning and performing at high levels involves awareness of inconsistencies and methods to resolve them (Behr & Harel, 1990; Steffe, 1990; Tirosh, 1990; Vinner, 1990). Inconsistencies often arise due to partial understandings or misconceptions. Various research studies have identified students as having misconceptions (e.g., Clement, 1989; Shaughnessy, 1977, 1981, 1982; Swedosh, 1996, 1998; Swedosh & Clark, 1997, 1998, 2000) or less pejoratively, alternative conceptions (e.g., Mevarech & Kramarsky, 1997).

Cognitive conflict refers to one's awareness that an experience or idea is not consistent with one's existing concepts (Behr & Harel, 1990; Dreyfus, Jungwirth & Eliovitch, 1990; Ellerton & Clements, 1990; Hewson & A'Beckett-Hewson, 1984; Tirosh, Stavy & Cohen, 1998; Watson & Moritz, 2001b). Such conflict may prompt one to modify one's concepts if the conflict creates dissatisfaction with an existing conception, and a new competing conception is available which is intelligible and appears initially plausible (Posner, et al., 1982; Strike & Posner, 1992).

Group-work is often advocated as an appropriate setting for learning by exposure to other students' ideas that are conflicting, intelligible, and without the authority status of a teacher (Sfard, Neshler, Streefland, Cobb & Mason, 1998; Webb, 1991). To promote intelligibility of new ideas, theorists such as Vygotsky have emphasised social aspects of learning, suggesting that students are ideally grouped with others who have slightly higher-level ideas that are intelligible; the range of intelligible ideas is described as the *Zone of Proximal Development* (Forman & McPhail, 1993; Goos, 2000). Some researchers have explored the resolution of student uncertainty, acknowledging different authorities to which students can appeal in reaching a resolution (e.g., prior experience, empirical data, a knowledgeable person, or a text) (Clarke & Helme, 1997) and different meta-cognitive activities that

assess whether the uncertainty has been resolved satisfactorily where no clear authority is available (Goos, 1998).

Collaboration has been described as involving conversation including components of (1) *New Ideas* proposed, and (2) *Local Assessments*, that is, evaluations of stated ideas (Goos, 1994; Goos & Galbraith, 1996). These ideas and assessments may be *explicit*, such as stating the idea or acknowledging the conflicting cognitions, or may be *implicit*, not referring to the ideas of others previously stated, and hence providing no evidence of being aware that the new idea conflicts with others. Using this distinction, Moritz (1996) proposed five theoretical conversational sequences in response to an initial idea, as shown in Figure 3.01.

Passive Agreement provides no clear evidence of reasoning, with an explicit assessment of agreement and an implicit idea. *Active Agreement* involves a student explicitly agreeing and explicitly restating or paraphrasing the idea. Collaboration is most clearly evident in *Collaborative Building*, when a student explicitly states a new idea, based upon implicit agreement to a previous idea. *Basic Disagreement* involves explicit assessment of disagreement. *Replacement Arguments* involve a student expressing a difference of opinion by an explicit idea. Distinctions between these sequences are blurred in practice. Classroom collaboration, and research of collaborative group-work, commonly involves personal group dynamics, and cognitive conflict may rarely be observed if collaboration is not optimal (e.g., Watson & Chick, 2001).

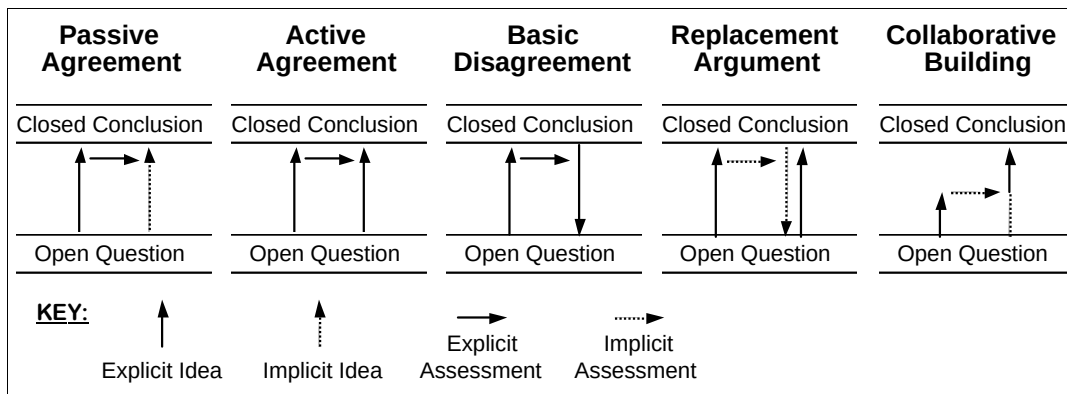


Figure 3.01. Five possible response sequences.

To emphasise conflicting cognitions in the classroom, some mathematics educators (e.g., Shaughnessy, 1977; Swedosh & Clark, 1998) have recommended that students should first clearly express their own ideas, and then teachers can ensure students are exposed to, or confronted with, cases where these conceptions fail. In attempting to promote conceptual change, Dreyfus et al. (1990) suggested science teachers should lead their students through three stages of scientific inquiry: prediction, observation, explanation. Statement of a prediction involves a clear assertion of knowledge based on an existing conception. Careful observation should then allow any conflict between the observation and prediction to become apparent, as both have been made explicit. Considerable research in statistics education has asked students to state their ideas, but little has then provided students with conflicting views of others and asked them to decide their preferred response, ensuring consistency in aspects of “the other” (e.g., familiarity with the person, gender, and age) and “the other’s idea” (e.g., terminology).

Some researchers have employed clinical interviews to prompt conceptual change. Posner and Gertzog (1982) described ideas for kinds of tasks to elicit students’ concepts and potential conceptual change: *general open-ended questions*, *problem-solving tasks* with “think aloud” strategies, *garden-path tasks* designed to

lead students' thinking to reveal the counterintuitive nature of some students' conclusions, and *comprehension tasks*, in which students attempt to reconstruct the reasoning that another student would have used to arrive at a given solution. More recently, some researchers (e.g., Macbeth, 2000) have based designs on garden-path tasks by interviewing students about their knowledge, and then "presenting them with practical tasks and exercises designed to disrupt, challenge, and perhaps even change their presenting understandings" (p. 237). This method "provides the possibility of witnessing conceptual change, on camera" (Macbeth, p. 237). Other researchers (e.g., Maher & Martino, 2000) have employed group settings to get students to explore their own reasoning by explaining to other students, and in the process, reinventing some cognitive structures. The tension between forced prompting and free exploration was described by Freudenthal (1991, cited by Maher & Martino, 2000, p. 269) as, "Guiding means striking a delicate balance between the force of teaching and the freedom of learning."

Drawing on both aspects of (a) prompting with conflicting situations and (b) social contexts involving communicating ideas and understanding those of others, some studies (e.g., Watson, 2002; Watson & Moritz, 2001a, 2001b) have asked students for their initial responses, and then presented them with video extracts on computer of other students' responses at higher and/or lower levels, selected by the interviewer to prompt cognitive conflict. The views in the video extracts were presented identically every time, thus meeting the demands of a controlled research study. The method attempted to retain the virtue of students learning from other students, without issues of authority status if researchers or teachers were to present their views. There were also no pre-existing social connections between the students expressing viewpoints, which might otherwise influence research outcomes. These

studies found many students did not change their minds, however some did agree to higher-level prompts. In one study (Watson & Moritz, 2001a) involving both longitudinal and cognitive conflict investigations, the proportion of students changing their preferred choice from incorrect to correct using the cognitive conflict technique was greater than that for students assessed longitudinally over three or four years. Although it is acknowledged a preferred choice among options in the cognitive conflict setting does not guarantee the student could subsequently supply a similar response without this support, these studies generally concluded that this research technique showed promise for exploring how students learn to reason at higher levels.

3.04 DIMENSIONS OF RESEARCH METHODS OF STUDENT ASSESSMENT

Various methods of assessment of school students' understanding differ according to their scope and focus to achieve different purposes. Educational systems often employ large-scale tests, using multiple-choice tasks with reliable psychometric properties to assess student ability of a broad learning area. Teachers often employ open-response paper-based tests, and then assign scores for responses to each task. Researchers have employed both of these methods, but are increasingly exploring in detail dialogue from classroom discussion or structured interviews to investigate aspects of the educational setting or process in relation to understanding of individual students or groups. These methods may be considered as part of a spectrum from large-scale quantitative methods broadly sampling the student population and student understandings to provide a summative assessment of student ability, to small-scale qualitative methods that explore more deeply detailed understandings or interactions of a small number of students. It is important to note,

however, that the strength of quantitative methods for reliability, consistent administration, and generalizability, and the strength of qualitative methods detailing rich phenomena, are not mutually exclusive, but rather two dimensions of breadth and depth (see Patton, 2002, pp. 227-228). For example, the Trends in International Mathematics and Science Study (TIMSS, 1999, 2003) incorporated into a large-scale international study detailed methods such as qualitative coding of open-response tasks and of video extracts of many classroom lessons. All assessment methods are constrained by pragmatic considerations, including the student time devoted to the assessment, and the researcher time and resources to design assessment instruments, gather responses, mark them, and report on them. The current study employed survey techniques administered to classes of students, and interviews administered to individual students, in an attempt to provide both breadth of sampling across a range of students and responses to a range of tasks, and depth by considering interview dialogue of a subset of students. These complementary methods were used to strengthen the study, an example of triangulation approaches (Patton, 2002).

3.04.01 Issues in Task Design for Paper-based Questions

Multiple-choice tasks are often employed in large-scale assessments, particularly to permit easy marking of responses. Open-response tasks, however, are commonly advocated as preferable to permit students to exhibit their own conceptions of the relevant features of the task. Mevarech and Kramarsky (1997) noted that the open-ended nature of their task “may explain why some of the alternative conceptions diagnosed... were not identified in previous studies” (p. 255). In the current study, the representation tasks were designed to be set in realistic contexts, and the desire to allow for alternative conceptions led to providing students

with a blank page on which to draw their representation as done by Mevarech and Kramarsky, rather than providing labeled axes that offer structural support but constrain the form of representations as done by Bell et al. (1987b). The choices of variables and of axes on which the variables were placed were considered structural elements of interest.

In employing open-response paper-based questions, it is particularly important that tasks and rubrics are carefully designed. Multiple-choice tasks require clarity of task wording and marking is trivial: as Woolfolk (1993) commented, “All test items require skillful construction, but good multiple-choice items are a real challenge” (p. 547). In contrast, extended tasks or interviews often permit repeated paraphrasing or clarifying of the task with the student. Open-response paper-based questions must balance avoiding ambiguity about what is demanded by the task, with ensuring openness that permits students a degree of discernment. As Woolfolk (1993) commented, “The most difficult part of essay testing is judging the quality of the answers: but writing good, clear questions is not particularly easy, either” (p. 548).

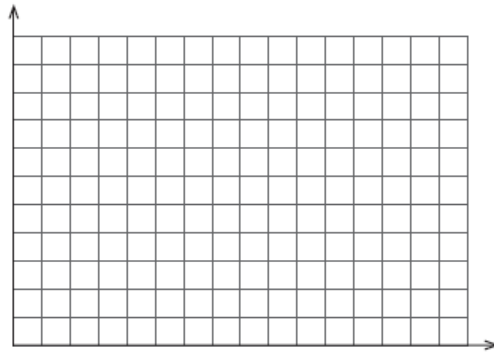
It is common practice in test development for scoring (or coding) schemes to be developed concurrently with tasks, in order to assist the task writer to refine wording of the task to correspond to the complete response that would demonstrate the understanding being assessed. If a task is poorly worded, students may respond in ways that minimally satisfy without demonstrating more complex skills or understandings that the task writer aimed to assess. This distinction has been referred to as functional versus optimal responses (Fischer & Knight, 1990): the task should optimally challenge students to demonstrate the depth of the understanding appropriate to the task, but also permit accessibility for students with partial or very

limited understanding to engage the task and demonstrate what they do know without being intimidated by a task they consider to be too challenging to offer a response. In exploratory studies, it may not be possible to anticipate in detail the richness of student responses, thus rubrics may be broadly defined in terms of partial and complete responses for the demands of the task (Woolfolk, 1993). It is also good practice to trial tasks with the target audience, to troubleshoot alternative interpretations. These practices were incorporated into the current study by consulting experienced researchers with draft tasks, and then piloting the tasks with a student before administering to larger numbers of students. Some tasks were also refined based on evidence from previous investigations.

An example of a graphing task and coding scheme, similar to one used in the current study, is shown in Figure 3.02. The task emphasised “realistic scale” in the task wording, which was a key feature of criteria in the coding scheme to differentiate correct responses from partial and incorrect responses. In contrast, graph form was not referred to in the task as it formed no part of the criteria apart from having axes that may be scaled. The coding scheme included different categories at each level of correctness; the code provided a diagnosis of why a level of correctness was assigned to the responses. The coding scheme included not only the wording of the criteria for each code, but also an example to assist those ratings responses. Explicit criteria evident in the response focused attention on what should be evident in the written response, rather than that which might presumed, often falsely, about the student understanding based on criteria not within the responses, such as awareness of a students’ grade level influencing the assigned level of the response.

Task:

Using the set of axes below, sketch a graph which shows the relationship between the height of a person and his/her age from birth to 30 years. Be sure to label your graph, and include a realistic scale on each axis.

**Coding:**

Code	Response
Correct Response	
20	All the following features correct: 1. Correct scales and labels on both axes: Age: 0 - 30 years Height: 0 - 200 cm OR 0 - 80 inches (0 - 7 ft) 2. The graph starts at approximately 50 cm (20 inches). 3. Maximum height is reached at a realistic age (14 to 20 years). 4. The graph is horizontal after age of maximum height. 5. Maximum height is reasonable.
Partial Response	
10	Incorrect start of graph. Other features correct. <i>Examples: Graph starts at height of zero.</i> <i>Graph does not start at year zero.</i>
11	Unrealistic age for maximum height. Other features correct.
12	Incorrect graph after age of maximum height. Other features correct. <i>Examples: Graph continuously increases in the range of 20 - 30 years.</i> <i>Graph decreases after age of maximum height.</i>
13	Includes incorrect scales or labels. Other features correct.
19	Other partial.
Incorrect Response	
70	Includes incorrect start of graph AND incorrect scales. Other features correct.
71	Includes incorrect start of graph AND incorrect graph after age of maximum height. Other features correct.
79	Other incorrect.
Nonresponse	
90	Crossed-out/erased, illegible, or impossible to interpret.
99	BLANK

Figure 3.02. A sample task and coding scheme from TIMSS (1995)

<http://isc.bc.edu/timss1995i/Items.html> (see C_Items.pdf#page23 - Question A10).

When categorising responses, some student responses prove difficult to categorise. Some students offer responses that suggest an alternative interpretation of the question such that the response does not exhibit features referred to in the coding

scheme – this issue is generally minimized by careful task design and trialling of tasks. Some student responses are difficult to rate because they exhibit features that might suggest assignment to more than one category. The magnitude of this issue is commonly measured by having two raters independently assign ratings, and then calculating an inter-rater reliability. In general, good coding schemes maximise the differences between categories and minimize the differences within categories, termed *external heterogeneity* and *internal homogeneity* (Patton, 2002, p. 465). A single category used for a large number of disparate responses provides less descriptive power, and suggests that the category should be divided into more restricted meaningful categories. Defining categories thus may involve iterative refinements to coding (Lesh & Lehrer, 2000; Maxwell, 1996; Miles & Huberman, 1994), a practice that reduces the relevance of inter-rater reliability in exploratory studies.

Preliminary investigations (Data Collections 1 and 2) in the current study employed written coding schemes and techniques of inter-rater reliability, providing evidence of high reliability, as well as the validity of the framework shown in examples. Responses to tasks in later investigations (Data Collection 3) were iteratively coded by the researcher and reviewed at times by the researcher supervisor, with examples shown in results or appendices providing evidence of the validity of the assessment framework.

3.04.02 Issues in Task Design for Interviews

The interview setting has a number of advantages over survey responses:

- students can speak detailed verbal descriptions more easily than write them,

- the researcher can listen to students speaking their thoughts as they develop with passing time, rather than merely observe the written product of their reasoning (Lesh, Hoover, Hole, Kelly & Post, 2000),
- the interviewer may intervene to clarify a simple issue and see whether such clarification enables the student to respond at a higher level, and
- video recording permits detailed and repeated observation of a student's gaze and pointing that may indicate the bases for misinterpretations or limited interpretations.

Qualitative interviews may range from informal conversation to standardized fully-scripted protocols, and often involve a combined approach (Patton, 2002, pp. 341-347). Standardized protocols minimize the variation among interviews, and thus permit stronger comparisons and simplify analysis. Standardized protocols also focus the dialogue to ensure interview time is used efficiently. Less structured dialogue, however, more readily promotes rapport with the participants. Less structure also allows more detailed exploration of selected ideas, particularly via pursuing natural follow-up questions, which further give interviewees a sense that they are encouraged to openly discuss their thinking, rather than giving brief definitive comments. These extremes and benefits of each approach are examples of the more general breadth versus depth issues in research design discussed in Section 3.03. Patton (2002) describes features of open-ended questioning, without bias, to encourage interviewees to express their views. In relation to probing follow-up questions he comments that

a kind of clarifying follow-up question is the *contrast probe* [...] The purpose of a contrast probe is to “give respondents something to push off against” by asking, “How does *x* compare to *y*?” This is used to help define the

boundaries of the response. [...] Probing is a skill that comes from knowing what to look for in the interview, listening carefully to what is said and what is not said, and being sensitive to the feedback needs of the person being interviewed. Probes are always a combination of verbal and nonverbal cues. Silence at the end of a response can indicate as effectively as anything else that the interviewer would like the person to continue. (Patton, 2002, p. 374)

Other issues involved in interviews concern the collection of data, often involving audio and possibly video recording. Provided sound recording is audible, this permits repeated replaying of dialogue to re-examine the dynamics throughout the dialogue, and it frees the interviewer from being concerned with taking notes. Privacy issues involved with recording exceed the simple consent to be interviewed, both for ethical research and for establishing rapport: interviewees may feel more willing to discuss their ideas openly if the audience of the recording is clearly confined to a small number of researchers for that purpose only, and that they, as interviewees, will remain anonymous. Coding of interview dialogue is a significant challenge to identify efficiently key emergent themes. Iterative coding is recommended with open consideration given to discern and refine emergent themes (Clarke, 1998; Lesh & Lehrer, 2000).

In the current research, semi-structured interviews were employed, using a series of questions, to allow students to explain personal responses and to evaluate the merits of other responses, and using natural follow-up questions, to encourage reflection in relation to conflicting ideas. In some cases the dialogue was readily coded in relation to agreement and disagreement with prompts. Other dialogue, however involved students' descriptive evaluation of various features that proved

difficult to code formally, but did indicate themes in reasoning among related concepts used in engaging statistical covariation.

3.05 DEVELOPING TASKS

Tasks described in this study were developed in four stages. The first stage involved tasks developed as part of a broader research program that served as starting points for the current investigations. The second stage involved trial of tasks. The third and final stages involved systematic design of a set of tasks exclusively to address the research of the current study. Although this study is presented according to the skills under investigation – Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation – the research may also be considered according to these four stages as the progressive development of the research agenda. These stages show an intentional progressive research from broad large-scale quantitative research methods to small scale qualitative methods, both extremes complementing each other to provide more complete appreciation of students' understandings and representations of statistical covariation.

3.05.01 Task Development – Data Collection 1

The first stage occurred in 1993 as part of a larger research project led by Dr Jane Watson at the University of Tasmania exploring student understanding of chance and data. Tasks developed for this project included a 20-item standard written survey, a series of 10 interview protocols designed similarly to the four-part Collis-Romberg problem-solving profiles (Collis & Romberg, 1992), and a 10-item media survey, involving tasks using newspaper extracts as stimulus pieces. The response set considered within this study included responses to two tasks as part of media survey,

as shown in Figures 3.03 and 3.04. The tasks and the responses they generated were of such interest that they formed a starting point at the commencement of these doctoral studies and were adopted into this study. The tasks were developed with a statistical literacy focus but were of an exploratory nature, without clear rubrics for coding responses. Some preliminary outcomes for the two tasks have been reported previously (Moritz & Watson, 1997; Watson, 2000). These two tasks were administered as described in Data Collection 1 in following sections.

Task 1 (Figure 3.03) was based on an extract of a newspaper article (“Family car,” 1991), as an example of where statistical literacy is required within a social context. The questions were left as open-ended as possible to permit students to exhibit their own conceptions of the relevant aspects to address. The first question to “Draw and label a sketch of what one of Mr. Robinson’s graphs might look like” did not specify which variables to represent, to allow students to identify the significant variables in the story. The instruction to label the graph encouraged students to specify the variables represented. The desire to allow for students’ alternative conceptions led to providing students with a blank page on which to draw their representation as done by Mevarech and Kramarsky (1997), rather than providing labeled axes as done by Bell et al. (1987b). It was felt that although axes would offer structural support, they might also constrain the form of representations. The second question, “What questions would you ask about his research?” was intended to assess students’ disposition and sensitivity to critique and question causal claims based on statistical association without being alerted to the issue by undue prompting (Watson & Moritz, 2000a). It was also worded in a general way so that students could identify any features that they considered might be needed to support the verbal claims made.

Family car is killing us, says Tasmanian researcher

Twenty years of research has convinced Mr Robinson that motoring is a health hazard. Mr Robinson has graphs which show quite dramatically an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles. Similar relationships are shown to exist between lung cancer, leukaemia, stroke and diabetes.

Q1. Draw and label a sketch of what one of Mr. Robinson's graphs might look like.

Q2. What questions would you ask about his research?

Figure 3.03. Task 1: Heart deaths versus use of motor vehicles
(*The Mercury*, June 11, 1991, p.2).

Task 2 (Figure 3.04) was based on an extract of a newspaper advertisement showing how telephone call rates are related to phone call duration. The graph has a number of interesting features. The most apparent is the presentation of unusual pictures of squashed telephones within the shaded region of a bar graph. Each of the variables of the graph involves potential for misunderstanding. Rate of the telephone call involves the proportional notion of cost per unit time, which may be confusing because time is the other variable – the rate-value confusion referred to by previous researchers (e.g., Mokros & Tinker, 1987). Further, the rate is not directly presented in the graph as a cost per minute, but rather as a percentage off the standard rate, and thus the percentage values reported are negatively related to the actual charge rate. The minutes of the call are discrete time intervals of unequal spacing. The data involve a negative association of five values, complicated by the scaling of the graph, which is uneven on both axes and does not begin at zero on the vertical axis. In short, this graph from a newspaper advertisement is authentic to the many complexities of graph interpretation to which students may be exposed in social

contexts. Students were asked to explain the general meaning of the graph verbally (Q1), to identify unusual features (Q2), and use the graph to determine numerical values for call costs at simple (Q3) and compound (Q4) levels.

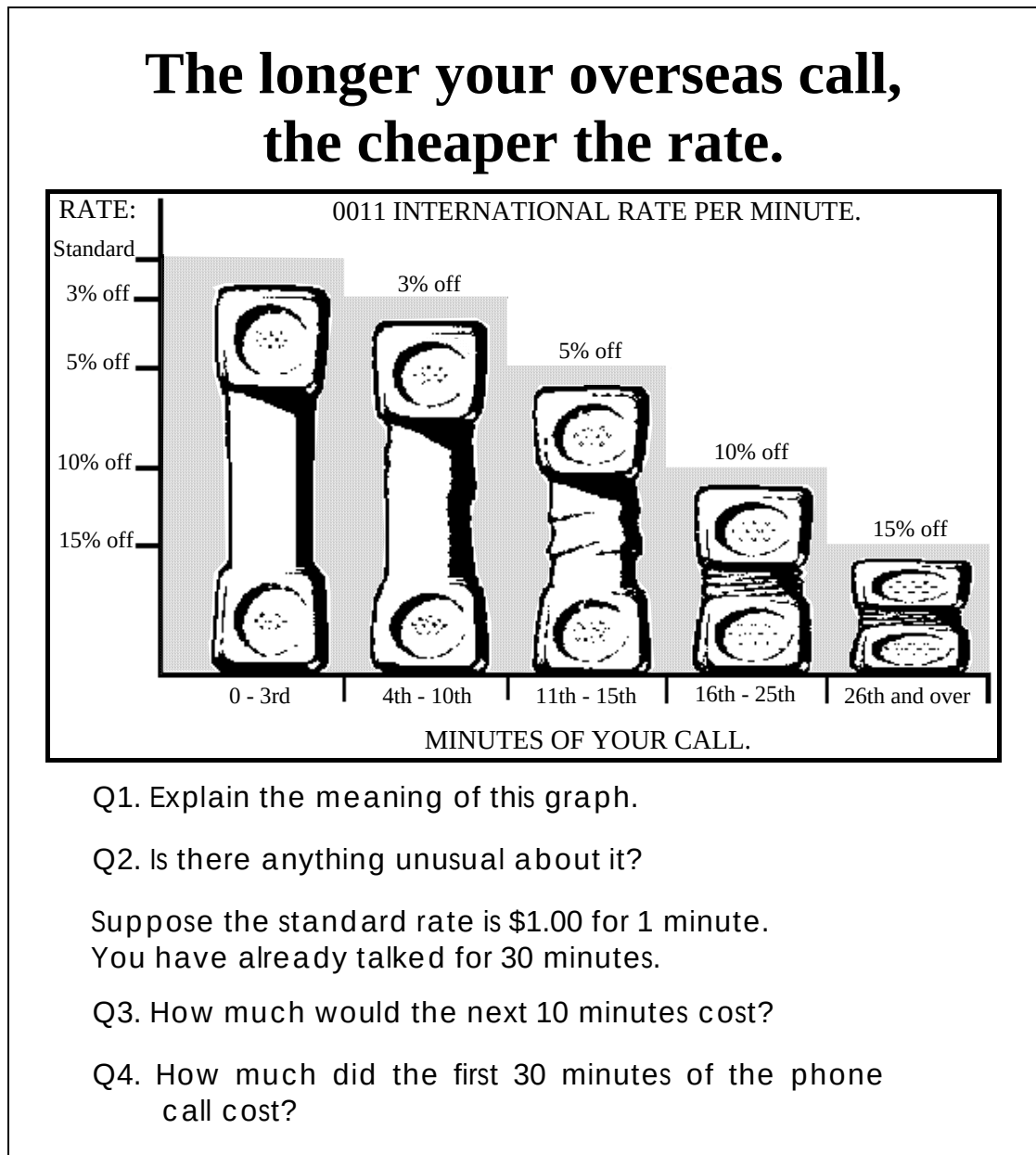


Figure 3.04. Task 2: Telephone call rate versus call time, based on a newspaper extract (*The Mercury*, 22 July, 1993, p.17).

3.05.02 Task Development – Data Collection 2

The second round of task development was related to a pilot study undertaken in 1998 that was also part of a broader research study. The broader study included two

survey tasks exploring students' understanding of variation in certain contexts. Task 3, shown in Figure 3.05, was trialled. The context of height versus age has been used to effect by researchers of primary school students (e.g., Ainley, 1995) and secondary school students (e.g., TIMSS, 1995). Q1 and Q2 in Figure 3.05 show the types of tasks often asked, which involve representing curvilinear covariation of height and age based on common topic knowledge. Q3, however, introduces a third, categorical, variable of gender to explore how students represent the two levels of this third variable using two-dimensional space.

One day at school, the class began to talk about how tall people are. The teacher measured everyone's height. Then they began to talk about their brothers and sisters. Some were taller, and some were shorter.

David said, "People grow taller as they get older."

Q1. Draw a graph to show what David is saying. Label the graph.

Mary agreed, "People do grow taller. But when you are 20 years old, you stop growing."

Q2. Draw a graph to show what Mary is saying. Label the graph.

The teacher said, "For 10 year olds, girls and boys are about the same height. But men usually grow to be taller than women."

Q3. Draw a graph to show what the teacher is saying. Label the graph.

Figure 3.05. Task 3: Height versus age.

3.05.03 Task Development – Data Collection 3

The third round of task development was undertaken in 2001 as the major focus of this study. Based on the responses obtained to Task 1 (Heart deaths) and Task 3 (Heights), not only was there great interest in tasks involving what was yet to be termed Speculative Data Generation, as a means of providing a window to student

understanding of covariation and statistical literacy. There was also interest in how the generally poor responses obtained from junior secondary students to Task 1 (Heart deaths) contrasted with the generally accomplished responses by senior primary students in response to Task 3 (Heights). The differing success of students across these tasks deserved further investigation of Speculative Data Generation in a context not directly intuitive with a positive association as Task 3 but not as complex as the multivariate and complex variable measurement in Task 1. It was acknowledged that these tasks differed in the familiarity of the contextual variables, in the clarity of the appropriate variables in either bivariate or multivariate settings, and in the complexity of data measures and thus covariation between them. It was planned that interviews be conducted to gain more evidence from students' verbal comments than their written responses to these tasks alone offered.

It was acknowledged, however, that survey administration prior to interviews was a means of data collection that permitted (a) gaining responses from a wider variety of student understandings for less school disruption and less researcher time than responses drawn during interviews, and (b) that these survey responses provided a means of strategically selecting a range of student understandings to sample for interview. Thus it was planned to develop further survey tasks involving Speculative Data Generation that varied the contexts and data complexity to explore the effect of these aspects.

In addition, it was acknowledged that Speculative Data Generation responses might for some students be limited by graph production skills. To assess conceptual development across Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation, a set of tasks, with comparable data set complexity, was required, with tasks assessing each skill. One

means of controlling complexity was to fix six data points within various tasks (see Tasks 4, 5, and 6 following), and to select simple wording and contexts that were likely to be familiar to most students.

Initial drafts were developed by the researchers, as found in Appendix 3A. Some of these emphasized natural mapping of the vertical axis, for example (a) representing the height and horizontal distance of a ladder placed at various angles against a vertical wall, or (b) representing the height of water in a bath tub for various events, as used by Swan (1985). Other draft tasks emphasized statements from actual newspaper articles, with the intention of using interviews to discuss with students their own graphs compared with those shown in the news articles.

A consultation among the researcher, the supervisor, and two other mathematics curriculum consultants with extensive experience in teaching, identified the following issues:

- assist reading for grade 3,
- use a san-serif font, especially for grade 3,
- on front page, delete PhD title,
- on front page, delete grade box, and type each grade form separately,
- delete words like “evidence,” either by deleting the question, or by rewording, and
- focus attention on a few tasks with similar complexity of data sets.

A pilot test of these tasks was conducted with a male third-grade student. The intention of this pilot was to ensure tasks were comprehensible, accessible and engaging at this year level, and to judge timing for administration. The student was described by a parent as not the best mathematics student, but generally of quite good ability. These tasks, along with the responses obtained, are seen in Appendix

3B. The student drew an increasing trend of height versus age, although acknowledging some difficulty graphing in the absence of data; as this is central to Speculative Data Generation, the task was not modified, although it was acknowledged that students may need encouragement to realise that the tasks permit them to make up data to show the required aspects. For other tasks, the student showed evidence of comprehension. To graph test scores versus study times, for example, the student wrote, “Student A studied quite a lot and got 10/10; student B didn’t study much at all and got 3/10; student C studied quite hard and got 9/10; student D studied quite hard but got 7/10,” and also used topic knowledge to comment “a student might study hard but the wrong subject, and so get a low score.” In general, the tasks were successful in terms of accessibility and timing; the graphing tasks took 10-15 minutes each.

3.05.04 Survey Tasks – Data Collection 3

The final survey included tasks concerning Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation. The surveys are presented in Appendix 3C. The graph production tasks were placed before interpretation tasks to ensure exposure to the printed graphs in the interpretation tasks did not suggest a graphing method, and similarly Speculative Data Generation tasks preceded the graph production tasks to ensure exposure to types of data sets did not suggest a method of Speculative Data Generation.

Three tasks were used to assess Speculative Data Generation:

- (1) Task 3 - Heights, as previously presented in Figure 3.05.
- (2) Task 4 – Test scores, shown in Figure 3.07, and
- (3) A revision of Task 1 – Heart deaths, shown in Figure 3.06.

Task 1 – Heart deaths was revised to provide greater focus for students about which variables were involved, as shown in Figure 3.06. The text of the news extract was further edited to focus attention simply on heart deaths and motor vehicle use, and the question more specifically referred to the verbal statement to be graphed. In addition, the second question was refined in clarifying the claim and the basis for the claim, and asked students to comment on the reasoning. It was expected that these modifications would raise student performance from the initial form of Task 1 – Heart deaths.

**Family car is killing us,
says Tasmanian researcher**

Mr Robinson believes motoring is a health hazard. He has graphs that show an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles.

Q1. Draw a graph to show:

“an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles.”

Label the graph.

Mr Robinson says “Motoring is a health hazard.”

His reason is based on the graphs that show “an almost perfect relationship.”

Q2. Do you think the graphs are a good reason to say this?

☐ YES or ☐ NO

Please explain your answer.

*Figure 3.06. Task 1 (revised): Heart deaths versus use of motor vehicles
(The Mercury, June 11, 1991, p.2).*

Task 4 – Test scores, shown in Figure 3.07, was developed using a context with variables likely to be familiar: study time and academic grades are experiential for students, and were used by Mevarech and Kramarsky (1997). The context was also chosen so students would expect a positive covariation between the variables, but the task described a negative covariation, forcing students to rely on the data rather than prior beliefs. The negative covariation in the task was intended to avoid success by drawing a stereotypical positive slope line graph. Interestingly, although there is a common perception of a positive association, one which teachers might often wish to reinforce to students, a curvilinear relation has been observed on the National Assessment of Educational Progress, in which performance peaked at 15 minutes average homework time for students in grades 4 and 8, as shown in Figure 3.08.

Anna and Cara were doing a project on study habits. They asked some students two questions: • “What time did you spend studying for the spelling test?” • “What score did you get on the test?” Anna asked 6 students. She used the numbers to draw a graph. She said, “People who studied for more time got lower scores.” Q1. Draw a graph to show what Anna is saying for her 6 students. Label the graph.
(Positive association format) ... She said, “People who studied for more time got higher scores.” Q1* Draw a graph to show what Anna is saying for her 6 students. Label the graph.
She met another student named Sarah. Sarah got a low score on the test. Anna thought, “I bet Sarah studied for a long time. If she studied for less time, she would have got a higher score.” Q2. Does Anna have a good reason to think these things? ☐ YES or ☐ NO Please explain your answer.
Cara asked 6 students. She used the numbers to draw a graph. She said, “Students studied different amounts of time, but the times were not related to the scores they got.” Q3. Draw a graph to show what Cara is saying for her 6 students. Label the graph.

Figure 3.07. Task 4: Test scores versus study times

Note: Third- and fifth-grade males received Q1* in place of Q1; Ninth-grade males only received Q3.

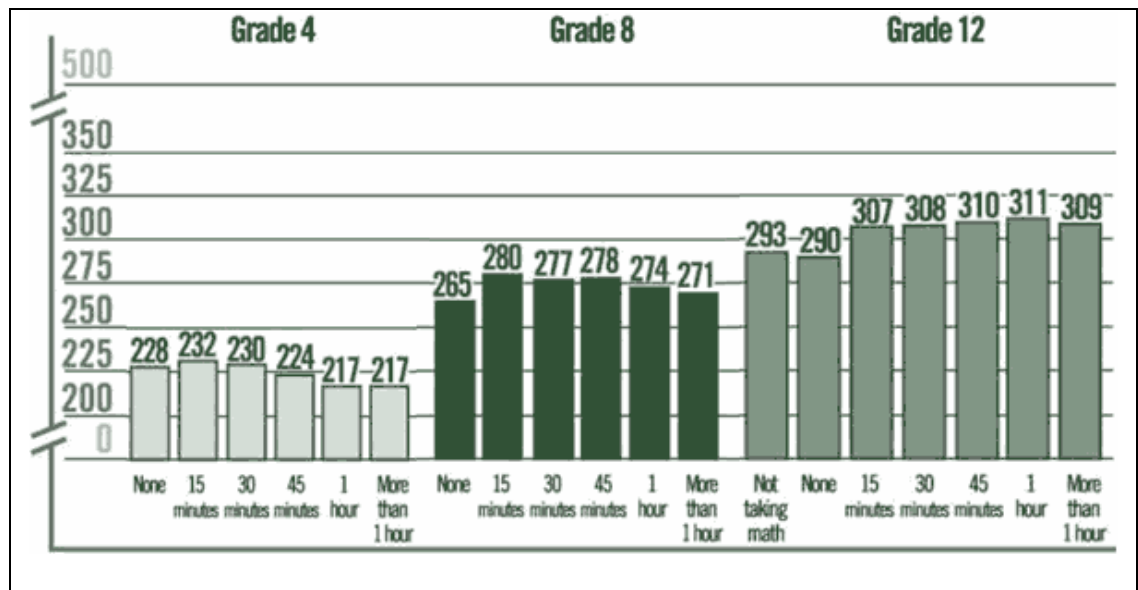


Figure 3.08. Graph of average mathematics scores by students' report on time spent daily on mathematics homework at grades 4, 8, and 12 from the National Assessment of Educational Progress, 2000 Mathematics Assessment, available at <http://nces.ed.gov/nationsreportcard/mathematics/results/homework.asp>.

Task 4 (Test scores) was also administered in a positive covariation form instead of the negative form to third- and fifth-grade males. The different forms were designed to explore whether students might respond differently due to their prior beliefs about the covariation. The tasks were worded to support a statistical context for covariation, such as awareness of the data collected and of possible variability from a perfect linear fit. No axes were provided, to permit students to decide the numbers and types of variables to represent and to develop their own form of representation.

Whereas three tasks were developed to explore Speculative Data Generation in detail, a single task was used to investigate Coordinate Graph Production. For Task 5 – Temperatures, shown in Figure 3.09, students were presented with a context involving temperature in relation to events involving a heater and a window, and with a data table including six numeric values of temperature and time. Students were asked to draw a graph to show how the temperature changed over time. This task was revised from the initial draft, which originally did not provide a numerical table of data to graph, in order to ensure a task to graph given data would assess graph production skills without the additional need for skills of Speculative Data Generation. The context of temperature was chosen as likely to be familiar from viewing television graphs (e.g., Figure 2.02). The narrative context, which described turning the heater on and off in a room and opening the window, was used not only to aid interpretation of the task, but also to allow students to decide which features of the narrative or the data table were most important to represent in a graph to show how the temperature changed over time. The data table involved discrete data points, but the context indicated these data were sampled from measures that were likely to be recognized as continuous, providing clues about intermediate values to represent gradual change. Six data points were chosen to match other questions. The data included repeated temperatures for different times to ensure students were required to distinguish two variables, following studies reporting student difficulties with the constant function (Nemirovsky & Tierney, 2001; Sherin, 2000).

A science class was studying temperature. They used a thermometer to measure the room temperature every 5 minutes for 30 minutes.

First they turned a heater on for 15 minutes.

Next they turned the heater off for 10 minutes.

Lastly they opened the window for 5 minutes.

They wrote down these numbers.

Time (Minutes)	5	10	15	20	25	30
Temperature (°C)	15	20	25	25	25	15

Q1. Draw a graph to show how the temperature changed over time.

Q2. What temperature do you think the room was before they put the heater on? (Even if you are not sure, please estimate or guess.)

Please explain your answer.

Figure 3.09. Task 5: Temperature change over time.

Verbal Graph Interpretation and Numerical Graph Interpretation were assessed using one task, Task 6, as shown in Figure 3.10. Two other graph interpretation tasks were developed and administered, but were not analysed as part of this study. Details are provided in Appendix 3C. Task 6 – Noise levels was developed using a familiar context. Noise level and number of people in a classroom, though rarely measured, are at least intuitively experienced by students in schools. The context was also chosen such that students would expect a positive covariation between the variables, but the graph showed a negative covariation so that students were forced to rely on the data rather than prior beliefs. The data were six cases, as for Tasks 4 and 5, and with numerical values in the range 10-80 similar in magnitude to Task 5 (Temperatures). The data included repeated values of each variable to explore the issue of imperfect statistical association. Verbal Graph Interpretation was assessed using Q1 and Q4. Q1 was worded in an open manner to avoid the assumption that an association exists (Donnelly & Welford, 1989). Because students may have avoided

comment on covariation in Q1, Q4* was included and then revised to Q4 to provide a more specific cue about covariation. Numerical Graph Interpretation was assessed using Q2 and Q3. Q2 involved reading a value, and Q3 was designed to identify whether students based interpolation on proximity to one or more of Classes A, C, and E, or on the general trend.

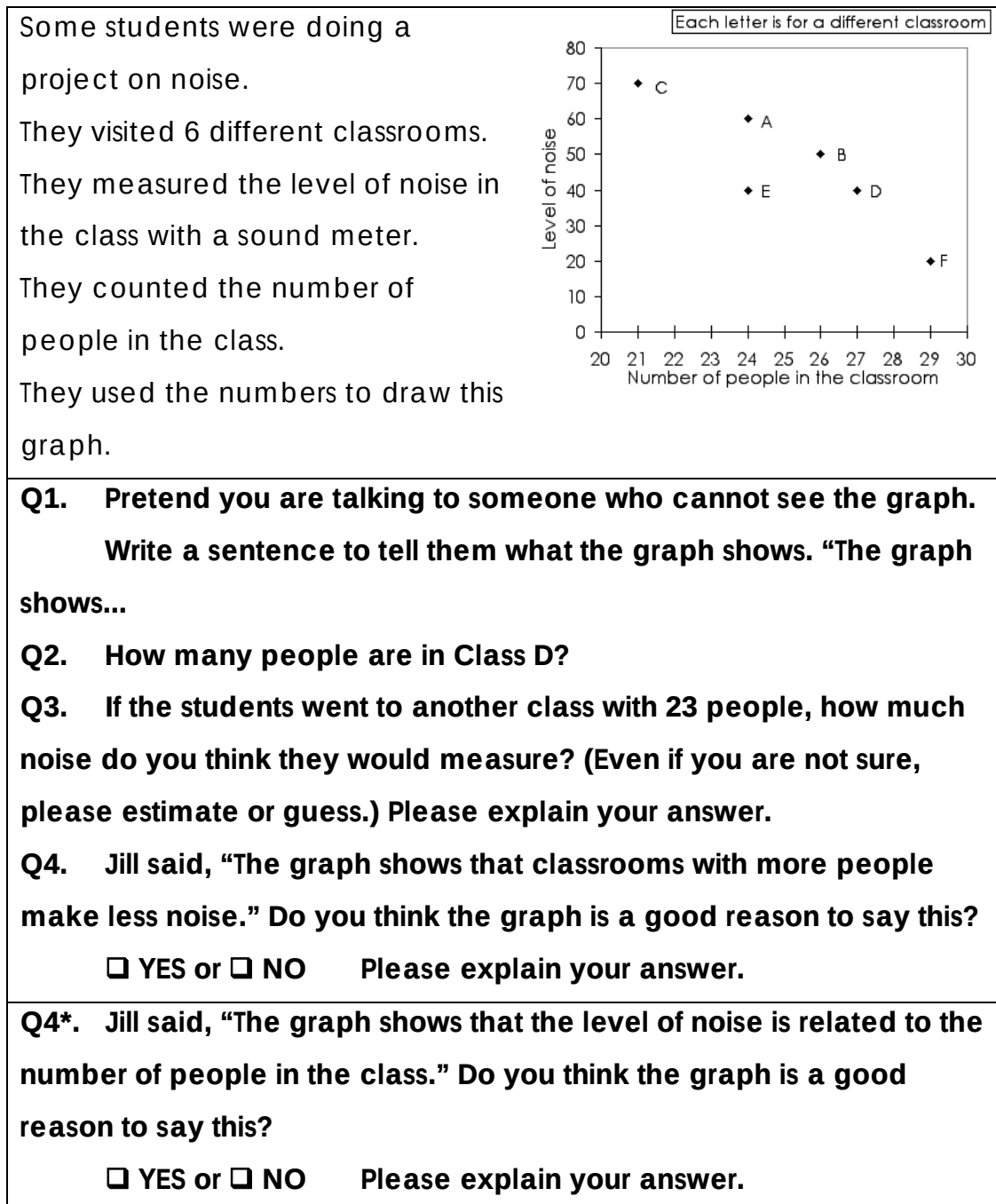


Figure 3.10. Task 6: Noise level versus number of people.

Note: Third- and fifth-grade males received Q4* in place of Q4.

3.05.05 Interview Tasks – Data Collection 4

Interview protocols were devised to follow survey administration. Interviews included brief explanations of students' own graphs, but most of the time was spent in responding to graphs drawn by others. The delay between survey administration and interviews was initially planned as being less than 24 hours to enable ready recall by students to explain how they arrived at their written response. Greater delay was needed, however, (a) to permit selection of students based on their survey responses, and (b) to permit appropriate parental approval to be obtained prior to removal of students from their class for individual interviews. The tasks used in interview were those used in the surveys for Data Collection 3, and included Tasks 1, 3, 4, 5, and 6, as seen in previous figures, with various prompt responses selected at various levels, as detailed in each interview-based investigation.

3.06 DESIGN AND PROCEDURES

Details of participants and procedures for four collections of student responses are provided in subsequent sections. The first three data collections involved administering individual student survey tasks to class groups for a cross-section of grades spanning from grade 3 to grade 11. The final data collection involved individual interviews with a subset of students surveyed.

3.06.01 Participants and Procedures Common Across Data Collections

Curricula in the schools involved in the study were individually adapted in general adherence to the Australia-wide curriculum documents cited earlier (AEC, 1991, 1994). In Tasmania, the mathematics curriculum to grade 8 was further

informed by guideline documents (Department of Education and the Arts, 1993), which specified in relation to graphing that students “need to develop skills to produce bar graphs, histograms, sector (pie) graphs, pictograms, line plots and stem and leaf plots” (p. 10). Four different levels of mathematics courses were available for each of grades 9 and 10 (Schools Board of Tasmania, 1993), and it was expected that all students take one of these courses. The three highest levels of these courses at grade 9 and at grade 10 included criteria about data handling related to developing skills to read and present information in graphical, tabular, and diagrammatic forms. The two highest levels also included criteria about generalizations involving the use of graphs to demonstrate simple linear relationships. Also at grade 10, the two highest courses included exploration of the basic notion of correlation. In grade 11, mathematics was not compulsory, but some courses included distinguishing statistical association from causal association.

In most cases, the duration of administration for surveys and interviews was aligned with school lesson durations, occasionally reduced by other school factors. Students’ survey representations were scanned into computer graphic files and their written responses were typed into a spreadsheet. Responses were categorized using clustering techniques (Miles & Huberman, 1994), successively refining categories and sub-categories by comparing and contrasting features of graphs or written responses. The results of coding responses informed the primary research aim, to develop assessment frameworks consistent with the structures evident in student responses. Frequency counts tabulated by grade level served the second aim to quantify proportions of students in various grades responding at each level to different tasks. For Aim 3, interviews were transcribed and repeatedly viewed by the researcher to identify themes and examples of student reasoning with concepts and

statistical language. Summaries of how students reacted to prompting with responses of others also informed this research aim. In cases where students' reactions to prompts indicated ambivalence or indifference, repeated viewing of the interview was undertaken to attempt to identify the predominant opinion. A final component to the research design was to investigate the conceptual features and empirical evidence across various tasks. The assessment frameworks developed were compared, frequency counts cross-tabulated, and correlations were calculated.

3.06.02 Data Collection 1

3.06.02.01 Participants

As part of a larger research project led by Dr Jane Watson at the University of Tasmania, responses were collected during 1993, 1995 and 1997 from students at 20 co-educational government primary schools, secondary schools, and matriculation colleges distributed throughout Tasmania. The numbers of students surveyed from different schools varied across years and grades due to availability of students. Approximately equal numbers of females and males were surveyed in each year at each grade level. Further details may be found in other reports (Watson & Moritz, 1998, 1999c).

For Task 1 (Heart deaths), survey responses of students who did not make any mark on the paper for this task ($n = 316$) were not included among the responses that were analysed as there was no evidence these students had read and attempted the question. Reasons for a student not reading the task were most likely that available class time was sometimes reduced, and that the task was one of the last on the written survey and the student may have devoted excessive time to responding to previous tasks. In support of the claim that no mark on the page was likely due to a

student not reading the task, it was observed that most students who made no mark were from secondary school grades; very few sixth-grade students made no mark, possibly because class schedules were more flexible and often longer, and the written survey involved fewer tasks. On the other hand students who wrote something, such as a simple mark, were deemed likely to have read the task, thus their responses were included in the analysis. The final response sample involved 1719 responses including included 369 sixth-grade students, 396 eighth-grade students, 604 ninth-grade students, 244 tenth-grade students, and 106 eleventh-grade students.

For Task 2 (Telephone call rates), more students (2251) were surveyed than for Task 1 (Heart deaths), as Task 2 appeared in the grade 6 version of the media survey in 1993, whereas Task 1 was administered to grade 9 only in 1993 and subsequently administered to students as young as grade 6. Task 2 appeared as a later task on the survey than Task 1, with fewer student attempts, however some students who did not mark the page for Task 1 responded to Task 2. Resulting numbers responding included 656 sixth-grade students, 434 eighth-grade students, 695 ninth-grade students, 248 tenth-grade students, and 218 eleventh-grade students.

3.06.02.02 Procedure

Tasks 1 and 2 were part of a written survey of student understanding of statistical concepts found in newspaper extracts. These tasks were (a) the 7th task of an 8-task survey for sixth-grade students, and (b) the 8th task of a 10-task survey for students of other grades. Other tasks in the survey included questions to interpret results of polls based on sampling, statements of risk, and average. One other task required students to interpret a graph: the 3rd task of the survey involved a pie graph that summed to more than 100% (Watson, 1997).

The survey was administered to whole-class groups for students to complete individually during class time, except for eleventh-grade students who were surveyed by voluntary participation during their free time in school hours. Durations of most classes was 40-45 minutes, although the researcher always took a few minutes to introduce the survey, and in some cases delays in scheduling reduced time to respond to approximately 30 minutes. Some sixth-grade classes were of an extended duration to permit most students to complete the survey. All students were informed that their participation was voluntary and that their responses had no impact on their school assessment. Students were, however, reminded that they should do their best to demonstrate their understanding because in so doing they would help researchers and teachers to ensure mathematics teaching would be most appropriate for students' learning.

During data collection for Task 1, some students expressed uncertainty about the task, including what their graphs were required to show, particularly which variables. In verbal instructions to individuals or the whole class, emphasis was given to the phrase "an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles." This was reinforced either by writing this phrase on the board at the front of the class, or by instructing students to underline this phrase in the news article on their surveys. As the task was the eighth of ten tasks on the survey, five to ten minutes before the surveys were collected, students were encouraged to respond to this task if they had not already done so. Preliminary analysis of responses (Moritz, 1999; Watson, 2000; Watson & Moritz, 1997) was used as the basis for sorting.

3.06.03 Data Collection 2

Responses to Task 3 were gathered from 23 fourth-grade, 25 fifth-grade, and 26 sixth-grade students at a private school for males, and from 18 fourth-grade, and 13 fifth-grade students at a private school for females. When requested by students, the researcher or classroom teacher offered assistance in reading the questions and encouraging students to do their best when they were uncertain about their responses. Of these 105 surveys, 93 students' complete responses were included in the analysis, and another 4 responses to Q1 and Q2 were deemed complete and also included; 8 surveys were not analysed because of non-response to Q2 and Q3, possibly due to time constraints and other class interventions.

3.06.04 Data Collection 3

3.06.04.01 Participants

Two single-sex private Tasmanian schools were approached, one for boys and the other for girls. Correspondence with these schools may be found in Appendix 1. The selection of these schools was based (a) on convenience, and (b) a willingness of school staff to participate. Both schools would be expected to draw students from a higher socio-economic status than the general school population in Tasmania. As a result, student achievement might be expected to be at a slightly higher level than from the broader school population.

In each school, class groups from the third, fifth, seventh, and ninth grades were surveyed, as shown in Table 3.01. Females described as fifth grade were from a composite class of fourth or fifth grade students, with 13 students at each grade level. Classes were selected by principals or other senior staff. Third, fifth, and seventh grades were selected for a willing teacher, with no reflection on student ability. In

contrast, ninth-grade students were from the lowest ability class in the girls' school, and from the highest ability in the boys' school, based on availability of these classes to undertake the survey with minimal interruption to their normal mathematics curriculum in mid-2001. It is noted that the small samples were whole class groups, that is students in a given group had a common teacher, and they were administered the surveys under slightly different conditions (See Section 3.06.04.02), hence it is not possible to draw conclusions confidently about gender ability, school performance, or even grade-level difference.

Table 3.01.

Numbers of Students Surveyed by Grade and Gender

Survey Task	Female Grade				Male Grade				Total
	3	4/5	7	9	3	5	7	9	
Q1. Task 3 (Heights)	26	26	25	15	19	21	25	27	184
Q2. Task 4 (Test scores)	26	26	22	5	19	18	24	27	167
Q3. Task 1 (Heart deaths)	-	-	25	0	-	21	25	27	98
Q4. Task 5 (Temperatures)	21	26	11	0	19	9	21	26	133
Q6. Task 6 (Noise levels)	13	23	20	12	13	7	12	21	121
Total (any task)	26	26	25	15	19	21	25	27	184

3.06.04.02 Procedure

After a brief verbal introduction to the purpose of the survey, students in class groups were asked to respond to the tasks in the survey. The researcher encouraged students to raise their hands during the class to discuss with the researcher any questions that were confusing or difficult to read. The researcher read the first question and other selected questions to students on a class or individual basis. Such assistance was required more commonly for students of lower grades.

It was intended that one class session of 40-45 minutes would be sufficient for most students to respond to most questions. Some questions included at the end of the survey were expected to be completed only by subsets of the classes due to time limitations, however their inclusion meant that students who otherwise would have completed the survey early used available time to provide additional information about their understandings in relation to statistical covariation. Students in higher grades were administered more questions than students in lower grades. Third and fifth grade classes were surveyed during class time, and seventh and ninth grade classes were surveyed during mathematics classes. For males, the duration was one class session of approximately 50 minutes (including 5 minutes of introduction). The duration for females depended on grade. For third and fifth grade classes, negotiation between the researcher and class teacher concerning timing generally resulted in a duration of approximately 60 minutes to allow most or all students to respond as they wished. Seventh-grade females were allowed an extended session lasting 70 minutes. Ninth grade females had about 25 minutes to respond due to a minor scheduling error; after about 15 minutes, the researchers directed these students to attempt Task 6 (question 6 on the survey) concerning interpreting a scattergraph of noise and class size. Thus, in various classes, non-response or an incomplete graph may be the result of time constraints rather than limited ability. The complete response set is presented in Appendix 4.

3.06.05 Data Collection 4 – Interviews

From the survey respondents in Data Collection 3, a subset of students was selected for interview. Initially it was intended that 6 to 8 students from each class would be interviewed. A preliminary examination of survey responses was made to identify students who collectively gave a variety of responses, including some who

gave unusual responses that might be explained by students in interview, and including those who gave low-level responses that might demonstrate greater appreciation when prompted with the responses of other students in interview. Letters were sent home to parents requesting consent to conduct individual interviews, advising them of the background to the study and that “Students have been selected for interview to include a variety of responses to these questions, and to include students who might be expected to learn from other students’ ideas.” Consent rates differed among classes; for example for female students, generally parental consent was obtained for students in grades 3 and 4/5; however for grade 9 students, consents were slow to return and an additional few students were selected. It is unknown whether this response difference reflected different attitudes of parents, the students, or of their teachers. Even once consent was obtained, absenteeism sometimes proved problematic, particularly for male students, and class scheduling and student time to get to the interview venue often resulted in interviews for less than the planned 45 minutes’ duration.

The resulting numbers of students interviewed for each of the tasks is shown in Table 3.02, along with the total interview time as recorded on video. The first two interview tasks (Task 3 - Heights, and Task 4 - Test scores) were asked to 34 and 33 students, that is all or almost all interviewed students, whereas other tasks were asked to 19 or fewer students. In addition, the first two interview tasks accounted for most of the total interview time of 16 hours 58 minutes; apart from the greater numbers of students, these tasks included greater numbers of prompts for students to consider, often to a number of questions within the task, whereas, for example, Task 6 – Noise levels involved no prompts but rather students explaining the reasoning for their responses. Transcripts of interviews are presented in Appendix 5A. Many

transcripts were not completed verbatim, and contain spelling and grammatical errors, however these transcripts were used to complement viewing of the video when coding interview dialogue.

Table 3.02.

Numbers of Students Interviewed by Gender and Grade

Interview Task	Female Grade				Male Grade				Total	Time (h:mm)
	3	4/5	7	9	3	5	7	9		
Int 1. Task 3 (Height)	5	6	6	5	3	3	3	3	34	8:08
Int 2. Task 4 (Test scores)	5	6	6	4	3	3	3	3	33	5:07
Int 3. Task 1 (Heart deaths)	0	0	3	1	0	3	1	2	10	1:15
Int 4. Task 5 (Temperature)	5	6	2	1	3	1	0	1	19	1:43
Int 6. Task 6 (Noise)	3	2	0	0	3	3	1	1	13	0:45
Students interviewed and total time	5	6	6	5	3	3	3	3	34	16:58

3.06.06 Presentation of Results of Investigations

Chapters 4, 5, and 6 present results of the analyses of student responses to the tasks described in this chapter. Students are identified by grade level for Data Collection 1, and by grade level and sex for Data Collections 2, 3 and 4. Coding of students' responses in assessments frameworks addresses Research Aim 1. Frequency counts tabulated by grade level address Research Aim 2 to quantify proportions of students in various grades responding at each level to different tasks. Coding of interview dialogue of students' reactions to prompts addresses Research Aim 3. In Chapter 7, the assessment frameworks developed are compared, frequency counts cross-tabulated, and correlations are calculated, to address Research Aim 4.

3.06.07 Summary of Research Design

A summary of the four data collections, with numbers and grade levels of students sampled, is shown in Table 3.03. This summary indicates the range of responses collected spanning grades 3 to 11, and includes large sampling for some tasks, and limited numbers of interviews across a variety of tasks. Table 3.03 also lists corresponding chapters and investigations, providing a sense of the investigations that appear in Chapters 4, 5, and 6, and the particular interest in Speculative Data Generation explored in Chapter 4.

Table 3.03.

Summary of Research Design

Task	Chapter	Investigation	Total Sample Size	Grades
Data Collection 1 (Surveys)				
1 (Heart deaths)	4	1	1719	6, 8, 9, 10, 11
2 (Telephone rates)	6	7	2251	6, 8, 9, 10, 11
Data Collection 2 (Surveys)				
3 (Heights)	4	2	97	4, 5, 6
Data Collection 3 (Surveys)				
3 (Heights)	4	3A	184	3, 5, 7, 9
4 (Test scores)	4	4A	167	3, 5, 7, 9
1 (revised) (Heart deaths)	4	5A	98	5, 7, 9
5 (Temperatures)	5	6A	133	3, 5, 7, 9
6 (Noise)	6	8A	121	3, 5, 7, 9
Data Collection 4 (Interviews)				
3 (Heights)	4	3B	34	3, 5, 7, 9
4 (Test scores)	4	4B	33	3, 5, 7, 9
1 (revised) (Heart deaths)	4	5B	10	5, 7, 9
5 (Temperatures)	5	6B	19	3, 5, 7, 9
6 (Noise)	6	8B	13	3, 5, 7, 9

CHAPTER 4. SPECULATIVE DATA GENERATION

4.01 SUMMARY

This chapter discusses eight investigations concerning students' Speculative Data Generation. Five investigations (1, 2, 3A, 4A, 5A) involved survey responses, and three (3B, 4B, 5B) involved interviews with students. Task 1 – Heart deaths, used in Investigation 1 and revised for Investigations 5A and 5B, asked students to generate speculative data related to a newspaper article stating a relationship between an increase in heart deaths and an increase in the use of motor vehicles. Task 3 - Heights, used in Investigations 2, 3A and 3B, involved graphing various relationships among height, age and gender. Task 4 – Test scores, used in Investigations 4A and 4B, concerned graphing how test scores covaried with study time.

Responses were described according to four levels of Speculative Data Generation: Nonstatistical (Level 0), Single Statistical Aspect (Level 1), Inadequate Covariation (Level 2), and Appropriate Covariation (Level 3). Responses to the different tasks were described using the same levels, however within each level, some categories were the same but others differed depending on the task. In addition, the proportions of students responding at each response level for each year level differed across tasks. Descriptions of the survey responses, and interview analyses, explored these differences in relation to accessibility and wording of the tasks.

4.02 INTRODUCTION

Speculative Data Generation involves proposing data that might lie behind a verbal statement of statistical covariation, as shown in Figure 1.01. In comparison to

Graph Production and Graph Interpretation, Speculative Data Generation had previously received less research, however it was the central focus of the current study, with eight investigations involving survey and interview responses to three tasks. This chapter also traces the evolution of the study described in Section 1.06, and across the three data collections described in Chapter 3. This evolution is evident in the discussion of the findings following Investigations 1 and 2, prior to the subsequent investigations presented in this chapter.

Characteristics of the levels are shown in Table 4.01, according to the structure and emphasis of correspondence and variation evident. Within this framework of four levels, various categories of response were observed, depending on the task requirements. The characteristics described in Table 4.01 refer to the nature of the data represented, but not the form of the representation, which was associated with the skill of Graph Production.

Table 4.01.

Characteristics of Four Levels of Speculative Data Generation

Level	Description
0. Nonstatistical	Responses represent either: (a) context in a narrative but without a data set of more than one value of one variable, or (b) graph axes or values, denoted by number or spatial position, but without a context indicating a data variable
1. Single Statistical Aspect	Responses represent either: (a) correspondence in a single bivariate case, or (b) variation of values for a single variable
2. Inadequate Covariation	Responses represent two variables but either: (a) correspondence is shown with inappropriate variation for at least one variable, such as one variable only has two distinct values (often categorical), or (b) variation is shown for two variables with inappropriate correspondence, such as not in the correct direction, or not between the appropriate variables
3. Appropriate Covariation	Responses represent both variables with appropriate correspondence between the variation of values for each variable

4.03 INVESTIGATION 1: HEART DEATHS VERSUS USE OF MOTOR VEHICLES

4.03.01 Introduction and Aims

Investigation 1 aimed to explore the ways in which students generate speculative data in a graph for a multivariate situation verbally described in a newspaper article, and to develop an assessment framework, aligned with Research Aim 1 of the wider study (Section 3.02). This exploratory investigation provided a strong impetus for the broader doctoral study, based on the range of responses observed and the challenges evident for many school students. The investigation also involved large sample of students to assess the proportion of students responding each way, aligned with Research Aim 2 of the wider study (Section 3.02).

4.03.02 Task

Task 1, shown in Figure 4.01, presented an extract of a newspaper article (“Family car,” 1991) that involved a verbal description of “an almost perfect relationship” among a number of variables, including “the increase in heart deaths and the increase in the use of motor vehicles” (Moritz, 1999; Watson, 2000; Watson & Moritz, 1997). Q1 asked students to sketch a graph based on the article, but did not indicate which variables to represent. The instruction to label the graph encouraged students to specify the variables represented. Q2 asked students to raise questions about the research reported in the news article. It was worded in a general way so that students might identify any features that they considered relevant. Only Q1 was analysed in this study, however examples of responses to Q2 are presented in Appendix 4B.

Family car is killing us, says Tasmanian researcher

Twenty years of research has convinced Mr Robinson that motoring is a health hazard. Mr Robinson has graphs which show quite dramatically an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles. Similar relationships are shown to exist between lung cancer, leukaemia, stroke and diabetes.

Q1. Draw and label a sketch of what one of Mr. Robinson's graphs might look like.

Q2. What questions would you ask about his research?

Figure 4.01. Task 1 to assess Speculative Data Generation based on a newspaper extract (*The Mercury*, June 11, 1991, p.2).

4.03.03 Participants and Method

The data collection followed the method specified for Data Collection 1 in Chapter 3. Students' representations were assigned to one of four levels based on the speculative data they generated to show covariation, that is, correspondence of variation. The relevant elements used in assigning responses to levels were (a) the correspondence between the variables, (b) the choice of variables, (c) variation shown for values of each variable.

Within the four levels, thirteen categories were described as shown in Figure 4.02, distinguished by the relevant elements for the task, namely the variables, the variation, and the correspondence. The first three categories are denoted by words, and the remaining ten categories are displayed with stylized icons to indicate common features of responses in the category. Form of graph and use of scale were not used to differentiate levels as these features were representational rather than related to the speculative data, and the statistical covariation could be represented

with different graph forms and with or without scales. Responses indicating scale units implying calendar years such as “1995, 1996, 1997,” or units implying the 20 years of Mr Robinson’s research such as “5, 10, 15, 20,” were categorized as if a time label had been included, which thus influenced assignment to a level. In cases where no values were indicated, ordinal values were assumed based on the graphing convention that values increase as one moves up and right on a graph.

Using all 1719 responses, the researcher made an initial clustering of students’ representations into categories similar to those reported in the results that follow, with 5% of responses remaining unclassified. A second researcher analysed this initial sort: 77% of responses required no further discussion, 18% were highlighted for possible changes of category, and the 5% unclassified were assigned levels. In a second round of coding, 5% of responses were highlighted; these were subsequently discussed and categorized by agreement. In two subsequent revisions by the researcher, categories were redefined, for example distinguishing Ordered Graphs (Category 1B) from Trend Graphs (Category 2B) (see Figure 4.02), and a further 4% of graphs were allocated to a different category, most of these being a result of minor redefinition of categories. This process was consistent with the cyclic coding procedure described by Miles and Huberman (1994, p. 61).

Results are presented according to a quantitative overview, followed by examples of each level and category of response. Response examples are annotated R1, R2..., for ease of reference between the text and figures presented.

Level of Speculative Data Generation	Category of Speculative Data Generation			
	Pictures are for illustrative purposes only, and are not indicative of all responses in the category.			
0. Nonstatistical	0A. No response	0B. Axes	0C. Labels	
	0D. Unlabelled Graph	0E. Labelled Axes		
		Value denoted by number or position, but with no variable label		Variable labels (one or more) on graph axes, but with no data
1. Single Statistical Aspect	1A. Single Comparison	1B. Ordered Graph	1C. Single Variable	
		Correspondence in a single bivariate data case		Ordered values denoted by number or position, no label
				Variation for one variable denoted by number or position
2. Inadequate Covariation	2A. Double Comparison	2B. Trend Graph	2C. Double Variable	
		Correspondence in two bivariate data cases with inadequate variation		Variation but correspondence is between inappropriate variables
				Variation for two variables but correspondence is not in the required direction
3. Appropriate Covariation	3A. Series Comparison	3B. Coordinate Variable		
		Two data series, each in a line or at least 3 cases, compared on the same axis dimension		A line or at least 3 bivariate data cases related on the opposing axes

Figure 4.02.

Levels and categories of Speculative Data Generation for Task 1.

4.03.04 Quantitative Results

Students from higher grades were more likely to respond at higher levels, as shown in Table 4.02. Level 1 was the modal level for sixth-grade students, whereas for ninth-grade students the modal response level was Level 3. Single Variable graphs and Series Comparison graphs were common student responses at many grade levels. A total of 343 of the 1719 response were from students attempting the task again after a two-year interval. Although 19% of students responded at a lower level of representing on the latter survey, most students retained (41%) or improved (40%) their level of response.

Table 4.02 also shows graph forms drawn by students from grades 6 to 11. Bar graphs were most common in grade 6, with increasing use of line graphs across grades to grade 11. Pie graphs and other representations were also drawn by students. Examples that follow illustrate that students choosing either the bar or line form of graph were able to do so at any of the levels of Speculative Data Generation, although bar graphs were more frequent at Level 1, and line graphs more frequent at Level 3.

Table 4.02.

Response Counts and Percentages of Responses by Grade at each Level and Category of Speculative Data Generation

Level of Speculative Data Generation	Response Count	Grade				
		6	8	9	10	11
	<i>N</i> =1719	<i>n</i> =369	<i>n</i> =396	<i>n</i> =604	<i>n</i> =244	<i>n</i> =106
Level 0. Nonstatistical	425	33	32	22	15	8
Level 1. Single Statistical Aspect	476	44	30	20	25	11
Level 2. Inadequate Covariation	299	13	17	19	18	20
Level 3. Appropriate Covariation	519	10	21	38	43	61
Category of Speculative Data Generation						
0A. No Response	196	11	15	11	9	4
0B. Axes	52	4	4	2	2	2
0C. Labels	14	3	1	0	0	0
0D. Unlabelled Graph	123	14	8	6	2	1
0E. Labelled Axes	40	2	3	2	2	1
1A. Single Comparison	144	17	9	4	6	2
1B. Ordered Graph	78	7	4	5	2	5
1C. Single Variable	254	20	15	11	17	5
2A. Double Comparison	49	4	2	2	3	3
2B. Trend Graph	194	5	12	14	12	14
2C. Double Variable	56	4	3	3	2	3
3A. Series Comparison	204	9	15	18	29	23
3B. Coordinate Variable	225	1	6	20	14	39
Form of Graph						
None	180	10	15	10	8	2
Bar	603	61	36	25	27	18
Line	673	8	31	52	53	71
Other	263	21	18	13	11	9

4.03.05 Level 0 – Nonstatistical

Student responses that were nonstatistical did not represent data as numbers with a context (Moore, 1990). Some students gave no response, some showed axes only or context only, and some displayed labels on graph axes without data. Others showed numbers in a graph form but without labels, and even if labels were added, an appropriate statistical covariation would not have been adequately represented. The five categories of Speculative Data Generation at Level 0 are shown in Figure 4.02.

Category 0A. No Response. Many students did not respond to the question, often indicating this with a simple dash mark on the page. One commented, “I can’t draw a graph because I don’t know any figures” (R1, Grade 9). These students responded to later survey items, so it would appear the students had an opportunity to engage the question, but failed to do so.

Category 0B. Axes. Some students showed a simple box or grid, and others an L-shape set of axes (R2, Figure 4.03a). Some drew a circle that may have been the beginning of a pie chart but without division to indicate numbers or proportions. It appeared these students did not know how to proceed to add meaningful data, one commenting, “Not enough information supplied” (R2).

Category 0C. Labels. Some students drew pictures or wrote words to indicate one of the variables of the context, but showed no evidence of representing data values. An example (R3) is shown in Figure 4.03b, in which the student attempted to show cars and health issues in pictorial story form.

Category 0D. Unlabelled Graph. Responses in this category were unlabelled and showed no evidence of an appropriate trend. Three examples (R4, R5, and R6) are shown in Figure 4.03c. Most responses showed a line graph (R4) or bar graph

(R5) form in which the heights of the line segments or bars possibly denoted values.

A few responses showed a general pie chart (R6). These responses offered no evidence that students intended to display an association of relevant variables.

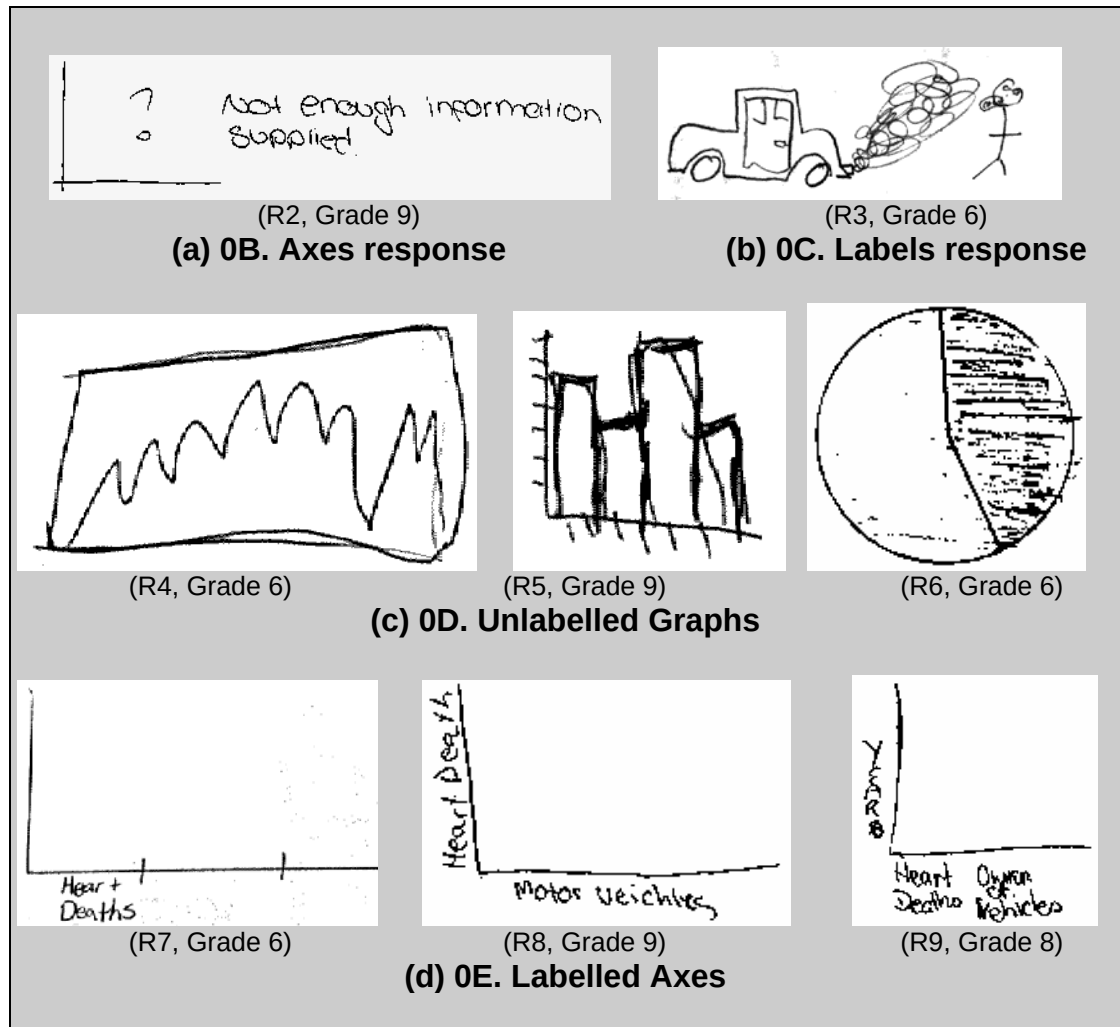


Figure 4.03. Student responses – Level 0 – Nonstatistical.

Category 0E. Labelled Axes. Some responses were labelled with a framework for representing values, but no values were represented. Three examples (R7, R8, and R9) are shown in Figure 4.03d. Some responses showed only one axis label (e.g., R7). Other responses showed car usage and at least one illness, some on opposing axes (e.g., R8), and others using the same axes to indicate possible comparison of values (e.g., R9).

4.03.06 Level 1 – Single Statistical Aspect

Student responses that involved a Single Statistical Aspect represented either correspondence of values without variation, or variation of values without correspondence. Three categories of response were observed, as stylized in Figure 4.02. Single Comparison responses (Category 1A) showed correspondence of values without variation in a single bivariate case comparing two values on the same axis. Variation of values without correspondence was shown in two ways: either as a set of ordered values unlabelled (Category 1B. Ordered Graphs), or as a single labelled variable (Category 1C. Single Variable). Labelling of variables in a graph was considered to be evidence of treating the variables in a statistical sense, as was ordering of values, a common data handling operation.

Category 1A. Single Comparison. Responses in this category represented a single value for heart deaths and for use of motor vehicles. Each of these measures was assigned to the same axis, or within a pie graph, giving the appearance that the values were being compared. Three examples (R10, R11, and R12) are shown in Figure 4.04a. Most responses represented the measures in a simple bar graph (e.g., R10 and R12), although some used a pie graph (e.g., R11). Some students displayed the measures with an equal value, perhaps in an attempt to represent the association, and some responses included a scale or number, perhaps to establish this comparison numerically (e.g., R11 and R12).

Category 1B. Ordered Graph. Some students drew a graph with a series of ordered values, but these values were not labelled to indicate which variables were shown. Two examples (R13 and R14) are shown in Figure 4.04b. Some students drew a bar graph showing each bar successively smaller than bars to the left of it

(e.g., R13). Other responses, almost all of which were line graphs, showed values in increasing order (e.g., R14).

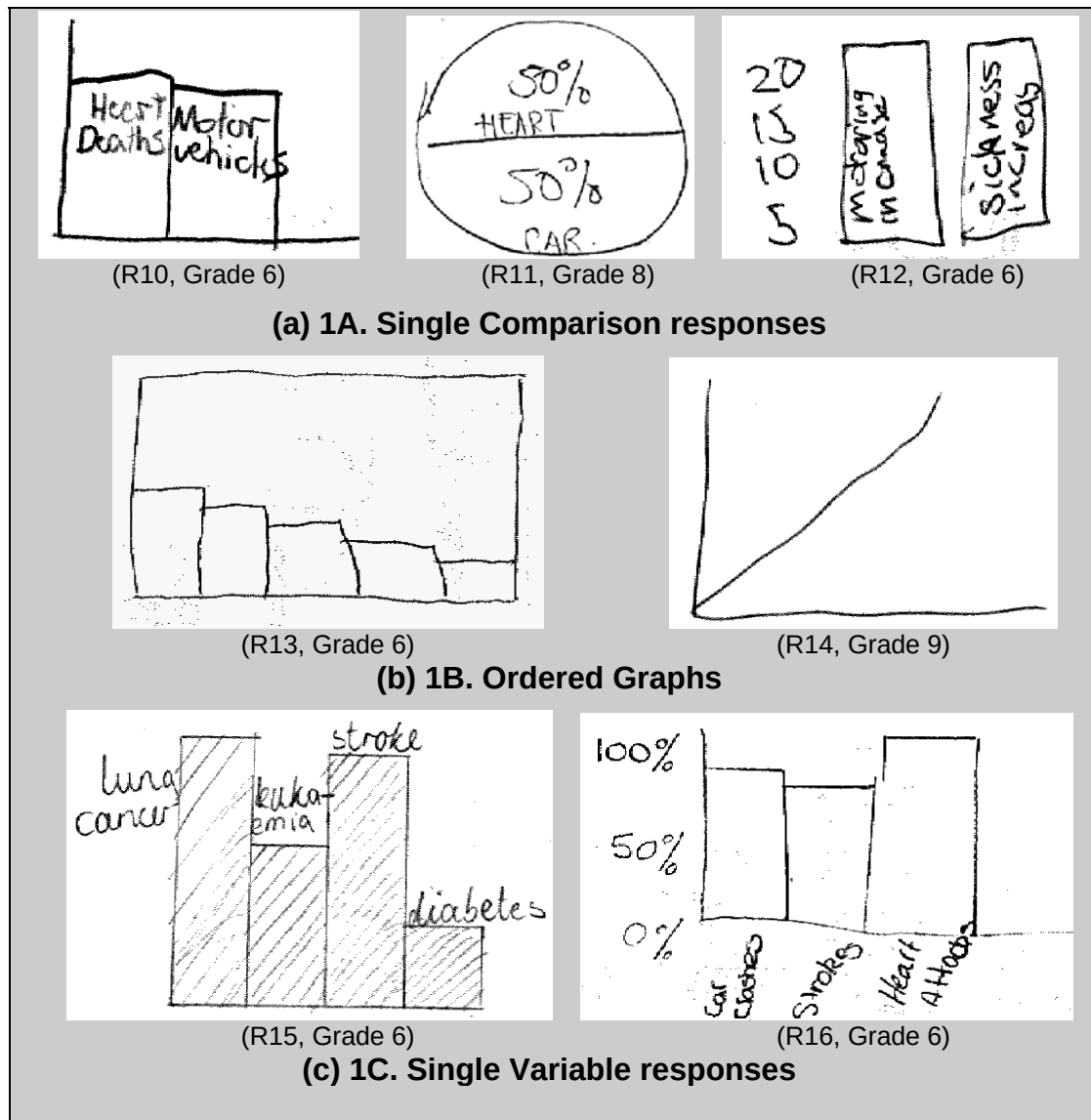


Figure 4.04. Student responses – Level 1 – Single Statistical Aspect.

Category 1C. Single Variable. Responses in this category included graphs showing variation of values for a single variable. At least one label or scale indicated motor vehicle usage, illness, or time. Two examples (R15 and R16) are shown in Figure 4.04c. Some responses indicated a number of different illnesses without indication of cars (e.g., R15). These responses displayed variation in values of illness across a number of cases labelled as different illnesses, as shown in the icon in

Figure 4.02 by the variable label (disease) and in some responses by a variable label such as “disease,” “cause,” or “condition.” Other responses represented a measure of motor vehicle use alongside the other diseases (e.g., R16), even adapting the measure to “car crashes” (e.g., R16) or “motoring deaths” to make the data conform to the single variable measure of deaths. Such responses were similar to Single Comparison responses (Category 1A) at this level in the respect of representing a single value for both car use and heart problems, however the additional values appeared to emphasize variation of values rather than the correspondence of two values.

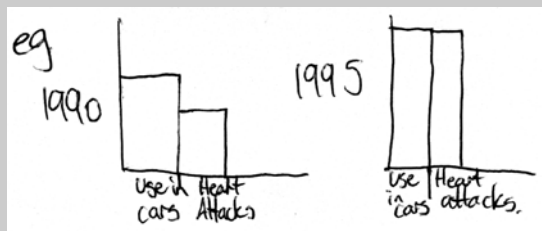
4.03.07 Level 2 – Inadequate Covariation

Student responses at this level were characterized by some success in representing the statistical covariation, but failure in representing all aspects of the association that were expected of the task set. The three categories of response at this level are indicated in Figure 4.02. Some students drew Double Comparison responses (Category 2A) that represented repeated values of the appropriate measures on the same axis for comparison, but these graphs did not represent data showing appropriate variation. Some students represented a Trend Graph (Category 2B) but between variables that were inappropriate or incomplete for the task, such as time and either heart deaths or motor vehicle use. Other students represented variation in two measures as Double Variable responses (Category 2C), but the correspondence was lacking.

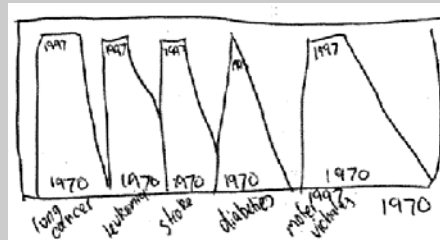
Category 2A. Double Comparison. Some responses displayed two values of both heart deaths and motor vehicle usage, demonstrating correspondence of the two variables. Two bivariate values, however, provided inadequate variation in each variable to represent appropriately “an almost perfect relationship.” The use of two repeated values indicated partial progress towards displaying a corresponding trend

of the measures. Two examples (R17 and R18) are shown in Figure 4.05a. The repeated measures frequently involved comparisons over time (e.g., R17 and R18), occasionally implicitly. For a few responses, the double comparison was over a variable with two levels (e.g., females versus males, or drivers versus non-drivers) placed as different conditions on the horizontal axis or placed as two separate graphs. A few students illustrated motor vehicle usage on the same axis as various illnesses, as for a Single Variable graph, but included values for two cases or times (e.g., R18).

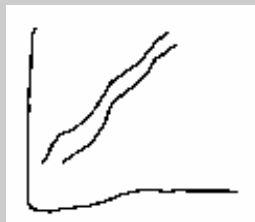
Category 2B. Trend Graph. Responses in this category showed evidence of variation in measured values (that is, *variables*), but the correspondence did not clearly indicate the measures expected of the task. Three examples (R19, R20, and R21) are shown in Figure 4.05b. Some responses were unlabelled dual-lines (e.g., R19), considered as having provided indication of representing two variables covarying over cases, in contrast to single lines in the Ordered Graphs (Category 1B). Some responses involved an association of an unlabelled variable with time (e.g., R20), some an association of one appropriate variable without the other variable being labelled, and some an association of one appropriate variable with time (e.g., R21). Responses were commonly line graphs (e.g., R19 and R21), although some were bar graphs (e.g., R20). In some responses, axes were inverted from convention (e.g., R21).



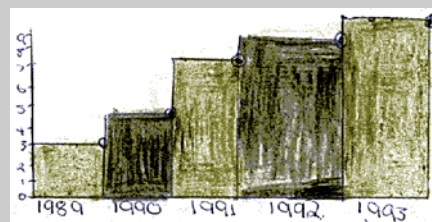
(R17, Grade 9)



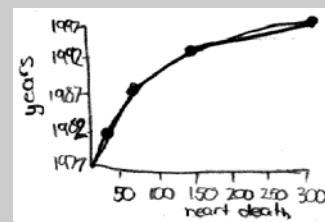
(R18, Grade 6)

(a) 2A Double Comparison responses

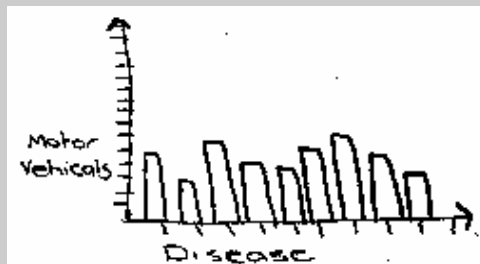
(R19, Grade 9)



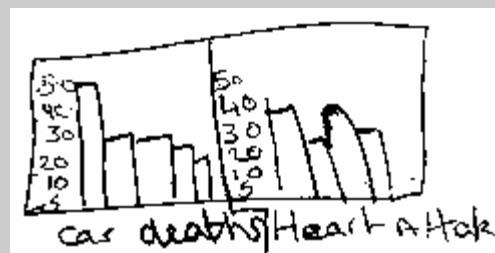
(R20, Grade 6)



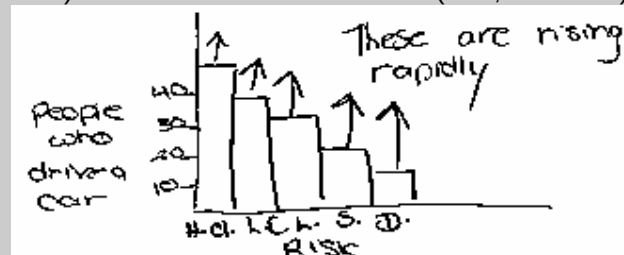
(R21, Grade 9)

(b) 2B Trend Graphs

(R22, Grade 9)



(R23, Grade 6)



(R24, Grade 9)

(c) 2C. Double Variable responses

Figure 4.05. Student responses – Level 2 – Inadequate Covariation.

Category 2C. Double Variable. Some students showed evidence of variation in values of variables, and the variables were appropriate, but the data did not show the positive covariation required: the covariation was negative, non-existent, or confusing to interpret. Three examples (R22, R23, and R24) are shown in Figure

4.05c. Some responses indicated the correct variables on opposing axes, but the variation in the data indicated a negative association or no covariation (e.g., R22). Other students represented two data series for heart deaths and motor vehicle use, variables involving at least three data values or a continuous line, but the correspondence between them was inappropriate (e.g., R23) or confused. Some other students illustrated various categorical illnesses along one axis, as for a Single Variable graph at Level 1, but included motor vehicle usage on the opposing axis in an attempt to represent an association (e.g., R24).

4.03.08 Level 3 – Appropriate Covariation

Two categories of response were identified that showed an appropriate correspondence in the variation for heart deaths and motor vehicle use, as shown in Figure 4.02. Series Comparison responses (Category 3A) involved a comparison of data series of two measures placed on the same axis that varied similarly over a number of cases, often time labels, and Coordinate Variable responses (Category 3B) involved representing a trend with relevant variables on opposing axes.

Category 3A. Series Comparison. Responses in this category displayed data series for both heart deaths and motor vehicle usage that involved similar variation as continuous lines or over at least three discrete cases. Three examples (R25, R26, and R27) are shown in Figure 4.06a. Most responses were line graphs, although some were bar graphs with three or more values to indicate “an almost perfect relationship.” The data represented commonly involved two straight lines with similar positive slope or two lines with similar fluctuations over a third variable. More than half of the responses explicitly displayed time as a variable (e.g., R25 and R26), although many responses omitted the time label as if it were implicit (e.g., R27). Some graphs inverted the axes to display time on the vertical axis (e.g., R26),

and some indicated scale units (e.g., R25). Most responses were double-graphs, which represented two data series on the one set of axes (e.g., R25 and R26), although other responses involved two graphs (e.g., R27).

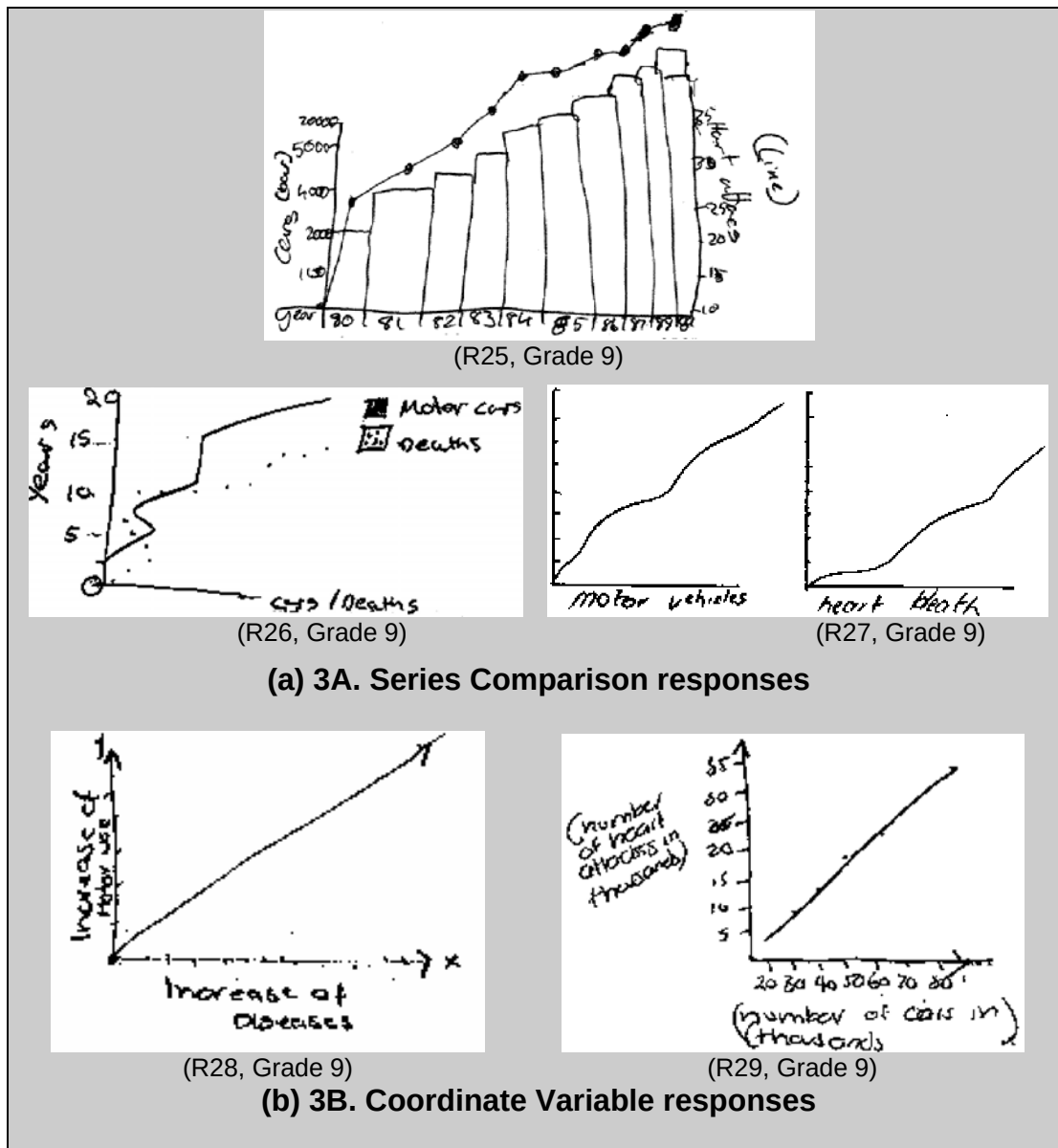


Figure 4.06. Student responses – Level 3 – Appropriate Covariation.

Category 3B. Coordinate Variable. Some students displayed Cartesian axes with heart deaths on one axis and motor vehicle usage on the other, and represented positive covariation in the data. Two examples (R28 and R29) are shown in Figure 4.06b. Approximately half of the responses in this category labelled motor vehicle

usage on the vertical axis (e.g., R28) and the other half of the responses labelled heart deaths on the vertical axis (e.g., R29). With the exception of a few bar graphs, responses were line graphs, often involving straight lines, though some displayed fluctuations or data points (e.g., R29) on the line. Some responses further illustrated influences of algebra teaching such as “x” and “y” labels (e.g., R28). Many responses indicated no scale (e.g., R28), although some responses showed basic or even detailed measurement scales (e.g., R29).

4.03.09 Discussion of Investigation 1

Four levels of Speculative Data Generation were described and illustrated with student responses, as shown in Figure 4.02. Levels indicated degrees of success in representing a data set that showed correspondence of variation for the appropriate variables of heart deaths and motor vehicle usage. The four levels related closely to the development of variable handling suggested by science educators (Donnelly & Welford, 1989; Ross & Cousins, 1993; Swatton, 1994; Swatton & Taylor, 1994), from no variables, to a single variable or data case, to two variables or data cases, and finally to multiple variables or two variables appropriately controlled. These levels also related to the general SOLO model of cognitive development (Biggs & Collis, 1982) concerning the structure of the observed learning outcome, involving prestructural, unistructural, multistructural, and relational levels.

Within the four levels of representing covariation, various categories were also identified, shown in Figure 4.02. At levels 1 and 2, categories differed in the presence and structuring of the relevant elements of the task, which included identifying appropriate variables, showing variation, and showing correspondence of values.

Single Comparison (Category 1A) and Double Comparison (Category 2A) responses illustrated correspondence of values but not appropriate variation for each variable. Responses that represented a single value for motor vehicle usage and for hearts deaths (Category 1A) were common for sixth-grade students (see Table 4.02) and were observed by Mevarech and Kramarsky (1997). Students drawing these graphs reduced the complexity of the task by focusing on local, individual values rather than on global trends (Bell et al., 1987a; Leinhardt et al., 1990). They indicated no variation in either measure, which made it impossible to indicate how the two measures vary together. One step toward representing covariation was observed in Double Comparison graphs (Category 2A), which represented two values for each of motor vehicle usage and heart deaths. This was similar to a partially correct strategy identified by Estepa and Batanero (1996) titled “correct interpretation of isolated points”: the addition of values made the points less isolated and permitted comparisons for two different values. Series Comparison graphs (Category 3A) that involved a comparison of two data series to show covariation appeared to be constructed either by adding data cases to Single Comparison and Double Comparison graphs, or by adding a data-series to time-based Trend Graphs (Category 2B). If from the latter, Series Comparison graphs could otherwise be known as Double Trend graphs. Single Comparison, Double Comparison, and Series Comparison graphs were similar to types of interpretation tasks introduced by Meyer et al. (1997). Comparison graphs also have the potential to retain covariate or case information on a separate axis (e.g., years, or in other contexts, names) and illustrate how bivariate data are the measured values for two attributes of that case. Case information retains stronger links with the source of the data for examination of outliers or sampling issues (Cobb et al., 2003).

Embedding the verbal description of bivariate association within a news article discussing multiple variables required students to identify appropriate variables as encountered in out-of-school contexts (Gal, 2002; Ross & Cousins, 1993; Tufte, 1997). Ordered Graphs (Category 1B) and Trend Graphs (Category 2B) illustrated corresponding variation of values, but not labeling of appropriate variables for which the covariation was intended. These were similar to the generalized graphs observed by Mevarech and Kramarsky (1997). At least one appropriate variable was not present in Trend Graphs. Even many Series Comparison responses (Category 3A) did not label time, although it seemed to be the implied covariate.

Single Variable and Double Variable responses illustrated variation of variables but not appropriate correspondence between the variables. The notion of ordering values of one variable, and then observing the corresponding order of values for another variable, has been described as the variational approach to teaching covariation (Nemirovsky, 1996a). Ordering values, described by Wavering (1989) as seriation, is an important data handling skill necessary to gain an appreciation of the variation present (Ross & Cousins, 1993).

Students from higher grades were more likely to respond at higher levels, and many surveyed two years later improved their response levels, however less than half of the students in any grade represented covariation appropriately. The following investigation aimed to examine Speculative Data Generation in a bivariate situation involving familiar variables, to explore the ways in which younger students might engage the task and demonstrate Speculative Data Generation under these conditions.

4.04 INVESTIGATION 2: HEIGHT VERSUS AGE

4.04.01 Introduction and Aims

In Investigation 1, Speculative Data Generation was considered in students' responses to Task 1 concerning a news article involving heart deaths and use of motor vehicles. The results indicated many students offered incomplete responses, such as Single Comparison responses (Level 1) that emphasized the correspondence of two values but did not illustrate variation in the measures. The task was considered to be made more complex by the multivariate description in the news article, and by the unfamiliar variables involved.

Investigation 2 explored students' Speculative Data Generation in a bivariate situation involving familiar variables, namely heights versus ages, as used by Ainley (1995), to expand the basis for the assessment framework for Speculative Data Generation, aligned with Research Aim 1 of the wider study (Section 3.02). It was expected that by employing familiar variables, and by simplifying the stimulus to a bivariate context, students would generally respond at higher levels, providing further evidence of the abilities of students and the challenges they face in representing statistical covariation, and aligned with Research Aim 2 of the wider study (Section 3.02).

4.04.02 Task

A survey item, shown in Figure 4.07, included two questions (Q1 and Q2) involving bivariate covariation and a third question (Q3) involving multivariate association. These questions were designed specifically to assess how students engaged the bivariate situation, and then to focus on any additional challenges that an additional variable posed. The constant function in Q2 was included following

research identifying this challenge for students (Nemirovsky & Tierney, 2001; Sherin, 2000).

One day at school, the class began to talk about how tall people are. The teacher measured everyone's height. Then they began to talk about their brothers and sisters. Some were taller, and some were shorter.

Q1. David said, "People grow taller as they get older."
Draw a graph to show what David is saying. Label the graph.

Q2. Mary agreed, "People do grow taller. But when you are 20 years old, you stop growing."
Draw a graph to show what Mary is saying. Label the graph.

Q3. The teacher said, "For 10 year olds, girls and boys are about the same height.
But men usually grow to be taller than women."
Draw a graph to show what the teacher is saying. Label the graph.

Figure 4.07. Task 3 to assess Speculative Data Generation.

4.04.03 Participants and Method

The task was administered to whole classes as a written survey, as described for Data Collection 2 in Chapter 3. Of 105 surveys administered, 93 students' completed responses to all questions were included in the analysis, and another 4 responses were deemed complete for the first two questions only; 8 surveys were not analysed because of non-response to all questions, possibly due to time constraints and other class interventions.

Each student's responses to Q1 and Q2 were treated as a composite response (Q1&2), which was assigned to a level and category based on criteria shown in Figure 4.08. Each response to Q3 was assigned a level and category based on criteria shown in Figure 4.09. In this investigation, Nonstatistical responses (Level 0) were not observed, in part because few items were in the survey, and non-responses were omitted from analysis. The three levels were comparable for responses to the

bivariate situation (Q1&2) and to the multivariate situation (Q3), and were also comparable to three of the levels observed in Investigation 1. Single Statistical Aspect (Level 1) responses failed to show covariation in the data, often omitting one of the relevant variables. Inadequate Covariation (Level 2) responses indicated an association between variables but not all critical features were displayed. Appropriate Covariation (Level 3) responses represented all critical features. For Q2, critical features included a curvilinear trend that showed increasing height with age and also constant height with changing age by or from age 20. For Q3, three critical features included (a) at age 10, females same height as males, (b) association of height with age (for females, males, or both), and (c) for some age greater than 10, males taller than females. The coding schemes were devised to assess at higher levels those responses indicating the student understood the covariation and had some coherent system for representing this, and not to penalise the student for violating a graphing convention, such as interchanging axes or reversing the order of an axis scale. Despite the task requirement for labelling, it was acknowledged that some students might have assumed height and age were obvious variables since they were written on the page just above where their graphs were to be drawn. Hence “indicative labelling” was credited: students were given the benefit of doubt if they provided some features of the graph or the data that indicated a distinction between height, age, and sex. Examples included a pictograph of people implying the vertical axis indicated height, or the number 20 indicating a significant age in response to Q2.

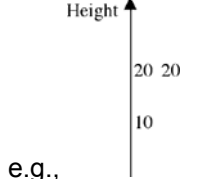
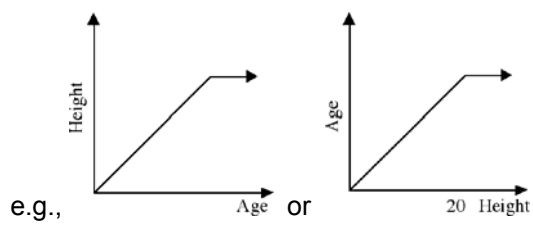
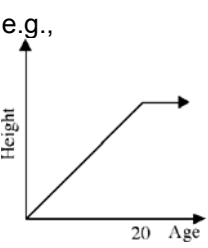
<u>Level 1: Single Statistical Aspect</u> One variable is omitted or two variables super-imposed. <u>Category 1: Single Axis</u> <i>Criterion (Q1&2):</i> failure to satisfy Level 2 criterion (Q1). A single axis measure is shown with one or two measures indicated with labels/scales, OR Cartesian axes are used but any data indicates only one unique measure.	
<u>Level 2: Inadequate Covariation</u> Some but not all of the features of the bivariate association are represented; errors/omissions may include labelling, scales, or failure to represent significant features of the association. <u>Category 2A: Increasing Trend</u> <i>Criterion (Q1):</i> a monotonic (non-decreasing) line OR ≥ 3 bivariate data points in a monotonic sequence OR ≥ 2 bivariate data points for which values are specified. If values are specified, an increase in one measure corresponds to an increase in the other. <i>Criterion (Q2):</i> failure to satisfy criterion for Curvilinear Trend Criterion (Q2). A graph with no constant/ decreasing segment.	<u>Category 2B: Partial Curvilinear Trend</u> <i>Criterion (Q1):</i> fulfilment of Increasing Trend criterion (Q1). <i>Criterion (Q2):</i> failure to satisfy criterion for Curvilinear Trend Criterion (Q2). A graph with a constant/decreasing segment. If values are specified, for that segment, an increase in one measure corresponds to no change or a decrease in the other measure. 
<u>Level 3: Appropriate Covariation</u> All relevant features of the bivariate association are represented. For this task, variable labels ("height" and "age") may be implicit in indicative features of shape and values (20) on the graph satisfying criteria. <u>Category 3: Curvilinear Trend</u> <i>Criterion (Q1):</i> fulfilment of Increasing Trend criterion (Q1), OR satisfy Curvilinear Trend Criterion (Q2). <i>Criterion (Q2):</i> (i) when "age" ≤ 20, criteria for Increasing Trend are satisfied, AND (ii) when "age" $\geq X$ (where $X \leq 20$), there is a straight line (or ≥ 2 bivariate data points) for which an increase (of scaled value or distance from origin) in "age" corresponds to no significant change in "height" (i.e., no change in comparison to other increase when age ≤ 20).	e.g., 

Figure 4.08. Coding scheme for levels and categories of response to Q1&2.

Note: Criteria for both Q1 and Q2 must be satisfied for assignment of the composite response (Q1&2) to the category.

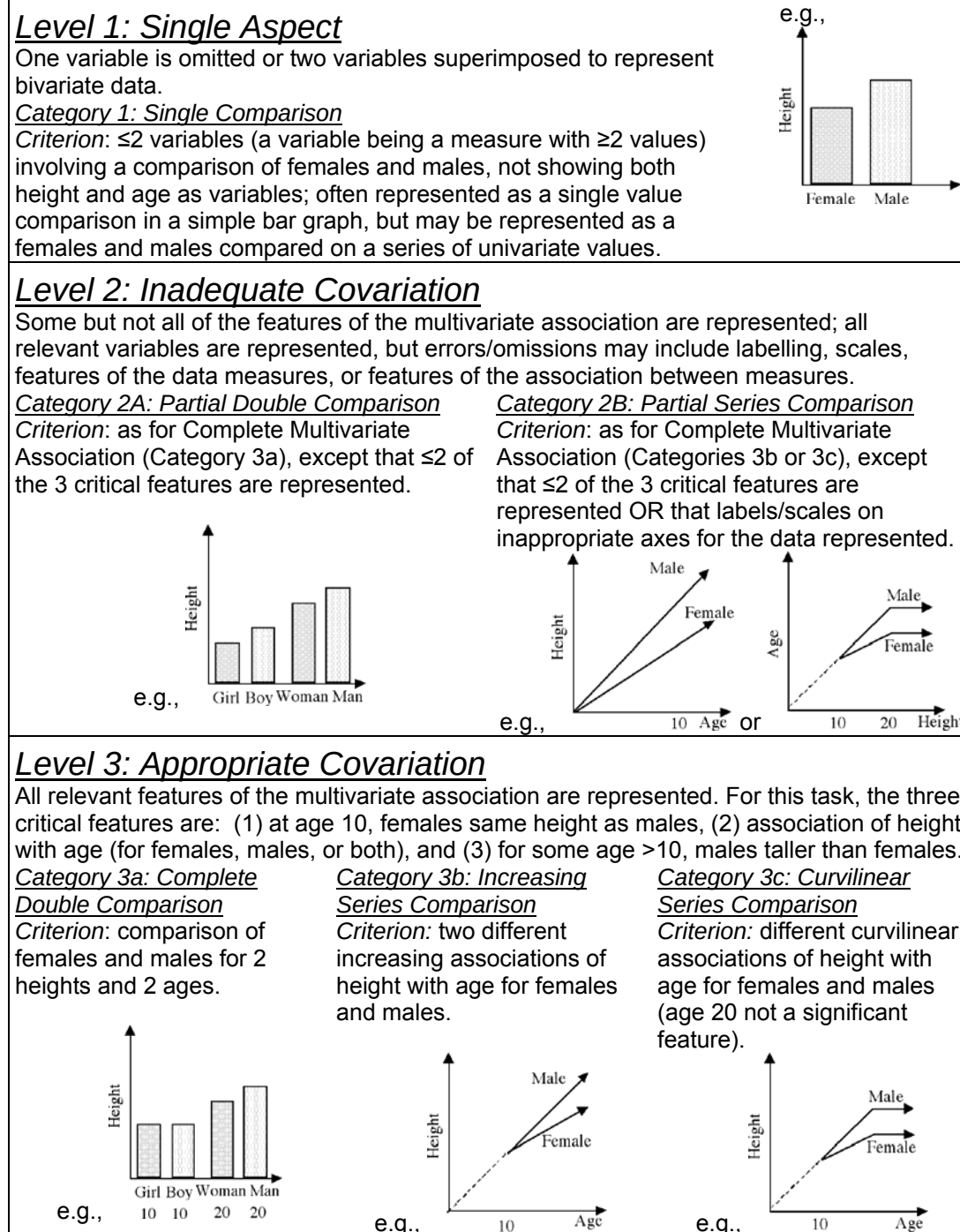


Figure 4.09. Coding scheme for levels and categories of response to Q3.

The researcher and an independent coder assigned responses to categories based on the coding schemes presented in Figures 4.08 and 4.09. Inter-coder reliability for coding of levels was 80% of responses to Q1&2 and 70% to Q3, and reliability for coding of categories was 69% for Q1&2 and 58% for Q3. Many of the discrepancies were due to the coder misunderstanding the criteria, and using the illustrative graphs intended to assist but not define coding; for example believing that form of graph (bar or line) was one of the criteria. Following joint discussion, agreement was reached on 98% of responses to Q1&2 and 96% for Q3. In remaining cases of disagreement—2 for Q1&2 and 4 for Q3—the researcher’s categorisation was preferred (see later discussion of R31, Figure 4.16). For these cases, students had represented features of the associations in unconventional ways that appeared to the researcher to represent relevant features. Compared to the researcher’s initial coding, the final coded data had no differences for Q1&2, and 2 responses coded differently for Q3.

4.04.04 Quantitative Results

Tables 4.03 and 4.04 show the percentages of response levels for Q1&2 and Q3 for each sex and grade. In response to Q1&2, more than 90% of students represented either Inadequate Covariation (Level 2) or Appropriate Covariation (Level 3). In response to Q3, about a third of grade 4 students responded with a Single Statistical Aspect (Level 1), but at grade 6 most students represented Appropriate Covariation (Level 3). For Q3, Series Comparison responses (Categories 2B, 3B, 3C) were more common than Double Comparison responses (Categories 2A, 3A). Comparison graphs, either Single (Category 1) or Double (Categories 2A, 3A) reduced the variation of ages represented for each sex. Comparison graphs were rarely observed in Q1&2 (see Table 4.02).

Table 4.03.

Response Counts and Percentages of Responses by Sex and Grade at each Level and Category of Speculative Data Generation for Q1&2

	Response	Female Grade		Male Grade		
	Count	4	5	4	5	6
Response Level	N=97	n=17	n=13	n=19	n=25	n=23
1. Single Statistical Aspect	4	12	0	5	4	0
2. Inadequate Covariation	39	53	31	32	48	35
3. Appropriate Covariation	54	35	69	63	48	65
Response Category						
1. Single Axis	4	12	0	5	4	0
2A. Increasing Trend	16	24	15	11	16	17
2B. Partial Curvilinear Trend	23	29	15	21	32	17
3. Curvilinear Trend	54	35	69	63	48	65

Table 4.04.

Response Counts and Percentages of Responses by Sex and Grade at each Level and Category of Speculative Data Generation for Q3

	Response	Female Grade		Male Grade		
	Count	4	5	4	5	6
Response Level	N=93	n=15	n=13	n=19	n=25	n=21
1. Single Statistical Aspect	19	33	8	37	16	10
2. Inadequate Covariation	34	33	23	37	48	33
3. Appropriate Covariation	40	33	69	26	36	57
Response Category						
1. Single Comparison*	19	33	8	37	16	10
2A. Partial Double Comparison	12	13	23	16	16	0
2B. Partial Series Comparison	22	20	0	21	32	33
3A. Complete Double Comparison	12	7	38	0	20	5
3B. Increasing Series Comparison	18	27	31	16	8	24
3C. Curvilinear Series Comparison	10	0	0	11	8	29

* One response from a Grade 4 female was not clearly Single Comparison, but otherwise hard to categorise.

4.04.05 Results – Q1&2 – Height Versus Age

4.04.05.01 Level 1. Single Statistical Aspect

Responses at this level indicated that students had difficulty coordinating the representation of two variables to show how height covaries with age. Some students superimposed the two variables on a Single Axis. If Cartesian axes were used, the data indicated either that there was only one unique measure, or that there was no association between the two measures. Four examples (R1-R4) are shown in Figure 4.10. A fifth-grade male (R1) showed vertical axes similar to the appearance of some growth charts, labelling various points likely to indicate ages of the same individual over time, with position denoting height. A grade 4 female (R2) listed years to denote possible age, and drew a line to denote height; however age and height were superimposed on a single axis such that at most one value was evident. A fourth-grade male (R3) represented a generic graph with no indication of the variables or of an association for Q1; for Q2 the labelling around 20 and change of direction showed awareness of the context, however the horizontal dimension could not be interpreted, thus the vertical dimension represented height and/or age. The bar graph by a fifth-grade male (R4) displayed six male names presented in reversed order on each axis, as if to emphasise the corresponding values by the case names, however the same values on each axis provided no indication of the two variables or of an association.

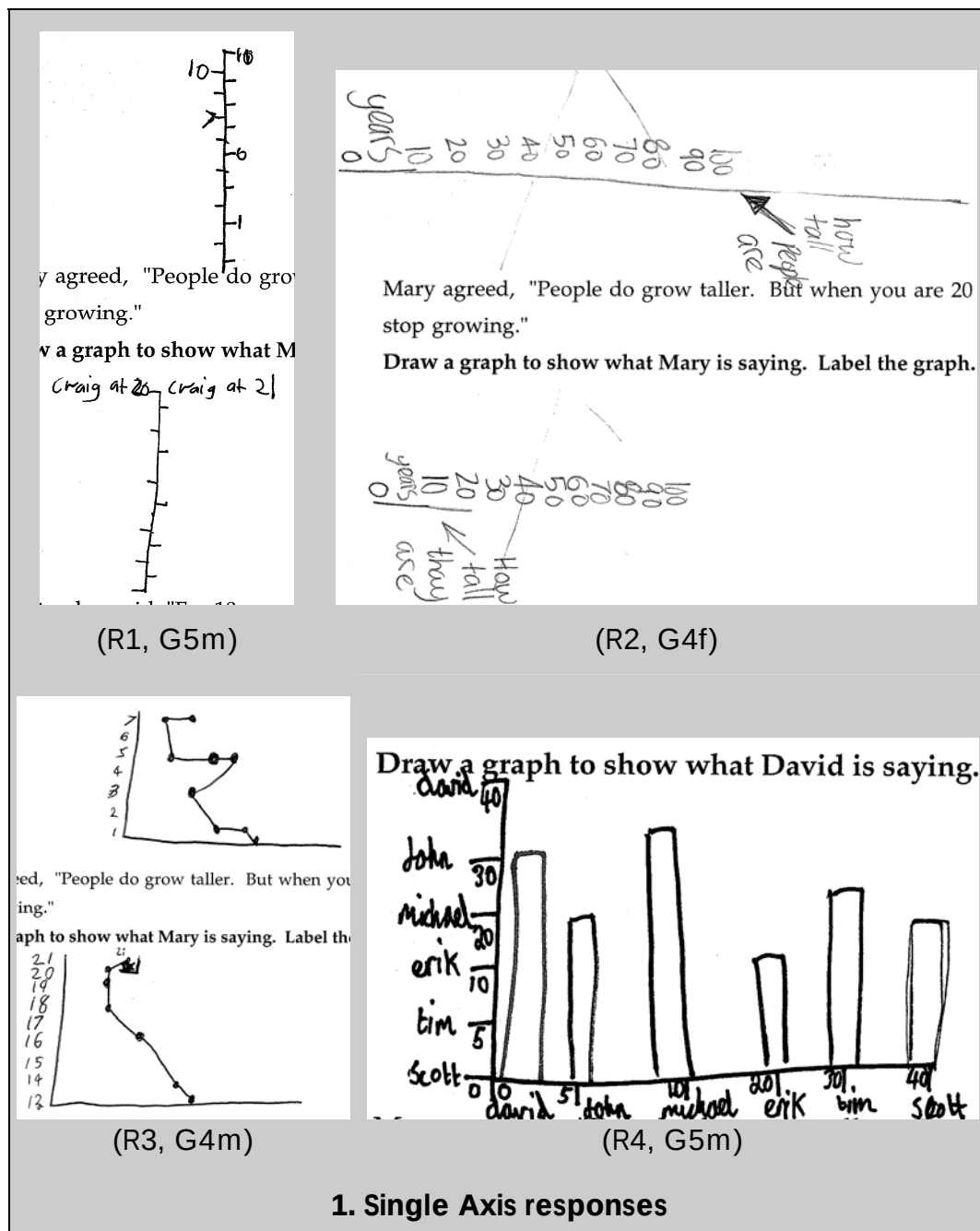
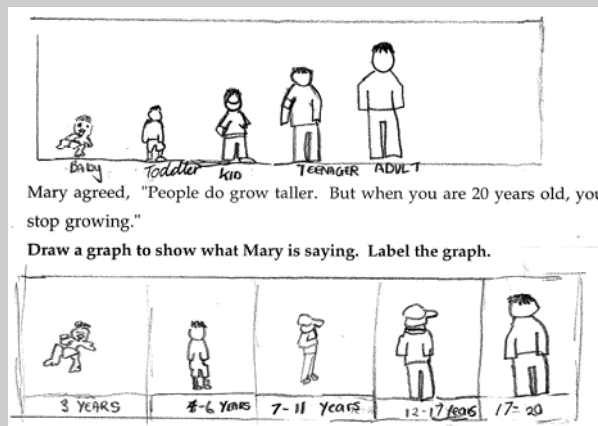


Figure 4.10. Student responses to Q1&2 – Level 1 – Single Statistical Aspect.

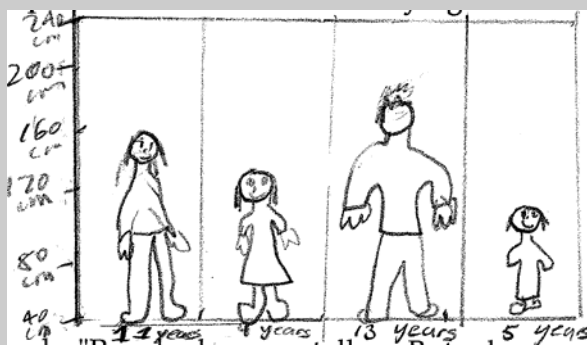
4.04.05.02 Level 2. Inadequate Covariation

Responses to Q1&2 at Level 2 displayed some but not all of the features relevant to the task. Students represented the increasing trend of Q1, but often had difficulties representing the curvilinear trend for Q2. These difficulties included

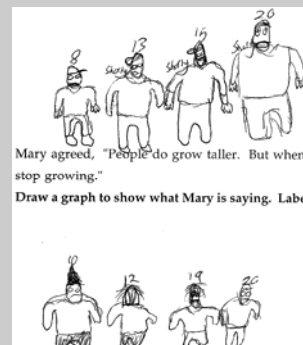
errors or omissions of labels or scales that resulted in failure to represent significant features of the covariation, such as failure to show constant height for changing age, or failure to show that this occurred by age 20. Five examples are shown in Figure 4.11 (R5-R9). A notable feature of four of these five examples is the use of the pictograph form: in three of these responses (R5, R7, R9), only the pictograph form indicated that height was being represented. R5, R6 and R7 show an Increasing Trend (Category 2A) with labelled ages. R6 did not order the cases, whereas R5 and R7 showed cases ordered by age, which, due to the covariation, also sorted them by apparent height. R7 included a pictograph that showed the Increasing Trend for Q1, but displayed a constant function for Q2 (Category 2B), thus failing to show a Curvilinear Trend required at Level 3. R8, from a fifth-grade male, apparently reversed the scale as well as putting labels on inappropriate axes in response to Q1&2 and thus he confused the representation as a Partial Curvilinear Trend (Category 2B). R9 displayed an Increasing Trend for Q1 but for Q2, a Partial Curvilinear Trend (Category 2B) because the fifth case, showing decline in height rather than constant height, was not clearly labelled with an age value over 20 years, as required criterion for category 3 in Figure 4.08.



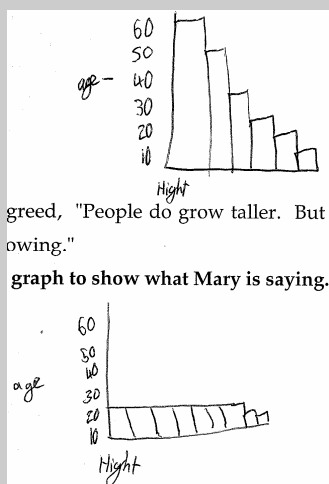
(R5, G5f)



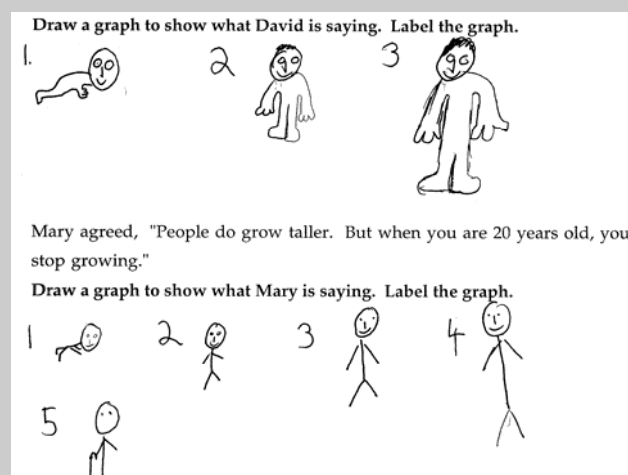
(R6, G5f)



(R7, G4m)

(a) 2A. Increasing Trend

(R8, G5m)



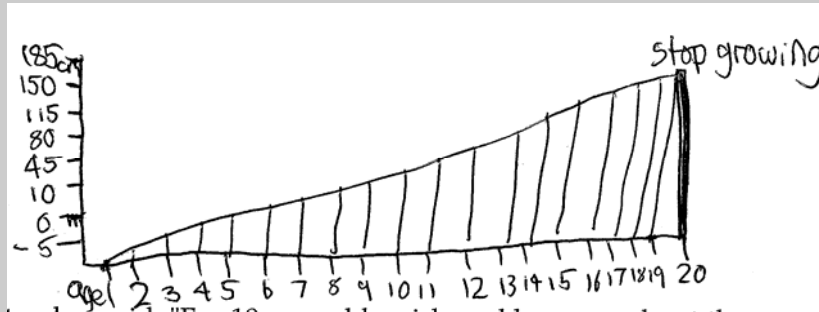
(R9, G4m)

(b) 2B. Partial Curvilinear Trend

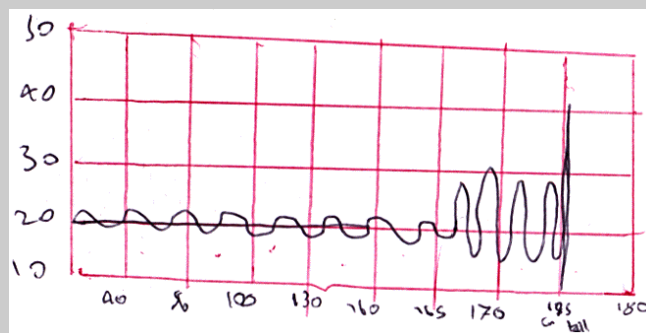
Figure 4.11. Student responses to Q1&2 – Level 2 – Inadequate Covariation.

Four additional examples (R10-R13) of the difficulties posed in representing the curvilinear trend for Q2 are shown in Figure 4.12. One idea used by some was simply not to represent ages greater than 20, as shown in R10. Others attempted to show the constant value by showing heights reaching a maximum value, but without clear linking of this maximum height to various ages greater than 20 (R11). Although many students drew pictographs which naturally aligned height on the vertical axis and thus various ages spread horizontally as different cases, some students who drew more conventional abstract graph forms placed age on the vertical axis and height on the horizontal axis and confused the relation (R11 and R12). R11, from a sixth-grade male, placed age on the vertical axis, which may have contributed to his difficulty in representing the curvilinear function. In R12, a sixth-grade male placed age on the vertical axis and height on the horizontal axis, which posed no problems for representation until age 20, beyond which he represented the constant function and showed no change in vertical axis value. The data series visually showed an increase and plateau, implying the axes were allocated reverse to that shown: in this case, it seemed that the axis allocation was an abstract and arbitrary decision, whereas the data series had some ikonic imaging similar to that shown in the pictographs in Figure 4.11. R13 was an unconventional representation, which included a vertical axis showing bivariate values (200/20, 185/15, 144/10) apparently indicating three heights in centimetres and corresponding ages in years: this axis constituted a table showing an increasing trend. The horizontal axis was used to show the three ages, and three data points were shown beside the vertical axis labels. The student appeared to have not yet developed the coordinate system, showing variation along vertical axis but establishing correspondence by double labelling rather than by

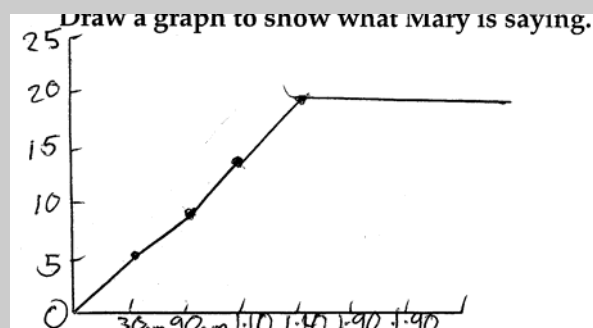
alignment with values along the horizontal axis. Issues in developing the coordinate system are considered further in Chapter 5 concerning Coordinate Graph Production.



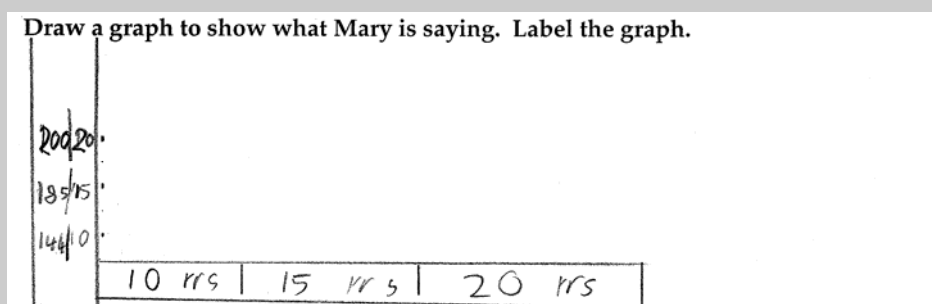
(R10, G5m)



(R11, G6m)



(R12, G6m)



(R13, G4f)

Figure 4.12. Student responses to Q1&2 – Level 2 – Inadequate Covariation.

4.04.05.03 Level 3. Appropriate Covariation

At this level, responses represented all relevant features of the task: a curvilinear trend that showed increasing height with age and that showed constant height with changing age by or from age 20. Examples are shown in Figure 4.13 (R14-R19). A feature common to all of these examples was the assignment of age values to the horizontal axis, and common to most of these was the use of pictures rather than of explicit labels to denote height. R14 used pictographs to represent the minimal features required to show the curvilinear trend: three points, showing both growth for ages less than 20, and constant height from some age range older than 20. In R15, a fifth-grade male showed evidence of learning in Q2 about the need for independent axes to display repeated values of the same height for different ages. In response to Q1, he drew a pictograph showing an Increasing Trend, with height implied in the pictures and age labelled with a vertical scale also. He then began drawing a similar graph arrangement to Q2 (left graph of R15, Figure 4.13), but found difficulty drawing a 23-year-old, since his system could not represent age differing when height remained constant. He crossed out this graph, and then developed an appropriate two-axis system.

In R16, a sixth-grade male reversed the horizontal scale and did not label the graph, but features of the graph data indicate for values 20 and 30 on the horizontal axis (age), there is a constant vertical value (height). R17 showed evidence of change of thought apparently prompted by having ages on the vertical axis, as for R15. This response illustrated confusion over which axis age applied to: an amendment to position of the vertical axis label of “25” to coincide on the vertical scale with 20 resolved the issue to represent in a Curvilinear Trend (Level 3). This solution resolved the visual level of the data series but not the vertical axis labels.

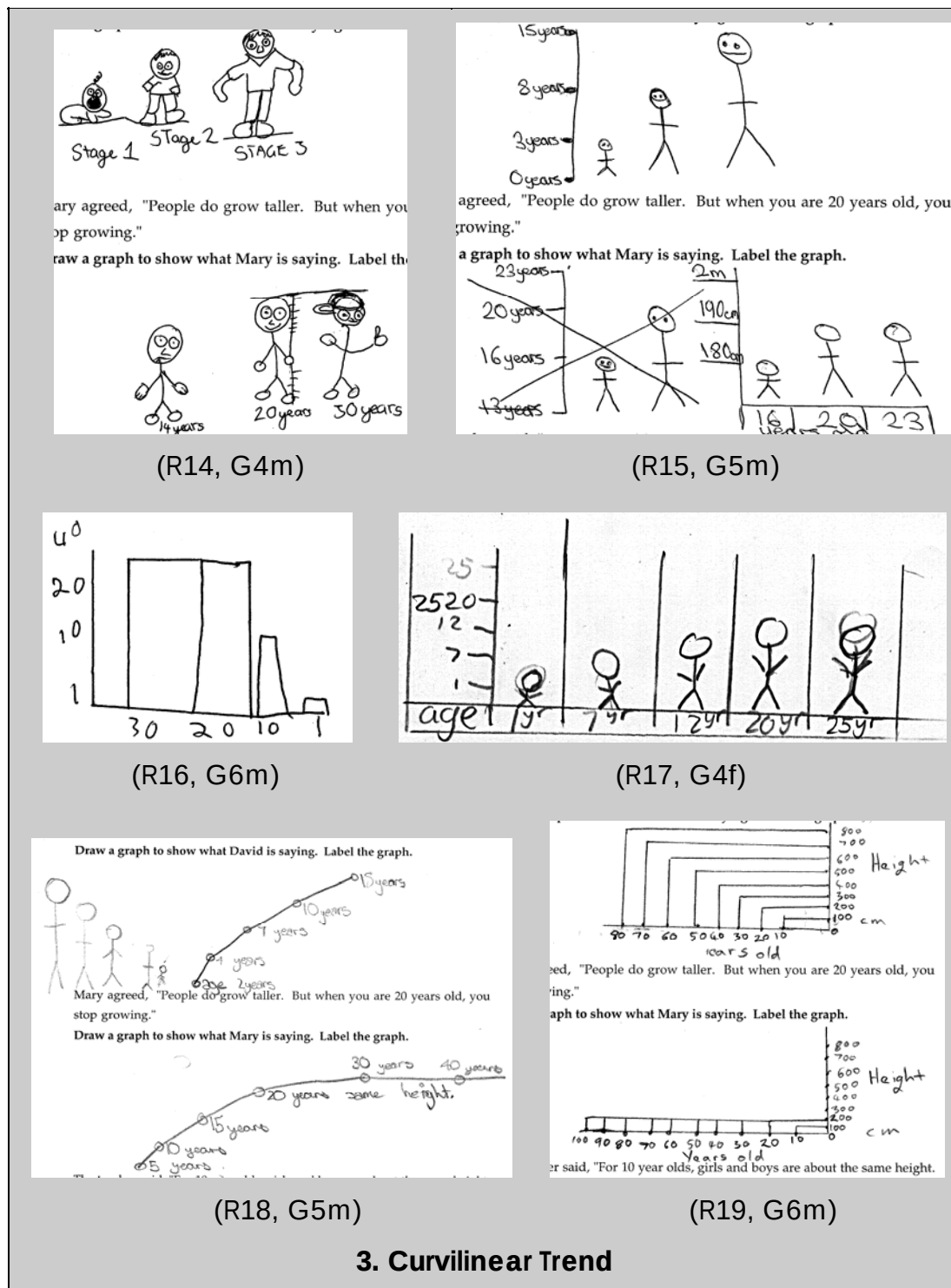


Figure 4.13. Student responses to Q1&2 – Level 3 – Appropriate Covariation.

R18 showed both a picture form and abstracted line graph form, without axes, for Q1, with only the abstracted form drawn for Q2. Height was implied but the shape of the line, and the picture graph for Q1 provided evidence the student

considered the representation this way, despite the lack of explicit height labels, values, or even a vertical axis of any sort. In contrast, R19 showed the Curvilinear Trend with complete labelling of variables, scale values and units, and explicit grid lines to emphasis the correspondence of values in reading the coordinate two-axis system.

4.04.06 Results – Q3 – Height Versus Age By Sex

4.04.06.01 Level 1. Single Statistical Aspect

Responses that were unsuccessful at representing the multivariate association omitted one variable or superimposed two variables to represent bivariate data. Three examples are shown in Figure 4.14 (R20-R22). R20 and R21 illustrated a method of omitting the variable of age to reduce complexity of the graph to a single comparison of the heights of one female and one male. R21 was drawn by the same student who drew R2 in Figure 4.10, a Single Axis response at Level 1 to Q1&2: both R2 and R21 reduced the complexity of the situation by reducing the number of variables to be shown. In R22, a fourth-grade female represented a constant function for Q3 just as she had for Q2, however for Q2 constant height was implied to be across varying ages, whereas for Q3, she annotated that the graph represented the same height across varying people at the given age 10.

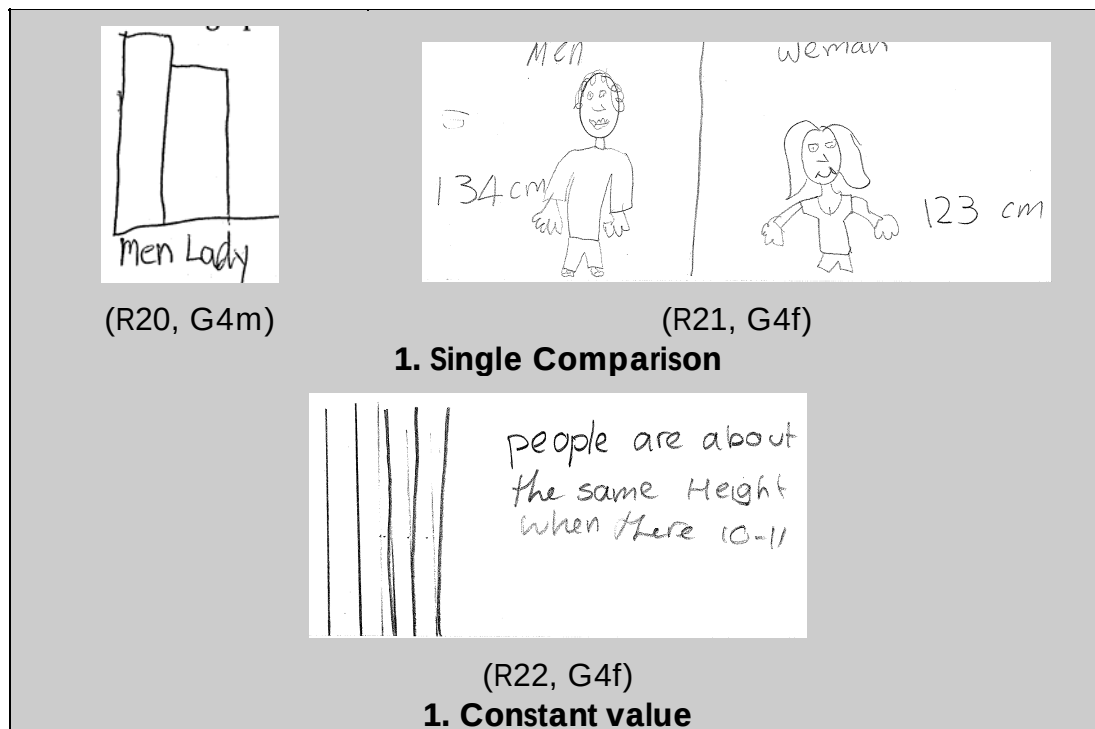


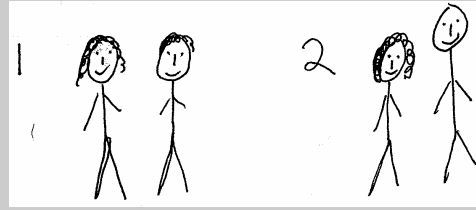
Figure 4.14. Student responses to Q3 – Level 1 – Single Statistical Aspect.

4.04.06.02 Level 2. Inadequate Covariation

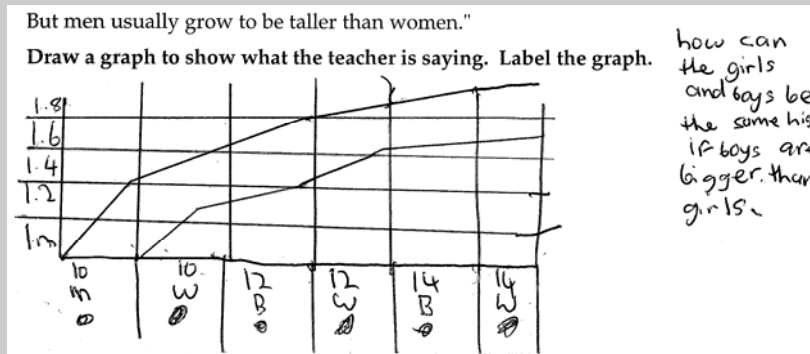
At Level 2, responses indicated all three variables—height, age, and sex—and displayed some but not all of the features relevant to the task. Five examples (R23-R28) are shown in Figure 4.15. In R23, a fourth-grade male, who responded at Level 2 to Q1&2 (R9, Figure 4.11), for Q3 represented a Double Comparison of female and male heights, indicating the same height at age “1,” and males taller at age “2,” but there was no indication of the critical feature of age 10.

The remaining four examples in Figure 4.15 (R24-R28) were Partial Series Comparison responses that showed growth for males and females, but did not establish equivalent heights at age 10. In R24, a student who offered an Increasing Trend at Level 2 for Q1&2 (R5, Figure 4.11) showed two trend lines, one male and one female, however the correspondence between these two was problematic. The offset between the data series was difficult to interpret, and even if the graph were

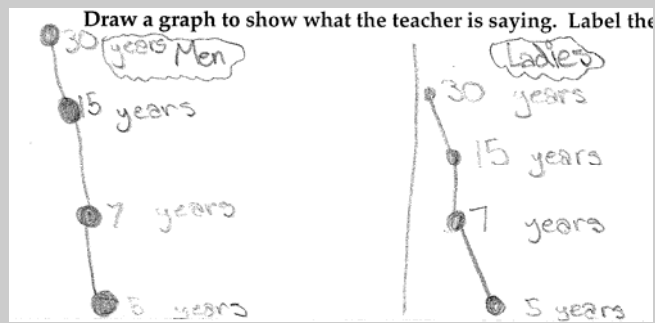
interpreted as shifting female data to overlay males at comparable ages, males and females were not shown as the same height at age 10. R25 also failed to show comparable heights at age 10, instead aligning them at ages 5 and 7, and otherwise showing relevant features, namely growth with age, and greater growth for males than females. This student had used the same graph form to effect in R18 (Figure 4.13), namely implying height by vertical position, and labelling data points with age labels, thus avoiding the need to use a formal horizontal age axis. R26 showed two vertical axes in a similar manner to his response to Q1&2 (R1, Figure 4.10), however in this case, the use of disparity between the scales for males and females indicated different variables. For example, for the label of 10, presumably age in years, males and females were at different height positions, hence although the responses did not indicate comparability at age 10, it did show growth in height over age for each sex with males finally taller than females. R27 also showed increasing heights for males and females, with greater growth for males, but not comparable heights at age 10. The problems labelling axes in R27 were evident in the previous response to Q1&2, seen in R17 (Figure 4.13): the vertical axis labels denoted ages, also shown on the horizontal axis to attempt to assert the correspondence, where the vertical axis positions, unlabelled, denoted heights.



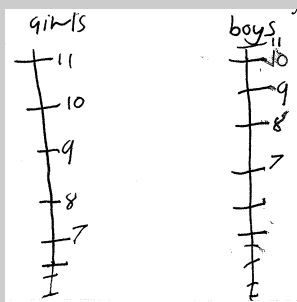
(R23, G4m)

2A. Partial Double Comparison

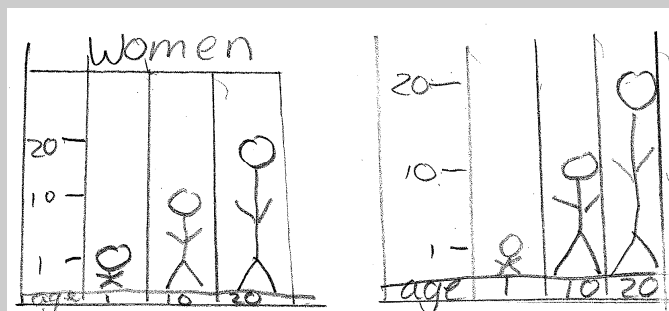
(R24, G6m)



(R25, G5m)



(R26, G5m)



(R27, G4f)

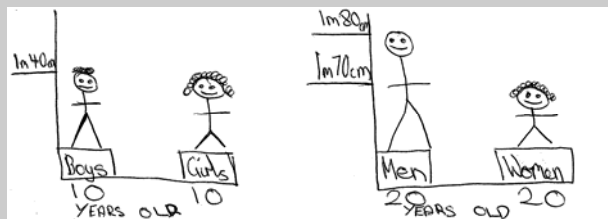
2B. Partial Series Comparison

Figure 4.15. Student responses to Q3 – Level 2 – Inadequate Covariation.

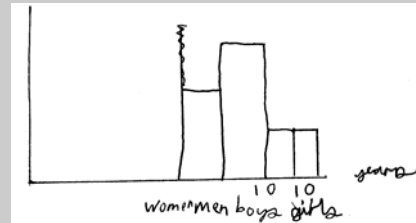
4.04.06.03 Level 3. Appropriate Covariation

Responses showing Appropriate Covariation represented all relevant features of the task: (a) at age 10, females same height as males, (b) height covarying with age (for females, males, or both), and (c) for some age greater than 10, males taller than females. Six examples (R28-R33) are shown in Figure 4.16. The fifth-grade male who developed his two-axis system in R15 (Figure 4.13) adapted it in R28 to produce a Complete Double Comparison, the horizontal axis denoting sex as a variable, and two graphs denoting two ages. Another student drew a Complete Double Comparison (R29) with a similar graph form to the Level 3 response the student offered to Q1&2 (R16, Figure 4.13).

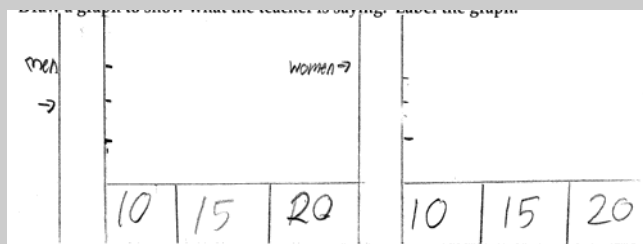
R30, drawn by the same student who drew R13 (Figure 4.12), highlighted some interesting features of the coding used in this investigation. In R13, the student's response showed growth by dual labels, but failed to show the constant height for increasing age required in Q2. In R13, the three heights read bottom to top corresponded consecutively with the three ages read left to right. This evidence was used to make informed sense of the student's Speculative Data Generation evident in response to Q3. In R30 the student dropped labelling on the vertical axis, thus vertical position was inferred as denoting height, by the principle of indicative labelling. Each of the three marks on the vertical axis were inferred as corresponding to the three ages on the horizontal axis, thus the response was assessed at Level 3, rather than consider the response as showing the three heights for age 10 (Level 2), which would effectively penalise the student for poor graph production skills.



(R28, G5m)



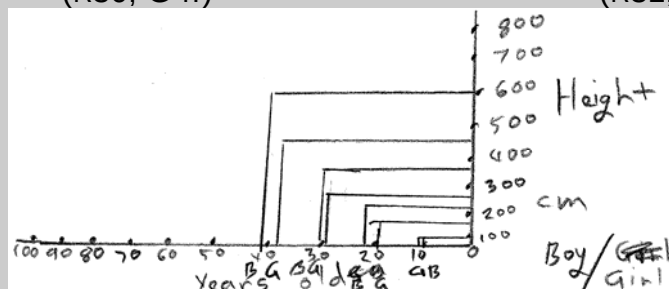
(R29, G6m)

(a) 3A. Complete Double Comparison

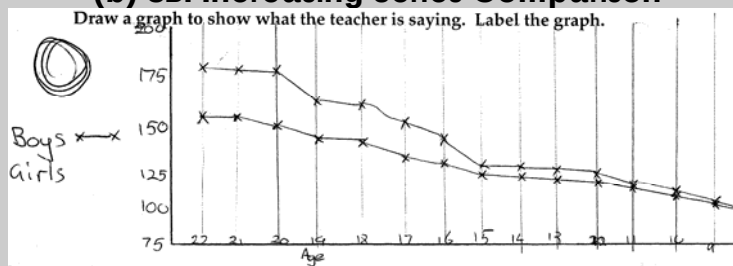
(R30, G4f)



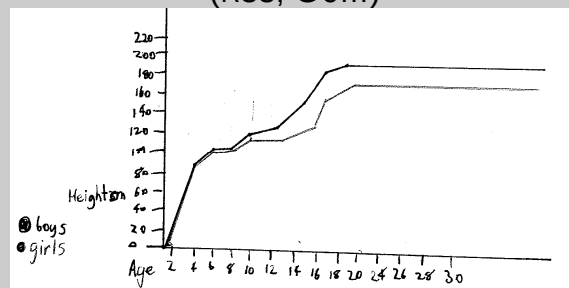
(R31, G5m)



(R32, G6m)

(b) 3B. Increasing Series Comparison

(R33, G6m)



(R34, G6m)

(c) 3C. Curvilinear Series Comparison

Figure 4.16. Student responses to Q3 – Level 3 – Appropriate Covariation.

The fifth-grade student who drew R8 (Figure 4.11) gave a response to Q3 that was difficult to interpret as shown in R31, Figure 4.16. He appeared to have used vertical position to indicate height as evident in Q2 (R8, Figure 4.11), and used “age 10-” not as an axis scale but as a key to the horizontal lines within each bar. Thus each of three marks in the bars indicated heights at ages 10, 20, and 30 years of females and males, with the same height at age 10, and males taller at age 20 and again at age 30. This response was one for which the second coder disagreed with the interpretation of the researcher. Interviewing the student may well have clarified the student’s intent, however interviews were not part of this investigation.

R32 showed height-age series for males and females effectively using the same form as the student had used in R19 (Figure 4.13). In R33, a sixth-grade student scaled the horizontal axis increasing right-to-left rather than left-to-right, but represented all relevant features of the task. R34 showed a more conventional representation, and included growth spurts within the data.

4.04.07 Discussion of Investigation 2

Three levels of Speculative Data Generation were described for bivariate tasks related to height versus age (Q1&2), and for multivariate tasks (Q3) also involving sex. These levels were comparable to levels observed in Investigation 1 for a different context (Task 1 – Heart deaths), however categories of response within each level differed across tasks. For Q1 and Q2, over 90% of students from grade 4 represented a trend successfully at Level 2 or Level 3, and for Q2 in particular, over 35% represented a curvilinear trend at Level 3. These levels for the bivariate task were in sharp contrast to the limited success of students’ in the multivariate context (Q3) of Investigation 1. They were also high success rates for primary school students in comparison to the eighth-grade students in the study by Mevarech and

Kramarsky (1997), of whom 55% and 38% respectively represented an increasing function and a curvilinear function. Their criteria, however, included labelling of axes. Indicative labelling was considered a strength of the current coding scheme for assessing Speculative Data Generation, in acknowledging student understanding of covariation implicit in representations, for example, the height label being implicit in a pictograph of a person.

A large number of responses were pictographs. Drawings of people implicitly show an iconic representation of height, which might be considered in the Ikonik mode of thinking according to the SOLO model (Biggs & Collis, 1991). This imagery, however, appeared to support representation of various heights, and also aligned height to the vertical orientation. For some students, labelling was incomplete and the resulting graph lacked clarity: for some, the issue was not resolved, due to co-alignment on the vertical axis, and for others, this issue supported correction to represent a coordinate system.

The task to represent a curvilinear association (Q2) posed difficulties of representation for over 30% students, who responded at the level of Inadequate Covariation (Level 2). Some of these students did not show a consolidated sense of a Cartesian coordinate system, which was perhaps not surprising at these grade levels. This did not mean that the task was inappropriate for these students; on the contrary, for some students, the difficulties posed in Q2 prompted revision of their representational system to display two independent axes, one each for height and age. The prompting provided by the task was similar to the design of garden-path tasks (Posner & Gertzog, 1982) and pre-empted the structured prompting for cognitive conflict used in Investigations 3B, 4B, and 5B in this study. The issues

involved in structuring an appropriate coordinate system are addressed in more detail in Chapter 5 in relation to students' Graph Production skills.

The task to represent multiple variables (Q3) involving comparison of bivariate associations posed difficulties for many students. The solution for over 30% of grade 4 students was to represent a comparison of heights of one female and one male, thus eliminating age as a variable. This strategy to reduce representational complexity has been observed by other researchers as “one-point” graphs (Mevarech & Kramarsky, 1997) and “single comparison” graphs (see Investigation 1). Responses at levels 2 and 3 were either in the form of Double Comparisons (R24, R28, R29), or in the form of Series Comparison graphs (R25, R27, R30-R33) as observed in Investigation 1.

4.05 DISCUSSION OF SPECULATIVE DATA GENERATION AS EVIDENT IN INVESTIGATIONS 1 AND 2

Investigations 1 and 2 (along with Investigation 7 appearing in Chapter 6) were conducted prior to the administration of other investigations, as mentioned in the study samples detailed in Chapter 3. The preliminary findings of Investigations 1 and 2 thus informed the design for the broader investigative structure of this study shown in Figure 1.01.

Investigations 1 and 2 both raised issues concerning the relationship between “understanding” covariation and representing it graphically, and the influence of the topic familiarity and complexity of the measures. The dominance of Single Comparison graphs from primary school students in Investigation 1 suggested that familiarity with a limited range of graph forms, particularly bar graphs, may have

restricted the ways students considered the data types, in particular the variation within each variable. In Investigation 2, for Task 3 (Figure 4.07), the familiarity of the context of height and age appears to have supported the alignment of the visual image of height with picto-bar graph forms and allowed students to conceive of representing multiple people with variable heights and ages.

Both investigations provided evidence of students considering discrete data cases, some with case labels. In Investigation 2, it was observed that the familiarity of height and age being attributes that correspond to a person – that is, the case of a person as an entity that establishes correspondence between data pairs (height and age values) or even data triples (height, age, and sex values) – may have allowed students to take as a given the correspondence of variables via the case labels, and thus sequentially construct the variation across cases. Evidence from Investigation 1 was also observed in reflecting on the responses of Double Comparison and Trend Graphs (Level 2), and Series Comparison graphs (Level 3): a feature of many of these responses was the representation of time in addition to heart deaths and motor vehicle use. Time values, usually years, often provided case labels that established the correspondence that appeared to be the basis for attempting to represent the covariation. Some students confidently structured Series Comparison graphs with an implicit common time basis, without a label or values for time. For others less confident, however, it is easy to see how thinking could be restricted to Single Comparisons. Consider possible reasoning behind a Double Comparison: “in 1970, hearts deaths had value H1, and motor vehicle use had value M1, whereas in 1990, hearts deaths had value H2, and motor vehicle use had value M2.” The dialogue would be rather problematic to resolve with the year references. A more subtle word such as “when” would be needed to establish the status of two corresponding pairs of

values as opposed to four distinct values (e.g., "when heart deaths were H1, motor vehicle use was M1, and when heart deaths were H2, motor vehicle use was M2").

Case information of times may also be the reason that functional thinking over time appears to be so supportive of students' understanding of covariation (Krabbandam, 1982). Time acts as both the ordered covariate and the case label: the case label is unique in functional situations, and time is a common identifier of events in natural language using verb tense ("Grace will run, is running, or has run to the shop") or more specifically with temporal clauses (e.g., "In the morning").

These issues are explored in following investigations, in part using the same tasks as the basis for discussions with students in interviews. Revisions and generation of subsequent tasks and coding methods, as discussed in Chapter 3 and the following investigations, was informed by these issues of multivariate complexity, discrete cases from collected data, and teasing out students' Speculative Data Generation as distinct from the Graph Production skills.

4.06 INVESTIGATION 3A: HEIGHT VERSUS AGE (SURVEYS)

4.06.01 Introduction and Aims

Investigation 3A involved a replication of Investigation 2 to confirm the coding framework with a broader sample of students from grades 3 to 9, aligned with Research Aim 1 of the wider study (Section 3.02). It also aimed to provide survey responses as the basis for interviews with a subset of these students in Investigation 3B. The administration of this task to the same sample of students for other tasks (Tasks 4, 5, and 6, involving Speculative Data Generation, Coordinate Graph Production, and Graph Interpretation respectively) was also important to explore conceptual development across the various tasks, , aligned with Research Aim 4 of the wider study (Section 3.02) to compare evidence across skills, as discussed in Chapter 7.

4.06.02 Task

Task 3 was identical to that used in Investigation 2, as shown in Figure 4.07. The task was the first in the survey task booklets for Data Collection 3, as detailed in Chapter 3.

4.06.03 Participants and Method

Students described in Data Collection 3 were administered Task 3, as seen in Table 3.02. Q1 and Q2 were administered to students in grades 3, 5, 7 and 9, whereas Q3 was only administered to students in grades 7 and 9. Responses were assigned to levels and categories within the assessment frameworks in accord with the criteria detailed for Investigation 2 in Figures 4.08 and 4.09. To investigate further issues

arising from Investigations 1 and 2, graph form, and representation of continuity or of case labels, were also analysed.

4.06.04 Quantitative Results

Table 4.05 shows the percentages of responses levels to Q1 and Q2, coded as a single task (Q1&2) for each sex and grade. In response to Q1&2, more than 90% of students represented either Inadequate Covariation (Level 2) or Appropriate Covariation (Level 3), as observed in Investigation 2. In response to Q3, only a smaller subset of older students were administered the question, and most responded at high levels, as shown in Table 4.06. Consideration of response categories indicated similar results to Investigation 2: in response to Q1&2, approximately equal numbers represented Increasing Trend (Category 2A) or Partial Curvilinear Trend (Category 2B), and in response to Q3, Series Comparisons (Categories 2B, 3B, 3C) were more common than Double Comparisons (Categories 2A, 3A).

Tables 4.05 and 4.06 also show the percentages of responses by graph form to Q1&2 and to Q3 for each sex and grade. Pictographs tended to be drawn by students in lower grade levels (e.g., Grade 3), line graphs by students in higher grade levels (7 and 9), and bar graphs by students in middle grade levels (5, 7 and 9). Differences between males and females were likely to reflect the class groupings from which they were selected, as discussed in Chapter 3. Case labels with people's names were drawn by few students (10 for Q1&2, 1 for Q3; mixed year levels and sexes); more students drew continuous line graphs (20 for Q1&2, 20 for Q3), including seventh- and ninth-grade students only, particularly males. This may have reflected the common teaching experience of the class groupings.

Table 4.05.

Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 3 Q1&2

	Response Count	Female Grade				Male Grade			
		3	4/5	7	9	3	5	7	9
Response Level	N=183	n=26	n=26	n=25	n=14	n=19	n=21	n=25	n=27
1. Single Statistical Aspect	14	8	12	0	14	26	5	4	0
2. Inadequate Covariation	85	73	62	32	43	63	62	44	0
3. Appropriate Covariation	84	19	27	68	43	11	33	52	100
Response Category (selected)									
1A. Miscellaneous	6	0	4	0	14	11	0	4	0
1B. Single Axis	8	8	8	0	0	16	5	0	0
2A. Increasing Trend	46	46	42	20	7	16	43	16	0
2B. Partial Curvilinear Trend	39	27	19	12	36	47	14	28	0
2C. Group Comparison	1	0	0	0	0	0	5	0	0
Graph Form									
Pictograph	39	80	31	0	0	32	14	8	0
Bar graph	69	4	38	48	73	42	76	32	11
Line graph	43	0	0	32	13	0	0	44	81
Mixed/Other	32	16	31	20	13	26	10	16	7

One G9f gave an incomplete response to Q1 only, due to time constraints. The response showing Level 2 Increasing Trend was not considered in the analysis, as it is unknown whether Level 3 would have been achieved for Q2.

Table 4.06.

Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 3 Q3

	Response Count N=83	Female Grade		Male Grade	
		7 n=25	9 n=7	7 n=24	9 n=27
Response Level					
1. Single Statistical Aspect	3	4	14	4	0
2. Inadequate Covariation	16	28	57	21	0
3. Appropriate Covariation	64	68	29	75	100
Response Category (selected)					
2A. Partial Double Comparison	4	4	14	8	0
2B. Partial Series Comparison	12	24	43	13	0
3A. Complete Double Comparison	5	4	0	17	0
3B. Increasing Series Comparison	23	36	0	29	26
3C. Curvilinear Series Comparison	36	28	29	29	74
Graph Form					
Pictograph	2	0	0	8	0
Bar graph	28	56	41	38	7
Line graph	50	40	43	54	89
Other	3	4	14	0	4

4.06.05 Results – Q1&2 – Height Versus Age

Responses at various levels and categories generally were similar to those provided as examples for Investigation 2. Two examples (R1, R2) shown in Figure 4.17 were classified as Single Axis graphs at Level 1, Single Statistical Aspect. Both resemble the growth chart of an individual, R1 using the name "David," and R2 using similar face images.

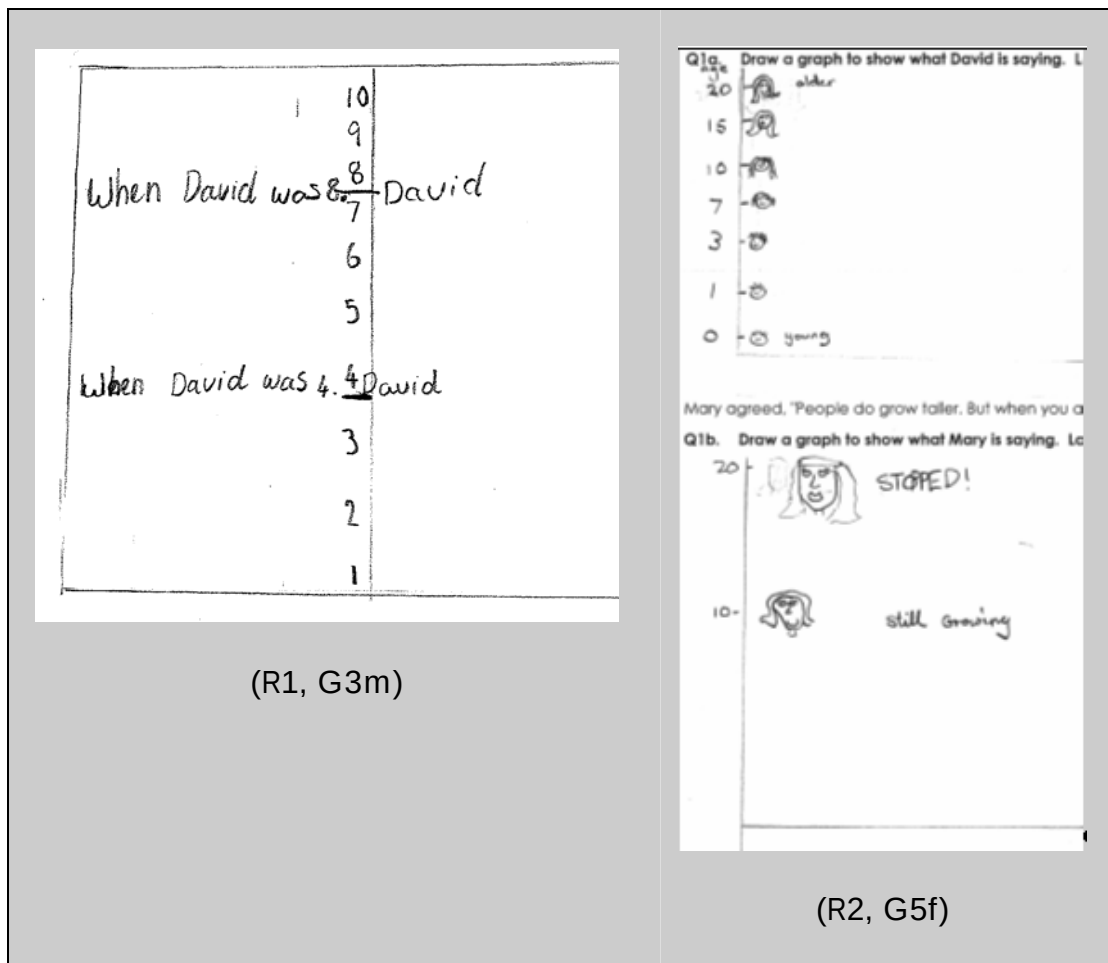


Figure 4.17. Student responses to Q1&2 – Level 1 – Single Statistical Aspect.

Responses at Level 2 displayed Inadequate Covariation. Five examples are shown in Figure 4.18 (R3-R7). R3 and R4 showed Increasing Trends (Category 2A) in the speculative data and exhibited unconventional features of graphs. The fifth-grade female who drew R3 showed unordered values, but showed a clear positive covariation between the values. In R4, a third-grade male displayed 10 people with names, with corresponding heights and ages showing an increasing trend. This response was similar at first glance to some Single Axis responses, however the dual labelling of the scale effectively produced a data table, hence satisfying criteria for Speculative Data Generation, although it was not displayed in a conventional coordinate system.

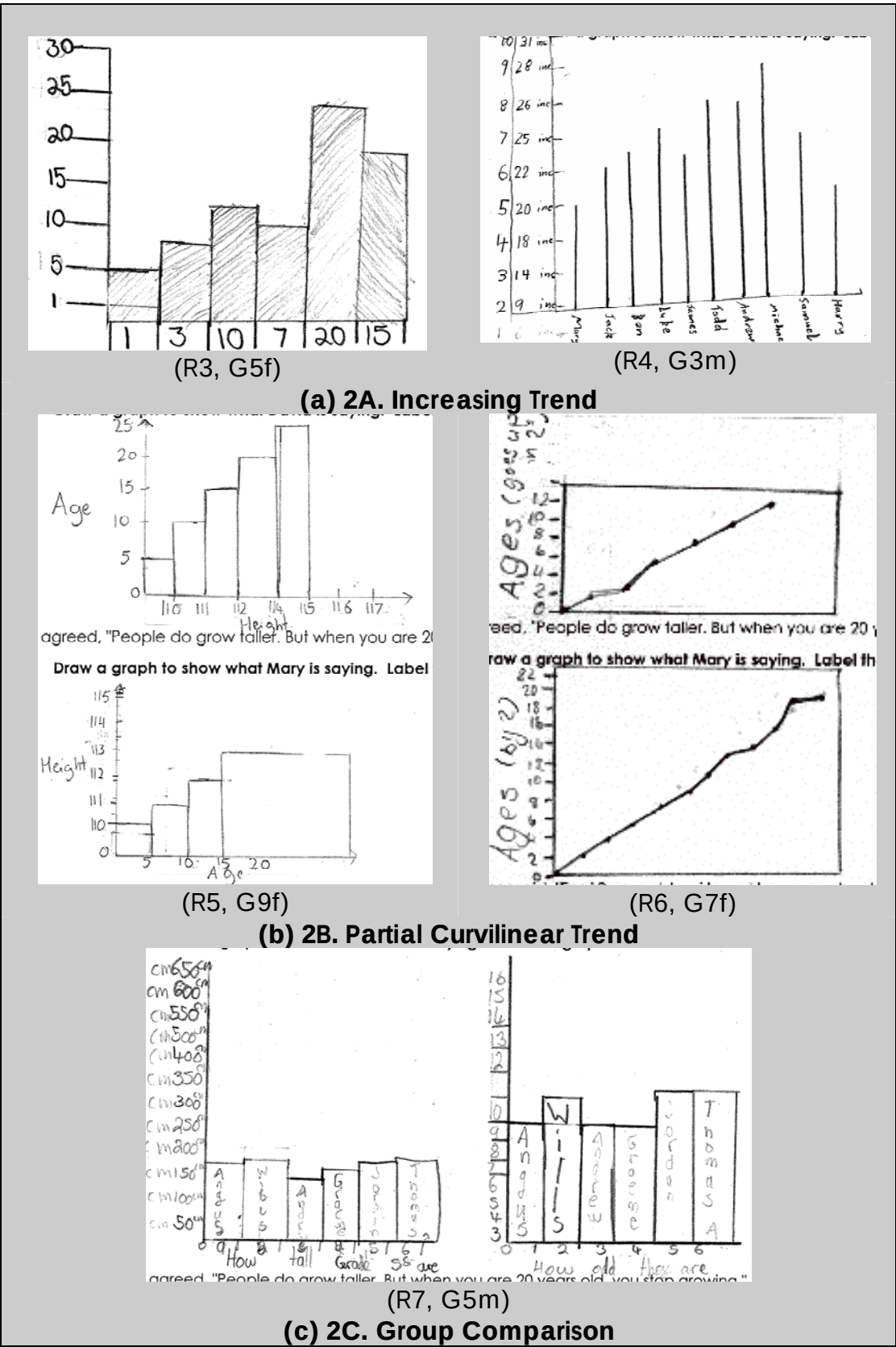


Figure 4.18. Student responses to Q1&2 – Level 2 – Inadequate Covariation.

Other responses at Level 2 showed a Partial Curvilinear Trend (Category 2B).

Two examples in Figure 4.18 (R5 and R6) displayed two common unsuccessful strategies in showing lack of growth from age 20, one representing no ages greater than 20 (R5), the other showing a flattening of the graph but with age rather than height on the vertical axis (R6). One fifth-grade male produced a Group Comparison (R7, Category 2C), a type not observed in Investigation 2. The use of dual graphs for age and height was uncommon, and although six names cases are shown, only two unique ages are shown (R7, right), for which those with the older age tend to have greater heights (R7, left).

Responses at Level 3 showed Appropriate Covariation. Two examples in Figure 4.19 (R8, R9) illustrated the sophistication two ninth-grade males brought to the task. In R8, the student titled the graph “How high someone is compared to their age,” using the language of comparison, with realistic units, growth spurts between dotted data values, and lines for both Q1 and Q2 showing the continuous nature of growth. In R9, the student titled the graph “Height for 100 surveyed people from 14 to 60,” and used a scattergraph form not only to show the curvilinear trend, but also to show population variation of heights for given ages.

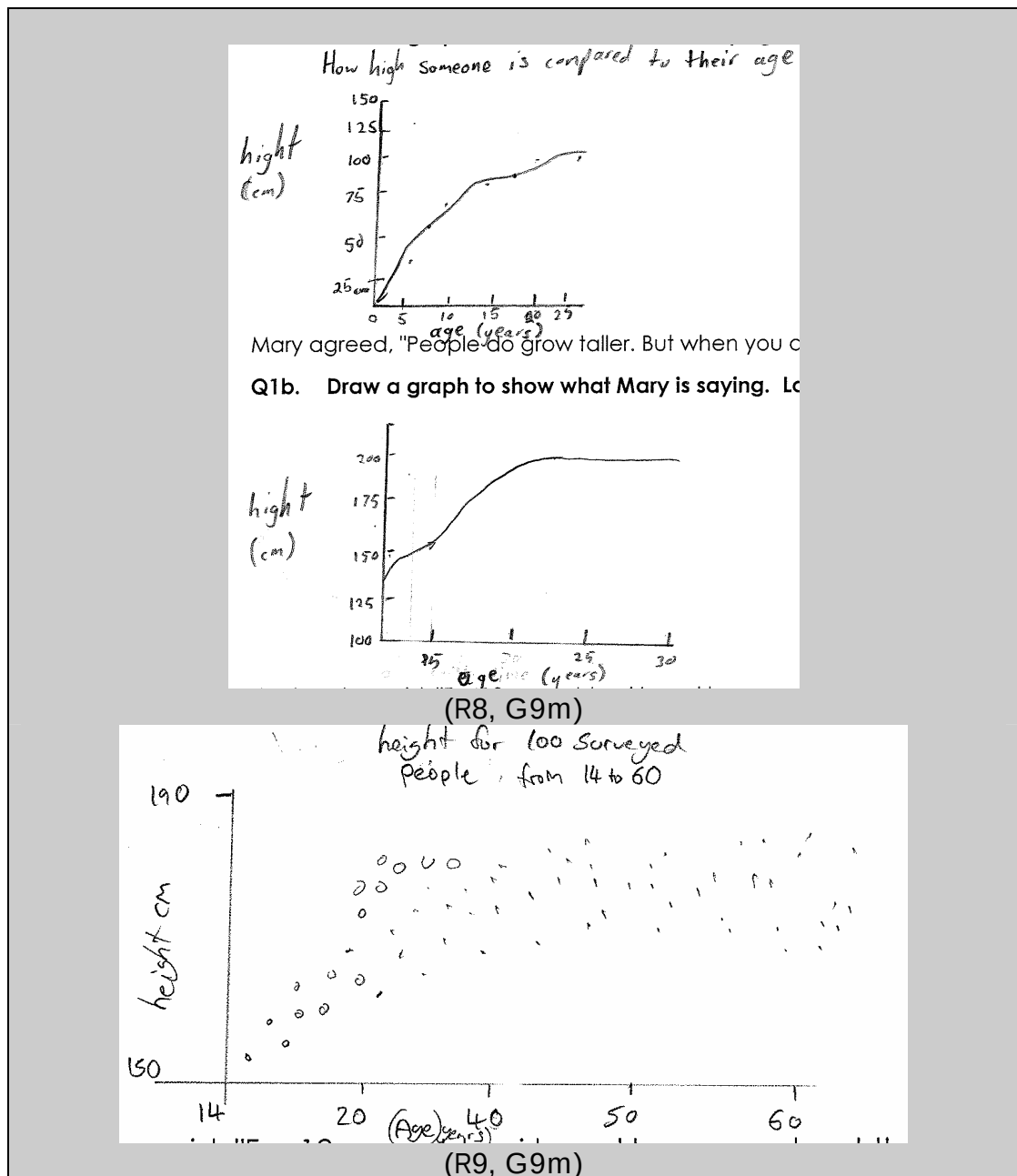


Figure 4.19. Student responses to Q1&2 – Level 3 – Appropriate Covariation.

4.06.06 Results – Q3 – Height Versus Age By Sex

Many responses showed Appropriate Covariation for Q3, as shown in Table 4.06. Some students, however, encountered difficulties showing the appropriate relation. Two examples are shown Figures 4.20 and 4.21. In R10, a seventh-grade male showed heights for a mixed-sex class of 10-year-old students, but showed no

indication of various ages; it was assessed at Level 1. In R11, a seventh-grade female drew an increasing series of heights (in feet) by ages for girls and boys, however the correspondence between the girls' and boys' graphs did not establish males being taller than females at age 20; it was assessed at Level 2.

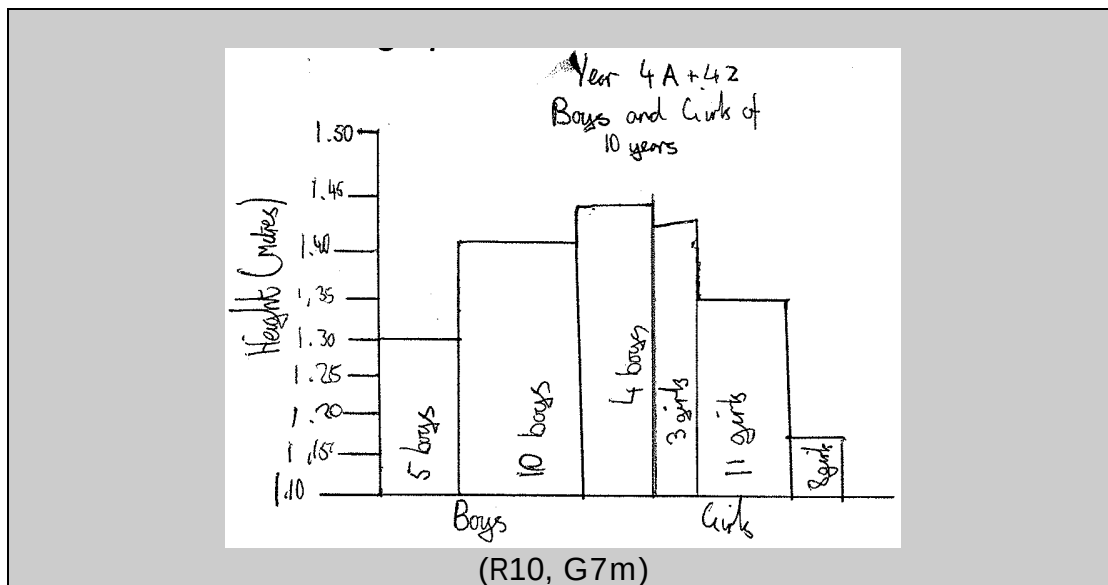


Figure 4.20. Student response to Q3 – Level 1 –Single Statistical Aspect.

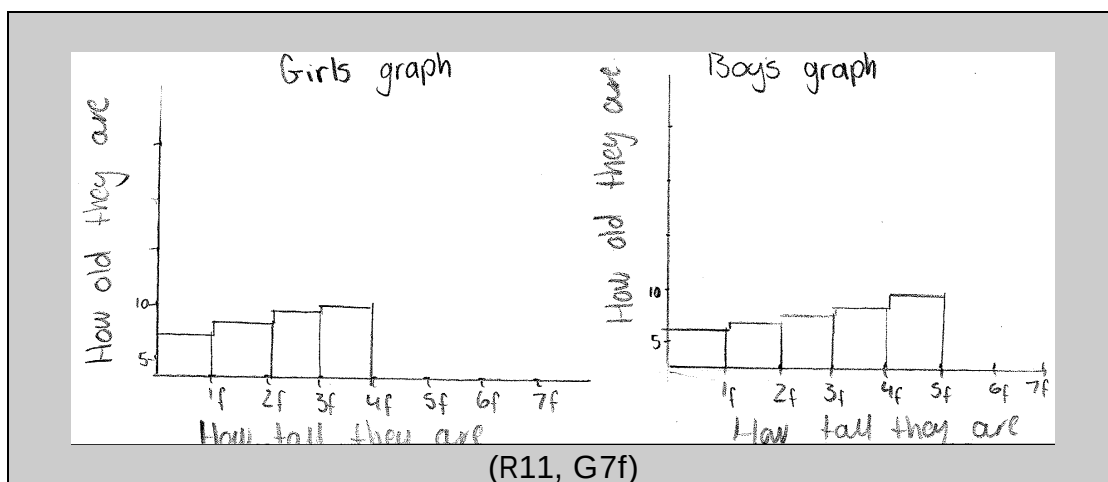


Figure 4.21. Student response to Q3 – Level 2 – Inadequate Covariation.

Two examples (R12, R13) of responses showing Appropriate Covariation (Level 3) are shown in Figure 4.22. In R12, a seventh-grade female showed six age ranges for each of females and males, and corresponding heights that were similar

for females and males at ages 6-9 and 9-11, and greater for males with older ages.

Pie graph sectors were increasing in area with age and height, however the covariation was established by the written values, thus constituting a tabular representation. In contrast, a ninth-grade male used the conventions of coordinates to show Appropriate Covariation using continuous lines (R13).

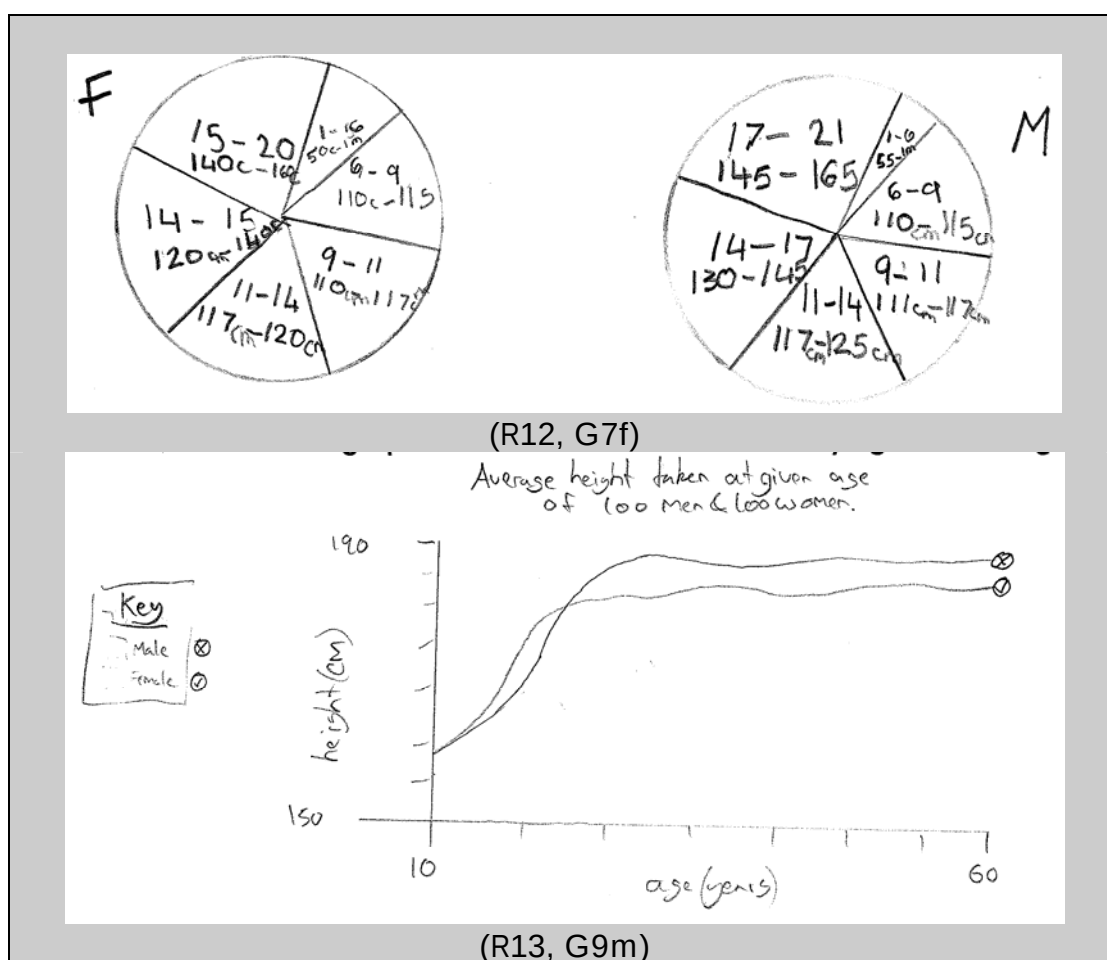


Figure 4.22. Student responses to Q3 – Level 3 – Appropriate Covariation.

4.06.07 Discussion of Investigation 3A

The coding framework used in Investigation 2 was used to code responses to students from a wider range of year levels. All responses were assigned to the framework of levels, although evidence of a Group Comparison category emerged for Q1&2, similar to those observed in Investigation 1 and Q3 from Investigation 2.

Many students responded at Levels 2 and 3, particularly seventh- and ninth-grade students. For Q3, Series Comparison graphs were more common by seventh- and ninth-grade students than Double Comparisons, suggesting familiarity in considering height versus age as a data series, as many students had established in Q1&2. Consideration of graph forms, case labels and continuous representation suggested that students in lower year levels often represented discrete data values, in pictographs or bar graphs, whereas students in ninth-grade tended to represent continuous change, such as in line graphs. Both of these approaches are supported by the context of age and height: the variables are known to be properties of individual people, especially as height is visible in pictures, and both variables are known to change gradually for individuals.

The difficulties students encountered most commonly related to the curvilinear relation (Q2), especially for lower year levels. Compared to Q2, fewer difficulties were observed for Q3, as it was not administered to these lower year levels. Investigation 3B explored further with students in interviews the nature of the challenges they faced and the reasoning they used, evident in their views of alternative responses presented to them.

4.07 INVESTIGATION 3B: HEIGHT VERSUS AGE (INTERVIEWS)

4.07.01 Introduction and Aims

Investigation 3B involved an interview protocol that asked students to explain their own graphs, and then to compare their responses with others, some of which were higher level, lower level, or simply alternative to their own responses. The investigation aimed to explore thought processes of students when verbally explaining their reasoning, and students' potential to develop reasoning via dealing with cognitive conflict, aligned with Research Aim 3 of the wider study (Section 3.02). In particular, it was explored whether students' reactions to prompts at various levels related to the levels of the students' survey responses.

4.07.02 Task

The interview protocol followed the general interview protocol format outlined in Chapter 3. Students were shown their own graph, and asked to explain what they had drawn and why they responded in that way to show the verbal statement. They were then shown a sequence of prompts, as seen in Figures 4.23, 4.24, and 4.25. Each prompt was presented on a separate A4 page, in colour where necessary, with the labels "Graph A," "Graph B," and so on appearing at the top left of the page. All three prompts to Q1 show an Increasing Trend (Category 2A), but the responses differ in labelling and graph form conventions. Prompts for Q2 included an Increasing Trend (Prompt A), a Partial Curvilinear Trend (Prompt B), and two showing a Curvilinear Trend (Prompts C and D, R14 and R16 of Investigation 2). Prompts for Q3 included a Single Comparison graph (Level 1,

Prompt A, R20 Investigation 2), a Partial Series Comparison (Level 2, Prompt B, R27 Investigation 2), and two responses at Level 3 (Prompts C and D).

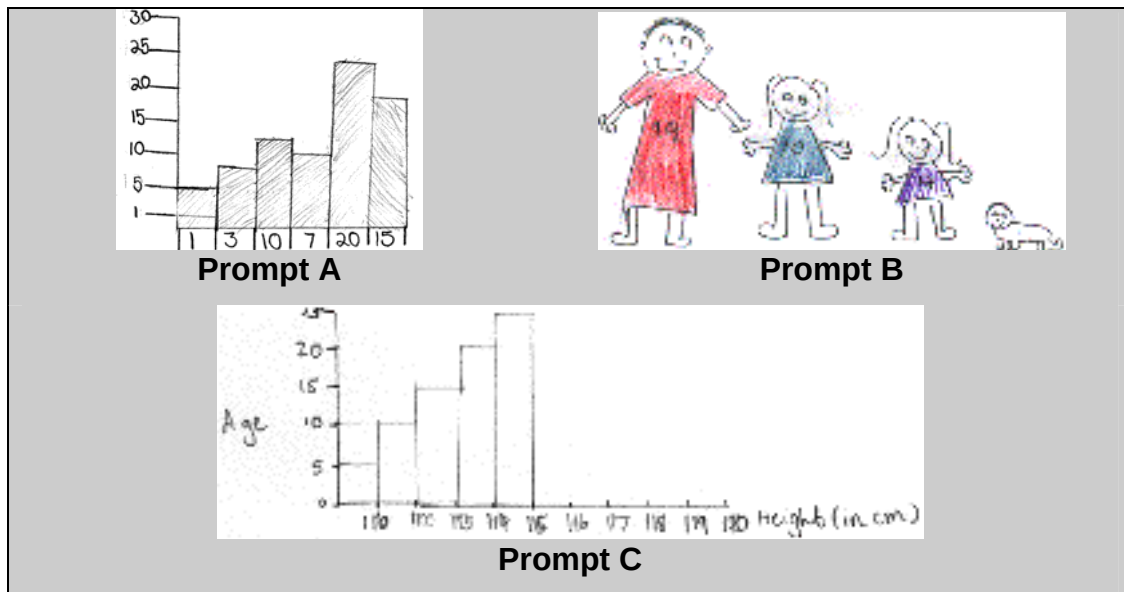


Figure 4.23. Prompts displayed to students to promote cognitive conflict for Q1.

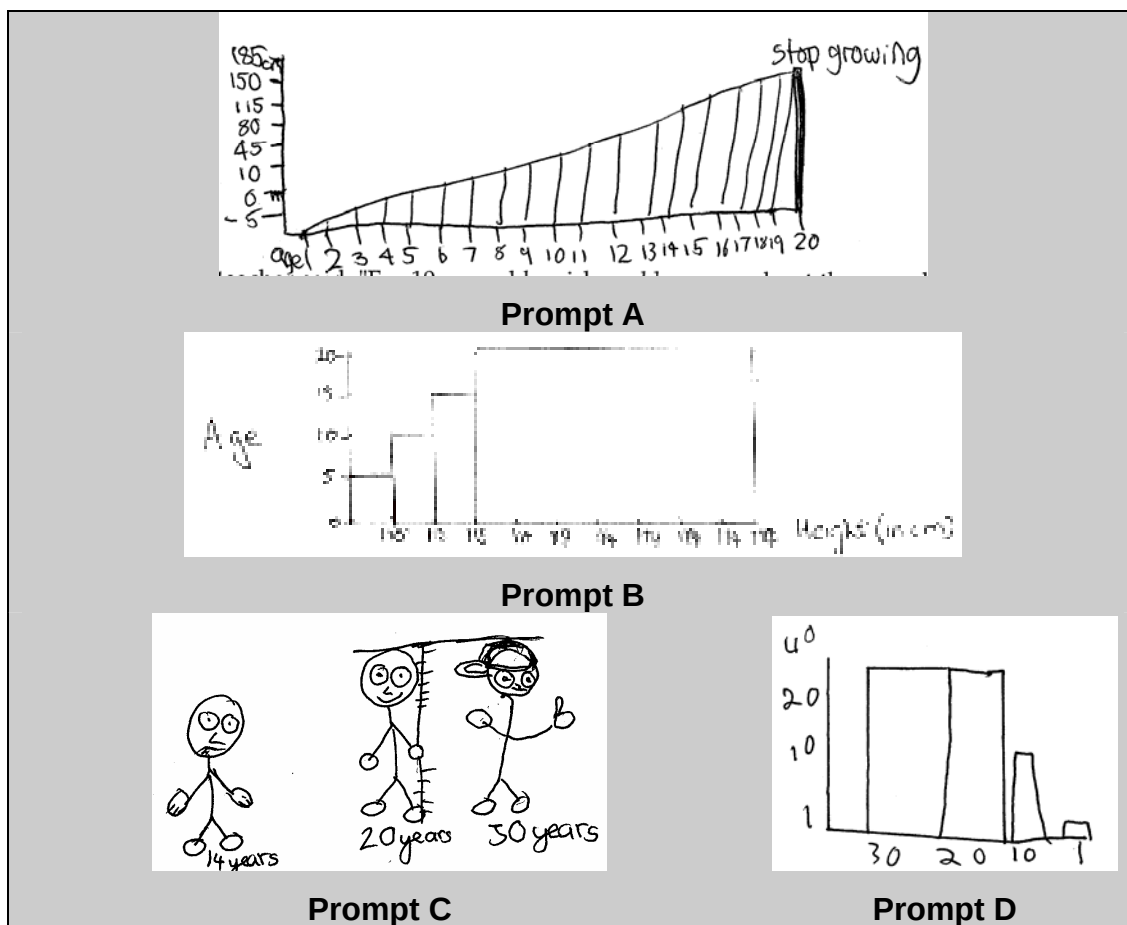


Figure 4.24. Prompts displayed to students to promote cognitive conflict for Q2.

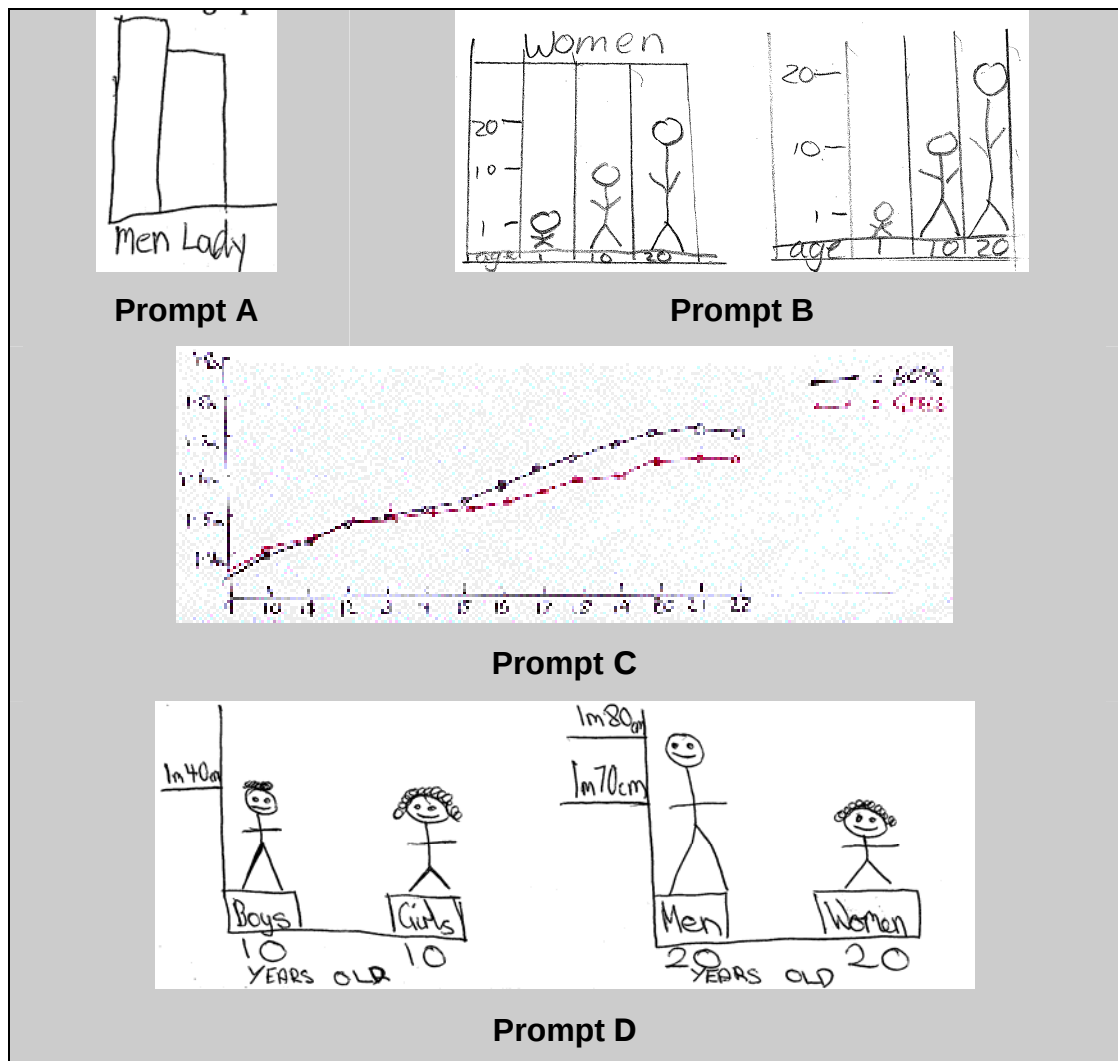


Figure 4.25. Prompts displayed to students to promote cognitive conflict for Q3.

Students were asked to comment on the each prompt graph, and to decide which response was preferred to show the required information appropriately and clearly. Prompt C for Q1 and Prompt B for Q2 were omitted for the student who drew them (G9f5).

4.07.03 Participants and Method

Participants were all 34 students interviewed as described for Data Collection 3, although only 19 of these explored Q3. Students were asked to explain their own graphs, and then to react to prompts, as per the general interview protocol as described in Chapter 3. Responses were considered in relation to their agreement,

disagreement or ambivalence to each prompt presented. Dialogue was also examined for emerging themes.

4.07.04 Results – Overview

A summary of the dialogue concerning Q1&2 for each of 34 interviews is shown in Table 4.07, and concerning Q3 for 19 interviews is shown in Table 4.08. Of those interviewed in relation to Q1&2, one did not respond in the survey, two offered Level 1 responses, 20 offered Level 2 responses, and 11 offered Level 3 responses to the survey. Of those interviewed in relation to Q3, four did not respond in the survey, two offered Level 1 responses, eight offered Level 2 responses, and five offered Level 3 responses.

Tables 4.07 and 4.08 show the reactions of general agreement (✓), disagreement (✗) or ambivalence (~) to each of the prompts. "N" was used to denote that a graph was not shown as a prompt. The summary counts shown in Table 4.07 indicated that for Q1&2, most interviewees disagreed with most prompts; those which attracted most agreement were Prompts B (Level 2) and C (Level 2) for Q1, and Prompt C (Level 3) for Q2. For Q1&2, of the 30 respondents who were coded with an initial survey response level, 6 agreed with a higher-level prompt, 2 agreed to a lower-level prompt, 15 agreed with a prompt at the same level, and 10 did not agree with any prompt. For Q3, most disagreed with Prompt A (Level 2), most agreed with Prompt C (Level 3), and opinions were mixed for Prompts B (Level 2) and D (Level 3). Of the 15 students who were coded with an initial survey response level, 10 agreed with a higher-level prompt, 4 agreed with a prompt at the same level, and 1 did not agree with any prompt.

No clear pattern differences were observed for reactions to prompts, according to the students' initial response levels. In particular, there were students

who initially responded in each of the three levels who agreed with Prompt C (Level 3) for Q2 (Table 4.07) and Prompt C (Level 3) for Q3 (Table 4.08).

Themes identified in many interviews concerned aspects of presentation including labelling, realism of values and growth spurts, neatness, and graph form. In some interviews, themes relating to Speculative Data Generation and/or presentation included limited ways of showing "stop growing," order, scale reversal, and axis allocation.

Table 4.07.

Summary Descriptions of Interview Dialogue for Q1&2

ID	Initial Response Category (Q1&2)	Prompt – Category of Prompt							Summary of Reactions to Prompts
		Q1			Q2				
		A – 2A	B – 2A	C – 2A	A – 2A	B – 2B	C – 3	D – 3	
G5f4	(none on survey)	✗	✓	✗	✗	✗	✓	✗	No initial response
G3f4	1. Single Axis	~	✓	✗	~	~	✓	✗	Agree Higher (1→3)
G5f3	1. Single Axis	✗	~	✓	✓	✗	✓	✗	Agree Higher (1→3)
G3f1	2A. Inc. Trend	✓	~	✗	✓	~	✓	✗	Agree Higher (2A →3)
G5f5	2A. Inc. Trend	✓	~	✗	✗	✗	~	✗	Agree Same (2A→2A)
G7f2	2A. Inc. Trend	~	✓	✓	✗	~	~	~	Agree Same (2A→2A)
G5m1	2A. Inc. Trend	✗	✓	✓	✓	✗	✗	✗	Agree Same (2A→2A)
G5m2	2A. Inc. Trend	~	~	~	~	~	✗	~	No change
G7f4	2A. Inc. Trend	✗	~	✓	✓	✗	~	✗	Agree Same (2B→2A)
G5f6	2B. Part. Curv.	✗	✓	✓	✗	✗	✗	✗	Agree Same (2B→2A)
G9f3	2B. Part. Curv.	✗	~	✓	✗	~	~	✗	Agree Same (2B→2A)
G7f5	2B. Part. Curv.	~	✓	✓	✗	✗	✓	✗	Agree Higher (2B→3)
G9f1	2B. Part. Curv.	✗	~	✓	~	~	✓	✗	Agree Higher (2B→3)
G9f5	2B. Part. Curv.	✗	~	N	~	N	✗	✗	No change
G3m3	2B. Part. Curv.	~	~	~	~	~	~	~	No change
G5f2	2B. Part. Curv.	~	✗	~	✗	✗	✗	✗	No change
G7m2	2B. Part. Curv.	✗	~	✓	✗	✗	✗	✗	Agree Same (2B→2A)
G7m3	2B. Part. Curv.	✗	~	✓	✗	~	✗	✗	Agree Same (2B→2A)
G3m1	2B. Part. Curv.	✓	✗	✗	✗	✗	~	✗	Agree Same (2B→2A)
G3m2	2B. Part. Curv.	✗	~	~	✗	~	✓	✗	Agree Higher (2B→3)
G3f2	2B. Part. Curv.	✗	~	✗	✗	✗	~	✗	No change
G3f3	2B. Part. Curv.	~	✗	✗	✓	✗	✗	✗	Agree Same (2B→2A)
G3f5	2B. Part. Curv.	✗	✓	✓	✗	✗	~	✗	Agree Same (2B→2A)
G5m3	3. Curv. Trend	~	✓	✓	✗	~	✓	✗	Agree Same (3→3)
G5f1	3. Curv. Trend	✗	✓	✓	~	✗	✓	✗	Agree Same (3→3)
G7m1	3. Curv. Trend	✗	~	✓	✗	✗	~	~	No change
G7f1	3. Curv. Trend	✓	✓	✓	~	✗	✓	~	Agree Same (3→3)
G7f3	3. Curv. Trend	~	~	✗	~	✓	~	~	Agree Lower (3→2B)
G7f6	3. Curv. Trend	✗	✓	✗	~	~	~	✗	No change
G9m1	3. Curv. Trend	~	✓	✗	✓	~	~	~	Agree Lower (3→2A)
G9m2	3. Curv. Trend	✗	~	~	~	~	~	✗	No change
G9m3	3. Curv. Trend	✗	~	✗	✗	✗	✗	✗	No change
G9f2	3. Curv. Trend	✗	~	~	~	~	~	~	No change
G9f4	3. Curv. Trend	✗	~	✓	✓	✓	✓	✗	Agree Same (3→3)
Total	Prompts ✗	20	3	11	16	16	9	26	
Total	Prompts ~	10	19	6	11	13	14	8	
Total	Prompts ✓	4	12	16	7	3	11	0	

Note: ✓ denotes agreement, ✗ denotes disagreement, ~ denotes ambivalence

Table 4.08.

Summary Descriptions of Interview Dialogue for Q3

ID	Initial Response	Prompt – Category of Prompt				Summary of Reactions to Prompts
		A – 2A	B – 2B	C – 3C	D – 3A	
G3m1	0. No response	✗	~	~	✗	No change
G5m1	0. No response	~	✗	✗	✓	Agree Higher (0→3A)
G9f3	0. No response	✗	~	✓	~	Agree Higher (0→3C)
G9f5	0. No response	✗	~	✓	~	Agree Higher (0→3C)
G7m2	1. Single Comparison	✗	~	✓	~	Agree Higher (1→3C)
G7f2	1. Single Comparison	✗	✓	✓	✗	Agree Higher (1→3C)
G7m1	2A. Partial Double Comp.	✗	~	✓	~	Agree Higher (2A→3C)
G7f3	2A. Partial Double Comp.	✗	~	✓	✓	Agree Higher (2A→3AC)
G9f4	2B. Partial Series Comp.	✗	✓	✓	~	Agree Higher (2B→3C)
G7m3	2B. Partial Series Comp.	✗	✗	✓	✗	Agree Higher (2B→3C)
G7f1	2B. Partial Series Comp.	~	✓	✓	~	Agree Higher (2B→3C)
G7f4	2B. Partial Series Comp.	~	✓	✓	~	Agree Higher (2B→3C)
G7f5	2B. Partial Series Comp.	✗	✗	✓	✗	Agree Higher (2B→3C)
G9f1	2B. Partial Series Comp.	✗	✗	✓	✗	Agree Higher (2B→3C)
G7f6	3A. Comp. Dbl. Comp.	✗	✓	✓	✗	Agree Same (3A→3C)
G9f2	3B. Curv. Series Comp.	✗	~	✓	✓	Agree Same (3B→3AC)
G9m1	3B. Curv. Series Comp.	✗	✓	✓	✓	Agree Same (3B→3AC)
G9m2	3B. Curv. Series Comp.	✗	~	✓	~	Agree Same (3B→3C)
G9m3	3B. Curv. Series Comp.	✗	✗	~	✗	No change
Total	Prompts ✗	16	5	1	7	
Total	Prompts ~	3	8	2	8	
Total	Prompts ✓	0	6	16	4	

4.07.05 Results – Interview Dialogue

The following extracts of interview dialogue illustrated common themes.

Each interviewee is uniquely identified corresponding to entries in Table 4.07 and in some cases, Table 4.08.

One fifth-grade female (G5f3) drew a Single Axis graph (Level 1, see R2 in Figure 4.17). She explained her own graph to Q1 coherently referring to ages and growing taller, but acknowledged the challenge posed by Q2.

G5f3: [Q1] Well here is like zero [?] And she is really short, and then here is 1 years old, and here's Karen taller and here's just the growing taller and then seven, taller and taller up to 20 when they stop growing there.

G5f3: [Q2] I wasn't really... didn't know really what to draw here.

I: What made it difficult to draw this one?

G5f3: I couldn't, like, I found it hard to set it out and show how it works stop growing and then you keep on growing.

Her reference to stop growing and keep on growing, without specifying whether the “growing” was in relation to growing taller or growing older, was the same ambiguity created by the Single Axis graph. For Q2, she considered Prompt A addressed the issue of stopping growing: “It is fairly clear. You can tell that this is the age and when you stop growing.” She preferred Prompt C, and found Prompts B and D hard to understand. Prompt C was preferred not only for ease of reading, but also because a key feature of relevant ages was shown.

G5f3: [Q2, Response to Prompt C] It is good way. It shows that when you're young you grow up and when you're 20 year-old you just stop. It is a good way I think. Yes it is a good way.

I: So do you think that's better than this way [Prompt A]?

G5f3: I think this way is more graphing like [Prompt A] but I think this is easier for people to understand this one [Prompt C]. It is best for little kids, is easy to understand.

I: When you were drawing this, did you think of drawing anything like this and then decided that that wasn't right to do it or...?

G5f3: No, because this is sort of like this, because they show it when you're 30 or older.

I: If you're going to show that on yours, what would you have done for the someone who is 30?

G5f3: Probably the same things as this person did [Prompt C] because more people older had stopped.

The question remained how she could have done so in the Single Axis framework used in her initial response (R2, Figure 4.16).

A seventh-grade male (G7m2) who drew a Partial Curvilinear Trend to Q1&2, and a Single Statistical Aspect graph to Q3 (see R10 in Figure 4.19) responded on presentation aspects as many other students did, but with his own emphasis on the frequency of students graphed and other intuitive notions irrespective of the verbal statement asserted.

G7m2: [Explaining own Q2] Well I was just trying to prove that she is actually wrong in some cases. Yes because most 20 year olds are still growing and other 20 year olds they shrink and of the 20 year olds just don't grow like she says. What she is saying is not even right, or not even wrong.

The ninth-grade female (G9f5) who drew Prompt C for Q1 with inverted axes (see Figure 4.22) represented ages to 25, and explained her response as follows.

G9f5: Well I did the 5, 6, again and when the person got to 15 and then to 20, they stopped, and I just, so they stop growing at 115 cm and I just made it to look like going along with them stopping going, just kind of the same thing.

She also drew Prompt B for Q2 (see Figure 4.23). Note that for Q1, ages were shown to age 25, however in Q2 the age scale went only to 20, and a large rectangle was designed to show repeated value of the same height along the horizontal axis. In discussing Q2, for Prompt A she was ambivalent, suggesting it “kind of” showed it, “Only because they wrote ‘stop growing’ on there.”

G9f5: Oh, you could say like have 21, the same height 22 the same height 23 the same height [indicating horizontal extension to age 20 bar] so you could get an idea that once you reach the age of 20, you stop growing.

For Q2 Prompt C, she believed additional heights would better show the trend. For Prompt D, she was able to use her understanding of the covariation to interpret the axis allocations, and then suggest swapping the axis allocation for her own response.

G9f5: [Q2, Prompt C] Not really, because umm.. Oh, well I suppose when it says 20 and 30 and it has got the same height but again you did not know how high it is just sketch. And then it started 14 they don't give you other ages to compare it with.

- G9f5: [Q2, Prompt D] It doesn't again, you wouldn't know where that is the age or the... Oh, that is the age probably [horizontal axis].
- I: Why do you think that would be the age?
- G9f5: Because that is 20 up there and then it goes to 40 [vertical axis] and so and then 20 and 30 here are the same height [pointing to horizontal axis and two left most bars] it doesn't give you any information as to which, what kind of like, age or anything.
- I: Yes so they haven't labelled it like you have, they haven't given very good units have they, like centimetres. So overall of the graphs you have seen would you say that yours is the best, you would say?
- G9f5: Yes.
- I: We were talking just here on the Prompt A about extending that out to age 21 . And things like that and everything. If you were to extend yours, maybe up until age 25 or something, how would you do that?
- G9f5: I would probably swap the ages and the heights around, so the age would be down the bottom and the height would be up there, so when you did get 20 you could keep the same line [motioning horizontal line].

Another ninth-grade female (G9f1) drew a graph with inverted axes, as shown in Figure 4.26. The inverted axes created a tension between values on the axes, showing repetition, and the vertical position of the data series, showing a horizontal plateau consistent with the intuitive mapping of height shown vertically. This tension was highlighted in dialogue for Q2 Prompt B, where the interviewer probed in relation to the ages shown.

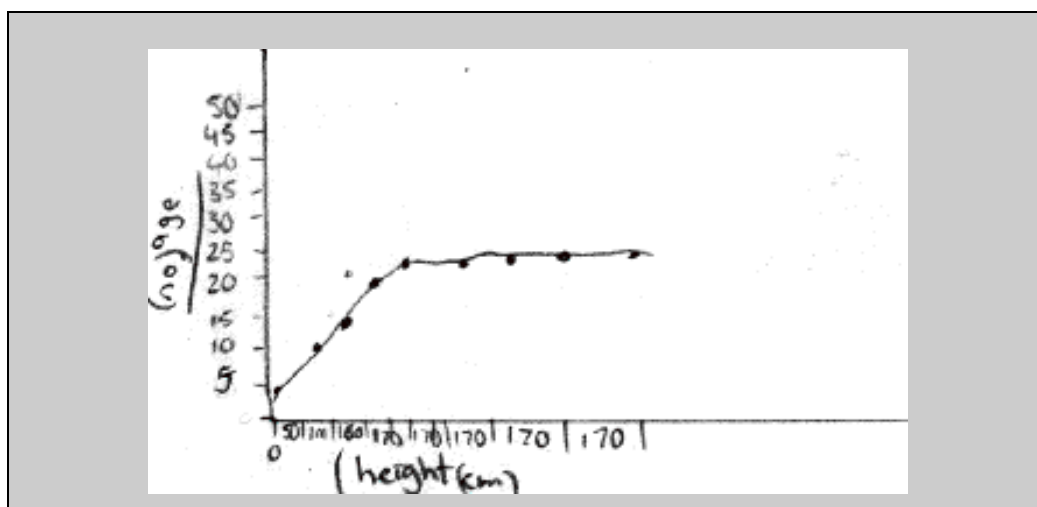


Figure 4.26. Response to Q2 with difficulty showing non-growth (G9f1).

- G9f1: [Q2, Prompt B] Okay at first glance, I didn't really understand that but I do now because you stop growing there and you keep keep on living but she does stop growing. And it is more like that [Student's own graph] because if I just did a bar graph, it would be the same as mine.
- I: So you think one of them is better for showing the information?
- G9f1: Maybe that one [Prompt B]. Because you can see it clearer that stop you there, that stays the same from then on, but maybe if they were to put separate bits, you could see that they stay the same, it might just look like one big graph thing that they have done.
- I: Can you tell how old the person would grow on this graph here?
- G9f1: 20. Yes.
- I: And on yours?
- G9f1: 50 that one confused me here, because I just kept going on 25 but yes it is meant to go up to 50.
- I: Is meant to go to 50. You don't have it quite like you would want it?
- G9f1: What they need to stay the same, but it does that... I'm a bit confused, does that show that that is 50 [Student pointing with two fingers raising vertical from horizontal axis and ruling horizontal line from vertical axis].
- I: You are the one telling the story, it is your graph.
- G9f1: I'm a bit confused. Because it is meant to go up to 50 years old, people 50 years old, but it looks like it just goes to 20 and then I just sort of...

In a lengthy dialogue, the student resolved to add bars to the right of the large rectangle in Prompt B, reproducing a graph very similar to her own.

In one rare case, a student (G9f3) used the word “corresponding” in relation to Prompt C for Q1, “probably the best way to display it, it goes, is it ‘corresponding,’ I’d better not use any big words and confuse myself. It is going up in the right order, and it shows the person getting taller as they get older.”

4.07.06 Discussion of Investigation 3B

Dialogue from interviews with students clarified some of the reasons why students drew their responses as they did. In some cases, students who did not represent the curvilinear trend appeared to have difficulty verbally expressing this trend. In other cases, explanations by students were procedural, thus not clarifying

the reasons for their representations or how their graphs showed the verbal statement. This may have reflected the interview process, which did not request laboured explanation in interview, but placed more emphasis on the prompts to follow to elicit discussion.

Coding of prompts was undertaken according to general agreement, disagreement or with ambivalence to each of the prompts. This coding generally reflected the student comments, although in some cases it was problematic, in that an initial reaction by a student was followed by other opinions, in some cases elicited as the interview probed various issues as part of the dialogue.

Analysis of reactions to Q1&2 prompts indicated students tended to disagree with prompts, often based on presentation reasons more than data structures related to Speculative Data Generation. This in part was likely to be due to the students interviewed. In particular, many students interviewed initially responded at Level 2, however most rejected Level 3 prompts (Prompts C and D for Q2). The criticisms for Prompt C concerned the pictograph not showing values for heights or not being a formal graph, and Prompt D was often rejected because the graph was not labelled. These students did not recognise the data structure and characteristic graph curve. In contrast, for Q3 prompts, Prompt C, a dual-line graph, was recognised by most students as appropriate. There was minimal evidence that reactions to prompts were related to the students' initial response levels. It is worth noting, however, that as the prompts and many students' responses were often similar at Levels 2 and 3 with respect to Speculative Data Generation, much of the dialogue and reaction to prompts concerned presentation aspects related to labelling, and graph form. The issue of axis allocation emerged as significant for the curvilinear trend, and was

apparently influenced by imagery of heights being represented vertically, even if age were labelled on the vertical axis.

Investigations 2 and 3 concerned students representing the familiar context of height growth with age, a context which has a natural visual mapping for pictographs and which has been reported previously (Ainley, 1995) as supporting high level responses. The following Investigations, 4A and 4B, involved surveys and interviews with a subset of the students involved in Investigations 3A and 3B, but in contrast used a task involving familiar variables with a counterintuitive covariation and without a natural visual mapping in pictographs.

4.08 INVESTIGATION 4A: TEST SCORES VERSUS STUDY TIME (SURVEYS)

4.08.01 Introduction and Aims

Investigation 4A further explored students' Speculative Data Generation using a task involving test scores versus study times, to expand the basis for the assessment framework for Speculative Data Generation, aligned with Research Aim 1 of the wider study (Section 3.02). The task was designed as part of the third phase of task design (see Chapter 3), aiming to provide a context of data collection with familiar variables and involving counterintuitive covariation. Assessing the proportion of students responding at the various levels to this task was aligned with Research Aim 2 of the wider study (Section 3.02). It was expected that students would demonstrate higher response levels than for Task 1 (Heart deaths), which had complexities of unfamiliar measures of multiple variables. It was also expected that the task would provide more challenges for Speculative Data Generation than Task 3 (Heights), which involved a familiar covariation that could be presented using imagery in pictographs.

4.08.02 Task

Task 4 was presented in different forms to different students, as shown in Figure 4.27. Q1 was administered in a negative, counterintuitive, covariation form to most students, and in an alternative positive form (Q1*) to third- and fifth-grade males. An additional question involving no covariation (Q2) was administered to ninth-grade males only. These different forms were designed to explore whether students might respond differently due to their prior beliefs about the covariation.

They were not administered to all students due to time constraints and the desire to sample student responses across a variety of tasks.

(Negative covariation form)

Anna and Cara were doing a project on study habits.
They asked some students two questions:

- “What time did you spend studying for the spelling test?”
- “What score did you get on the test?”

Anna asked 6 students. She used the numbers to draw a graph.
She said, “People who studied for more time got lower scores.”

Q1. Draw a graph to show what Anna is saying for her 6 students.

Label the graph.

(Positive covariation form)

... She said, “People who studied for more time got higher scores.”

Q1. Draw a graph to show what Anna is saying for her 6 students.*

Label the graph.

(No covariation form)

Cara asked 6 students. She used the numbers to draw a graph.
She said, “Students studied different amounts of time, but the times were not related to the scores they got.”

Q2. Draw a graph to show what Cara is saying for her 6 students.

Label the graph.

Figure 4.27. Task 4 to assess Speculative Data Generation.

4.08.03 Participants and Method

Subsets of students were asked to graph a negative covariation (Q1) or a positive covariation (Q1*). Coding was based on the four levels used in Investigations 1 to 3, as shown in Table 4.01. For this task, the coding categories identified are shown in Figure 4.28, which were similar to those used for Task 1 in Investigation 1 (see Figure 4.02). To be coded at Level 3, Appropriate Covariation, responses showed the correspondence of variation in two variables, in that (a) the variables were identified with adequate variation and (b) the direction of the

correspondence of variation was appropriately specified. Variables were considered adequate if two conditions were satisfied:

- (1) labels were explicit, or units (e.g., hours/minutes) or values (e.g., digital time format) were used that indicated which variable was denoted, using the notion of indicative labelling as employed in Investigations 2 and 3, and
- (2) the graph included adequate variation of at least three bivariate values; although the context described six data cases, three were considered sufficient to demonstrate the covariation.

The direction of the correspondence of variation was appropriately specified either by values at least ordinal in nature (e.g., “not at all,” “not much,” “a lot”) or by convention of height/sector angle. In addition to analysis by response levels and categories, responses were coded with respect to graph form used.

Level of Speculative Data Generation	Category of Speculative Data Generation		
0. Nonstatistical Numbers and context are not combined; if numbers are represented they are not ordered	0A. No response 0D. Unlabeled Graph Value denoted by number or position, but with no variable label	0B. Axes 0E. Labeled Axes Variable labels (one or more) on graph axes, but with no data	0C. Picture/Narrative
1. Single Statistical Aspect Graphs show either correspondence in a single bivariate data case without variation, or variation for single variable data	1A. Single Comparison Correspondence in a single bivariate data case	1B. Single Variable Variation for one variable denoted by number or position	
2. Inadequate Covariation Correspondence and variation are shown, but either the variation, the variables chosen, or the correspondence is inadequate	2A. Double/Group Comparison Correspondence in two bivariate data cases/conditions with inadequate variation	2B. Trend Graph Variation but correspondence is between inappropriate variables	2C. Double Variable Variation for two variables but correspondence is not in the required direction
3. Appropriate Covariation Correspondence of variation is shown	3A. Series Comparison Two data series, each in a line or at least 3 cases, compared on the same axis dimension	3B. Table Covariation is shown using written values rather than using spatial aspects of a graph.	3C. Coordinate Variable A line or at least 3 bivariate data cases related on the opposing axes

Figure 4.28. Levels and categories of Speculative Data Generation for Task 4.

4.08.04 Quantitative Results

Table 4.09 shows the percentages of response levels to Task 4 for each sex and grade. Of 167 responses, most students demonstrated Appropriate Covariation (Level 3, 96 responses), or Inadequate Covariation (Level 2, 35 responses). Of the 96 responses at Level 3, most (50) showed Coordinate Variable (Category 3C); Tables (18 responses) tended to be drawn by third- and fifth-grade students, and Series

Comparison (Category 3A) graphs (15 responses) by fifth- and seventh-grade students. Whereas there were similar numbers of Double/Group Comparison (Category 2A) responses as Double Variable (Category 2C) responses (16 each), there were very few Single Comparison (Category 1A) responses (2) in comparison to Single Variable (Category 1B) responses (16). Compared with the negative covariation task format, in response to the positive covariation format, fewer students gave Double Variable (Category 2C) graphs with an incorrect direction and more students gave a Single Variable (Category 1B) graph, as if both variables could be aligned into a single axis of corresponding or identical values.

In relation to graph forms, tables were drawn only by students in lower grade levels, bar graphs were drawn by students at all grade levels, and line graphs were drawn by students in higher grade levels only. Case labels were shown in 42 responses, whereas only 7 responses showed continuous data in a line. Case labels were coded for many bar graphs, and tables also showed discrete cases although they were rarely labelled with names. Discussion of responses is divided into the three forms of the questions: negative, 130 responses, positive, 37 responses, and no covariation, 27 responses.

Table 4.09.

Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 4

	Response Count	Female Grade				Male Grade			
		3	4/5	7	9	3 ^a	5 ^a	7	9
Response Level	N=167	n=26	n=26	n=22	n=5	n=19	n=18	n=24	n=27
0. Nonstatistical	18	23	12	5	40	11	11	8	0
1. Single Aspect	18	0	12	14	20	42	11	4	0
2. Inadequate Covariation	35	42	38	18	20	5	0	21	11
3. Appropriate Covariation	96	35	38	64	20	42	78	67	89
Response Category									
0A. No response	4	0	0	5	20	0	11	0	0
0B. Axes	2	4	0	0	0	5	0	0	0
0C. Picture/Narrative	5	15	4	0	0	0	0	0	0
0D. Unlabelled Graph	4	4	8	0	0	5	0	0	0
0E. Labelled Axes	3	0	0	0	20	0	0	8	0
1A. Single Comparison	2	0	4	0	0	0	6	0	0
1B. Single Variable	16	0	8	14	20	42	6	4	0
2A. Double/Group Comparison	16	27	8	9	0	5	0	8	7
2B. Trend Graph	3	4	8	0	0	0	0	0	0
2C. Double Variable	16	12	23	9	20	0	0	13	4
3A. Series Comparison	15	0	8	27	0	5	17	13	0
3B. Table	18	27	15	5	0	11	17	0	4
3C. Coordinate Variable	63	8	15	32	20	26	44	54	85
Graph Form									
None/Picture	10	19	0	5	20	0	17	0	0
Table	31	62	31	0	0	16	17	4	0
Bar	84	8	65	77	60	74	61	58	22
Line	25	0	0	14	0	0	6	29	52
Scatter	8	8	0	0	0	0	0	4	19
Other	9	4	4	5	20	11	0	4	7
Case labels	42	8	8	55	20	22	37	33	21

^a Third- and fifth-grade males were administered Q1* rather than Q1.

4.08.05 Results – Negative Covariation (Q1)

4.08.05.01 Level 0. Nonstatistical

Fourteen students responded with no evidence of a data set of covariation for test scores and study times. Examples (R1-R5) are shown in Figure 4.29. Two students gave no response. Five students identified the Narrative context (Category 0C) without a data set, such as a written narrative with names for individuals and a single test score of “10/10” (R1) or narrative explanation (R2). Four students drew an Unlabelled Graph (Category 0D) that gave no indication of the data set for the variables being measured and also failed to show six data cases. One fifth-grade female’s Unlabelled Graph (R3) showed ordered steps, a response close to being coded as a Trend Graph (Category 2B), however apart from lack of labels, the lack of six data values and the lack of values to clarify the direction of the trend resulted in coding at Level 0. Three students drew Labelled Axes (Category 0E) that identified each variable but without clear data points (R4, R5).

4.08.05.02 Level 1. Single Statistical Aspect

Eight students showed a Single Statistical Aspect, either correspondence or variation, in an attempt to show covariation. Examples (R6-R9) are shown in Figure 4.30. One student drew a bivariate data point as a Single Comparison (Category 1A), presented in a rudimentary table of data (R6). Seven students represented a Single Variable (Category 1B): five responses showed six data cases ordered by values of the unlabelled single variable (e.g., R7) and two showed test scores without indication of study times (e.g., R8, R9).

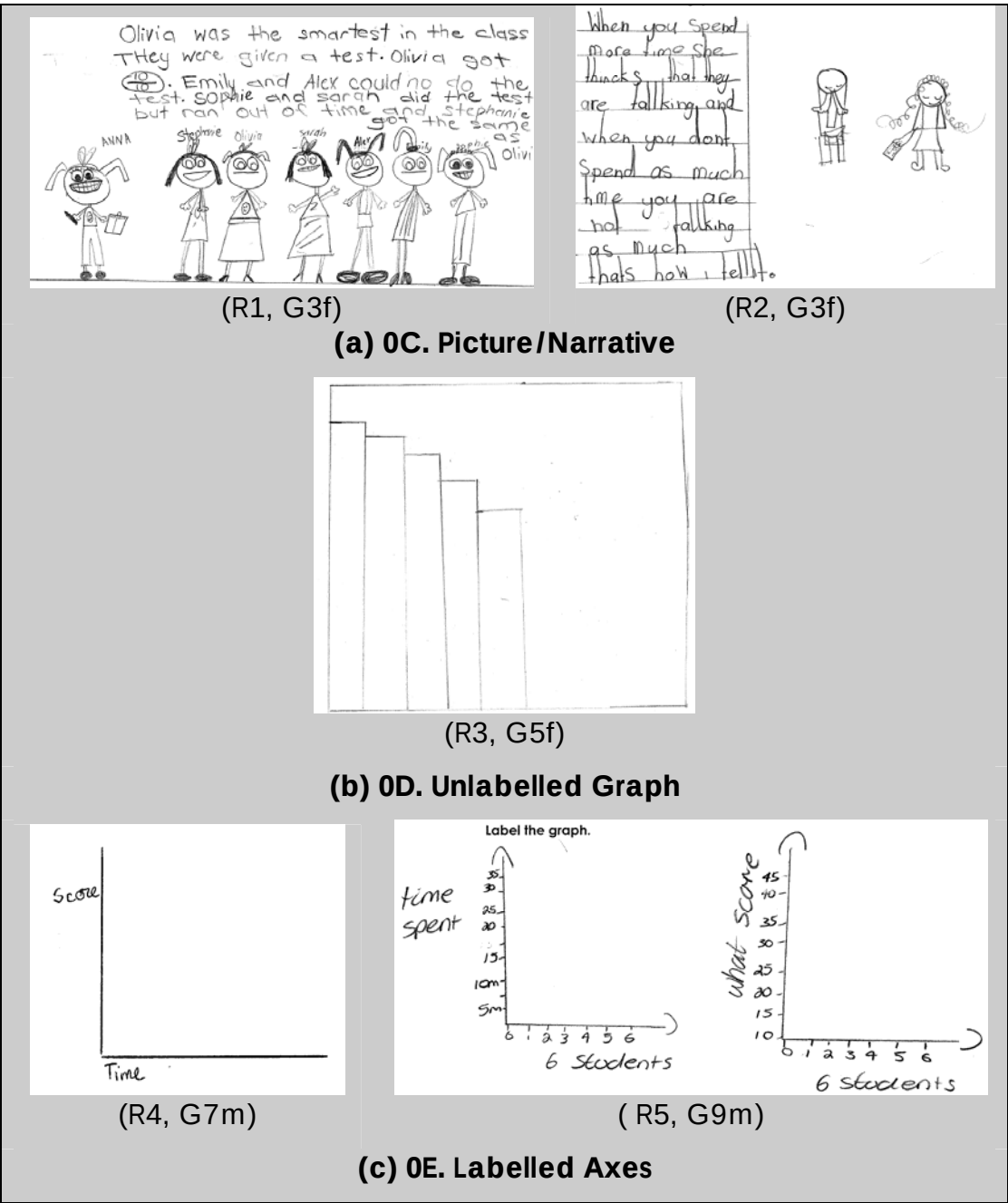
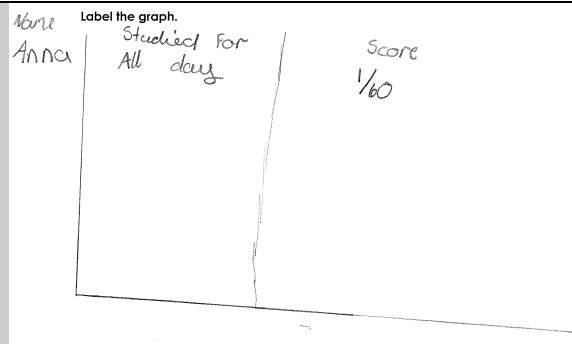
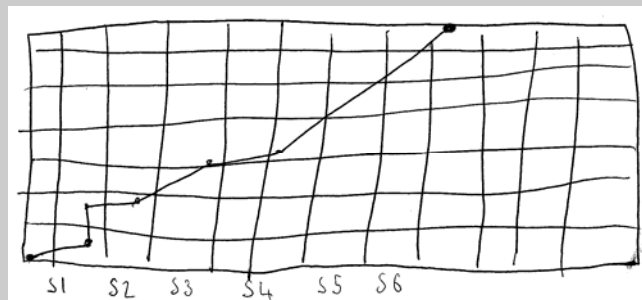


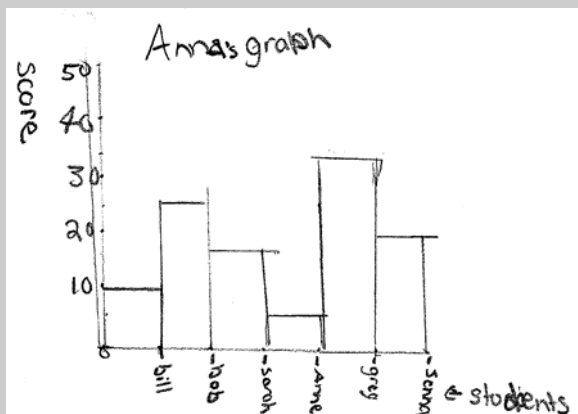
Figure 4.29. Student responses – Level 0 – Nonstatistical.



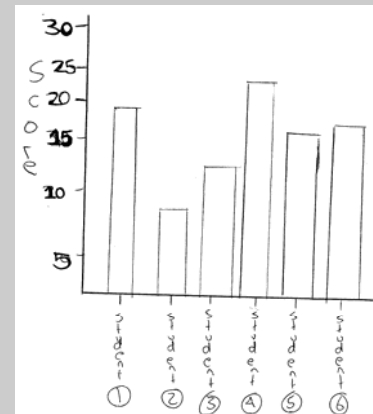
(R6, G5f)

(a) 1A. Single Comparison

(R7, G7m)



(R8, G7f)



(R9, G9m)

(b) 1B. Single Variable

Figure 4.30. Student responses – Level 1 – Single Statistical Aspect.

4.08.05.03 Level 2. Inadequate Covariation

Thirty-four students showed some features of the required negative covariation but lacked either appropriate variation or appropriate correspondence. Fifteen students presented study time as a binary variable in the form of a Double or Group Comparison (Category 2A). Examples (R10-R14) are shown in Figure 4.31.

Five students drew a Double Comparison involving two bivariate pairs. One third-grade female did so in a picture clearly showing two bivariate values (R10), whereas others did so in more conventional graph forms (e.g., R11, R12). Ten students represented a Group Comparison including test scores of six students for only two levels of studying (e.g., R13, R14).

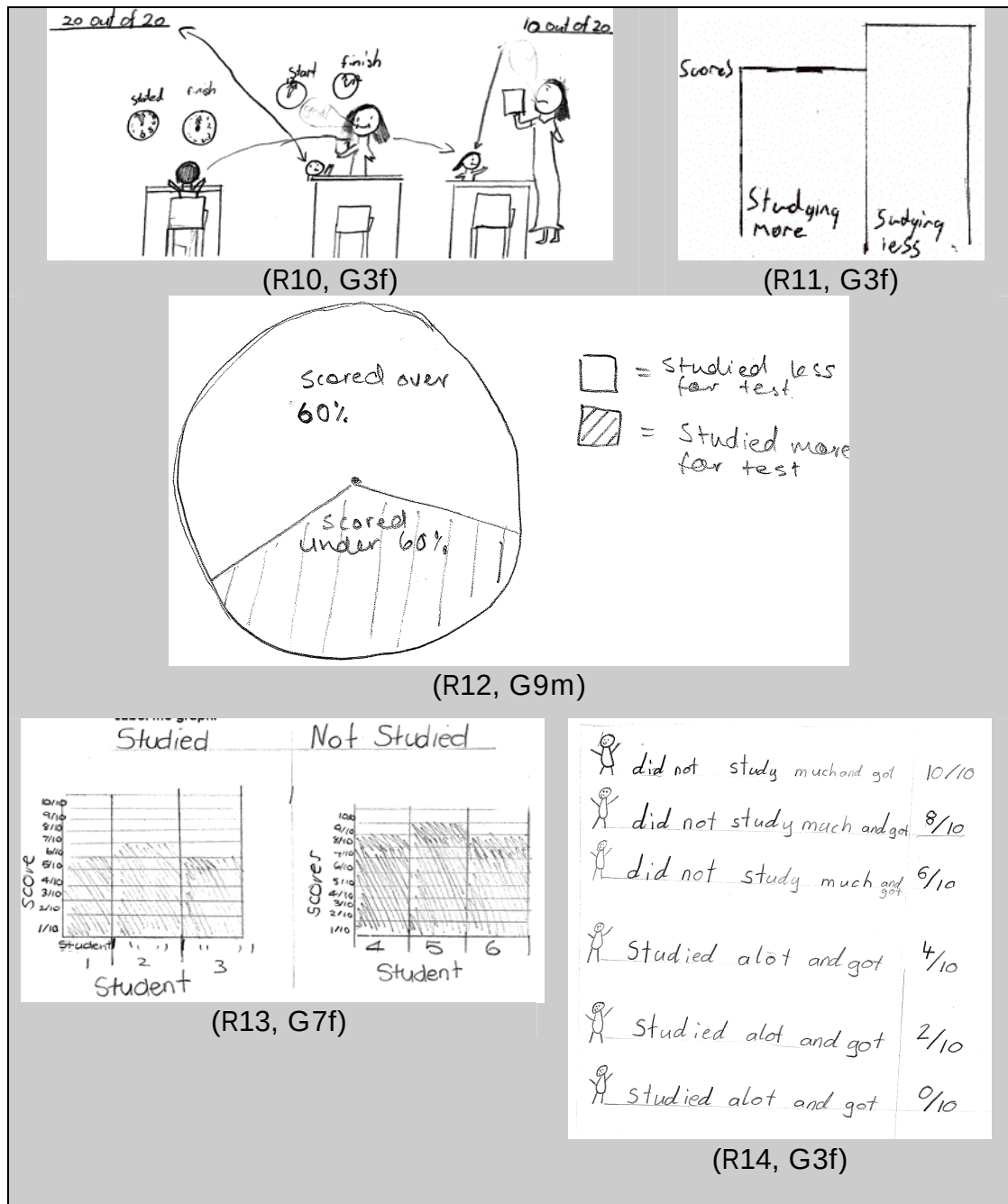


Figure 4.31. Student responses – Category 2A. Double and Group Comparison.

Three students showed a Trend Graph (Category 2B) with some explicit numbers but without labels or units to indicate the variables. A fifth-grade female drew a pie graph with larger sectors corresponding to labels of smaller percentage values (R15, Figure 4.32); it is possible that the percentage values refer to scores and the sectors refer to time intervals, however even with the principle of indicative labelling, the “%” symbol was not considered to indicate scores unambiguously.

Sixteen students drew Double Variable (Category 2C) responses that did not adequately show the direction of covariation. Nine students failed to indicate clearly any covariation (e.g., R16, R17, Figure 4.32), and seven represented a positive covariation, that is, in the wrong direction to that specified in the question (e.g., R18, Figure 4.32).

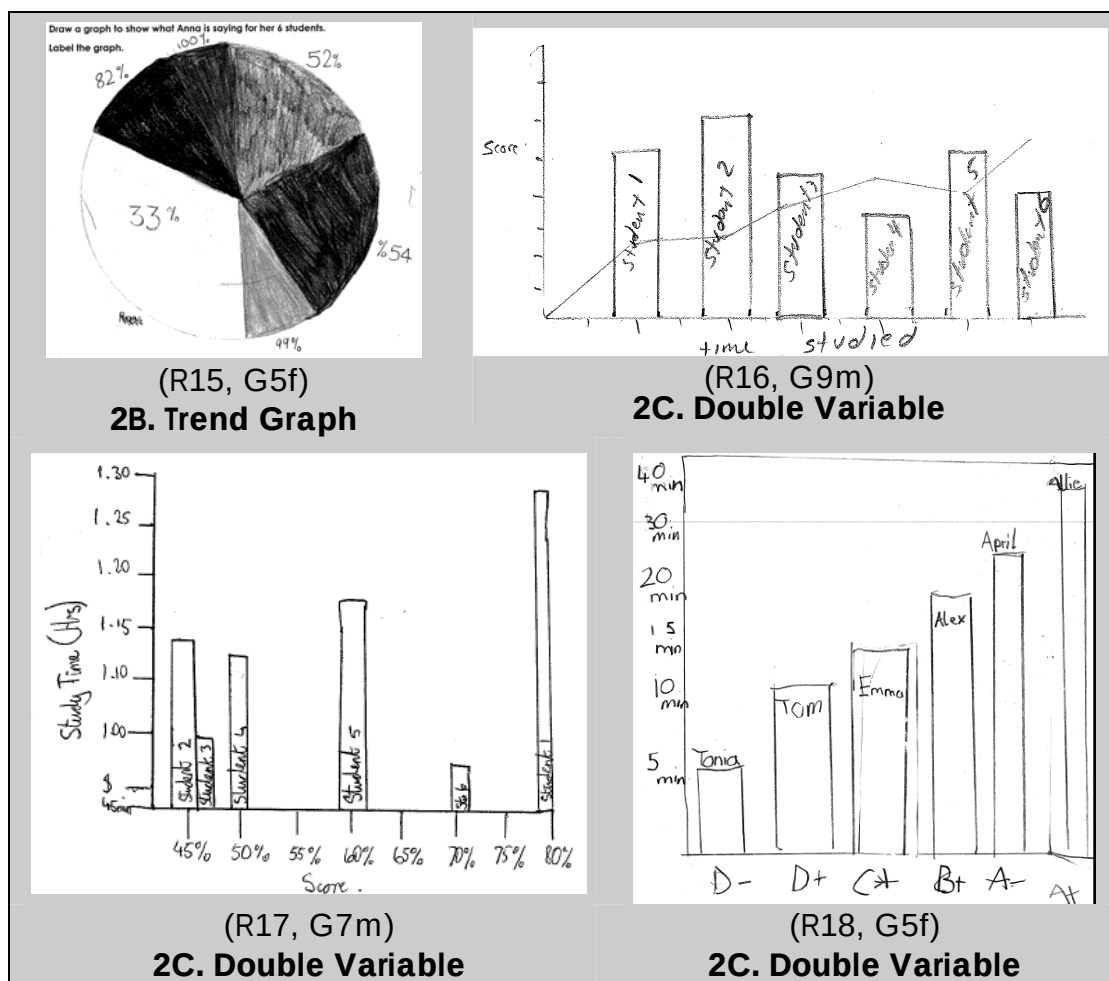
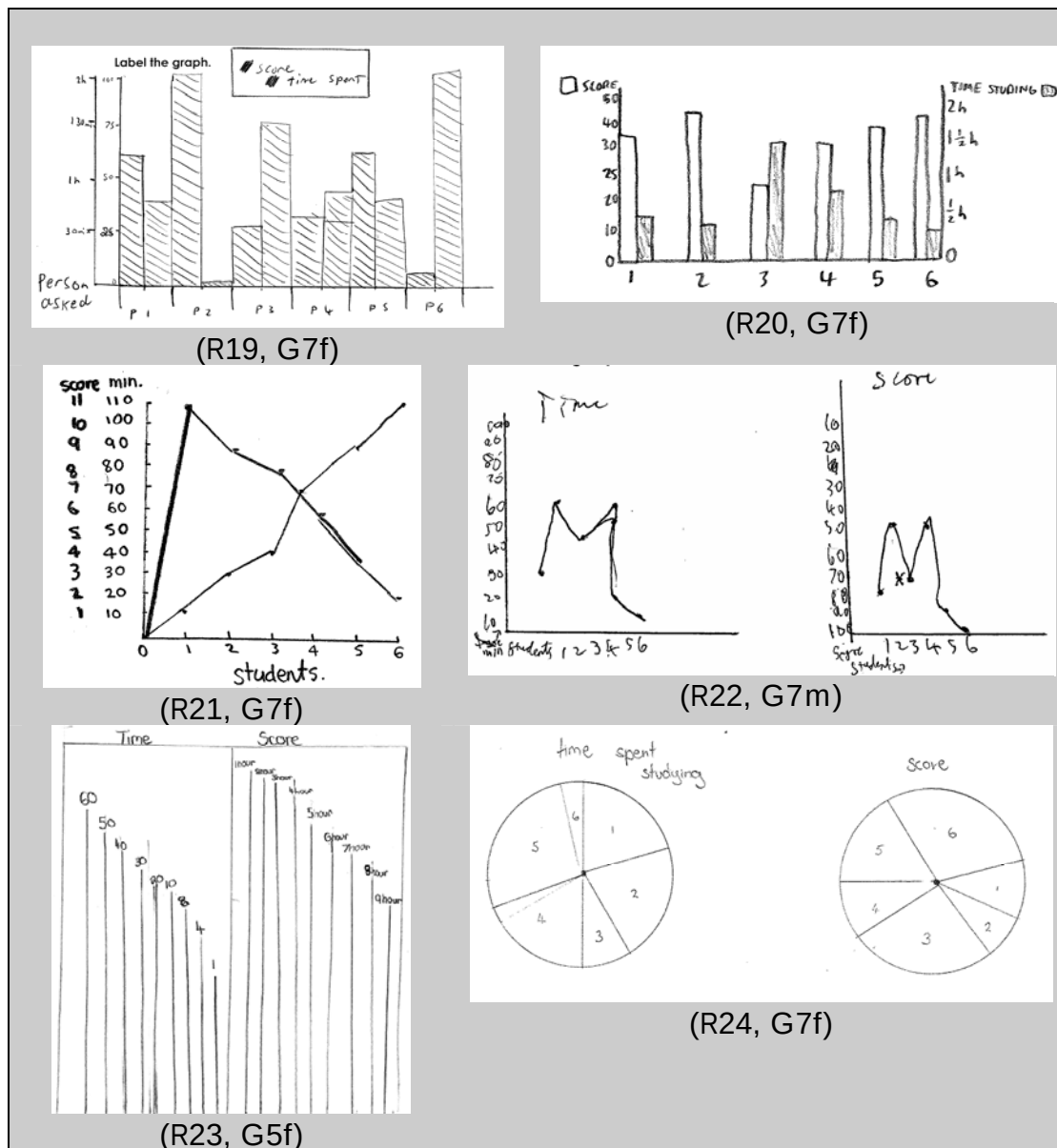


Figure 4.32. Student responses – Level 2 – Inadequate Coordinate.

4.08.05.04 Level 3. Appropriate Covariation

Seventy-four responses provided data for study times for which higher values were associated with lower test scores, with the conditions that at least three bivariate data points were shown and study time was not a binary variable. Eleven students drew Series Comparison (Category 3A) graphs for which one axis (or graph feature) represented the six students that Anna asked, and the other axis (or graph feature) displayed study times and test scores. The covariation was displayed either in two graphs or superimposed in one graph, often with two scales. Seven of these were bar graphs, three line graphs, and one was a double pie graph; seven graphs were unordered on the horizontal axis, and four were ordered on one variable. Examples (R19-R24) are shown in Figure 4.33. Three examples by seventh-grade females show a single graph with two data series (R19, R20, R21), with numbering of the six data cases. R20 illustrates an unordered horizontal axis, although after the first two cases, the student appears to order the remaining cases. With the use of order, R21 supports a literal reading, by moving from person 1 to 6, as the time goes up, the score goes down. The three other examples (R22, R23, R24) in Figure 4.33 show use of two graphs in various forms. R22 illustrates the power of the Series Comparison graph form: the M shaped graphs initially appear to show a positive association, however closer inspection shows that scores are labelled on the vertical axis in the reverse direction to convention, thus resulting in the negative covariation expected for the task. R23 displayed more than six data cases, however the labels and data values showed clear representation of the data expected of the context. The seventh-grade female who drew R24 displayed six numbered data cases for each of time and score in pie graphs in which sector size was used to denote value, and larger sectors for time corresponded with smaller sector for scores.



people, the other axis used to show one variable, and the second variable was drawn within the bar as a written value. The use of the written values were the basis for categorising as Tables (Category 3B), indicated by the lack of a numerical scale expected of Coordinate Variable responses (Category 3C): in R27, the vertical dimension was only used to correspond the data table to case labels (Person 1 to 6), and in R28, the horizontal axis was ordered by case label and not time values, which were written in the bars.

Fifty students represented Coordinate Variable (Category 3C) with the variables on opposing axes. In some cases axes were unlabeled, but units made clear the variable measured on at least one axis. Thirty represented study time on the horizontal axis and scores on the vertical, whereas 20 interchanged the axes. Forty students used conventional ordering of values on the axes, that is, increasing value as one moves up or right; seven reversed the values on one axis (giving the visual impression of a positive covariation); and three showed values unordered in bar graphs (giving the visual impression of no covariation). Thirty-one responses appeared to indicate a perfect linear fit with values of equal spacing on each variable, and the other 19 showed some variation from a perfect linear fit. Students differed in the form of graph used: bar graphs (25), scattergraphs (7), line graphs (5), and line graphs of connected dots (13).

Examples (R29-R34) are shown in Figure 4.35. The first four examples (R29, R30, R31, R32) show dotted lines graphs, in some cases with fewer (R29) or more (R30, R31, R32) than six data cases suggesting the line was the aspect of the graph showing the covariation. Two other examples in Figure 4.35 show scattergraphs with case labels (R33), in some cases using meaningful units that produced scales that appeared reversed (R34).

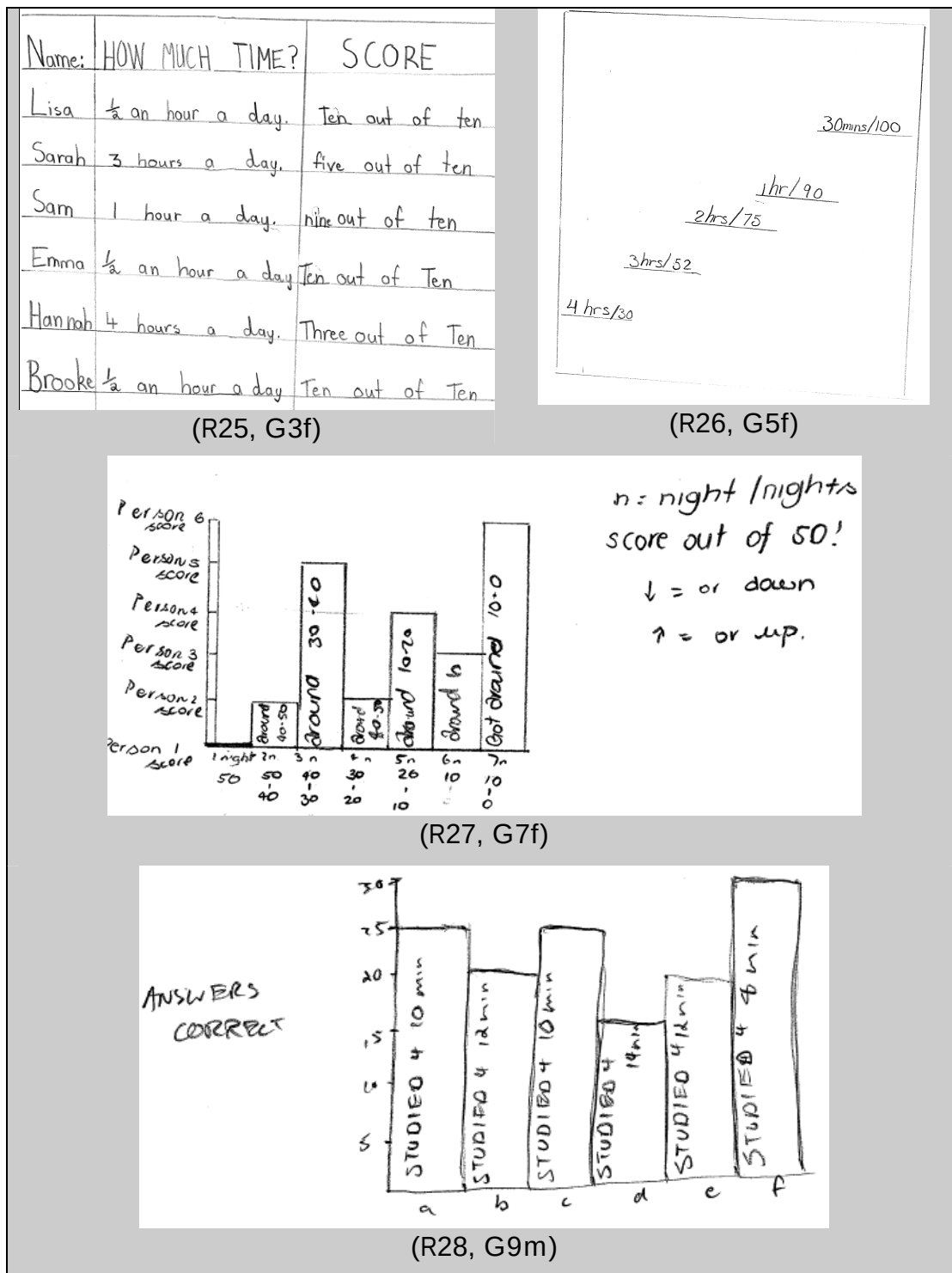
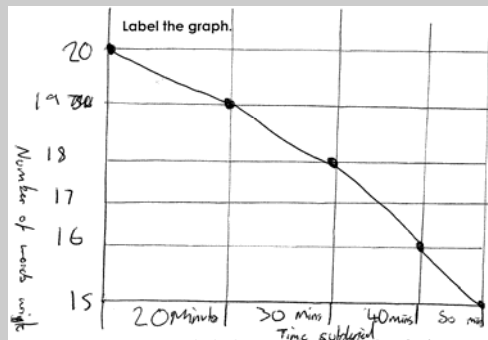
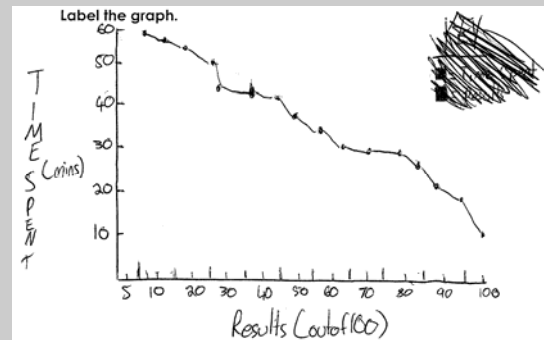


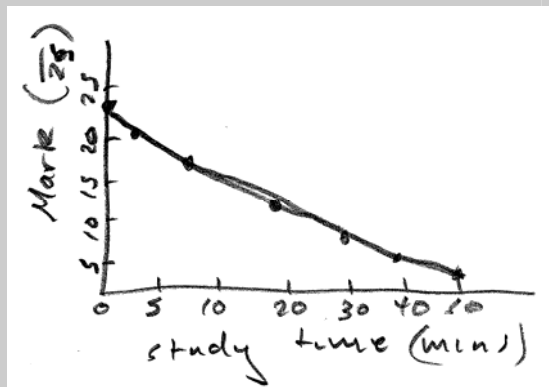
Figure 4.34. Student responses – Category 3B. Table.



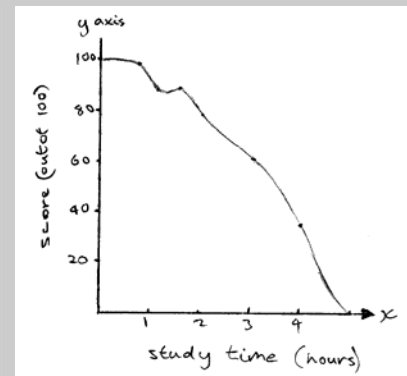
(R29, G7m)



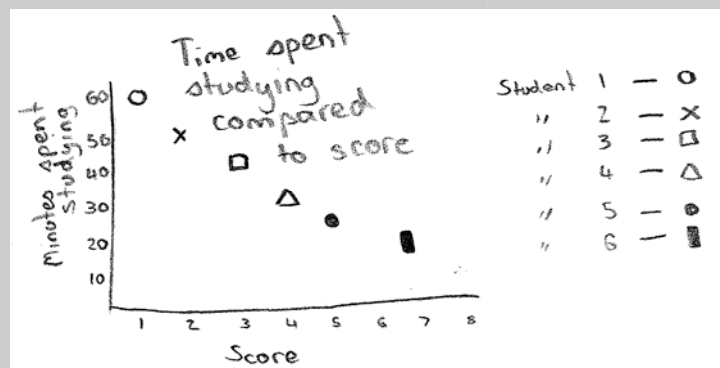
(R30, G7m)



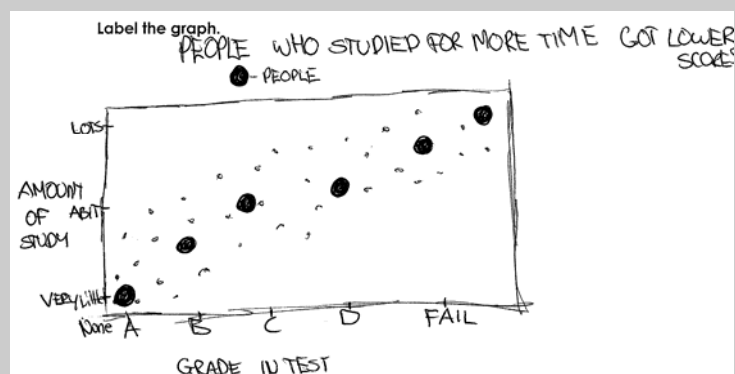
(R31, G9m)



(R32, G9m)



(R33, G9m)



(R34, G9m)

Figure 4.35. Student responses – Category 3C. Coordinate Variable.

4.08.06 Results – Positive Covariation (Q1*)

A positive covariation is consistent with prior beliefs. Table 4.09 showed that for the third- and fifth-grade males administered the task in this form, most responses were either at Level 1 or Level 3. Three examples (R35-R37) of Single Variable (Category 1B) graphs are shown in Figure 4.36. One third-grade male (R35) showed scores with units of points “e.g. 30 p” with six tally marks along the scale. Two other students (R36, R37) used bars with labelled names to denote a Single Variable (Category 1B). R37 was considered to show a single variable of study time, although if the student had indicated that position on the horizontal axis denoted score, the response would have been coded at Level 3.

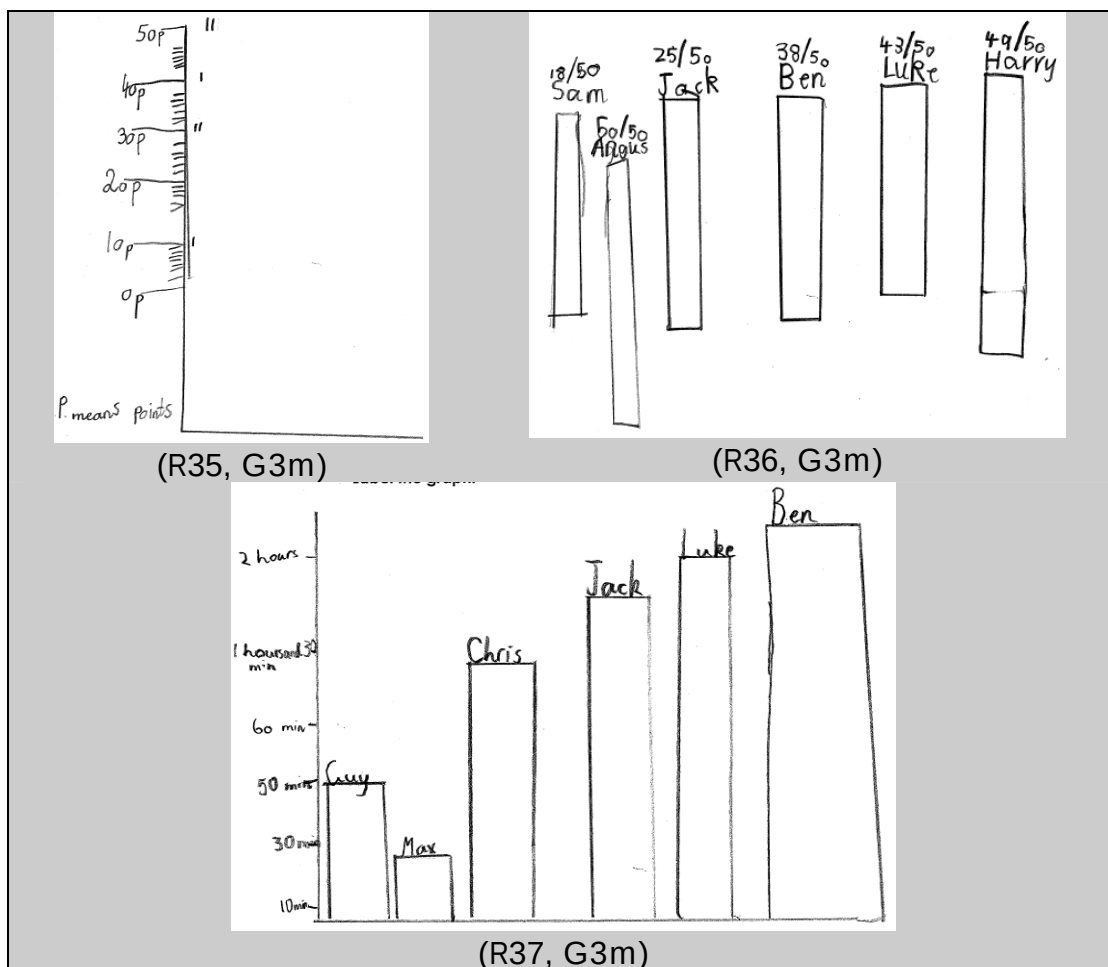


Figure 4.36. Student responses – Level 1 – Single Statistical Aspect.

Fewer responses were offered at Level 2, in part because the covariation was positive and because it was consistent with intuition, thus students did not confuse the covariation direction. An example (R38) of a Group Comparison (Category 2A) is shown in Figure 4.37. The third-grade male showed test scores as bar heights for 6 people who studied for 25 minutes, and for the same six names who studied for 50 minutes achieving higher test scores. The use of case labels in this way indicated a paired-sample Group Comparison (Category 2A), which preserved the score relations among students: those who scored more than others with 25 minutes study also scored more than others with 50 minutes study.

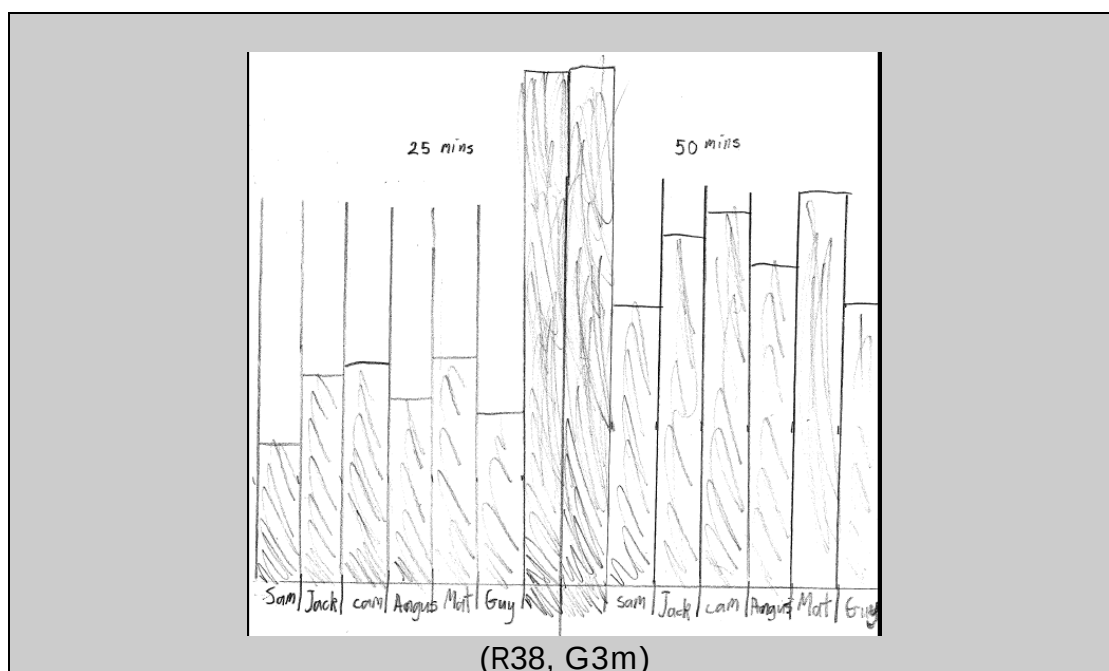


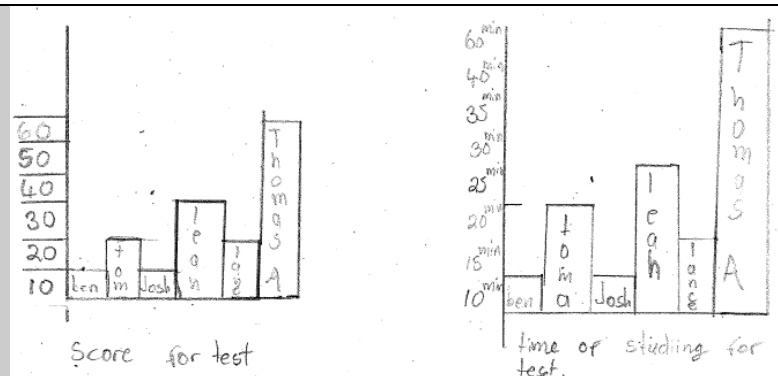
Figure 4.37. Student responses – Level 2 – Inadequate Covariation.

Responses at Level 3 showed similar response categories as observed in the negative covariation task format. Figure 4.38 shows examples of Series Comparisons (Category 3A) (R39), Tables (Category 3B) (R40, R41), and Coordinate Variable (Category 3C) (R42, R43). The use of case labels in Series Comparison graphs (Category 3A) was evident in R39. Pictures of people (R40) and clocks (R41) were used to create cases upon which to attach bivariate values, without the use of formal

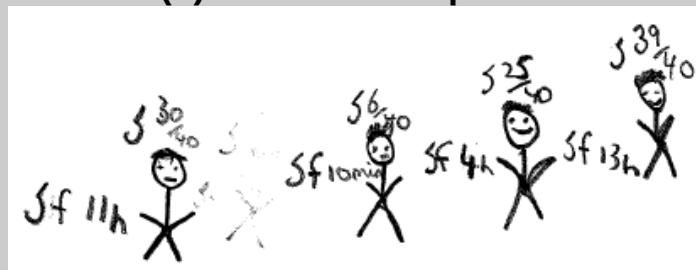
tabular rows and columns. Coordinate Variable (Category 3C) responses did not always employ graph conventions, as shown by R42 and R43. In R42, the student showed identical values for times and score (40, 38, 36, 32, 30, 28, 26, 24) with times on the vertical axis, scores on the horizontal axis in descending order left-to-right, and student names labelled on both axes to support the reading of corresponding values. In R43, the representation was a hybrid between a bar graph and a table: the horizontal axis values are evident and the vertical axis values are ordered and positioned not along the axis but repeated in the bars: if dropped to the bottom, this would form a simple two-row table.

4.08.07 Results – No Covariation (Q2)

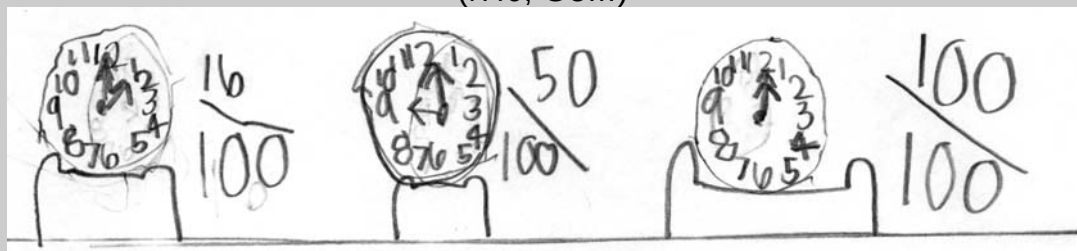
Of the 27 ninth-grade males who were administered the no covariation format task (Q2), as shown in Figure 4.27, 22 offered Level 3 responses, 3 at Level 2, and 1 at Level 1, and 1 no response. Examples (R44-R47) are shown in Figure 4.39. The Single Variable (Category 1B) response (R44) was similar to those observed for Q1. An example of a Group Comparison (Category 2A) (R45) involved study times of “short” and “long,” and a low, medium and high score for each. Two examples shown of Coordinate Variable (Category 3C) (R46, R47) indicate some students drew scattergraphs (R46), some line graphs (R47), and others were bar graphs. Of the 22 responses at Level 3, all showed variation in both measures without any correspondence of the variation; none showed a constant value across variations in another measure.



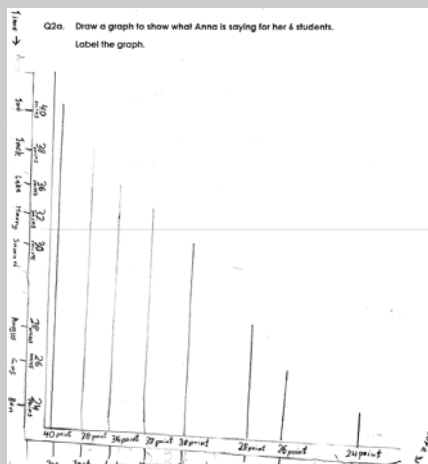
(R39, G5m)

(a) 3A. Series Comparison

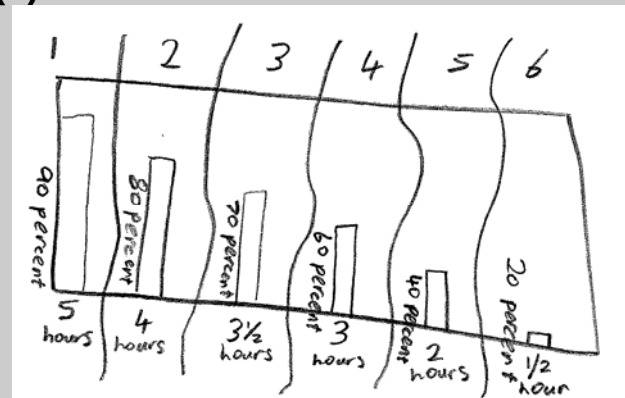
(R40, G5m)



(R41, G3m)

(b) 3B. Table

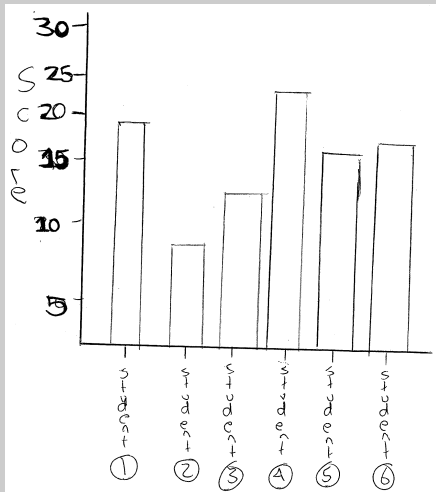
(R42, G3m)



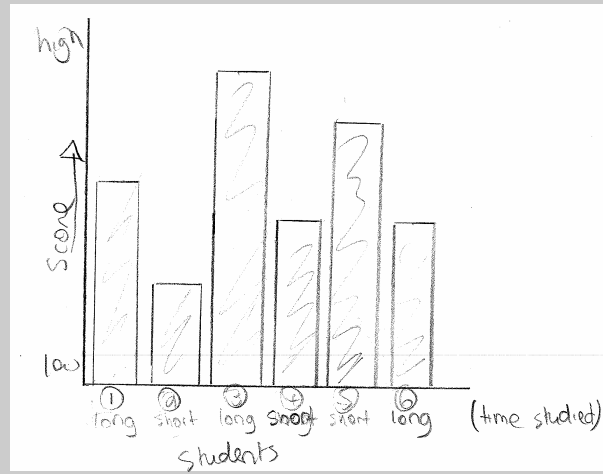
(R43, G5m)

(c) 3C. Coordinate Variable

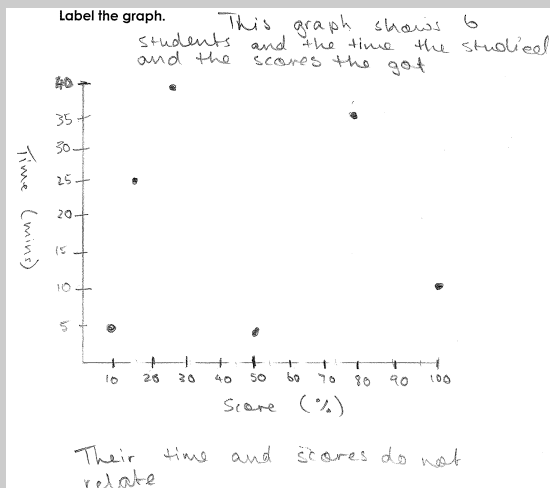
Figure 4.38. Student responses – Level 3 – Appropriate Covariation.



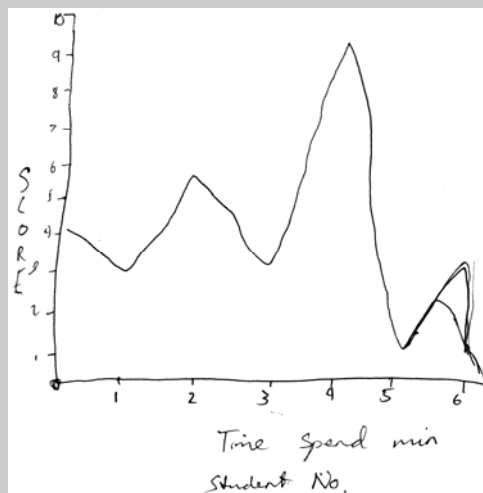
(R44, G9m)

1B. Single Variable

(R45, G9m)

2A. Group Comparison

(R46, G9m)

3C. Coordinate Variable

(R47, G9m)

3C. Coordinate Variable

Figure 4.39. Student responses to No Covariation task format (Q2).

4.08.08 Discussion of Investigation 4A

Responses to a task involving negative, counterintuitive covariation between familiar variables with six data cases were described according to the four levels of Speculative Data Generation used in Investigations 1, 2 and 3. The categories of responses observed were similar to those from Investigation 1.

Students responded at higher levels than Investigation 1 for comparable year levels, and at similar levels to Investigation 3A for the same students. Cross-comparisons are discussed in more detail in Chapter 7. It is noted, however, that compared to Investigation 1, in this investigation there were few Single Comparison responses, and a higher proportion of responses showing Coordinate Variable, and showing Tables. This result occurred despite a negative, counterintuitive direction of covariation for the task in this investigation. One feature of the task which may have contributed to these higher levels of response was the requirement to produce six data cases. Many responses showed consideration of the covariation case-by-case, either in Tables or in other responses including case labels. Even some Single Variable responses, which showed no correspondence between two measures for each data case, did show correspondence between one measure and case labels.

4.09 INVESTIGATION 4B: TEST SCORES VERSUS STUDY TIME (INTERVIEWS)

4.09.01 Introduction and Aims

Investigation 4B involved an interview protocol that asked students to explain their own graphs for Task 4 shown Figure 4.27, and then to compare their responses with others, some of which were higher level, lower level, or simply alternative to their own responses. The investigation aimed to explore thought processes of students when verbally explaining their reasoning, and students' potential to develop reasoning via dealing with cognitive conflict, aligned with Research Aim 3 of the wider study (Section 3.02). In particular, it was explored whether students' reactions to prompts at various levels related to the levels of the students' survey responses.

4.09.02 Task

The interview protocol followed the interview protocol format outlined in Chapter 3. Students were shown their own graph, and asked to explain what they had drawn and why they responded in that way. They were then shown a sequence of prompts at a variety of levels, as seen in Figure 4.40 and Figure 4.41. Students were asked to comment on each prompt graph, and to decide which response was preferred to show the required information appropriately and clearly. For the negative covariation format prompts (Figure 4.40),

- Prompt A involved a Double Comparison (Category 2A),
- Prompt B, a Group Comparison (Category 2A),
- Prompt C, a tabular Double Variable (Category 2C) response with no covariation,

- Prompts D and F, Series Comparisons (Category 3A), and
- Prompt E, a Coordinate Variable response (Category 3C).

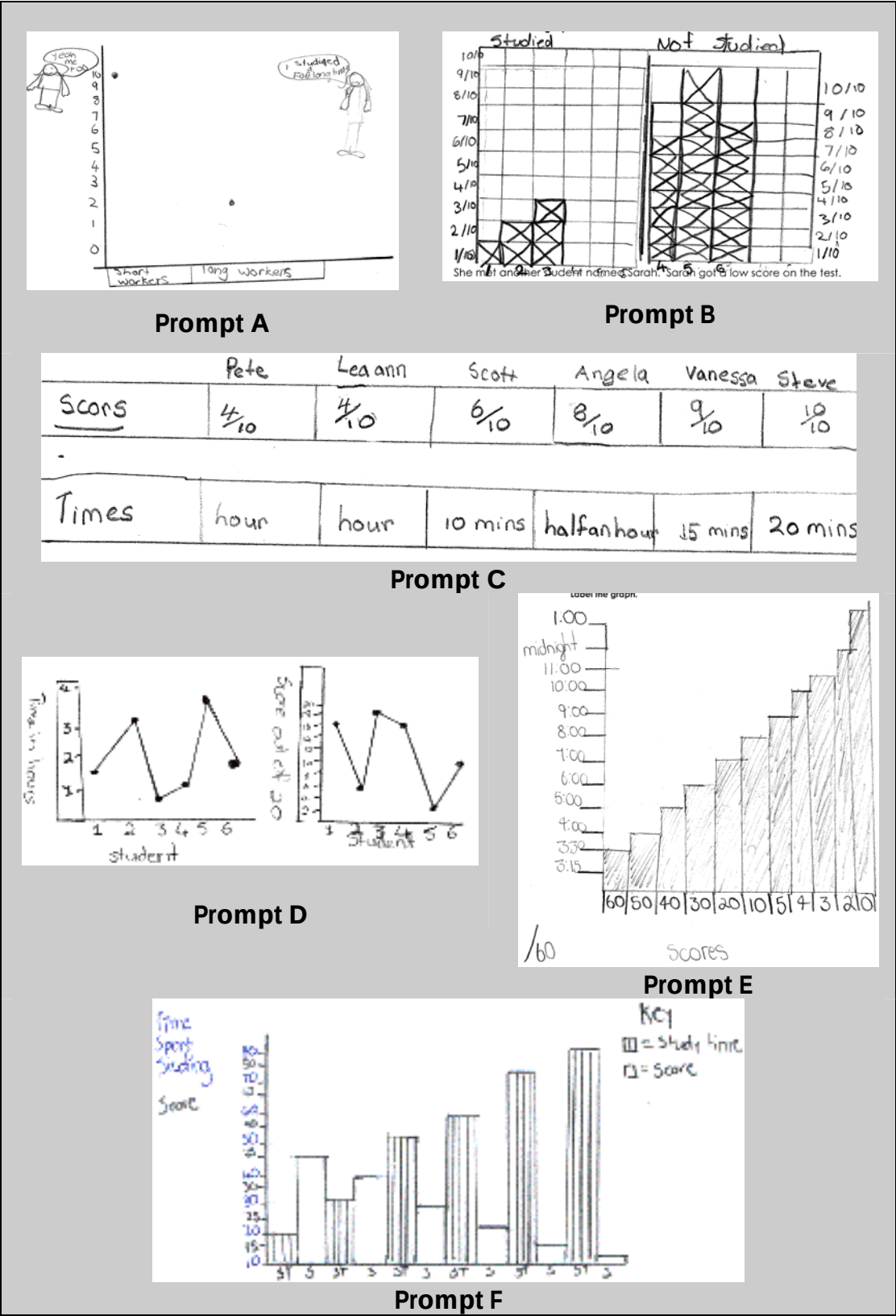


Figure 4.40. Graphs used as prompts (Negative covariation).

For the positive covariation format prompts (Figure 4.41),

- Prompt A involved a Single Variable response (Category 1B),
- Prompt B an Coordinate Variable response (Category 3C) with unordered scales,
- Prompt C an Coordinate Variable response (Category 3C) with a reversed scale, and
- Prompt D a Series Comparison response (Category 3A).

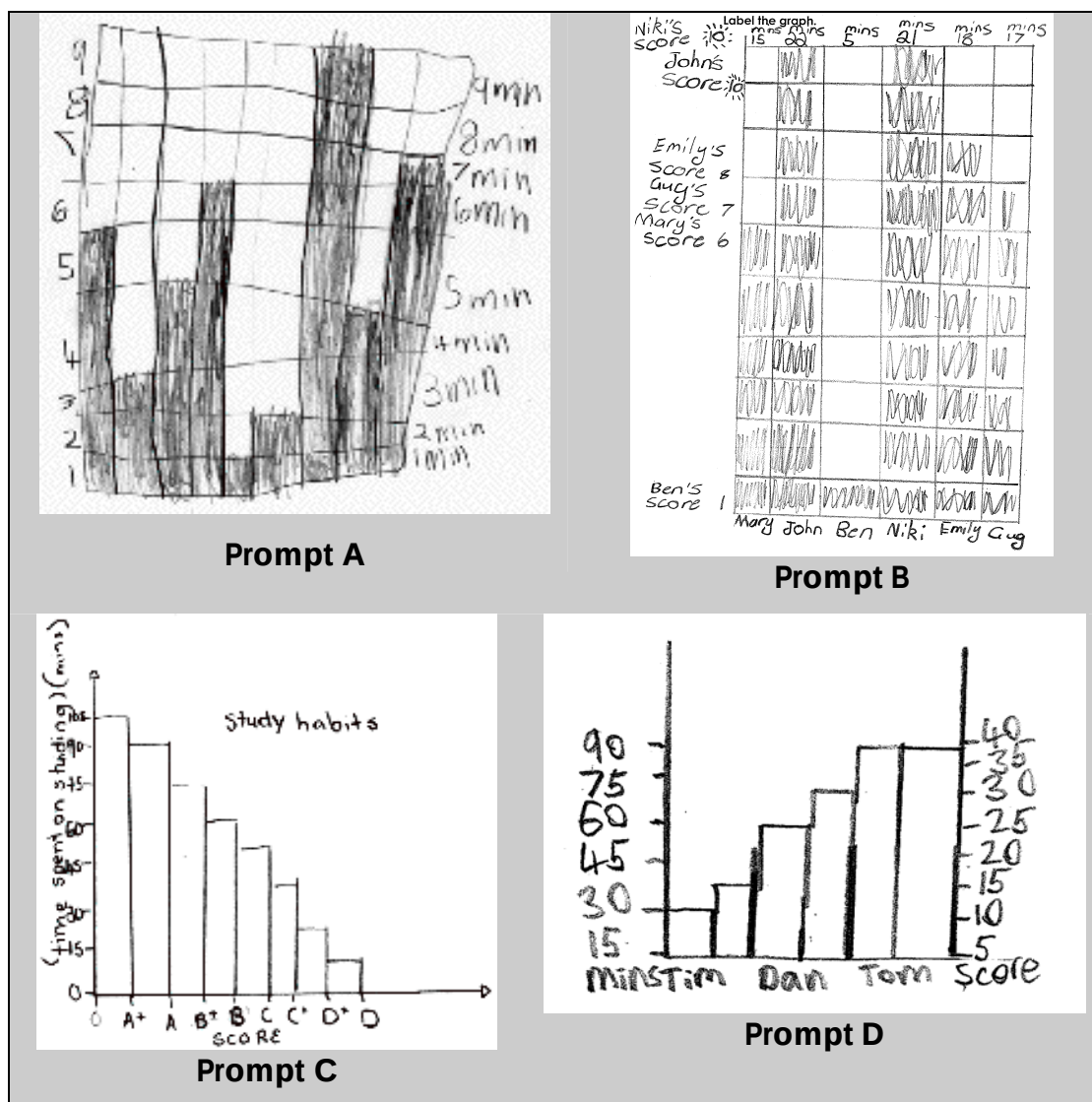


Figure 4.41. Graphs used as prompts (Positive covariation).

4.09.03 Participants and Method

Participants were a subset of 33 of the 34 students interviewed as described for Data Collection 3. Students were asked to explain their own graphs, then shown prompts relevant to the task they were administered: for most students, the negative covariation form and prompts in Figure 4.40, and for third- and fifth-grade males, the positive covariation form and prompts in Figure 4.41. Interview dialogue was analysed in relation to general agreement (✓), disagreement (✗) or ambivalence (~) to each of the prompts. Analysis explored these reactions by initial response levels, and for emergent themes in dialogue.

4.09.04 Results – Overview

The results of the 33 interviews are shown in Table 4.10 (27 interviews for negative covariation) and 4.11 (6 interviews for positive covariation task format). For the negative covariation format, most students disagreed with most of the six prompts, with at most 7 out of 27 students agreeing to any prompt, as shown near the bottom of Table 4.10. Of the 3 students who initially responded at Level 0, 2 students agreed to prompts at Level 2. Of the 2 students who initially responded at Level 1, 1 student agreed to a prompt at Level 3. By contrast, of the 11 students who initially responded at Levels 2, only 2 students agreed to higher level prompts (Level 3). In general, across both Tables 4.10 and 4.11, of students who initially responded at Levels 2 or 3, most agreed to prompts at the same level, or did not change their preference. Thus despite the general disagreement to prompts, there was weak evidence that students reacted with more agreement to responses closer to the level of their own response.

Common themes in dialogue included general comments about graphs being confusing or hard to understand. Units or values attracted comment, particularly for Prompts A and E (Figure 4.40). Prompt C (Figure 4.40) was identified by some as having inappropriate values, demonstrated by some students graphing the values.

Table 4.10.

Summary Descriptions of Interview Dialogue (Negative Covariation format)

ID	Initial Response Category	Prompt – Category of Prompt						Summary of Reactions to Prompts
		A – 2A	B – 2A	C – 2C	D – 3A	E – 3A	F – 3C	
G9f3	No attempt	✗	✗	~	✗	✓	~	Unknown initial level
G9f4	No attempt	✓	✗	~	~	~	N	Unknown initial level
G9f5	No attempt	✗	✗	✗	N	✗	N	No change
G5f3	0C. Picture	~	✗	✗	N	✗	~	No change
G3f4	0D. Unlabelled	✓	✓	✗	✗	✗	✗	Agree Higher (0D→2A)
G5f6	0D. Unlabelled	✓	✓	✓	N	~	✗	Agree Higher (0D→2AC)
G7f2	1B. Single Variable	✗	~	~	N	~	✗	No change
G7f4	1B. Single Variable	✗	✓	~	✗	✓	✗	Agree Higher (1B→3A)
G7m1	2A. Double Comp.	✗	✗	✓	N	✗	N	Agree Same (2A→2C)
G7m3	2A. Double Comp.	✗	~	✓	✗	~	~	Agree Same (2A→2C)
G5f4	2A. Group Comp.*	✗	✓	✓	✗	~	✗	Agree Same (2A→2C)
G5f5	2A. Group Comp.	~	✓	✗	✗	~	~	Agree Same (2A→2A)
G9m3	2A. Group Comp.	✗	✗	✗	~	✓	✗	Agree Higher (2A→3A)
G3f2	2C. Double Var.	✗	✗	✗	N	✗	✗	No change
G3f3	2C. Double Var.	✗	✗	✗	✗	✓	✗	Agree Higher (2C→3A)
G5f1	2C. Double Var.	✓	✗	✓	N	~	✗	Agree Same (2C→2C)
G5f2	2C. Double Var.	✗	✗	~	N	✗	✗	No change
G7m2	2C. Double Var.	✗	✗	~	✗	~	~	No change
G9f2	2C. Double Var.	✗	~	~	✗	✗	~	No change
G7f6	3A. Series Comp.	✗	✗	✓	N	✗	✓	Agree Same (3A→3C)
G3f1	3B. Table	✗	✗	~	✗	✓	✗	Agree Same (3B→3A)
G3f5	3B. Table	✗	✗	~	✗	~	✗	No change
G7f1	3C. Coord. Var.	✗	~	✗	✗	~	~	No change
G7f3	3C. Coord. Var.	✗	✗	✓	✗	✗	✓	Agree Same (3C→3C)
G7f5	3C. Coord. Var.	✓	✗	✗	✗	✓	✗	Agree Same (3C→3A)
G9m1	3C. Coord. Var.	~	~	~	~	~	~	No change
G9m2	3C. Coord. Var.	✗	✗	✗	✗	✓	~	Agree Same (3C→3A)
Total	✗	19	17	10	15	9	13	
Total	~	3	5	10	3	11	9	
Total	✓	5	5	7	0	7	2	

Note: ✓ denotes agreement, ✗ denotes disagreement, ~ denotes ambivalence, N denotes not administered, * denotes drawn in interview

Table 4.11.

Summary Descriptions of Interview Dialogue (Positive Covariation format)

ID	Initial Response Category	Prompt – Category of Prompt				Cognitive Conflict Summary
		A – 1B	B – 3C	C – 3C	D – 3A	
G3m2 2A Group Comp.		✓	✓	✓	~	Agree Higher (2A→3C)
G5m1 3A Series Comp.		✗	✗	~	✗	No change
G3m3 3B Table		~	~	✗	✗	No change
G3m1 3C Coordinate Var.		✗	✗	✗	✗	No change
G5m2 3C Coordinate Var.		~	✓	~	~	Agree Same (3C→3C)
G5m3 3C Coordinate Var.		✗	✗	~	~	No change
Total ABCD ✗		3	3	2	3	
Total ABCD ~		2	1	3	3	
Total ABCD ✓		1	2	1	0	

4.09.05 Results – Interview Dialogue

The following extracts of interview dialogue illustrated common themes.

Each interviewee is uniquely identified corresponding to entries in Table 4.10.

One fifth-grade female (G5f3) in Table 4.10 drew a picture of a book with a single value, “2/100, very bad!” She explained her difficulties in responding to the task, found Prompts A and B confusing, and commented on specific values from other prompts.

- G5f3: [Own graph] I think with this one I had a little bit of trouble so I ended up putting something that I didn't think was very good which is that. [...] I haven't even put the time.
- G5f3: [Prompt A] Yes I don't think that's very good because they sort of... it's hard to understand. It's good but I just think it could have been like labelled 10 minutes or 10 hours or 10 days.
- G5f3: [Prompt B] That's hard to understand. I don't know what this bottom part means. I think it would be... because it shows like one day, it's got one out of 10, two, three, six, nine, it's gone back to seven out of 10 or 6 hours or whatever.
- G5f3: [Prompt C] This is good but to chose two [data cases]: 10 minutes, it says 10 minutes and they got 6 out of 10, and it's got 20 minutes and 10 out of 10 and it's supposed to be the people who studied for more time got lower scores.

G5f3: [Prompt E] This is good. It doesn't show... What does it say, "midnight"... but does that mean the hours?

G5f3: [Prompt F] What it does S and ST stand for? [...] Yes it's a bit puzzling that it goes up down the score I think the study times should have been that there somewhere.

One third-grade female (G3f3) in Table 4.10 drew a Double Variable (Category 2C) as a table of six values, as shown in Figure 4.42. Four values showed exactly matching values for times and scores (9 minutes, 9/10; 8 minutes, 8/10; 7 minutes, 7/10; 5 minutes, 5/10), thus a positive covariation, with two exceptional data points (6 minutes, 1/10; 1 minute, 10/10). Her explanation was slightly confused, and she disagreed with most prompts, including Prompt C, despite being remarkably like her own, which she claimed was in the wrong direction. She approved only of Prompt E, but for the wrong reason, as she referred to the single extreme point, "1:00" and low score, as evidence of a positive covariation, not noticing that the time scale on Prompt E indicated that 1:00 was the clock time representing the greatest study time.

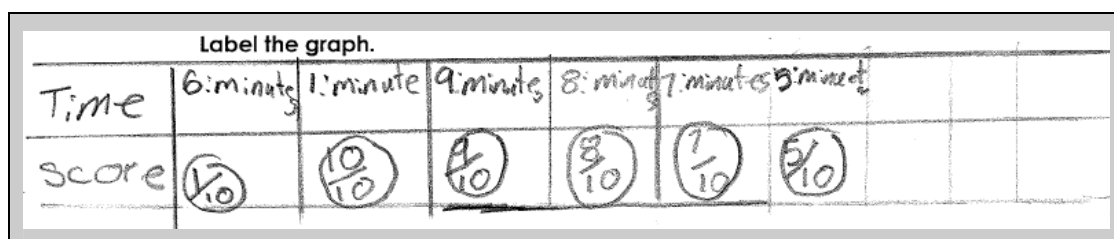


Figure 4.42. Student response – Double Variable (G3f3).

G3f3: [Own graph] Well if you got like 6 minutes you only got 1/10, that is a high time studying. And 1 minute, because that is done in fast time, I get a 10 out of 10.

G3f3: [Prompt C] It is a good one, but it doesn't really show it, just says that, names and stuff and how long they studied for, and if they had a high time they would get high score and it said if he studied for a long time you would get lower score.

G3f3: [Prompt E] That is a good one [...] What it says like 1:00 then you've got a lower score [points to zero on horizontal axis] that is the time that is the score.

With the counterintuitive direction of the covariation, some students commented this was surprising, and others seemed to have drawn graphs based on their own views, irrespective of the task to represent the stated verbal statement. One seventh-grade male (G7m2) referred to different reasons for drawing different data points which resulted in his Double Variable graph of no association (Level 2, see R17, Figure 4.32).

G7m2: Well I'm using Student 1 as me, because I'm not 100 percent accurate spelling. So when there's a spelling test, I usually study really hard for it... maths, biology and science, chemistry, I study really hard for it. Because my mom makes me. For something like this, she would make the study flat out until I had done like heaps of exercises out of these books that we buy. Yes and I used 1 and 5 and 6, these three are the main ones. I just drew these for decoration but these are what I'm really focusing on. I'm using her theory with Number 6. That he didn't study for a long time but he got a high score. And I'm using Student 5 as her actual theory: he/she/it studied for a while and it got a lower score.

I: So do all of the students in there fit that theory, that people who studied for more time got lower scores?

G7m2: Yes and no. These, all of these do [...] that is just my theory.

I: And what would your theory be?

G7m2: My theory is that, well, it depends how you hard you try for the test.

Another seventh-grade male (G7m3) drew an unusual Double Comparison (Category 2B), as shown in Figure 4.43. The graph shows two data lines in different colours, indicating two conditions of study time, "Studied Longer" and "Studied Shorter." The student appears to have compared these two conditions on the left-hand side in the vertical dimension according to study time or memory, whereby "Studied Longer" appears higher on the graph. In moving from left to right, labelled "Forgetting Time," the two lines cross. At the right-hand side, the two conditions are compared vertically, implicitly on test scores, such that the condition of "Studied Longer" is subject to "Forgetting" (sic) and appears lower.

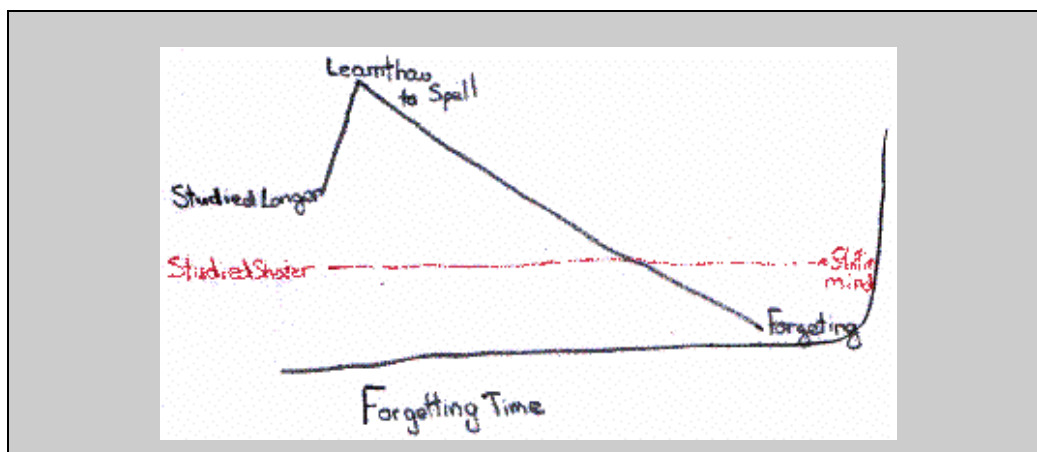


Figure 4.43. Student response – Double Comparison (G7m3).

G7m3: If they studied longer, they learned heaps and heaps. And then they start forgetting while they're learning more. They forget other things. And studied shorter [referring to line in graph], they just keep it in their mind, because they don't go over the top, more than their memory bank can hold, so they just stop and they will still have it.

The interviewer probed regarding the vertical axis to examine whether it represented test scores, given that forgetting time and two conditions of study time were shown, and the student still did not make explicit the scores, but implied them in relation to “learning lots and lots.”

I: So when this is going from here up to here [pointing to vertical axis], what is that telling us about?

G7m3: That they're getting, they're learning lots and lots and lots, and they're not overdoing it, but then until there [middle of graph, line crossing] they start to learn way too much [pointing to data series going down].

One ninth-grade female (G9f2) drew a Double Variable graph (Category 2C) with positive covariation, which showed the negative trend visually but with a reversed horizontal axis. This response is shown in Figure 4.41 as Prompt C, and because of the positive direction, was described as a Level 3 response if offered to the positive covariation task. She immediately recognized the problem with her graph being in the wrong direction. She acknowledged various prompts as showing the information, but generally with ambivalence, noting that had she reversed the

direction of her graph, then her graph would be clearer than the prompts shown. For Prompt D, the two graphs were considered to show two different things, that is lacking correspondence in pairs, and interview probing elicited the student understood the use of labels (“Student 1, 2, ..., 6”) to show the correspondence. For Prompt F she verbalised the corresponding values for both extreme cases shown, and then provided a verbal statement of covariation.

- G9f2: [Explaining own response] Umm... An A-plus, they studied for 105 minutes, and an A... I didn't read it properly, because it says the longer you studied, the lower you scored.
- I: What have you shown?
- G9f2: I've shown that the longer you study, the higher the score you get.
- G9f2: [Prompt B] Umm... It's a bit everywhere around the page, it's not very readable but it does show the information.
- G9f2: [Prompt C] It's not really set up in a graph, but it shows it. It just says the times and the scores, so it's not actually set up in a graph, it's just the information.
- I: Does the data in there match the sentence?
- G9f2: Yes.
- G9f2: [Prompt D] It just shows how many hours they spent on it, and what score they got. It's just two different graphs, showing two different things.
- I: Does it relate them at all and show that the people who studied more time got lower scores?
- G9f2: Umm... yes, because it has got Student 1, 2, 3, 4, 5 and 6. That's where the dots are, so...
- G9f2: [Prompt F] Well it just says that the person has studied for less time and got a bigger score. And this person studied for more time and got lower score. And the more you study, the less score you get, that's what they have shown.

The only student interviewed who successfully used a Series Comparison to show the negative covariation (G7f6, see Prompt D, Figure 4.39) explained case-by-case how her graph linked to the verbal statement. She liked Prompt F, also a Series Comparison response, and liked Prompt C, not recognising the data were not in a consistent covariation direction.

- G7f6: The time that they studied, and this person for an hour and a half and they got a pretty high score. This one studied for a long time, and got

low score. And someone studied for a small amount time and really high score.

G7f6: [Prompt A] Not quite as accurate. [...] Because of the ways it is presented.

G7f6: [Prompt B] It is not very clear.... [long pause] it is pretty accurate.

G7f6: [Prompt C] It is a lot clearer than the other two. And it shows what the girl is saying. [Yours or C?] This way [Prompt C].

G7f6: This one [Prompt C] is more clear than [Prompt E]. [...] This one has amount time that they studied [Prompt C] it is not just got the time that they were studying for [Prompt E].

G7f6: [Prompt F] This one is pretty accurate. And clear to read and stuff.

4.09.06 Discussion of Investigation 4B

When asked to explain their own graphs, some students identified weaknesses in these graphs, such as the wrong direction of covariation, and others explained they had intentionally graphed something different from the verbal statement of the task. The basis of these explanations was similar to findings of previous research concerning influence of prior beliefs overriding consideration of the data, and concerning selective attention to various data points or features of graphs. Compared with students' dialogue in Investigation 3B, these explanations in Investigation 4B more clearly revealed whether students' understandings of Speculative Data Generation for covariation were based on data or on prior beliefs. Contributing factors to this may have been (a) that during the 45-minute interviews, this task followed that of Investigation 3B and thus the students had gained more confidence, (b) that students felt the graph, including counterintuitive covariation, was less self-evident than height growth with age, and (c) that discrete data cases, as indicated in the task, supported conversation.

Students' reactions to prompts were frequently in disagreement, referring to confusing issues in prompts, such as values or units. There was weak evidence that students' initial response level was related to the prompts they found favourable. Of

the few students who initially responded at Level 0 or Level 1, some agreed with prompts of Double or Group Comparison at Level 2. Of the students who initially responded at Level 2 or Level 3, most of the agreement shown was with prompts at Level 3. A Double Variable (Level 2) prompt in tabular form, however, was found favourable by students who initially responded at various levels. Part of the appeal of this prompt may have been its tabular presentation which supported easy reading of data values and corresponding pairs, rather than confusing values, units, or graph structure. Indications of dialogue referring to specific cases of corresponding data values in support of the verbal statement, further support a conclusion from Investigation 4A, that specifying data cases was a significant feature of the task design supporting students' reasoning about covariation.

4.10 INVESTIGATION 5A: HEART DEATHS VERSUS USE OF MOTOR VEHICLES (SURVEYS)

4.10.01 Introduction and Aims

Investigation 1 explored Speculative Data Generation evident in students' responses to Task 1 concerning a news article involving heart deaths and use of motor vehicles (Figure 4.01). The results indicated many students offered incomplete responses, such as Single Comparison (Level 1) responses that emphasized the correspondence of two values but did not illustrate variation in the measures. The task was considered complex due to the multivariate description in the news article, and the unfamiliar nature of the variables and how they were measured. Investigations 2, 3, and 4 explored students' Speculative Data Generation in bivariate situations involving familiar variables, and students demonstrated greater success in generating speculative data to represent covariation.

Investigation 5 revisits the task of Investigation 1 in two new respects. The findings of Investigation 1 included the possibility that some students were not clear which variables were expected in the task. Although identifying appropriate variables in a multivariate situation is an aspect related to Speculative Data Generation, the task in Investigation 1 did not clearly specify the appropriate variables in writing, and in most classes, verbally instructions were provided to clarify the question regarding the verbal statement to represent. Investigation 5A was an attempt to correct this aspect of task design. Investigation 5A involved a survey task that was modified from Investigation 1 to remove some of the extraneous context of the news article and to specify more clearly the verbal statement involving two variables that was to be shown in a graph. It aimed consolidate the basis for the

assessment framework for Speculative Data Generation, aligned with Research Aim 1 of the wider study (Section 3.02). Investigation 5A also aimed to provide a new sample of students from which a subset of students could be drawn for interview in Investigation 5B.

4.10.02 Task

The graphing task, shown in Figure 4.44, was modified from Task 1 in two ways. The first was reducing the complexity of the news article. A comparison of this task with that in Investigation 1 shows the deletion of the final sentence concerning other illnesses, and the simplifying of earlier sentences, such as removing the mention of “twenty years of research.” The second modification to the task, Q1, was the specification to draw a graph to show the verbal statement of a relationship about heart deaths and motor vehicles, in comparison to the vaguer instruction to “Draw what one of Mr Robinson’s graphs might have looked like.” It was expected that these modifications would reduce the number of non-responses by students due to the reduced reading demands of the news article, and that they would reduce the number of responses involving irrelevant variables, while still allowing for selection of variables including time, implicit in the phrases involving the term “increase.” Q2 was administered, but not analysed in this investigation.

**Family car is killing us,
says Tasmanian researcher**

Mr Robinson believes motoring is a health hazard.
He has graphs that show an almost perfect
relationship between the increase in heart deaths
and the increase in use of motor vehicles.

Q1. Draw a graph to show:
**“an almost perfect relationship between the increase in heart
deaths and the increase in use of motor vehicles.”**
Label the graph.

Mr Robinson says “Motoring is a health hazard.”
His reason is based on the graphs that show “an almost perfect
relationship.”

Q2. Do you think the graphs are a good reason to say this?
☐ YES or ☐ NO
Please explain your answer.

Figure 4.44. Task 1 (revised), to assess Speculative Data Generation
(*The Mercury*, June 11, 1991, p. 2).

4.10.03 Participants and Method

A subset of students described in Data Collection 3 was administered this survey item: females in seventh-grade ($n = 19$), and males in fifth-grade ($n = 20$), seventh-grade ($n = 25$), and ninth-grade ($n = 27$). Third-grade students were not administered this item due to its complexity. Fifth-grade males were administered the item, but as will be seen, they encountered difficulties. For this reason, and time constraints, grade 4/5 females were not administered this item. Ninth-grade females were administered this item in their survey task booklet, but due to time constraints,

none attempted it. The administration of the survey task followed the method described as part of Data Collection 3 in Chapter 3. Responses were coded according to the framework of levels and categories described in Investigation 1.

4.10.04 Quantitative Results

All 91 students' responses were classified according to the framework of levels and categories as described in Figure 4.02 for Investigation 1, although not all categories of response were observed. The numbers of students responding in the various levels and categories are shown in Table 4.12. Students from higher grades responded at higher levels, and generally performed at higher response levels than Investigation 1 at comparable grades. Most common response categories were Coordinate Variable (Category 3B), Series Comparison (Category 3A), and Single Comparison (Category 1A). There were few Trend graphs (Category 2B), and no Single Variable (Category 1C) or Double Variable (Category 2C) graphs, both categories frequently observed in Investigation 1. These findings were an expected outcome of modifying the task (a) to eliminate the other illnesses and (b) to make explicit those variables to include in the graph.

Most students drew line graphs or bar graphs. The number of students who displayed case labels was similar to the number who drew continuous data.

Table 4.12.

Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Speculative Data Generation for Task 1 (revised)

	Response Count	Female Grade		Male Grade	
		7	5	7	9
Response Level	N=91	n=19	n=20	n=25	n=27
0. Nonstatistical	22	42	55	12	0
1. Single Statistical Aspect	12	11	20	20	4
2. Inadequate Covariation	6	11	5	8	4
3. Appropriate Covariation	51	37	20	60	93
Response Category					
0. No response	11	5	45	4	0
0. Unlabelled graph	4	16	5	0	0
0. Labelled axes	7	21	5	8	0
1A. Single Comparison	12	11	20	20	4
2A. Double Comparison	4	11	5	4	0
2B. Trend	2	0	0	4	4
3A. Series Comparison	26	26	10	48	26
3B. Coordinate Variable	25	11	10	12	67
Graph form					
None	12	32	25	4	0
Bar	26	32	30	44	11
Line	41	21	20	40	85
Other	12	16	25	12	4
Continuous	17	5	0	12	48
Case labels	19	32	20	32	4

4.10.05 Results – Survey Responses

Students responses at Level 0 included categories observed in Investigation 1, including non-response, unlabeled graphs (axes with values shown), and labelled axes (without values shown). Two examples (R1, R2) are shown in Figure 4.45. In both responses, no data values were shown.

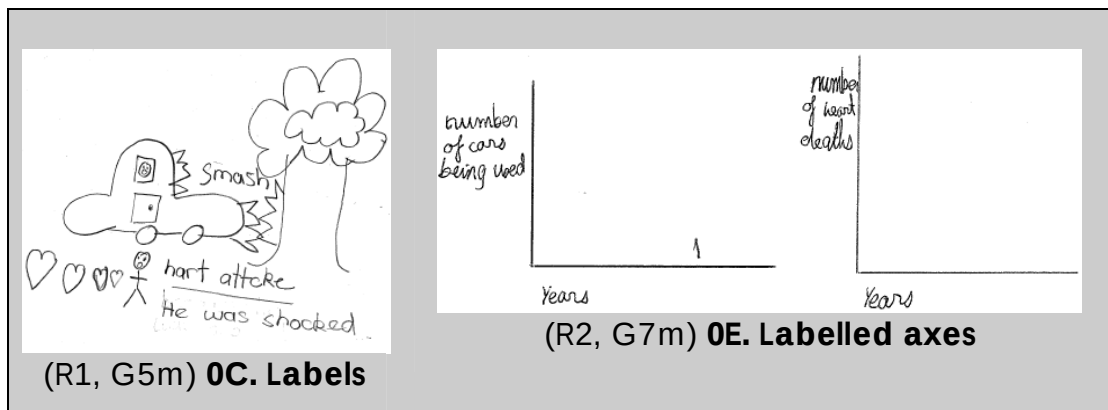


Figure 4.45. Student responses – Level 0 – Nonstatistical.

Students at Level 1 drew Single Comparison graphs (Category 1A). Two examples (R3, R4) are shown in Figure 4.46. Both show “an almost perfect” equality of values for motor vehicles and heart deaths, illustrating the emphasis on correspondence of values and absence of variation that is characteristic of the Single Comparison category. R4 indicates the comparison need not appear within a graph framework, provided values were shown.

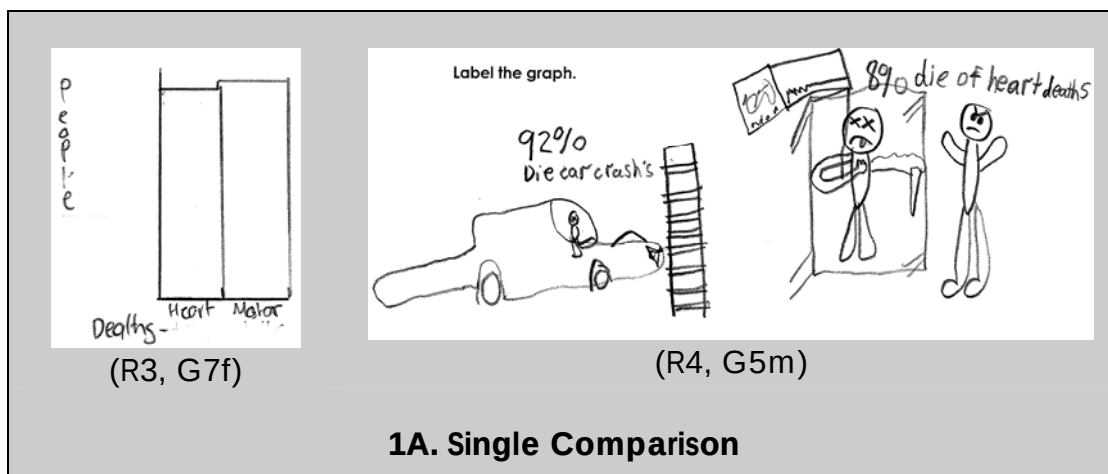


Figure 4.46. Student responses – Level 1 – Single Statistical Aspect.

At Level 2, students drew Double Comparison graphs (Category 2A) and Trend graphs (Category 2B). Examples (R5-R7) are shown in Figure 4.47. R5 shows two pie graphs with year labels. R6 is a common bar graph Double Comparison. R7

shows an increasing trend over time, however the labels of cars and heart along the same line were not clear whether two distinct variables were both shown increasing over time.

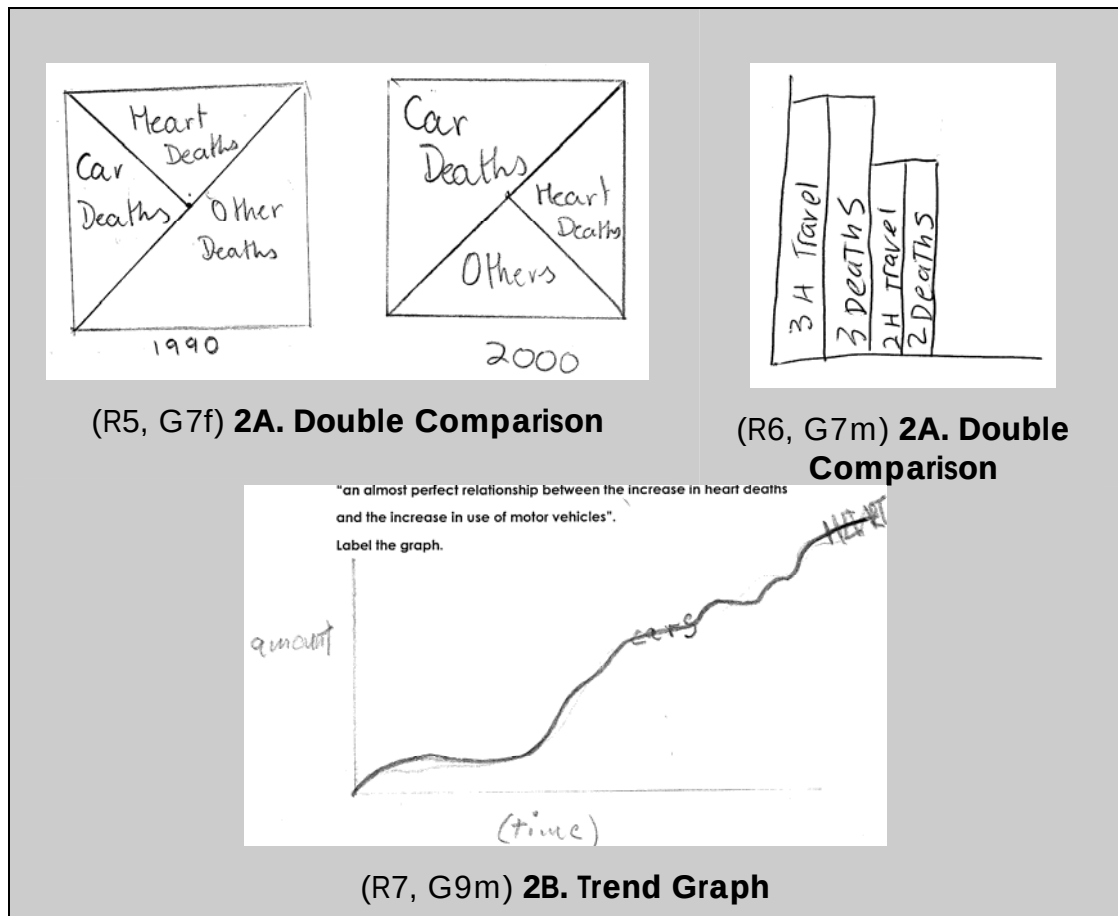


Figure 4.47. Student responses – Level 2 – Inadequate Covariation.

Students at Level 3 drew Series Comparison graphs (Category 3A) and Coordinate Variable graphs (Category 3B). Examples (R8-R12) are shown in Figure 4.48. The examples of Series Comparison responses (R8-R10) all show comparisons with case labels for years. R11 and R12 show Coordinate Variable, R11 a common line graph representation, and R12 a rarely seen scattergraph.

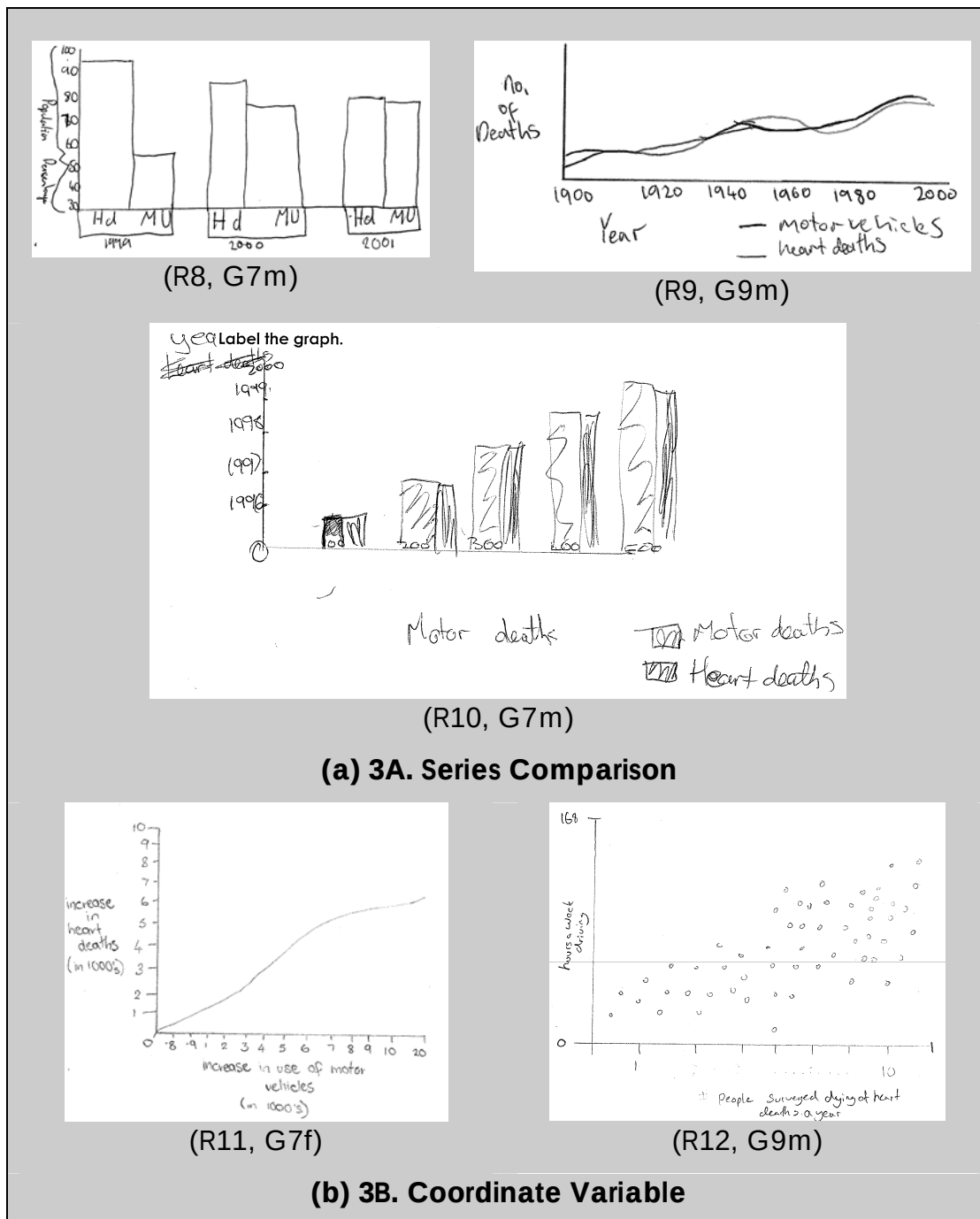


Figure 4.48. Student responses – Level 3 – Appropriate Covariation.

4.10.06 Discussion of Investigation 5A

In response to the revised task used in this investigation, students responded with higher response levels than those in comparable year levels from Investigation 1. Revising the task wording to focus on relevant variables may have contributed to

the absence of Single Variable and Double Variable graphs. Most students emphasized correspondence of values, but some had inadequate notions of variation for each variable, such as shown in Single Comparison (Category 1A) and Double Comparison (Category 2A) responses. Some students represented case labels or continuous data. For this task, case labels were time values, such as years, hence time was displayed either at specified points or as continuous change. This was evident despite the removal to the stem “20 years of research,” suggesting the phrase “increase in” was used as the signal to represent time.

4.11 INVESTIGATION 5B: HEART DEATHS VERSUS USE OF MOTOR VEHICLES (INTERVIEWS)

4.11.01 Introduction and Aims

Investigation 5B involved an interview protocol that asked students to explain their own graphs, and then to compare their responses with others, such as a Single Comparison, a Double Comparison, a Series Comparison, and an Coordinate Variable response. The aims of these interviews were to explore students' reasoning about representing covariation and to explore their potential to develop reasoning via dealing with cognitive conflict, aligned with Research Aim 3 of the wider study (Section 3.02).

4.11.02 Task

The interview protocol followed the general interview protocol format outlined in Chapter 3. Students were shown their own graph, and asked to explain what they had drawn and why they responded in that way. They were then shown a sequence of four or five prompts, as seen in Figure 4.49. Students were asked to comment on the each prompt graph, and to decide which response was preferred to show the required information appropriately and clearly. The prompts shown were at a variety of levels:

- Prompt A was a Single Comparison response (Category 1A),
- Prompt B was a Double Comparison response (Category 2A),
- Prompt C was a Series Comparison response (Category 3A),
- Prompt D was a Series Comparison response (Category 3A),
- Prompt E was a Coordinate Variable response (Category 3B).

Prompt C was omitted for the student who drew it (G5m2), and for others who gave lower level initial responses to avoid creating confusion.

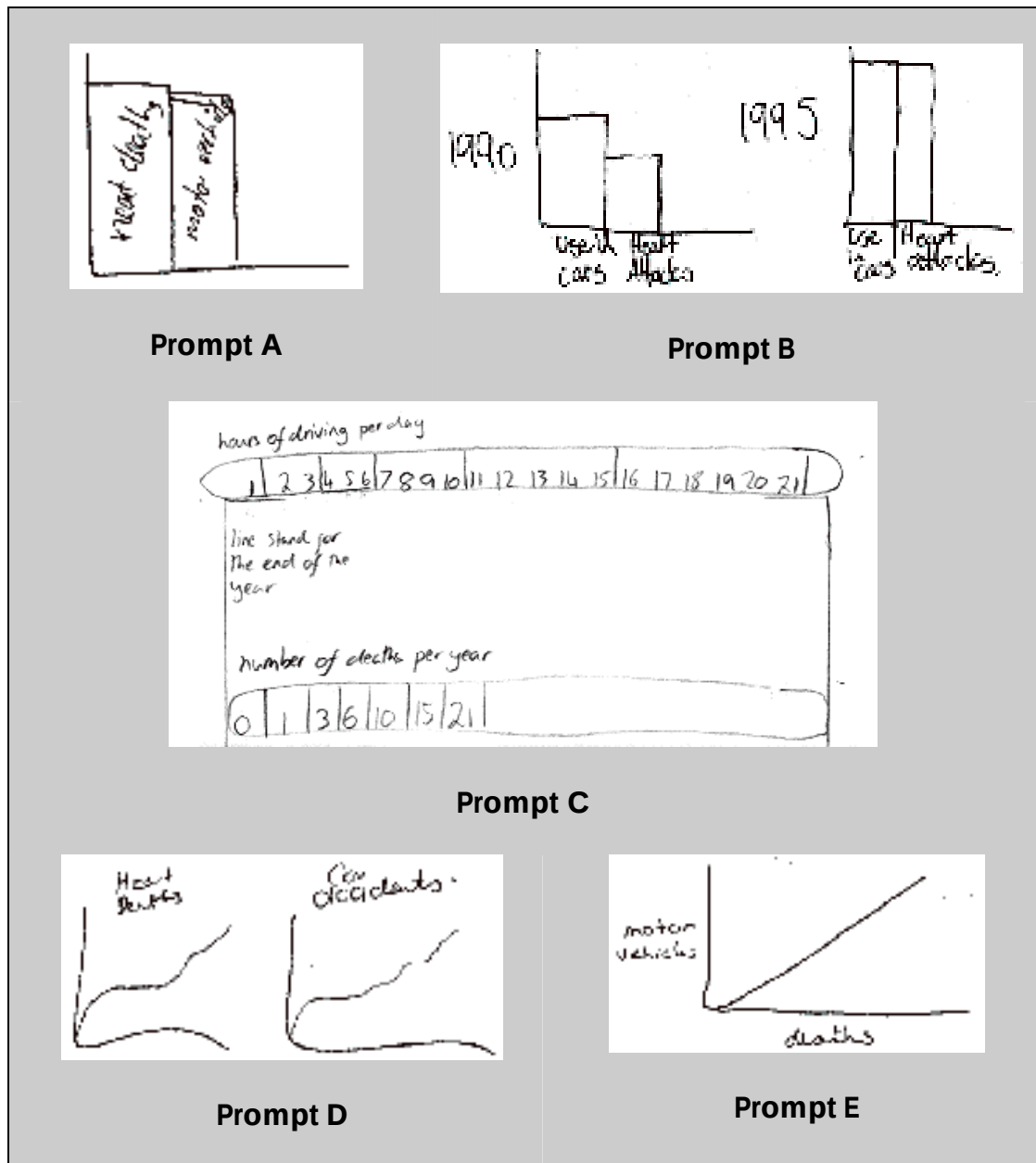


Figure 4.49. Responses used as prompts to create cognitive conflict.

4.11.03 Participants and Method

Participants were a subset of students interviewed as described for Data Collection 4. Students interviewed for this task were limited to those administered the two preceding interviews tasks (Investigation 3B: Height vs Age, and

Investigation 4B: Test Scores vs Study Times). Some students were not interviewed for this task in cases where the subsequent task (Temperature vs Time) took precedence in the interviewer's judgement as being likely to produce more fruitful discussion in the remaining interview time. Such a judgement was made with respect to whether the student had drawn graphs for this task or the next one, and a spontaneous judgement by the interviewer about the ability of the student to engage the tasks. As such, a selection bias occurred, in that students who spent additional time on the first (easier) tasks were less likely to be interviewed for this task, and thus it is likely that the participants for this task had performed quite well in responding to the preceding interview tasks. The ten students interviewed for this task included three fifth-grade males, four seventh-grade students (one male, three female), and three ninth-grade students (two male, one female). Two students (G5m1, G9f2) were interviewed who did not attempt the survey task.

4.11.04 Results – Overview

A summary of the interviews for the ten students is shown in Table 4.13. In general, most prompts received reactions of disagreement. Prompts A and B were agreed to by those who initially responded at Level 0 or 1. Two students who drew Coordinate Variable graphs approved of Prompt E, itself a Coordinate Variable graph. Hence from the few students interviewed, there was weak evidence that students favoured responses close to the level of their initial response. Interviews are discussed in the order of the level of graph the student chose to draw in the survey.

Table 4.13.

Summary Descriptions of Interview Dialogue

ID	Initial Response Category	Prompt – Category of Prompt					Summary of Reactions to Prompts
		A – 1A	B – 2A	C – 3A	D – 3A	E – 3B	
G7f1	0. No response	✓	✓	✗	~	✗	Agree Higher (0→2A)
G7m2	1A. Single Comp.	✓	✓	~	~	✗	Agree Higher (1A→2A)
G7f3	1A. Single Comp.	~	✓	✗	✓	✗	Agree Higher (1A→3A)
G7f6	2A. Double Comp.	✗	✗	N	✗	✗	No change
G5m1	3A. Series Comp.*	✗	✗	✗	✗	✗	No change
G5m2	3A. Series Comp.	✗	✗	N	~	~	No change
G5m3	3A. Series Comp.	✗	✗	✗	~	~	No change
G9f2	3B. Coord. Var.*	~	~	✗	✗	✗	No change
G9m1	3B. Coord. Var.	✓	~	✗	~	✓	Agree Same (3B→3B)
G9m2	3B. Coord. Var.	✗	✗	N	✗	✓	Agree Same (3B→3B)
Total	✗	5	5	6	4	6	
Total	~	2	2	1	5	2	
Total	✓	3	3	0	1	2	

Note: ✓ denotes agreement, ✗ denotes disagreement, ~ denotes ambivalence, N denotes not administered, * denotes drawn in interview

4.11.05 Results – Interview Dialogue

At Level 0, one of the interviewees (G7f1) offered no response in the survey. When asked, “What made that difficult, do you think, more than the other questions?” she responded, “Because you had to, like, explain two things, kind of.” This explanation appeared to explain the interview dialogue that showed the student had a clear preference for a comparison style graph that was simply presented. The interviewer did not encourage the student to attempt to draw the graph in interview. She accepted Prompt A, commenting “It explains it,” and Prompt B, commenting, “Yeah, it’s like showing how things have changed, because of the more technology and stuff.” Prompt C was considered “kind of confusing.” Prompt D was acknowledged as being similar to Prompt A, and “Probably that one [Prompt A]

would be easier than that one [Prompt D], like clearer and stuff.” The student appeared to be confused by Prompt E.

G7f1: [Prompt E] Is that the deaths, that line there [horizontal axis]?

I: Yeah, I think so. Does that make sense how they’ve done it?

G7f1: Kind of. Not as good as the other ones though.

At Level 1, two students (G7m2, G7f3) offered a Single Comparison (Category 1A) in the survey. An example is shown in Figure 4.50. One student (G7f3) conceived of the correspondence as same value. She acknowledged Prompt A as similar to her own, and used the “same value” notion in judging Prompt B.

G7f3: [Explaining own response] I didn’t understand it very well, so I just thought, maybe, that heart deaths and motor vehicles were nearly had the... they had the same amount of deaths in it.

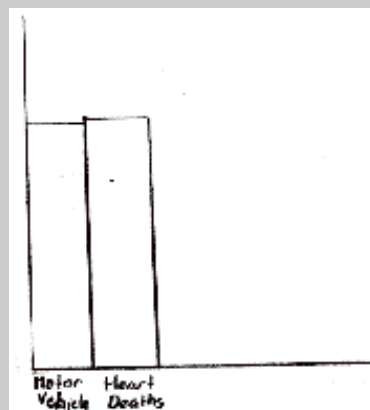
G7f3: [Prompt B] Yeah, it shows it okay, that it’s increased alot, the heart attacks and car things, they’ve both increased [I: Mmm] more, and they’re nearly the same now.

I: Mmm. I notice in this one you were talking about them having the same sort of value here, and in this one they’re a bit different [left graph] and there [right graph] the same. Do you think when we’re talking about a relationship between two things like heart deaths and motor vehicles, or when we’re talking about a relationship between study time and scores, that it’s important that they do have the same sorts of values, or don’t they need the same values?

G7f3: I don’t think they need, really, the same values all the time.

I: Okay. So which one do you prefer for showing that relationship?

G7f3: Probably this one [Prompt B], it shows how it’s increased as well.



(G7f3) **1A. Single Comparison**

Figure 4.50. Survey response of an interviewee at Level 1.

At Level 2, one student (G7f6) offered a Double Comparison (Category 2A) in the survey, as shown in Figure 4.51. She disagreed with the prompts shown, Prompt A as being too simple, and even Prompt B, a Double Comparison response.

G7f6: [Prompt A] It's not very clear, and it doesn't have, it doesn't really show very well, like it might have been that and that for lots of years, but they couldn't tell.

G7f6: [Prompt B] It's more accurate, but it's still not quite, very clear.

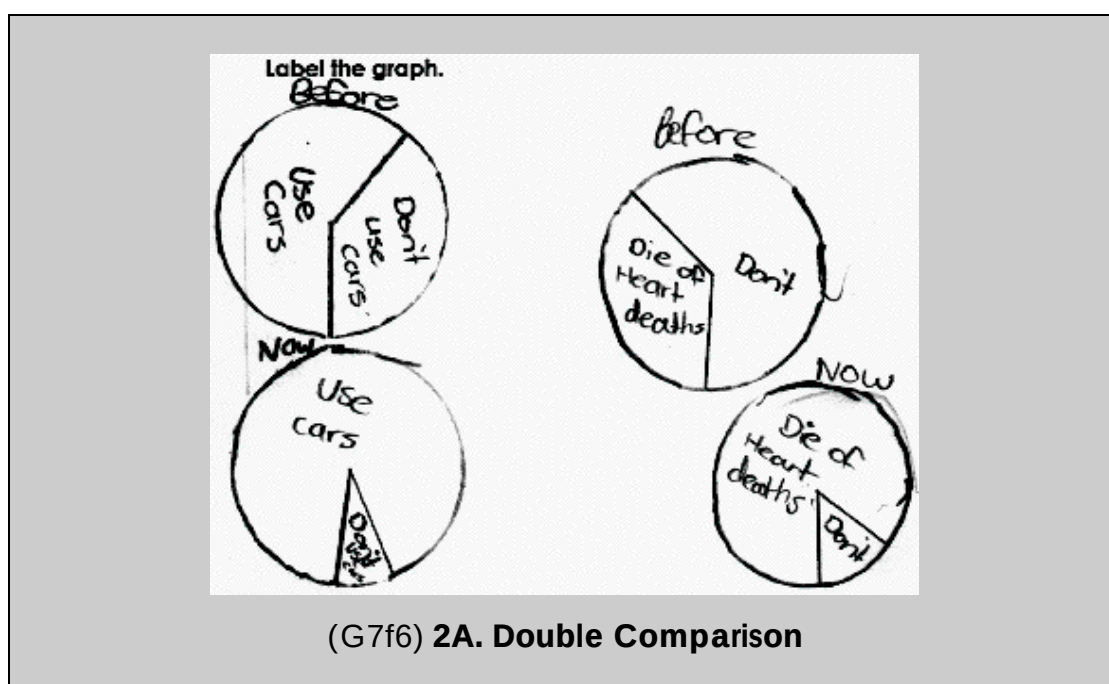


Figure 4.51. Survey response of an interviewee at Level 2.

At Level 3, three fifth-grade students (G5m1, G5m2, G5m3) drew Series Comparison graphs (Category 3A), and three ninth-grade students (G9m1, G9m2, G9f2) drew Coordinate Variable graphs (Category 3B). Three examples are shown in Figure 4.52. Those offering Series Comparison graphs showed remarkable similarity in their patterns of considering their own graphs similar to Prompt D, and preferring these to Prompt B, which in turn was preferred over Prompt A. These students thus appeared capable of generating speculative data to show covariation, and of observing the inadequate variation in Single Comparison and Double Comparison

graphs. They were, however, sensitive to the form of graph, in that all were slightly confused by Prompt E, but gave some indication that with specific values listed, Prompt E might have potential to show covariation. Some of this confusion may have been due to the Cartesian system, the line graph which did not display specific values, and the omission of the time basis which corresponded to the phrase involving “increase.”

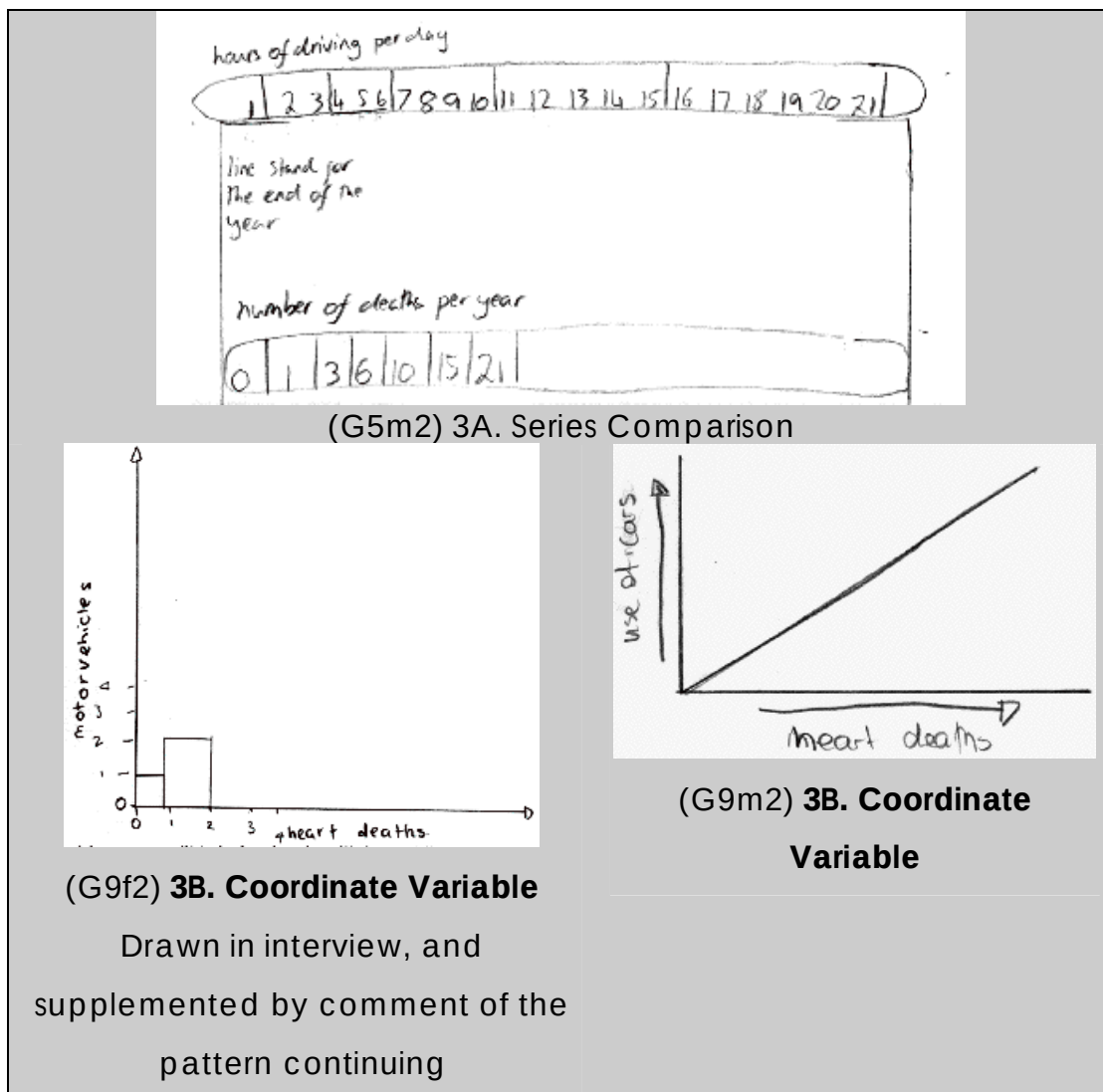


Figure 4.52. Survey responses of interviewees at Level 3.

The most illuminating interview dialogue of the three came from the student (G5m2) who drew the unconventional Series Comparison response used as Prompt

C. He had simple reasons to reject Prompt A, and offered more involved reasoning for Prompt B.

G5m2: [Prompt A] Because they've just got motor vehicles and heart deaths, and it doesn't say how much motor vehicles there are, because it doesn't say it increases or anything. Just at one time how much there is.

G5m2: [Prompt B] That's probably better than the other one we saw [Prompt A]. And it shows that in 1990 there was that much use of cars... and there was a bit more use of cars than there was heart deaths. But this one is exactly the same. So what it kind of gives the impression that, at 1995, there was equal use of cars and heart deaths. So I think it shows that heart deaths has gone up alot, and use of cars has kind of stopped. That doesn't really make much sense, because if the cars are causing it, and the cars stop, then heart deaths should go down, not up. So it makes it a bit confusing.

In describing Prompt D, he used the phrase "they've both gone up" on three separate occasions, believing the horizontal axis to refer to time by adding in "you don't know how long." He indicated Prompt D could be improved, "if they'd used more numbers, this [Prompt D] would be a really good way." For Prompt E, however, the design of one variable on each axis appeared to confuse the student.

G5m2: [Prompt E] Once again it's sort of saying, as this [Prompt D], except it says motor vehicles on the side, except it doesn't say how much, whether it's gone up or not, it could be going down and heart deaths could be going up, because I think that line shows the heart deaths. Because they've only got one line so it's not very clear, if that's [horizontal axis] the deaths line, then you don't know how much there has been use of the motor vehicles, you don't know how much, many deaths. So it only really gives you one answer, and you don't know, and even then, it would be nearly impossible to know which one it was unless you asked the person.

I: If they included numbers along here, like 1, 2, 3, 4, up to 21, like yours, across here, and they included the numbers of hours driving, from 1, 2, 3, 4, 5, right up to say 21 again, do you think that would help the graph to show the relationship or would it still not show it?

G5m2: I think it would help alot to show it. To show that there has been a relationship, because it shows that the hours and how much the motor vehicles have been used goes up, and you can see that the deaths have gone up as well.

Of the three students who drew Coordinate Variable graphs, the ninth-grade female (G9f2) considered covariation to involve exact value correspondence. The

graph she drew in the interview used values $\{ (1,1), (2,2) \}$ and she described continuation of this, thus warranting the assessment as Coordinate Variable (Category 3B). She considered Graphs A and B to show the correspondence, but preferred her own as “more in-depth.” For Prompt D, she commented on the lack of numbers, “Well it has no numbers, so you have no idea what they’re doing,” and when asked, she suggested the horizontal axis provided a time basis, “The hours and the deaths [frowning], I suppose or... no, the years and the deaths... Probably. Just do the same amount of years for each one, and the same amount of deaths.” Although she had drawn a Bivariate Association graph in a bar graph form, she rejected the line graph form in Prompt E.

G9f2: [Prompt E] Well it’s comparing them, not only because of deaths, but there’s only one line. If they had numbers on there to show the comparance [sic] between...

I: Yeah. Is it a bit like yours, or quite different?

G9f2: The same, except a line.

Another student (G9m2) who drew an Coordinate Variable graph almost identical to Prompt E agreed with Prompt E, and rejected Prompts A and B. He had more difficulty, however, in engaging Prompt D, which he considered to lack labels. He made repeated reference to “going up,” but repeatedly suggested the two variables should be on opposing axes. Throughout the interview there was no indication he understood “goes up” to refer to “over time,” and he even rejected the interviewer’s suggestion of the time basis.

G9m2: [Prompt D] Well, yeah, that shows it I think... they both go up... actually it doesn’t really, ... it sort of, the heart risk, it needs to be, no, it doesn’t really. Both of these need to be on a different axis to show that they are related. You can’t really relate these to one another, except for they’re both going up [...]

I: Yeah. What would be the axes labels here, do you think? If they were to say that this is measuring something and this is measuring something?

- G9m2: I don't know... I'm not sure, perhaps if they would have put heart risk, heart deaths here [left graph, vertical axis] and car [pointing to left graph, horizontal axis]... car accidents, what do they mean by that, I think they misunderstood it. [...] perhaps if they labelled the axis like I did, with the heart deaths here and the car usage here.
- I: When you were saying it's going up, I suppose you're saying that heart deaths is measured up here [Prompt D, left graph, vertical axis], do you mean that time is measured across here?
- G9m2: Umm. It sort of shows that car usage going up [sic], I mean, say we had heart deaths here [writes H on vertical axis] and car usage there [writes C on horizontal axis], it shows that as the car usage goes up, heart deaths goes up as well. Like that [draws line].

4.11.06 Discussion of Investigation 5B

Student dialogue showed evidence of reasoning in relation to statistical covariation, including reasoning of correspondence based on “same value,” and reasoning based on continuous variation. Some students’ verbal descriptions of “going up” illustrated that the task verbal description of “increase” was important in students’ understanding of covariation, which they referred to as a criterion for whether certain prompts illustrated appropriate variation in each measure. Some students could not read a line graph as a set of bivariate values (continuous versus discrete).

Few students were interviewed, and more than half of them had initially responded at Level 3, however there was weak evidence that students reacted most favourably to prompts near the level of their original response. As with previous investigations, such a finding was only weakly supported, and some students disagreed with prompts similar to their own response, possibly because they favoured their own design presentation.

4.12 GENERAL DISCUSSION – SPECULATIVE DATA GENERATION

The framework of four levels for Speculative Data Generation was applied with similar responses categories across three tasks (Task 1 – Heart deaths, Task 3 – Heights, and Task 4 – Test scores). The levels of response were supported by both survey-based and interview-based investigations. Further, there was weak evidence that students exposed to prompts agreed to responses at similar levels to their own. Within the framework of levels, categories of response were distinguished by the variables shown or labelled, the variation of values for a variable, and the correspondence of values between the variables. Stability of most categories across the tasks provided broader evidence of the usefulness of the categorisation assessment frameworks.

Student performance across the tasks differed according to the task presented. Task 3 involving heights and ages, yielded high response levels, attributable to the familiar variables, and the intuitive mapping permitting pictographs as a means of representing heights. Task 1, involving unfamiliar measures of heart deaths and motor vehicle use, over time, in a complex newspaper context, yielded lower response levels. Task 4, involving a context requiring six data cases with counterintuitive covariation between test scores and study times, yielded quite high response levels. Further, interviews noted that in some cases, especially for Task 4, students' survey responses reflected a simple misreading of the task or other beliefs contrary to the verbal statement, as opposed to inability to appreciate and represent the given verbal statement of covariation. It was suggested that for Task 4, the description of six data cases was influential in achieving high response levels,

despite the counterintuitive covariation, as evident in students' common use of case labels to construct covariation case-by-case.

Further comparison across tasks is presented in Chapter 7. The next two chapters turn to other skills for translating among representations of statistical covariation: Coordinate Graph Production (Chapter 5), and Verbal Graph Interpretation and Numerical Graph Interpretation (Chapter 6). Both include tasks involving six bivariate data cases, to consider how this task feature contributes to responses for these skills.

CHAPTER 5. COORDINATE GRAPH PRODUCTION

5.01 SUMMARY

This chapter discusses two investigations of students' productions of coordinate graphs. Investigations 6A and 6B involved survey and interview responses respectively, each for a task to represent numerical data about temperature change over time.

Investigation 6A involved surveys of 133 students in grades 3, 5, 7, and 9 using a written task to construct a graph. Four response levels described the degree to which students transformed the table of data values into a coordinate graph.

Nonstatistical responses did not display the data, showing either the context or a graph form only. *Single Statistical Aspect* responses showed data along a single dimension, either in a table of corresponding values, or a graph of a single variable.

Inadequate Covariation responses showed bivariate data in two-dimensional space but inadequately showed either spatial variation or correspondence of values.

Appropriate Covariation graphs displayed both correspondence and variation of values along ordered axes, either as a bar graph of discrete values or as a line graph of continuous variation.

Investigation 6B involved interviews with 18 students from among participants of Investigation 6A. The researcher asked students (a) to explain how their graph represented the information and why they chose to represent the data that way, and (b) to compare their own response with four graphs drawn by other students to decide which was preferred and why. Interview dialogue was analyzed with reference to students' survey responses and whether students' preferred graph changed after exposure to other students' ideas. Students tended to accept prompts similar to their own response, with respect to both the level and the emphasis on

either correspondence or variation. In some cases students engaged a new idea presented in a prompt, but not in a way that allowed them to build upon it to develop a coordinate system.

5.02 INTRODUCTION

Graph production involves using data to provide a graphical response, as shown in Figure 1.01. Most graphs use the principle that position denotes value, but in some cases, such as pie graph or bar graphs, area is used to denote value. This principle has been discussed with reference to Playfair’s historical development of graphs (Section 2.01.03). Colloquial verbal dialogue often involves positional terms to denote value, such as referring to “higher” or “lower” values, rather than involving magnitude, such as referring to larger or smaller values. Position denoting value is a component of any place-value number system, for which the position denotes the multiple of the base (e.g., 10) and the digits denote the value for the multiple. Analogue scales on many measurement devices, including clocks, thermometers, weight scales, speedometers, and oven temperature dials, provide a set of labels showing the value at each position along the scale, whereas digital equivalents of these measurement devices rely on the digit labels rather than a scale position.

Coordinate graphs employ the general graphic feature that position denotes value and apply this feature in two-dimensional space to represent values of two variables. Coordinate graphs represent both (a) the data points, emphasizing the *correspondence* of values of two variables, and (b) general trends, emphasizing *variation* of the two variables due to the ordination of the values along each axis. These two aspects, which featured in student responses in relation to Speculative Data Generation (Chapter 4), were also relevant features in student responses for Coordinate Graph Production.

The framework used for analysing student responses to tasks of Coordinate Graph Production is shown in Figure 5.01. Four response levels described the degree to which students transformed a table of data into a coordinate graph. At Level 0, *Nonstatistical* responses did not display the data, but included either elements of the Context (Category 0A) or a Graph Form (Category 0B). At Level 1, *Single Statistical Aspect* responses represented at least one data series, either as a Table of corresponding values (Category 1A) or as a graph of a Single Variable (Category 1B). *Inadequate Covariation* representations (Level 2) represented both sets of values but did not use position in two dimensions to denote values of the two variables, either showing a Diagonal Graph (Category 2A) with correspondence but lacking an ordered scale to show variation, or showing a Spatial Variation graph (Category 2B) with variation but lacking a direct correspondence of variables. At Level 3, *Appropriate Covariation* representations used coordinate position in two dimensions to denote values of the two variables in Bar graphs (Category 3A) or Line graphs (Category 3B).

A number of categories within this framework differed from frameworks for tasks of Speculative Data Generation (Figure 4.02, Figure 4.27). These differences emerged due to the expectations of the Coordinate Graph Production task differing from the expectations of the tasks for Speculative Data Generation. These differences are discussed further in Section 7.02.

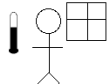
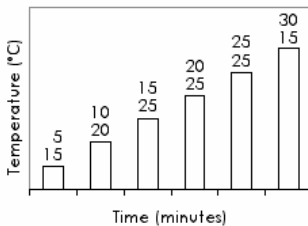
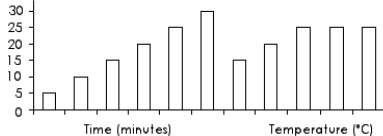
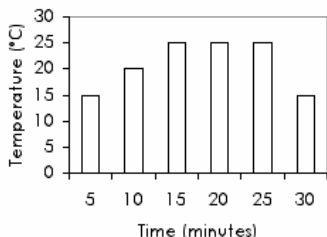
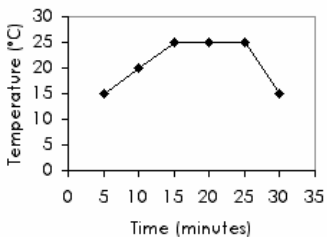
Level of Coordinate Graph Production	Category of Coordinate Graph Production														
0. Nonstatistical	0A. Context Picture or narrative shows the context 	0B. Graph Form Axes or Lines drawn; some show irrelevant values 													
	1. Single Statistical Aspect Data are aligned in a single dimension, such as a table <table border="1" data-bbox="456 707 847 754"><tr><td>Time (Minutes)</td><td>5</td><td>10</td><td>15</td><td>20</td><td>25</td><td>30</td></tr><tr><td>Temperature (°C)</td><td>15</td><td>20</td><td>25</td><td>25</td><td>25</td><td>15</td></tr></table>	Time (Minutes)	5	10	15	20	25	30	Temperature (°C)	15	20	25	25	25	15
Time (Minutes)	5	10	15	20	25	30									
Temperature (°C)	15	20	25	25	25	15									
2. Inadequate Covariation	2A. Diagonal Graph Graph shows corresponding values but lacks spatial variation 	2B. Spatial Variation Graph Graph shows both variables but lacks correspondence 													
	3. Appropriate Covariation Graphs coordinate bivariate data	3A. Bar Graph Bar graph uses coordinates to show discrete values  3B. Line Graph Line graph uses coordinates to show continuous variation 													

Figure 5.01. Levels and categories of Coordinate Graph Production.

5.03 INVESTIGATION 6A: TEMPERATURE CHANGE OVER TIME (SURVEYS)

5.03.01 Introduction and Aims

The review of historical development and research on graphing in Chapter 2 included issues such as tabular emphasis, reduction to a single variable, continuity, and axis allocation. Investigation 6A aimed to explore the structures that students produce when graphing bivariate data in the context of temperature change over time, and to describe responses within an assessment framework, aligned with Research Aim 1 of the wider study (see Section 3.02). Assessment of responses gave particular attention to how students structured relevant elements in their attempts to construct coordinate graphs. Of interest were alternative structures to coordinate graphs, and other features and contrasting emphases including representing narrative context versus data values, representing in tables versus graphs, focusing on one variable, adopting conventions such as ordering, spacing, and labelling, representing discrete values versus continuous lines, and allocating variables to axes. A second aim concerned the proportions of school students in grades 3, 5, 7, and 9 that draw coordinate graphs and various alternatives, aligned with Research Aim 2 of the wider study (see Section 3.02).

5.03.02 Task

Students were presented with a context involving temperature in relation to various events involving a heater and a window, and with a data table including six numeric values of temperature and time, as shown in Figure 5.02. They were asked to draw a graph to show how the temperature changed over time. The context of temperature was chosen as likely to be familiar from viewing television graphs, as

shown in Figure 2.01. Time was also chosen as likely to be familiar, particularly with respect to awareness of ordered continuity. The narrative context, which described turning the heater on and off and opening the window, was used not only to aid interpretation of the task, but also to allow students to decide which features of the narrative or the data table were most important to represent in a graph to show how the temperature changed over time.

A science class was studying temperature. They used a thermometer to measure the room temperature every 5 minutes for 30 minutes.

First they turned a heater on for 15 minutes.

Next they turned the heater off for 10 minutes.

Lastly they opened the window for 5 minutes.

They wrote down these numbers.

Time (Minutes)	5	10	15	20	25	30
Temperature (°C)	15	20	25	25	25	15

Draw a graph to show how the temperature changed over time.

Figure 5.02. Task 5 to assess Coordinate Graph Production.

The task was designed to present a context and data set that supported students choosing to represent both individual data points and continuous variation. The data table involved discrete data points, but the context indicated these data were sampled from measures that were likely to be recognized as continuous, providing clues about intermediate values to represent gradual change. Six data points were chosen to match other questions in the survey. The data included repeated temperatures for different times to ensure students were required to distinguish two variables, following studies reporting student difficulties with the constant function (Nemirovsky & Tierney, 2001; Sherin, 2000). Times were evenly spaced as being realistic within the context, and both data sets were intended to show realistic and yet simply rounded integers. Providing evenly-spaced values permitted students to

represent times as categorical labels without being aware of spatial placement along an interval scale. The analysis of student responses was based on how students designed axes and graph forms, and how they represented values by positions along an ordered scale for temperature; finer distinctions of proportional scaling and ordering of unordered values were not considered in this investigation.

5.03.03 Participants and Method

Responses from 133 students from grades 3, 5, 7 and 9 were gathered, following the method specified for Data Collection 3 in Chapter 3 (Section 3.06.04). Student responses were scanned into computer graphics files. Qualitative analysis of responses followed iterative-clustering processes (Miles & Huberman, 1994). Survey responses were first clustered according to visual aspects alone. Many responses were clearly recognisable as pictures, tables, coordinate bar graphs, or coordinate line graphs. About 50 remaining responses were in forms that modified, combined, or omitted features from one of these basic forms, and these responses required careful consideration in coding. Most of the significance of the coding framework concerned identifying features to provide a lens to make sense of students' responses, as evidence of their understandings. Elements of the responses identified as relevant to the task included (a) conveying clearly and accurately the corresponding values of each data set, and (b) consistent use of spatial position to show variation of values. Correspondence and variation emerged as overarching themes for coding, and sub-coding involved more specific issues such as graph form or allocation of variables to axes.

Clusters were successively refined, particularly for responses in which elements of conventional coordinate graphing were mixed with inconsistencies or omissions, such as poor labelling. A few graphs (e.g., R11 and R12, Figure 5.05)

were initially clustered with coordinate graph responses (Level 3), consistent with Speculative Data Generation analyses. The nature of the Coordinate Graph Production task, however, was to transform the table of data values provided to show graphically the covariation of temperature change over time, hence these responses were assessed at Level 2 for Coordinate Graph Production. Common characteristics of clusters were identified to define categories of response. These categories are described in Section 5.02, and summarized in Figure 5.01.

5.03.04 Quantitative Results

Table 5.01 shows the percentages of response levels and categories to Task 5 for each sex and grade. Students from higher grades tended to respond at higher levels. All ninth-grade students (26 students) offered coordinate graphs, most as line graphs (24) as shown in Table 5.01. Most seventh-grade students (22/32) drew coordinate graphs, either as bar graphs (13) or line graphs (9). Most fifth-grade students responded at Level 1 (9/35) or Level 2 (15/35), and most third-grade students responded at Level 0 (21/40) or Level 1 (14/40). Levels and categories were those described in Figure 5.01. The character of these responses is described in the following results, and illustrated with examples.

Table 5.01.

Response Counts and Percentages of Responses by Gender and Grade at each Level and Category of Coordinate Graph Production

	Response Count	Female Grade			Male Grade			
		3	4/5	7	3	5	7	9
Coordinate Graph Production								
Response Level	N=133	n=21	n=26	n=11	n=19	n=9	n=21	n=26
0. Nonstatistical	33	71	19	18	32	33	10	0
1. Single Statistical Aspect	26	19	35	0	53	0	14	0
2. Inadequate Covariation	21	10	42	9	5	44	10	0
3. Appropriate Covariation	53	0	4	73	11	22	67	100
Response Category								
0A. Context	12	43	0	0	5	22	0	0
0B. Graph Form	21	29	19	18	26	11	10	0
1A. Table	18	19	15	0	53	0	0	0
1B. Single Variable	8	0	19	0	0	0	14	0
2A. Diagonal Graph	8	0	23	9	0	0	5	0
2B. Spatial Variation Graph	13	10	19	0	5	44	5	0
3A. Bar Graph	19	0	4	55	11	11	33	8
3B. Line Graph	34	0	0	18	0	11	33	92

5.03.05 Level 0 – Nonstatistical

Thirty-three students gave Nonstatistical responses, in which the numerical data given were not displayed (see Figure 5.01). Twelve students offered *Context* responses (Category 0A), ten in the form of a picture (e.g., R1, Figure 5.03) and two as a narrative. Twenty-one students drew *Graph Form* responses (Category 0B) in which a graph or table format was represented but not with the given data. Of these twenty-one, five drew axes or a graph frame without values (e.g., R2, Figure 5.03) and six included unlabeled inappropriate values. The remaining ten Graph Form responses included (a) five bar graph variants with each bar related to the on/off status of the heater and bar height implying temperature values that were not

specified or were inappropriate (e.g., R3, Figure 5.03), (b) three attempts at coordinate graphs but with inappropriate or uninterpretable values, and (c) two tables of inappropriate values (e.g., R4, Figure 5.03). Graph Form responses did not show the data, but did show evidence of adopting the mechanics of graphing that are foundational for representing statistical data. Some responses showed weak evidence of emphasising correspondence or variation but without use of the data provided; for example, the confused correspondence of two temperatures for given times in R2, or the visual impression of variation shown in R3.

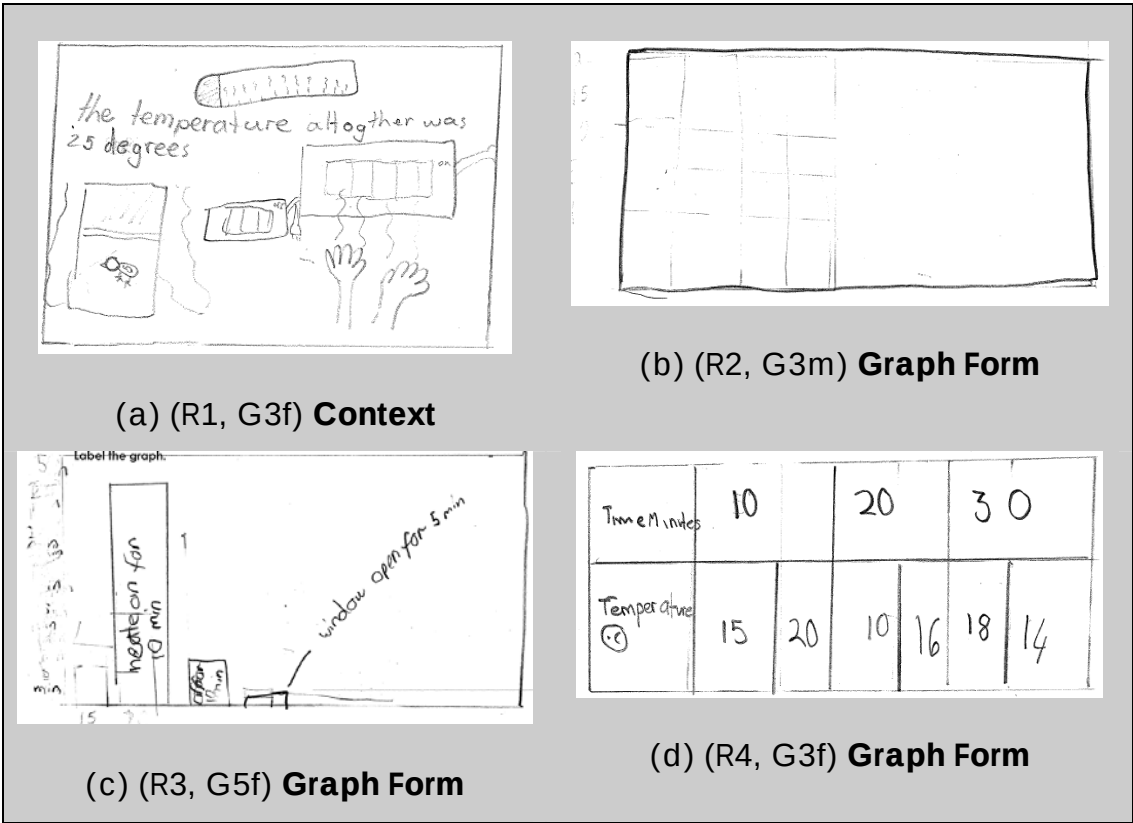


Figure 5.03. Student responses – Level 0 – Nonstatistical.

5.03.06 Level 1 – Single Statistical Aspect

Twenty-six students represented numeric values of temperature, time, or both, but written values were aligned along a single dimension (see Figure 5.01). Two categories of response were evident: a Table (Category 1A), reproducing that given in the task stimulus, or a Single Variable (Category 1B) response. In all these responses, values were written numbers, that is, position was not used to indicate value.

Eighteen students produced a *Table* of the given data (Category 1A), even though a graph was requested. Eight responses involved minimal change from the table given, six of which represented only the six data cases (e.g., R5, Figure 5.04), and two of which annotated information. Ten students transposed the table such that it was aligned vertically rather than horizontally (e.g., R6, Figure 5.04). Of these ten students, (a) five aligned the corresponding values on each side of a centre line appearing as an axis (e.g., R7, Figure 5.04), (b) two wrote the values on the same side of a vertical axis, one interleaving corresponding times and temperatures (R6), the other listing all times then all temperatures, (c) two wrote the values within a grid structure like the table but aligned vertically, and (d) one wrote the values on two separate vertical axes but with no data series.

temperature	15	20	25	25	25	15
minutes	5	10	15	20	25	30

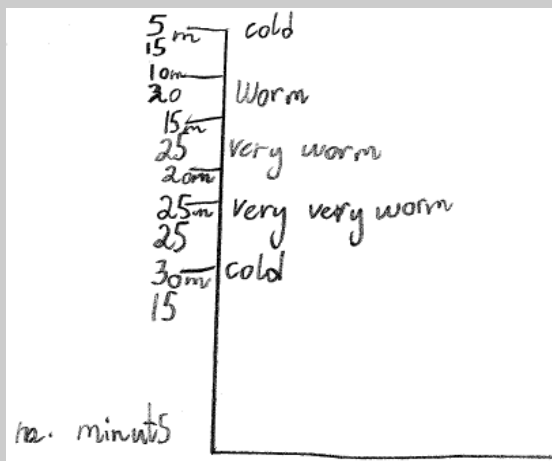
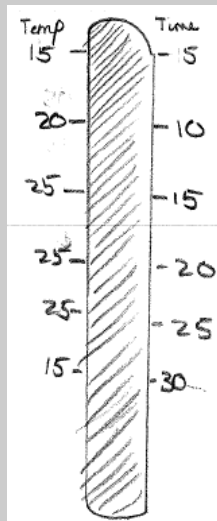
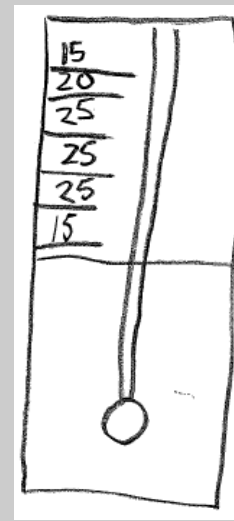
(a) (R5, G3m) **Table**(b) (R6, G3m) **Table**(c) (R7, G5f) **Table**(d) (R8, G5f) **Single Variable**

Figure 5.04. Student responses – Level 1 – Single Statistical Aspect.

Eight students represented a *Single Variable* (Category 1B), seven showing temperature values (e.g., R8, Figure 5.04), and one showing time values. Of these eight students, (a) three drew rectangular axes but showed values for temperature only without accuracy in spatial positioning, (b) two drew rectangular axes with the values of one variable written along one axis without a data series or values on the other axis, (c) two drew thermometers (R8, Figure 5.04), one showing only distinct values of 15, 20, and 25, and (d) one drew a pie graph of distinct sector sizes with distinct temperature values colour-coded.

5.03.07 Level 2 – Inadequate Covariation

Twenty-one students drew a two-axis system for representing the data, either as a Diagonal Graph (Category 2A) lacking consistent use of space to denote ordinal variation, or as a Spatial Variation graph (Category 2B) lacking clear correspondence between variables (see Figure 5.01). Because given time values were ordered and equally spaced, it could not be determined whether some responses were using position to denote time values, or merely as convenient space in which to write a list of categorically labelled values.

Eight students drew a *Diagonal Graph* (Category 2A) in which bivariate data points were shown as a series of increasing bars with written values. These students preserved the correspondence of values with a notion of two dimensions, however at least one dimension was not ordered, and thus written values rather than position were used to denote value. Two students wrote the values for both variables on bars of the data series (e.g., R9, Figure 5.05), that is, they reproduced the table data in a diagonal layout, but within a graphical framework of two axes. One student wrote the values for times on the vertical axis, and the values for temperatures on the bars. The remaining four students wrote the values on the axes with increasing bars and values of temperature not in value order but in the order they appeared in the table; two placed temperature on the vertical axis (e.g., R10, Figure 5.05b), and two on the horizontal axis.

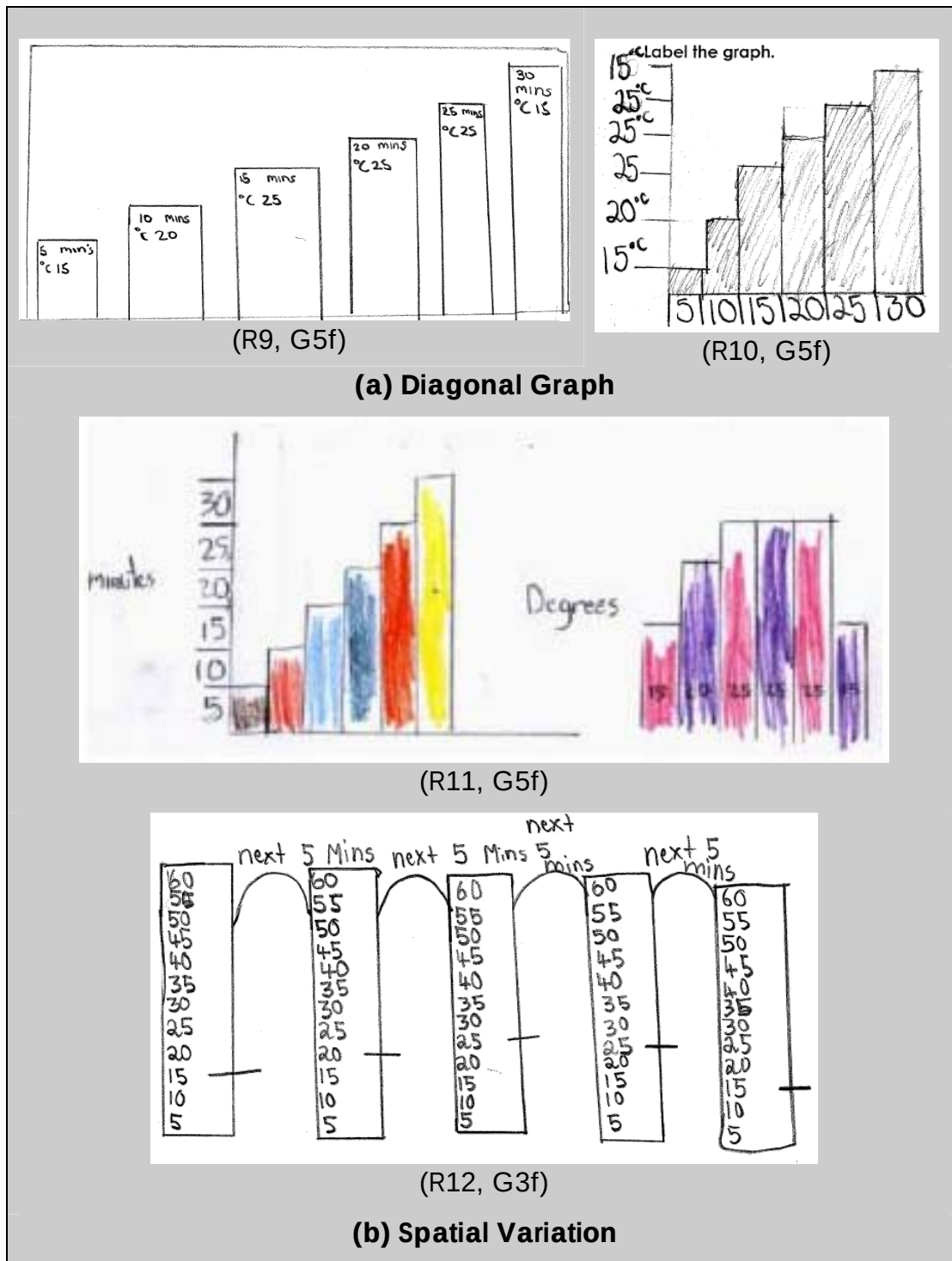


Figure 5.05. Student responses – Level 2 – Inadequate Covariation.

Thirteen students offered a *Spatial Variation Graph* (Category 2B) in which temperature variations were shown spatially in the vertical dimension as a data series, and time was represented but not directly corresponding with the

temperatures. Six students drew series-comparison graphs in which each variable was graphed as a data series (e.g., R11, Figure 5.05); the display of temperature variation gave the spatial appearance of a coordinate representation, however the correspondence with time was weak, relying on cross-referencing the ordinal position of the bar in two graph frames. Seven other students drew graphs resembling coordinate axis systems, but involving errors of (a) inconsistent spacing such as baseline violations or (b) labelling problems that meant that time values could not be read with corresponding temperatures. In R12, a third-grade female drew repeated scales resembling thermometers to represent temperatures, however one data value was missing and times were poorly represented by sequential references (“next 5 mins”) rather than values (e.g., “15 minutes”).

5.03.08 Level 3 – Appropriate Covariation

Fifty-three students employed a coordinate system to represent the bivariate data, in which each variable was ordered along an axis for which position was consistently ordered to denote value. Bar Graphs (Category 3A) showing discrete data emphasized correspondence of values whereas Line Graphs (Category 3B) emphasized continuous variation of temperature with passing time (see Figure 5.01).

Nineteen students drew a *Bar Graph* (Category 3A) within a Cartesian coordinate system, in which six discrete bivariate values were represented. All except one represented temperatures on the vertical axis and times on the horizontal axis. Fourteen labelled “Time” and “Temperature” as well as data values (e.g., R13, Figure 5.06), whereas five had minor labelling omissions (e.g., R14, Figure 5.06), but the values or units distinguished the variables. Bars were adjoining for thirteen responses, whereas the remaining six had a small gap between bars. Fifteen labelled

times in the middle of bars (e.g., R14, Figure 5.06), and four on right hand edge of the bar (e.g., R13, Figure 5.06).

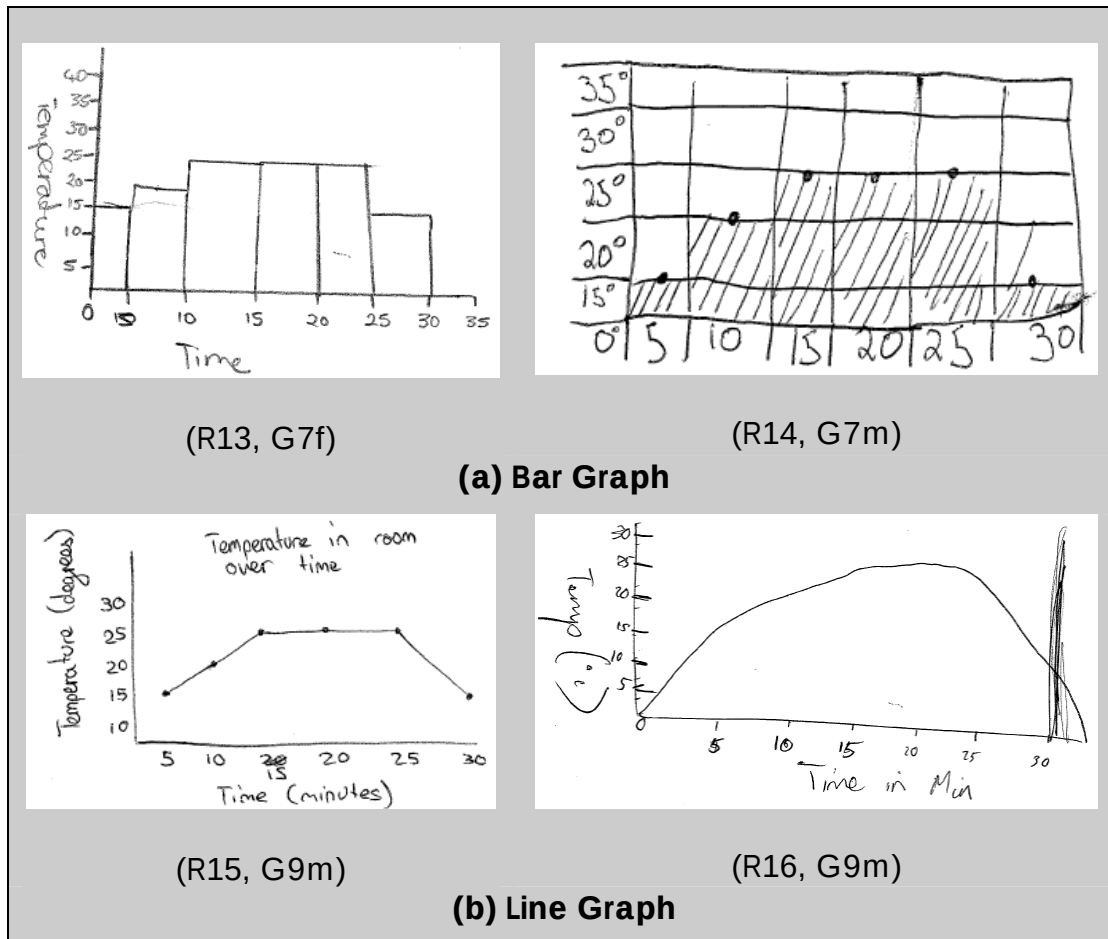


Figure 5.06. Student responses – Level 3 – Appropriate Covariation.

Thirty-four students drew a *Line Graph* (Category 3B) within a coordinate system, in which the six bivariate data values were shown, and were joined by line segments. All thirty-four placed time on the horizontal axis. Twenty-nine of these emphasized the data points along the line (e.g., R15, Figure 5.06), whereas five drew lines without evidence of points (e.g., R16, Figure 5.06). Eleven responses used a line to indicate appropriate temperatures from time 5 minutes to 30 minutes, that is without making assumptions of temperatures outside of these times (e.g., R15, Figure 5.06). Fourteen responses assumed the line should start at the origin, six of

which also made assumptions about the temperature after time 30 minutes (e.g., R16, Figure 5.06). Nine other students made another assumption about the temperature at time 0, five indicating a temperature of 10°C, three 15°C, and one 5°C.

5.03.09 Discussion of Investigation 6A

Students constructed a range of graph structures other than using coordinates appropriately. These responses were assigned to Levels 0, 1, and 2 of the framework in Figure 5.01, indicating those who did not represent the data (Level 0), those who represented either correspondence or variation in the data but not both (Level 1), and those who inadequately represented both correspondence and variation (Level 2).

At Level 0, some students represented the context in a picture or narrative (Category 0A). Some represented the events involving the heater and window, attempting to show all they knew about the narrative context (Nemirovsky & Tierney, 2001). Other students drew a graph form (Category 0B). In some cases, students displayed data that appeared fictitious, with little resemblance to the data values given. It is possible this was a consequence of preceding questions in which students were provided verbal statements and asked to represent speculative data.

Students who drew Tables of data (Category 1A) either were unable to draw a graph or did not see a purpose in graphically representing data when a table would suffice to show the values (Roth & McGinn, 1997). Some students drew a Single Variable (Category 1B), a phenomenon observed previously (Chick & Watson, 2001; Mevarech & Kramarsky, 1997). Most of these represented only temperature; apparently this was considered to be the measured variable of interest, whereas time was implicit or irrelevant.

Diagonal graphs (Category 2A) showed the correspondence of labelled bivariate values within two-dimensional space, but they did not use the graph space

to show variation in values. Similar responses have been reported by others (Mevarech & Kramarsky, 1997) who observed that students represented increasing functions irrespective of task descriptions. Spatial Variation graphs (Category 2B) violated some aspect of correspondence, such as omitting some data values or labels that could have clarified the correspondence of the variables or representing the temperature and time values in separate series of bars, previously described as Series Comparison graphs (Chapter 4) or ordered case-value bar graphs (Konold & Higgins, 2003). In this study, such responses were considered inadequate for showing coordinates as they failed to show the correspondence of the n th bar in one series with the n th bar in the other series. In other tasks, a linking case identifier was available or invented that established the correspondence, such as using students' names to identify cases (Konold & Higgins, 2003), or calendar years to link the incidence of heart deaths with data about the use of motor vehicles (Watson, 2000).

The reasons that students drew either Bar Graphs (Category 3A) or Line Graphs (Category 3B) are not evident from their responses alone: for example, perhaps each student was familiar with only the one graph form of the response. The clear majority of ninth-grade males who responded with line graphs, however, probably indicated the effects of a common teaching experience in mathematics and science classes. Most responses at Level 3 represented temperature on the vertical axis and time on the horizontal axis, probably due to exposure to this allocation, particularly with time on the horizontal axis, in news reports (e.g., Figure 2.02) or other experiences such as from science classes. In the current study, a few responses at various levels showed temperature in a thermometer form, giving some indication of one reason for assigning temperature to the vertical axis.

Another reason for the preference in axis allocation may be that students drawing coordinate graphs appreciated the nature of the covariation evident in the

constant temperature segment or in the implied dependence of temperature on time. This hypothesis is supported by the result that this preference was less evident at lower levels where covariation was lacking. Of students who drew inadequate coordinates (Level 2), many spatial variation graphs assigned both variables to the vertical dimension, some diagonal graphs assigned variables to the data series rather than axes, and the few who did assign one variable to each axis were evenly split in their allocation. Diagonal graphs at Level 2 notably did not visually acknowledge the constant segment, and spatial variation graphs did not make clear the dependence between the variables. It may be that for students who drew coordinate graphs, the constant segment or other repeated temperature value cued students to represent the repetition horizontally rather than vertically. Horizontal repetition retains the common visual property of line graphs involving no vertical repetition, that is, temperature as a function of time such that each time had only one corresponding temperature value.

Coordinate graphs were drawn by most secondary students and a few primary students, as shown in Table 5.01. This result is in line with curriculum expectations that recommend the teaching of coordinate graphing in upper primary school or junior secondary school (AEC, 1991; NCTM, 2000). The proportions of students drawing Appropriate Covariation graphs at Level 3 were similar to those for drawing test scores versus study times (Investigation 4A), slightly lower for the primary students than comparable students observed for the context of height versus age (Investigation 3A; Ainley, 1995), and slightly higher for the secondary students than comparable students observed for plotting height of a ball's rebound versus height of drop (Brasell & Rowe, 1993).

Student responses appeared to depend on both cognitive aspects concerning ways of structuring the complexity of series of data values for multiple variables, and

purpose-driven aspects concerning decisions about what is relevant to represent. In some cases, students judged the contextual narrative to be more significant to represent than the data values, and in other cases a data table was preferred to show the exact values.

The following investigation, involving discussions in interviews, explored further students' reasons for drawing graphs for tasks concerning time-based functional relationships or discrete value covariation. The analysis of the interview dialogue aimed to inform further the assessment framework for Coordinate Graph Production.

5.04 INVESTIGATION 6B: TEMPERATURE CHANGE OVER TIME (INTERVIEWS)

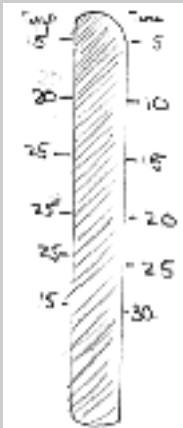
5.04.01 Introduction and Aims

A number of questions were raised from the results of Investigation 6A, related to the difficulty of discerning with confidence students' intentions or abilities from their survey responses alone. Investigation 6B explored explanations for why they drew these responses, aligned with Research Aim 3 of the wider study (Section 3.02), and provided further evidence concerning the assessment framework in relation to Research Aim 1 of the wider study. Investigation 6B aimed to gather evidence of the reasons for drawing the context only, tables of data, a single variable, line graphs rather than bar graphs, and graphs with conventional axis allocation. A second aim concerned exploring students' reactions to graphs drawn by other students, aligned with Research Aim 3 of the wider study (Section 3.02). Use of prompts to promote cognitive conflict explored the potential of students to acknowledge merit in other responses. Of particular interest was whether students'

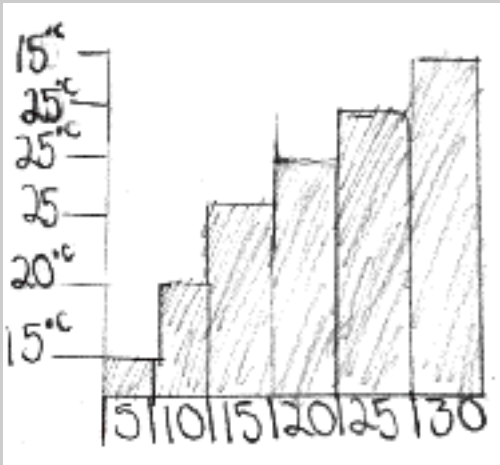
responses to prompts of other graphs depended on the differences between the levels of students' original responses and the levels of prompts.

5.04.02 Task

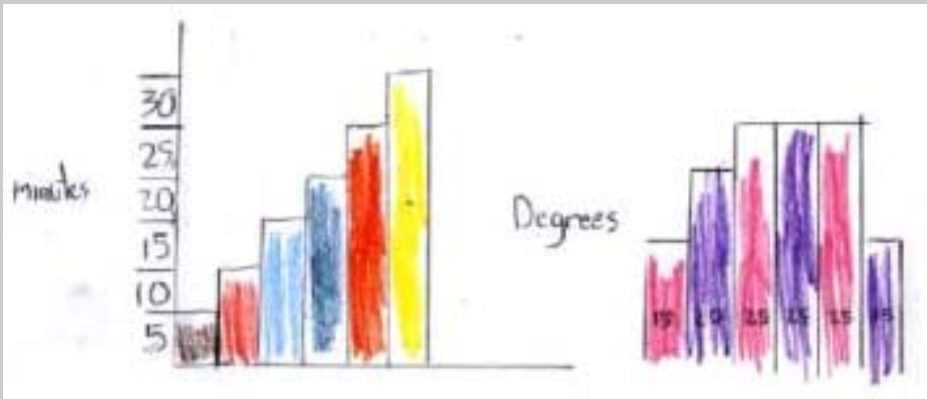
Students were asked to explain their own graph and then to comment on four graphs drawn by other students. Structured interviews were designed to allow students to talk about their own graphs and to prompt conflicting cognitions (Moritz, 1998) as per the general method described in Section 3.05.05. Four prompts were selected for this investigation, as shown in Figure 5.07, reproduced from R7 in Figure 5.04, and R10, R11 and R12 in Figure 5.05. As with other graphic prompts, criteria for selection of individual graphs as prompts included readability of labels and general presentation, and the set of prompts was designed to include various categories of response that had differing strengths and weaknesses. Prompt A was a Table (Category 1A) pictured as a thermometer, which emphasized correspondence at Level 1. Prompt B was a Diagonal Graph (Category 2A) that also emphasized correspondence. Prompts C and D were two different expressions of Spatial Variation Graph (Category 2B) at Level 2. Notably, no Level 3 prompts were offered, however three of the prompts were at Level 2. Students were hence not offered a complete resolution to issues of coordination, but were prompted with various inadequate attempts, creating the opportunity to suggest modifications that would show coordination. In particular, Prompts C and D were initially coded as being at Level 3 (See Section 5.03.03), and they displayed the general visual shape of the series of temperature values, which permitted comments about this feature and the option for students to make minor corrections to produce a Level 3 response.



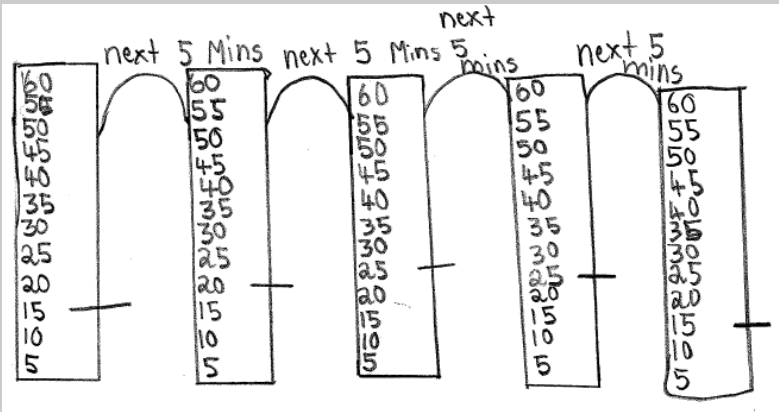
Prompt A



Prompt B



Prompt C



Prompt D

Figure 5.07. Graphs used as prompt in interviews.

5.04.03 Participants and Method

Investigation 6B involved interviews with 18 students from among those who had previously drawn a graph as part of Investigation 6A, as part of the general method for data Collection 4 (See Section 3.06.05). The 18 students interviewed from grades 3, 5, 7, and 9 were from among the 34 students interviewed for a variety of tasks, as detailed in Section 3.06.05: only 18 interviewed students engaged this task due to time constraints. The interviewed students offered a range of initial responses including all levels and categories excepting the Bar Graph (Category 3C), as detailed in the results that follow.

5.04.04 Results – Overview

A summary of the 18 interviews is shown in Table 5.02. Initial response categories refer to the categories labelled in Figure 5.01. For Prompts A, B, C, and D, the notations ✓, ✗, or ~ are used to denote approval, rejection or ambivalence/indifference about prompts, respectively. It should be noted that these summaries were not always clear cut: for example, some students approved of a prompt but commented about it lacking elements such as labelling. Where it was possible to discern, the allocation of approval or rejection was based on the structure of the graph for showing coordination, and not based on labelling or presentation.

Table 5.02.

Summary Descriptions of Interview Dialogue

ID	Initial Response Category	Prompt – Category of Prompt				Summary of Reactions to Prompts
		A – 1A	B – 2A	C – 2B	D – 2B	
G5m1	No Response	✗	✓	✗	✗	No initial response
G3f1	0A. Context	~	✓	✓	✗	Agree Higher (0→2AB)
G3f3	0A. Context	✗	✗	~	✗	No change
G7f2	0B. Graph form	~	~	~	✗	No change
G3f2	0B. Graph form	✗	✗	✗	✗	No change
G5f2	0B. Graph form	~	~	✗	✗	No change
G3m1	1A. Table	~	✗	✗	✗	No change
G3m2	1A. Table	✓	✗	~	~	Agree Same (1A→1A)
G5f4	1A. Table*	✗	✗	✗	~	No change
G3m3	1A. Table	✓	~	✓	✗	Agree Higher (1A→2B)
G5f3	1A. Table	N	✓	~	~	Agree Higher (1A→2A)
G5f5	1B. Single Variable	✗	✗	✗	✓	Agree Higher (1B→2B)
G5f1	2A. Diagonal Graph	✓	✓	✗	✗	Agree Same (2A→2A)
G7f1	2A. Diagonal Graph	✓	✓	✗	✗	Agree Same (2A→2A)
G3f5	2B. Spatial Variation	✗	✗	✗	N	Agree Same (2B→2B)
G5f6	2B. Spatial Variation	✓	✓	✓	✗	No change
G9f2	3B. Line Graph*	✗	~	✗	~	No change
G9m1	3B. Line Graph	~	~	✓	✓	Changed prompts to 3B
Total	✗	7	7	10	11	
Total	~	5	5	4	4	
Total	✓	5	6	4	2	

Note: ✓ denotes agreement, ✗ denotes disagreement, ~ denotes ambivalence, N denotes not administered, * denotes drawn in interview

The summary counts at the bottom of Table 5.02 indicate that the students interviewed were divided in opinions about Prompts A and B, but most rejected Prompts C and D. Most students who initially offered Level 0 responses did not accept the prompts at Levels 1 and 2, often rejecting them on the basis of not showing contextual information about the heater. Of students who initially offered Level 1 responses, predominantly Tables, some accepted Prompt A, which was

similar to their own responses, and one other prompt at Level 2. Students initially offering Level 2 responses commonly accepted Prompts A and B emphasizing correspondence at Level 1 and 2, and rarely saw strength in the prompts emphasizing variation (C and D). At Level 3, the two students responded in different ways, one accepting no prompts, and the other agreeing to the two Spatial Variation prompts (Prompts C and D) with modifications resulting in a higher level representation.

Students tended to accept or reject prompts based on a similar emphasis to their own initial response. Students whose initial responses emphasized variation often commented on temperature variation, whereas those who emphasized correspondence often read pairs of values from the prompts. Further details are discussed for each student interviewed in the order presented in Table 5.02.

5.04.05 Initial Response Level 0 – Nonstatistical

Of the eighteen students interviewed, six offered Nonstatistical responses to the survey task. One was no response, and the other five drew responses as shown in Figure 5.08.

One fifth-grade male (G5m1) offered no response for this task, even when given an opportunity in interview. When asked what made this task hard, he commented “it has, like, four things they did,” referring to turning the heater on, turning it off, opening the window, and the data recording. He also commented, “It’s kind of already a graph,” and thus found little purpose in re-presenting the data graphically. In minimal responses to the four prompts, he presented weak evidence of understanding by occasionally reading single values. He found Prompts A, C, and D “confusing” or “not really that clear.” He read a value from Prompt B, commenting, “Yeah, that makes sense. It’s quite clear. Like it’s 15 degrees....” This reading of a

single temperature value was not mentioned with a corresponding time, and in general the student appeared not to appreciate the features of Level 1 or 2 prompts.

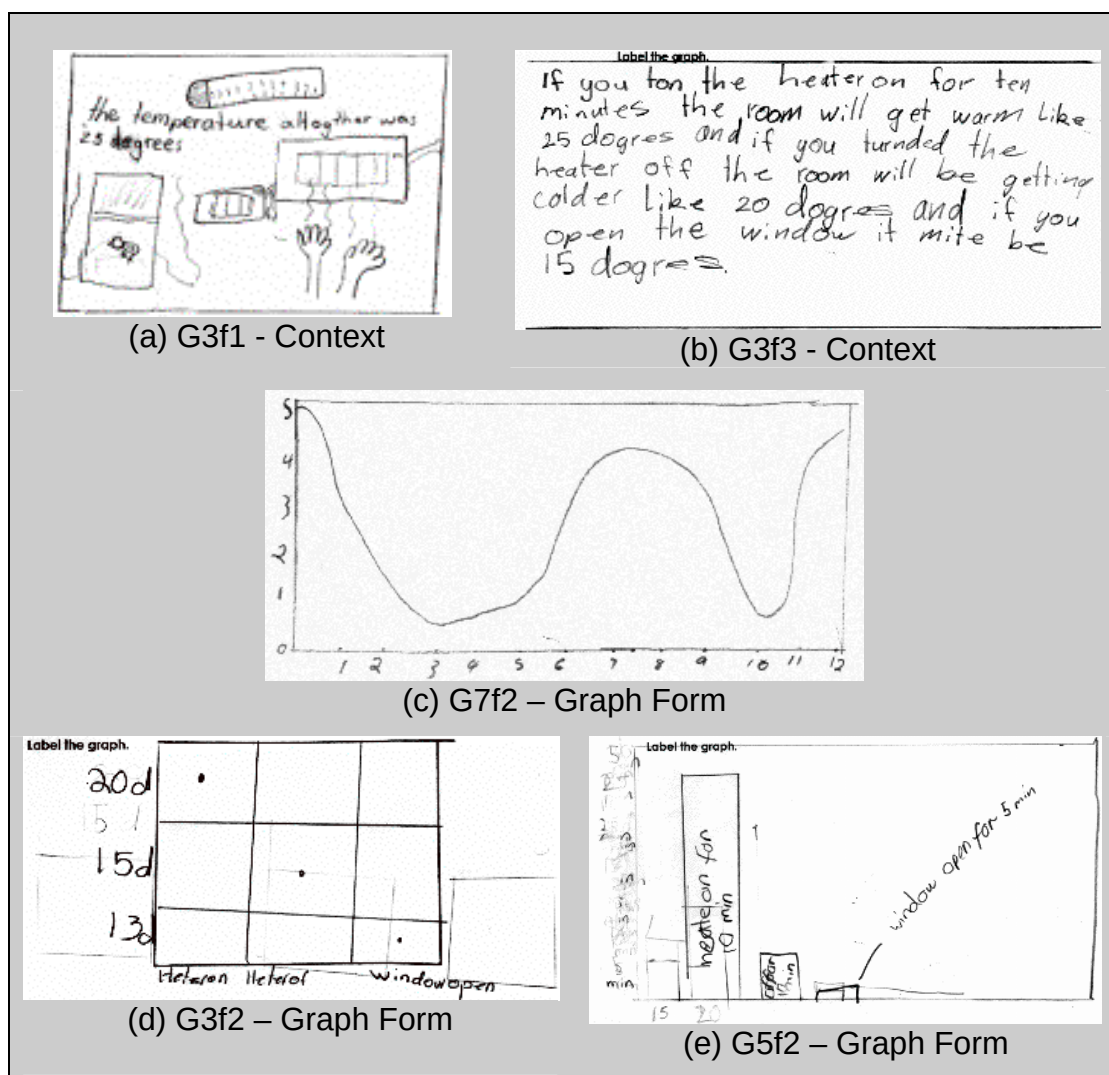


Figure 5.08. Survey responses – Level 0 – Nonstatistical.

Two third-grade females (G3f1, G3f3) offered contextual responses, as shown in a picture (Figure 5.08a) and a narrative (Figure 5.08b). Both indicated temperature values and contextual events without any reference to corresponding time values, and to the interview prompts, both offered responses focusing on these aspects. Both found Prompts A and B initially confusing, particularly the inadequate labelling of B. Both partially appreciated the up and down variation of temperatures shown in the right graph in Prompt C, but were concerned about the contextual events of the heater and window status more than the time values.

- G3f1: [Prompt C] That one's probably better than that one, I suppose [...] Well, they have, well that's going up, up, up, down, so like, they must mean 15 degrees there [Prompt C, right graph, rightmost column], and then here [Prompt C, left] they turned it, I think they opened the window for 5 minutes, and then they've got the heater, I think, and then they've got, they must have done something else, and then it goes up to 30, so that must mean 30 minutes or something.
- G3f3: [Prompt C] No, because you don't know what's opened or anything, so you don't know whether the heater's on, or that or anything. [...] Well, at first it's going up, and then goes down, back down to 15 [pointing to top of bars in Prompt C right, reading columns left to right].

Despite these promising responses to Prompt C, both G3f1 and G3f3 found Prompt D confusing. G3f3 was so concerned with the contextual elements of temperature change that in reading the markers on Prompt D, she believed, despite the data in the table, that the temperatures should keep increasing, "I don't really get it, because if it gets hotter, it should have kept going up instead of staying on the same number." In summary, for Prompt C (Level 2), these students extracted only the single variable (Level 1) temperature, and showed interest in temperature variation, but not in correspondence with time.

The three other students interviewed who gave Nonstatistical (Level 0) written responses included females from the seventh-grade (G7f2; Figure 5.08c), the third-grade (G3f2, Figure 5.08d), and the fifth-grade (G5f2; Figure 5.08e). They each presented a Graph Form (Category 0B) with some fictitious data, and did not accept any of the four prompts, finding them either confusing or not showing the contextual elements they considered important. The student emphasizing variation (G7f2) represented a continuous line of unlabelled waves going up and down from 0 to 5 across values 0-12, presumably intended as temperature variations. She continually referred to the temperature going up and down, often motioning with her hands the shape of her graph, and thus did not accept Prompts A and B.

- G7f2: [Prompt A] Umm.. yeah, it's a good way of showing it, but to me it doesn't show much. It's just showing a whole heap of lines.

Compared from mine, it actually shows where the level is [hand in air, motioning rounded waves of students' graphs], but that one there, it just doesn't appeal to me.

She more simply rejected Prompt B, not appreciating values or temperature variation. She considered the right graph in Prompt C to be "okay" because it "shows you what rate you're at." This reference and that in the extract to Prompt A were her only references to static value levels. She never read a numerical value, and when initially asked about the times on her own graph, claimed they referred to 1 o'clock, 2 o'clock, etc. Most of her language concerned changing temperatures going up and down, however she failed to appreciate the spatial variation in Prompts C and D. Overall, her primary concern was the contextual awareness of temperature variations, with minimal interest in values for time or temperature, nor in realistic graphical representations of temperature change. In this respect, she had more in common with interview responses of others emphasising variation (G3f1, G3f3) than she did with those emphasising correspondence (G3f2, G5f2).

Two other students representing a Graph Form with fictitious data (G3f2, G5f2) appeared to represent three temperature values in bars corresponding to the states of heater on, heater off, and window open. G3f2 rejected all prompts as not showing whether the heater was on or off, despite acknowledging that the vertical axis of Prompt B showed temperatures and the horizontal axis times when asked by the interviewer what these axes might be showing. G5f2 initially accepted then rejected each of Prompts A and B, and rejected Prompts C and D, claiming none made sense to her. For Prompts A and B she read the numbers as pairs, "30 with 15, 25 with 25...", but did not appreciate the correspondence, claiming neither prompt made sense to her. For Prompt B, she claimed that the top value on the vertical axis "should have been the highest, because if it was on for 30 minutes, then it should be more like 25 or something [...] 30 or something, 40 or something like that, higher,"

that is, an overriding concern for corresponding number values in a pattern irrespective of the data available. She was totally unaware of the true variations in the data, namely the drop in temperature evident in the table.

In summary, these six students in the interviews maintained their overall lack of reference in the data set presented in the table. Some read coordinates, but did not appreciate them as important. More often they were interested in showing only temperatures, occasionally with variations in temperatures, but with reference to contextual events (e.g., heater on/off) rather than time values.

5.04.06 Initial Response Level 1 – Single Statistical Aspect

Six students interviewed had offered responses in the survey attempting to display the data but involving a Single Statistical Aspect. The six responses are shown in Figure 5.09. One of these displayed the temperature values only in a pie graph, and the others showed a table of values aligned along a single axis similar either to the table given or to Prompt A. The emphasis of five of these six responses was representing temperature values and corresponding times as numbers, without using graphic space to show variation of values. The six included three third-grade males and three fifth-grade females, one of whom drew Prompt A (Figure 5.09e).

Two third-grade students (G3m1, G3m2) often gave surface agreement to prompts, but then admitted they found them confusing, often reading isolated values. G5f4 showed concern with context and with correspondence in her own graph (Figure 5.09c), and evaluated prompts based on these aspects as well as presentation aspects concerning colour and labelling.

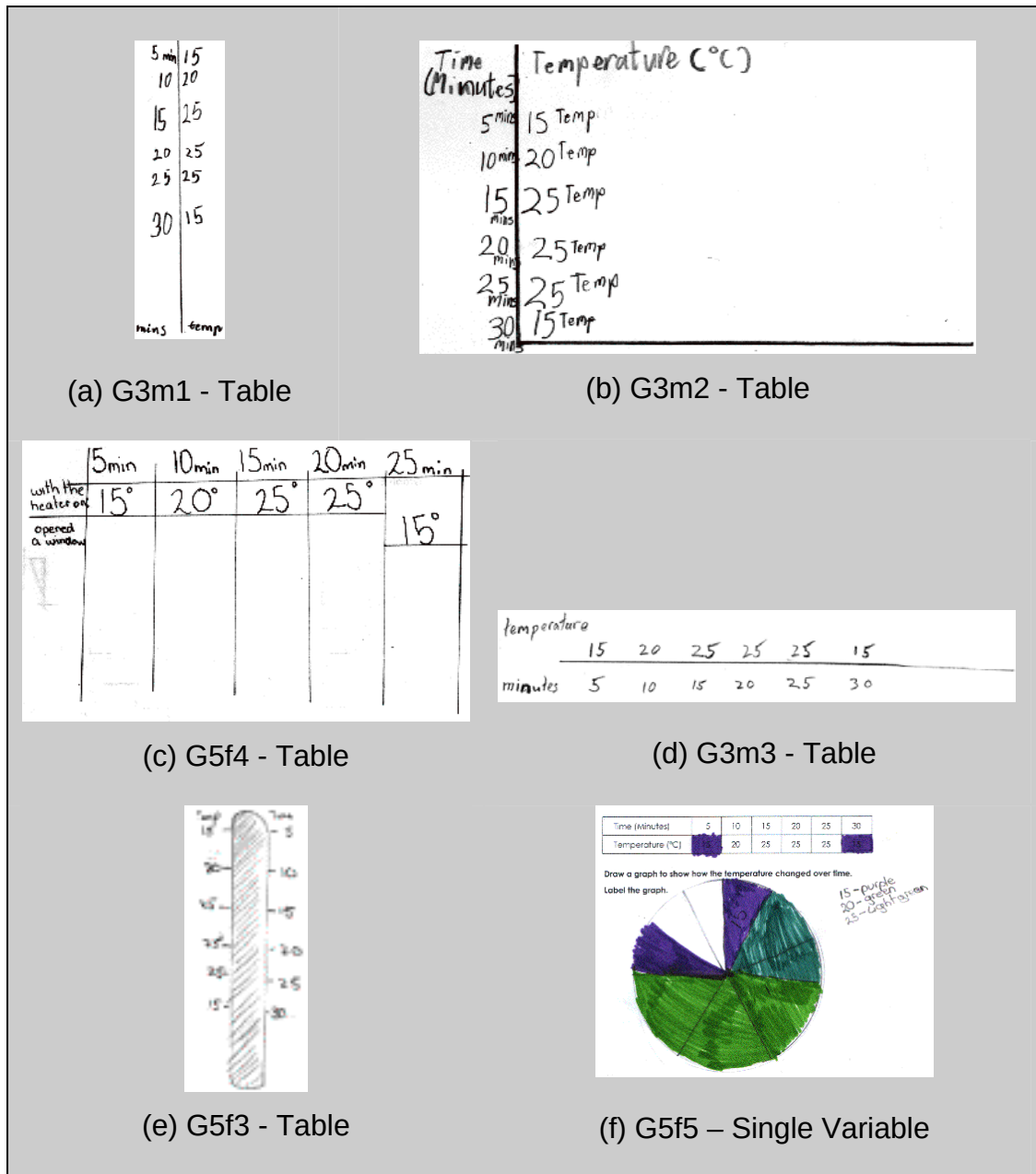


Figure 5.09. Survey responses – Level 1 – Single Statistical Aspect.

One third-grade male (G3m3), who drew Figure 5.09d, considered the original table, his own graph, and Prompt A as about the same, although A was like a thermometer. For Prompt B, he cross-referenced each data pair with the table, and then began to address spatial aspects.

G3m3: [Prompt B] But why would you go 15, 20, 25, 25, 25, 15 [...] I thought you might go up in numbers like 15, 15, 20, 25, 25, 25.

His comments about Prompt C followed a similar pattern, first cross-referencing values, and then addressing spatial concerns. He noted that in the right-hand graph, “that 15 is higher than that 15,” thus compromising the scale as a device for graphic position to denote value. He liked how the right graph was set out more than his own, indicating where the temperature went up and down, and that this showed an extra way to just writing the values.

G3m3: [Prompt C] So it does show it two ways, that way [vertical motion for heights of bars] it is quite clear, and the temperature is the writing [written values]. They could make the writing a bit clearer, like put it underneath or something.

I: ...Do you think it helps to use the bars like that, or it doesn't really?

G3m3: Well, it does show an extra way, for the temperature to rise and go up.

When examining Prompt D, his major interest remained in written values. The numbers up to 60 were mistaken for a scale of minutes, and even once corrected, he lapsed back into this conception. The concern for no duplication of the “25 minute” value showed a focus on individual data points, but with no recognition that the markers displayed a shape as in Prompt C displaying the rise and fall of the temperature. In summary, this student was able to engage, briefly, the idea that “position denotes value.” The focus, however, was often to ensure correspondence of each pair of data values, and not to represent shape in two dimensions.

The fifth-grade female (G5f3) who drew Prompt A (also Figure 5.09e) explained “Because thermometers measure the heat in the room, and so I put one of them down.” She considered Prompt B to be a good way to draw it, but preferred her own graph as “more easier to understand, for me. That’s probably because I was the one that did it.” For Prompt C, her interest in correspondence led to attempts to improve this aspect, verging on development of a coordinate system. She found it initially confusing, but then in cross-referencing the values from the left and right graphs as pairs (“5, 15; 10, 20”), she had an idea to overlay the written values of the right

graph on the left graph horizontal axis, an idea which she rejected, possibly confusing which axis was being overlaid, and could not resolve this issue to the level of Appropriate Covariation (Level 3).

G5f3: [Prompt C] I think this is very hard to understand, because there's one part over here, and another part over there, so you've got to turn your head and... it does make sense but...5, 15: 10, 20 [reading matching columns of left and right graphs] if they'd put this part [left graph] sort of down here [horizontal axis of right graph] like just the 15, 20, 25, oh no, that wouldn't work...

I: Why wouldn't it work?

G5f3: Because then the 5 minutes and the 15... but if they put, like, that thing there [right graph, vertical axis] and put like 5, and then 10, and then 15 [tries to overlay times into vertical axis of right graph] and that would have... somehow got 30 [that is, height on bars in right hand graph would not go high enough]

For Prompt D, she also showed promise, commenting it was good but lacked a third value of 25 degrees, and it lacked actual minute values. In trying to demonstrate how the graph could be improved, she again became confused about the variables.

The student (G5f5) who displayed a single series of temperature values used a pie graph of various sector sizes and with times absent. Her reason for drawing the pie graph was idiosyncratic.

G5f5: Well, with temperature and things, my mum said to draw a circle, so that it's easy to understand. And like the circle of life and everything, and the circle of temperature, so I just kind of did that.

She found Prompts A, C, and D confusing, and although she could read corresponding values in Prompt B, she did not appreciate it as any better than her own graph showing only temperature values. She asserted, "It doesn't matter what shape it is, because any shape, it could be square, oblong, any kind of shape, it's still a graph. So really it doesn't matter." Once explained by the interviewer, she preferred Prompt D, not for showing variations in temperature as a variable, but rather as a set of values and "I noticed that it's got 15, and there's a marker, so it's like a thermometer stuck in the mouth." Thus although interested in presentation

aspects, she was not concerned with time values or spatial variation, but only in representing the temperature data.

In summary, the six students who drew Level 1 responses were concerned to represent the data values, often aware of presentation aspects that made reading data confusing, such as inadequate labels.

5.04.07 Initial Response Level 2 – Inadequate Covariation

Four students initially drew Inadequate Covariation (Level 2) response. These responses are shown in Figure 5.10.

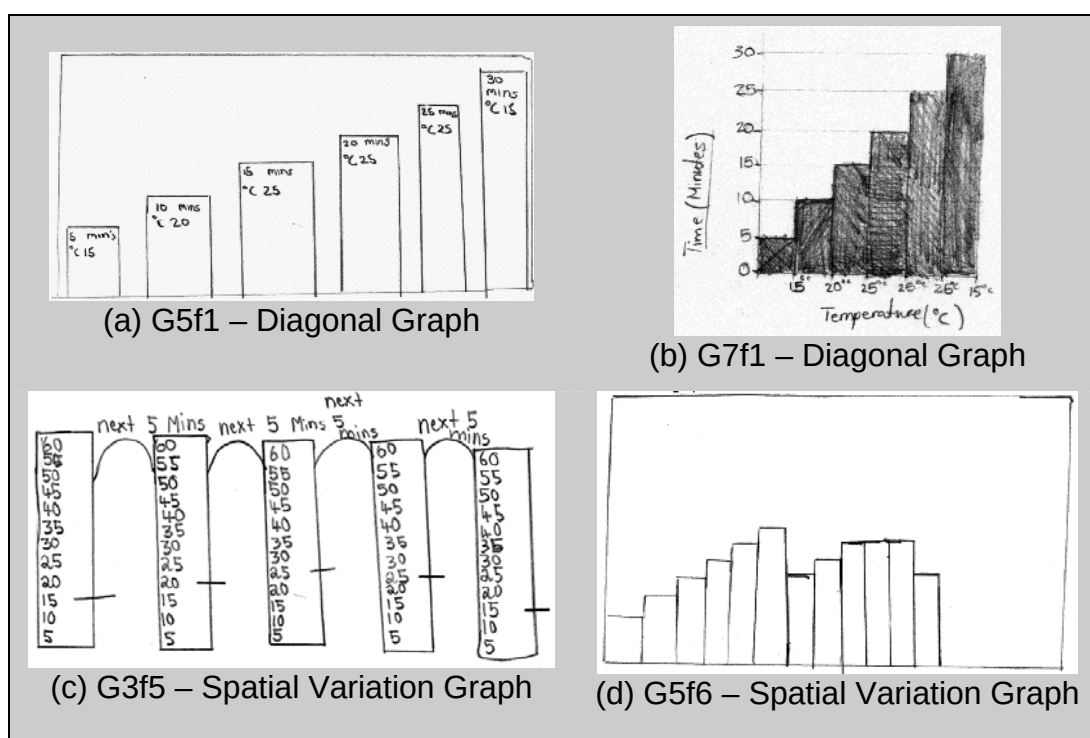


Figure 5.10. Survey responses – Level 2 – Inadequate Covariation.

Two students (G5f1, G7f1) drew Diagonal Graphs, as shown in Figures 5.10a and 5.10b. Both provided similar response patterns in explaining their own graphs and in responding to the four prompts. In explaining their own graphs, they both provided very simple justifications that read the first coordinate and then made some comment about the progression.

G5f1: Well, um 5 minutes, it's got 15 degrees, because it's down lower [leftmost column] and it's [rightmost column] up higher because it was for 30 minutes, and it was 15 degrees after they had opened the window.

G7f1: Well, in 5 minutes it was only 15 degrees Celsius, and it kept getting hotter as they left it on for longer.

Both tentatively commented Prompt A was good, although G7f1 noted the alignment of pairs was not completely accurate. Both preferred their own graphs. For Prompt B, both considered it fairly similar to their own, although G5f1 slightly preferred Prompt B without clear explanation why. Prompts C and D were considered confusing and not as good as their own.

G7f1: [Prompt C] Well it does show the information, [inaudible] I suppose it'd be a bit easier to read if they were both in one. Because this way, you have to look at one then look at the other. [...] Well, it shows like the degrees, how they go up then down and stuff.

This inkling of use of space to denote value, such as “up and down and stuff” variations appeared to border on acknowledging these points.

Two students (G3f5, G5f6) drew Spatial Variation graphs. Both explained their graphs emphasizing progression or variation in “it,” implying temperature, and did not refer to times or time values explicitly.

G3f5: Well, when they opened the... when they turned on the heater, it went up to 15, probably because it was hot, and then it's, sort of like, going up in 5s, and then going back down again.

I: And, in how you drew that, you've shown that it goes back down again, how do you show that?

G3f5: Well, there's 15, then 20, 25, 25, and then 15.

G5f6: It goes up [pointing to leftmost “time” columns] and then it goes down [pointing to rightmost “temperature” columns].

G3f5 who drew Prompt D gave minimal response to other prompts, generally finding them confusing. G5f6 also gave minimal responses to explain acceptance of Prompts A and B, though suggesting they were clear, possibly due to the labelling

not present in her own graph. Prompt C was similar to her own graph, and she suggested it was good, though just as for her explanation of her own graph, she referred to the graph going up and then down, with confusion or ignorance as to the left hand part of her graph or of Prompt C.

G5f6: [Prompt C] It's good. That [left graph] one's not very clear,

I: What makes it not very clear?

G5f6: Umm [8 second pause]... Because it goes up, and it's not like that one [right graph] [inaudible] it gets taller and then it gets smaller again.

She found Prompt D confusing. As noted in Investigation 6A (Section 5.03.03), in preliminary coding of responses Prompts C and D (see Figure 5.07) and other similar responses (e.g., G5f6, see Figure 5.10d), there was consideration of whether these responses (a) failed to show correspondence of values (Level 2) or (b) showed correspondence unconventionally (Level 3). Examination of the interview dialogue of two students, one offering Prompt D and the other G5f6, indicated that their intention was to show variation in temperatures and that correspondence of values was lacking, hence supporting the classification at Level 2.

In summary, these four students did not use the prompts at Level 2 (Prompts B, C, and D) to resolve the issue of a coordinate system. Their responses were often brief but mentioned some sense of comparing values or progression of variation, such as “going up.” Similar comments were sometimes observed from students who drew responses at Level 0 or Level 1, but at Level 2 they were characterised by the lack of attention to data values or contextual information. This move away from attending to values, or corresponding pairs, towards considering the series of data, showed some development of features more clearly expressed by students at Level 3.

5.04.08 Initial Response Level 3 – Appropriate Covariation

Two ninth-grade students (G9f2 and G9m1) drew conventional coordinate graphs at Level 3, as shown in Figure 5.11.

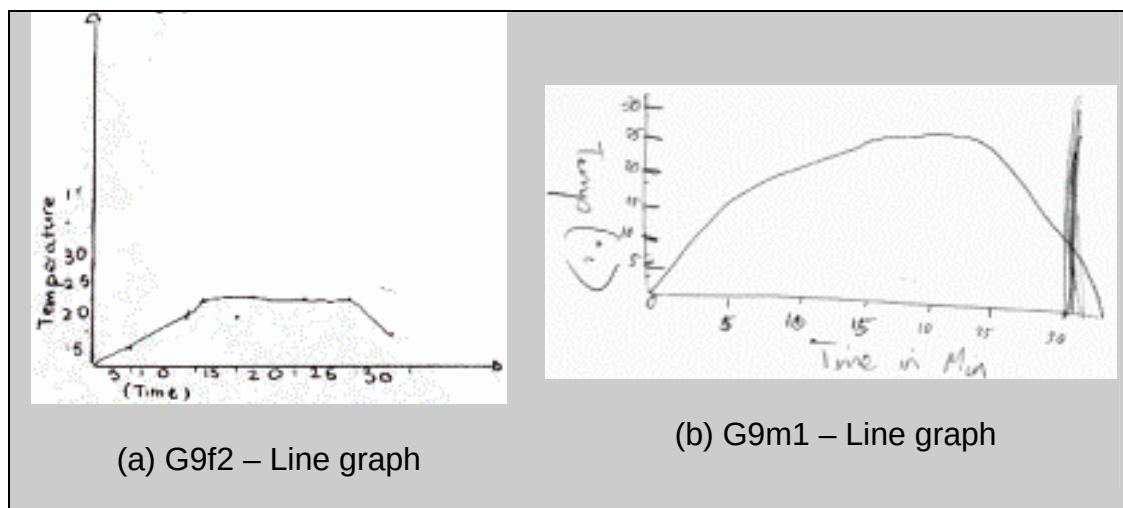


Figure 5.11. Survey responses – Level 3 – Appropriate Covariation.

The ninth-grade female (G9f2) had not had time to respond to this item in the survey, and did so in interview. Her response provided an interesting insight into the construction of Diagonal Graphs. She began by ruling the axes, labelling them, numbering the time axis, and then numbering the temperature axis using the results. She began plotting points, then after the third point, paused and erased the vertical axis numbers and rewrote them, and then completed the points and line, which still showed remains of the “15” at the top of the vertical axis (see Figure 5.11a). The interviewer asked her to explain, “when you were doing the axis up there, what made you change your mind there?” and she responded “Umm, because I wrote the results instead of the actual just, degrees, to show...” In response to Prompt B, a Diagonal Graph, she commented as follows.

G9f2: [Prompt B] Umm. Well, it just says 5 minutes was 15 degrees, 10, 20, so it's just a bar graph [...] But... [5 second pause] it's just the results, and the time across the bottom, but usually you'd put just 10, 15, 20, 25, 30 [vertical axis], and then do it along those numbers, instead of actually putting the results up the side.

She considered Prompts A, C, and D were not clear ways of showing the data, although she saw some merit in Prompt D, “it shows the information, you don't know what's what when you first look at it.” This student seemed to be on the transition

from Level 2 to 3, and thus could not recognise the data shape in Prompt D and the right graph of Prompt C as being the basic required graph element.

The ninth-grade male (G9m1) demonstrated a command of individual values but a greater interest in the global pattern, repeatedly referring to “inclines,” “planes,” and “dropping back down.” He was able to appreciate merits of Prompts A and D: Prompt A was described as “another picture way of showing it,” and Prompt D as “pretty simple, you’ve got to look at it a bit though.” He also suggested modifications to prompts that would generate conventional coordinates: for Prompt B, inverting the axes and the scale of times to form an inverted U shape, and for Prompt C, combining the times onto the temperature graph.

5.04.09 Discussion of Investigation 6B

Many students found difficulty articulating the reasons why they drew their own graphs as they did. They appeared to assume their reasons were self-evident, and focussed on the steps of drawing the representation. These responses gave little evidence of sources of graph form exposure that were not already expressed in earlier interview investigations such as drawing height versus age. For this task, many students commented on the presence of relevant information, whether contextual, time values, or data values. These emphases in dialogue, and lack of comment on other aspects, perhaps accounts for why many students offered Context or Table responses. Language related to correspondence was mostly implicit: students offering responses in Category 1C rarely articulated the aspect of corresponding values built into their tables. The language of variation, however, was hinted at by students who had drawn graphs at various levels, for example comments describing it getting hotter such as “it goes up.” The implicit references to time were translated to the Single Variable graphs showing only temperature, as seen in

Investigation 6A, and the language of height (“up”) to refer to the variation in the value of the explicit measure (temperature) may explain the strong preference for allocating “it” (temperature) to “up” (vertical axis) in Investigation 6A.

Students’ responses to prompts were not always clear cut. Dialogue often included a combination of praise and criticism about various aspects of the prompts, and changes in opinion after confusing features were comprehended. For students who initially responded at Level 0 and Level 1, often their interest in contextual elements and data values respectively was their basis for judging prompts. In general they did not appreciate the design features of Level 2 prompts (B, C, and D), often finding them confusing, or not particularly valuing these features. In a few cases, students who drew Level 1 responses attempted but failed to resolve the issues in these prompts to develop a coordinate system. For the four students who offered Level 2 responses, the prompts at Level 2 did not contribute ideas that helped them resolve coordination issues either, whereas a student who offered a Level 3 response did contribute ideas to improve Level 2 prompts to Level 3.

Students often could recognise the merit of representations emphasising either correspondence or variation consistent with their own response, but they had difficulty seeing the other aspect of covariation. Thus although students rarely articulated clearly aspects of correspondence or variation, they appeared to notice these aspects within the four prompts. This finding offers qualified support for importance of these two aspects of covariation for coordinate graph production.

5.05 GENERAL DISCUSSION – COORDINATE GRAPH PRODUCTION

The assessment framework for Coordinate Graph Production, shown in Figure 5.01, involved four levels matching those for Speculative Data Generation (see Table

4.01). Survey responses were presented from Investigation 6A that illustrated graphic responses within this framework, and verbal interview dialogue from Investigation 6B explored further this framework. The levels of response were supported by both investigations: interview dialogue often showed emphasis of contextual information without concern for data values for students who drew graphs at Level 0, emphasis on data values for students who drew graphs at Level 1, and attempts to coordinate data for students who drew graphs at Level 2. Further, students exposed to prompts often recognised merits of responses at similar levels to their own, and only sometimes engaged ideas embedded in responses at higher levels (that is, Level 2).

The aspects of covariation – correspondence and variation – were considered as significant for coordinate graph production. In particular, Tables (Category 1A) emphasised correspondence, whereas Spatial Variation graphs (Category 2B) emphasised variation but without clear correspondence. In Investigation 6A, Level 1 responses represented data values, but most were in the form of a data table similar to that provided. Investigation 6B showed these students could rarely articulate the aspect of correspondence built into the table, but many verbally described comparison or variation of temperature values without any explicit reference to time. The role of correspondence and variation was most significant at Level 2 in the lack of resolution of both aspects to present a coordinate system.

The importance of correspondence and variation in constructing a coordinate system were aspects identified by others researchers (e.g., Clement, 1989; Nemirovsky, 1996; Wavering, 1989). In this respect, the framework more closely resembles frameworks from studies of interpreting and generating data-based covariation (Chapter 4) than other frameworks of graph construction (Chick & Watson, 2001; Jones et al., 2001). Correspondence and variation were used to assign both Tables of data (Category 1A) and Single Variable responses (Category 1B) to

Level 1, in contrast to Chick and Watson (2001) who considered univariate graphs at a higher level of graphing than lists of values.

Students who did not display the given data (Level 0) were found often not to be attempting to represent the given data, but were concerned predominantly with the context. An intentional feature of the task design was offering information regarding temperature values changing over time in relation to contextual events. A range of tasks for coordinate production have been used in previous research, some providing labelled axes requiring scaling and point plotting, some with unlabelled axes also requiring a decision about axis allocation and labelling (e.g., Bell, Brekke, & Swan, 1987a, 1987b), and others with a blank starting point but raw bivariate data without other contextual information (e.g., Swatton & Taylor, 1994). The task used in this investigation left open not just how to structure and present a graph, but also the decision of whether to represent information in a picture, table or graph, and what to represent. Many students drew tables of data, and when interviewed, these students often read corresponding values from tables or graphs such as Prompts A and B. For Prompts B, C, and D, in some cases they showed evidence of some interest in displaying variation, but could not resolve the problems of coordination in these Level 2 graphs.

A central purpose for Coordinate Graph Production is to aid data interpretation. The task in Investigations 6A and 6B supplied a set of six bivariate values readily interpreted without a graph. Six values were used to make data complexity across tasks comparable, and to reduce the time needed for manual graph production. The purpose for the graph production in this task was in the communicative expectation of drawing a graph to show how the temperature changed over time. The following chapter examines aspects of Graph Interpretation.

CHAPTER 6. GRAPH INTERPRETATION

6.01 SUMMARY

This chapter discusses three investigations of students' interpretations of graphs. Each investigation involved verbal interpretation and numerical interpretation of graphs that showed one variable decreasing as the other increased.

Investigation 7 involved survey responses to interpreting a newspaper advertisement including a bar graph with embedded pictures that showed how telephone call rates decrease for longer phone call durations. Questions concerned generalizing from the graph, identifying unusual features, and calculating numeric values based on the graph. Investigations 8A and 8B involved survey and interview responses to interpreting a scattergraph showing that the level of noise was generally lower for classes with greater numbers of students. Questions concerned generalizing from the graph, reading values, interpolating, and judging an appropriate verbal generalization.

The descriptive labels of four levels of Speculative Data Generation – Nonstatistical, Single Statistical Aspect, Inadequate Covariation, and Appropriate Covariation – were used to describe Verbal Graph Interpretation and Numerical Graph Interpretation. Nonstatistical responses referred to the context or visual features of graphs, such as dots. Single Statistical Aspect responses either referred to a single data point, such as reading a coordinate value, or referred to a single variable. Inadequate Covariation responses either (a) referred to selected points, for example extreme points or using nearby points to predict a value, or (b) referred to both variables but did not verbalise the association in the appropriate direction. Appropriate Covariation responses verbalised a trend appropriately, and made use of a trend to predict numerical values.

6.02 INTRODUCTION

Graph interpretation involves providing numerical or verbal responses based on a graph, and as such, it involves the reverse translation to either Coordinate Graph Production or Speculative Data Generation examined in previous chapters, as shown in Figure 1.01. In general, graph interpretation tasks involve the questioner in providing a specific graph from among many graphs of the same data that might differ in form, scale, or axis allocation. Graph interpretation tasks generate student responses, dependent on that graph that are verbal or numerical, with limited variations to classify, in comparison to tasks involving Speculative Data Generation or Coordinate Graph Production, in which students' graphical responses, using text, numbers and visual features arranged in graphic space, require a classification system that accommodates a range of responses across a number of aspects. Questions of Verbal Graph Interpretation involving covariation included (a) generating a verbal statement, and (b) judging whether a given verbal statement is appropriate. Questions of Numerical Graph Interpretation included (a) reading values, (b) comparing or computing with values, (c) predicting values by extrapolation or interpolation, and (d) interpreting unusual scales.

Pinker (1990) has suggested that graph comprehension involves two fundamental components: (a) comprehension of the axis framework and scale, and (b) comprehension of the data elements. Comprehending scale is necessary for reading numerical values, whereas the relative positions shown for data cases permit trend identification and qualitative comparison of cases without reference to the scale. This distinction is closely aligned to the distinction between skills of Numerical Graph Interpretation and Verbal Graph Interpretation as sustained in this chapter.

Frameworks for analysing student responses to tasks of Verbal Graph Interpretation and Numerical Graph Interpretation are shown in Table 6.01. Response levels for both Verbal Graph Interpretation and Numerical Graph Interpretation were characterised by the degree to which responses incorporated the components of covariation embedded within the graph to achieve more global interpretations. Verbal Graph Interpretation responses required reference to relevant variables and to an appropriate correspondence. For Numerical Graph Interpretation tasks, some questions only targeted Level 1 responses based on reading single values from a graph, whereas other questions sought numerical responses based on appreciation of the covariation presented in the graph, such as appropriate interpolation.

Within each level, various types of response to individual questions were observed, however distinct categories of response were not differentiated as they were for frameworks for Speculative Data Generation and for Graph Production. Distinctions between response categories did not emerge in preliminary coding of the responses, and were considered unlikely to be sustained due to the verbal and numerical response format, which was less specific in structure of correspondence and variation than graphic responses for Speculative Data Generation and for Graph Production. In general, however, it can be observed from the descriptions in Table 6.01 that Verbal Graph Interpretation questions tended to emphasise variation in verbal references to variables, whereas Numerical Graph Interpretation questions tended to emphasise correspondence, for example by reading one value from a given corresponding value.

Table 6.01.

Characteristics of Four Levels of Verbal and Numerical Graph Interpretation

Level	Verbal Graph Interpretation	Numerical Graph Interpretation
0. Non-statistical	Non-response, or refers to: (a) context but not variables or the association, or (b) visual features, e.g., “dots.”	Non-response, or refers to: (a) context based “guesses,” or (b) visual features, e.g., the maximum on the scale.
1. Single Statistical Aspect	Refers to either (a) a single data point, or (b) a single variable (dependent).	Refers to single data points, e.g., reads a value given a corresponding bivariate value. Interpreting scale concerns data values.
2. Inadequate Covariation	Refers to both variables but: (a) <i>correspondence</i> is noted by comparing two or more points without generalizing to the entire data set, or (b) <i>variables</i> are described but the correspondence is not mentioned or is not in the correct direction.	Reads values and interprets them within context. Interpolates within local range by referring to two or more data points, but without generalizing to the entire data set. Interprets scale with concern for direction.
3. Appropriate Covariation	Refers to both variables and indicates appropriate direction.	Interprets relationships between values. Interpolates or extrapolates with accuracy by referring to multiple points or general trend. Interprets issues with scales reversed or non-linear.

6.03 INVESTIGATION 7: TELEPHONE RATE VERSUS CALL DURATION

6.03.01 Introduction and Aims

Investigation 7 re-analysed responses from a previous study of graph interpretation (Moritz & Watson, 1997). Additional student responses collected in late 1997, not available at the time of the study by Moritz and Watson, increased the pool of student responses. This investigation advanced the analysis previously undertaken to clarify levels of response for Verbal Graph Interpretation and Numerical Graph Interpretation, as described in Table 6.01. This aim was to develop an assessment framework, aligned with Research Aim 1 of the wider study (Section 3.02). In addition, this analysis of levels, and the linking of student responses to responses by the same students to Task 1 (Heart deaths) supported consideration of association between Speculative Data Generation, Verbal Graph Interpretation and Numerical Graph Interpretation, as explored in Section 7.03.

6.03.02 Task

Prior to the commencement of this doctoral study, Moritz and Watson (1997) reported on a study of 1584 responses by sixth- to eleventh-grade students to Task 2 shown in Figure 6.01. The graph has a number of interesting features. The most apparent is the presentation of unusual pictures of squashed telephones within the shaded region of a bar graph. Each of the variables of the graph involves potential for misunderstanding. The rate of the telephone call involves the proportional notion of cost per unit time, which may be confusing because time is the other variable – the rate-value confusion referred to by previous researchers (e.g., Mokros & Tinker, 1987). Further, the rate is not directly presented in the graph as a cost per minute, but

rather as a percentage off the standard rate, and thus the percentage values reported are negatively related to the actual charge rate. The minutes of the call are discrete time intervals of unequal spacing. The data involve a negative association of five values, complicated by the scaling of the graph, which is uneven on both axes and does not begin at zero on the vertical axis. In short, this graph from a newspaper advertisement is authentic to the many complexities of graph interpretation to which students may be exposed in social contexts.

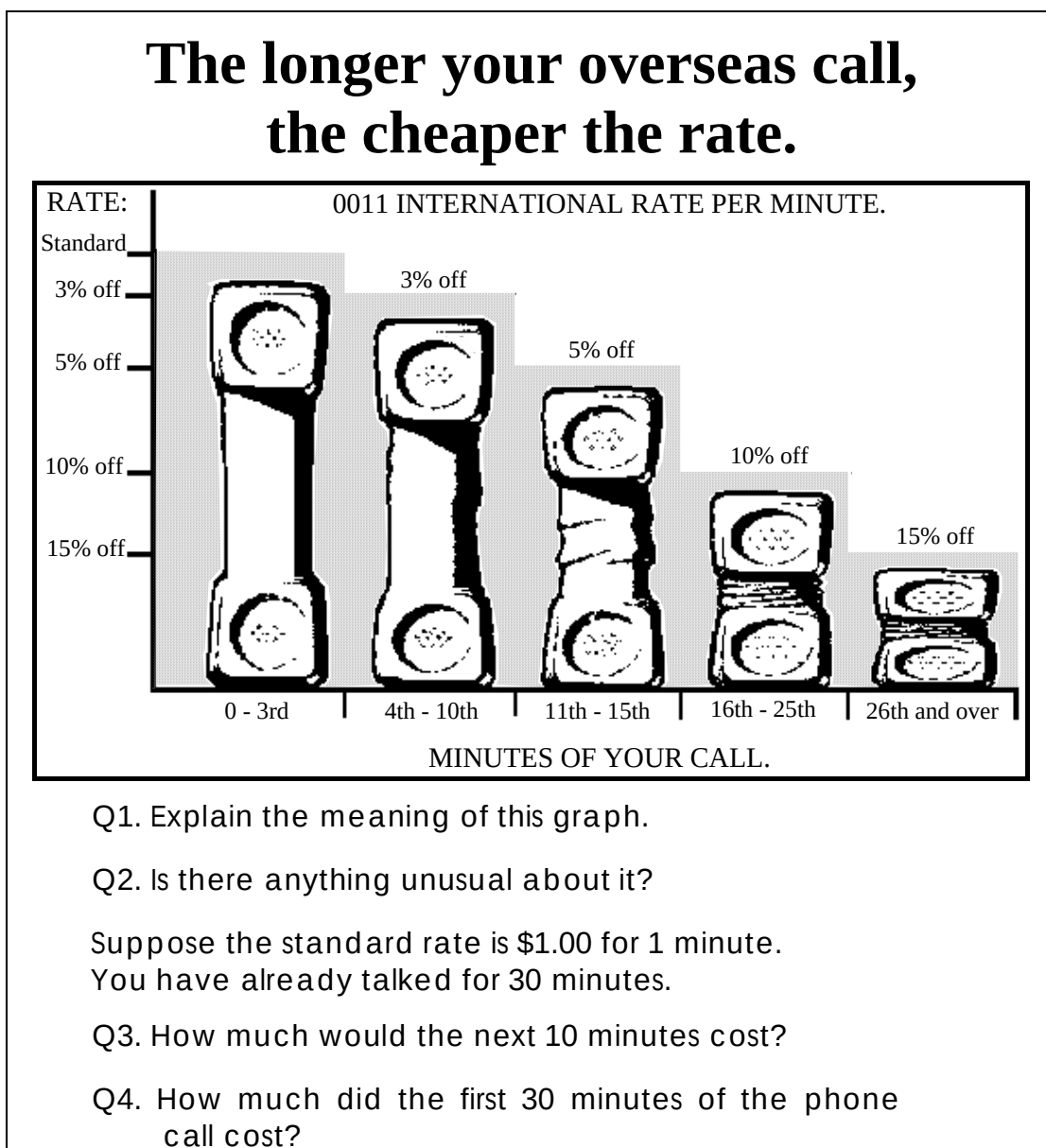


Figure 6.01. Task 2 to assess Verbal Graph Interpretation (Q1) and Numerical Graph Interpretation (Q3 and Q4; Q2) based on a newspaper extract (*The Mercury*, 22 July, 1993, p. 17).

6.03.03 Participants and Method

The data collection followed the method specified for Data Collection 1 in Chapter 3. Responses were coded by successive refinement into the four levels as described in Table 6.01. Q1 contributed to assessment for Verbal Graph Interpretation. Q3 and Q4, assessing Numerical Graph Interpretation, were considered together as a composite question due to their similar nature, in that Q4 was a more sophisticated interpretation dependent on the simpler interpretation of values assessed in Q3. Both Q3 and Q4 were considered sources to assess within Levels 0, 1, or 2, whereas Level 3 was only evident from the more sophisticated question, Q4. Q2 involved critique of the representation rather than traditional graph interpretation, however it was considered in relation to Numerical Graph Interpretation as a separate assessment for Q3 and Q4, due to the need to read the numerical scales to make relevant criticisms of the unusual scales, and the resulting effect on interpreting covariation in the graph.

6.03.04 Results – Verbal Graph Interpretation (Q1)

Q1 asked students to explain the meaning of the graph. The percentages of students by grade who responded at each of the levels of Verbal Graph Interpretation are shown in Table 6.02. Of 2251 responses, 947 were at Level 2, of which 623 were paraphrases of the title. The percentage of student who responded at Level 3 increased considerably from eighth-grade (17%) to ninth-grade (31%). The following discussion of results offers further detail concerning the nature of responses within each level, with examples.

Table 6.02.

Response Counts and Percentages of Responses by Grade at each Level of Verbal Graph Interpretation to Q1

Verbal Graph Interpretation Response Level	Response	Grade				
	Count N=2251	6 n=656	8 n=434	9 n=695	10 n=248	11 n=218
0. Nonstatistical	389	30	21	10	8	7
1. Single Statistical Aspect	392	33	13	13	6	6
2. Inadequate Covariation	947	30	49	46	48	43
3. Appropriate Covariation	523	7	17	31	38	44

6.03.04.01 Level 0. Nonstatistical

Of 389 responses coded as Level 0, 196 were categorised as non-responses, and 193 were categorised as irrelevant, in that they identified neither variable shown in the graph. Examples of student response included “I don’t know,” comments about surface features, such as “it is a column graph” and “it goes high to low,” and comments about the context without identifying a relevant variable, such as “calls overseas” and “the phones getting squashed.”

6.03.04.02 Level 1. Single Statistical Aspect

A total of 392 responses referred to a Single Statistical Aspect, most often a single relevant variable, but in some cases a single data point. A single variable related to the cost or rate was referred to in 243 responses, such as “how much you can save” and “it tells what the rate would be,” and time was referred to in 86 responses, like “how many minutes you talk on the phone.” A further 33 responses referred to only one relevant variable and one irrelevant variable: 22 suggested phone size such as “the bigger your phone the less amount you get off,” and 11 suggested distance of the call rather than duration as the interpretation of “longer,” such as “the

further overseas the cheaper the call.” Finally, 30 responses referred to a single point in the graph, some by reference to values such as “that in 26 minutes you will get 15% off,” and some by descriptive terms, such as “if you call for a long time you save money.” These responses mentioned time and cost, however they did not indicate variation of values that might denote a variable was involved.

6.03.04.03 Level 2. Inadequate Covariation

Many students at each grade level responded at Level 2, Inadequate Covariation. Of 947 responses at this level, 623 involved repeating or paraphrasing the title, such as “The longer your call, the cheaper the rate.” Such phrases referred to both variables, however, the repetition of the phrasing “the longer... the cheaper...” provided little evidence the students appreciated the nature of the variables and the covariation. Some of these 623 paraphrased the title but with misinterpretations of the cheaper rate, such as “The graph shows that the longer you talk the cheaper the phone call will be.”

The other 324 responses involved inadequate descriptions either of the variables or of the correspondence of the variation between them. Eighty-one responses offered one appropriate variable and one poorly-defined variable. In some cases the poorly-defined variable involved the numbers on the horizontal axis as dates, such as “it shows that it’s dearer to ring up earlier in the month” and “On which days how much you get off your phone call.” In other cases the horizontal axis, actually the duration of a given phone call, was misinterpreted as the number of calls or the duration across multiple calls, for example, “the more you call the less you pay.” A further 27 students emphasized the correspondence of the variables, at the expense of the variation of values, by comparing values. Some compared two durations and corresponding cost rates (e.g., “It means up until the 3rd you don’t get

much cut on your phone bill, from the 26th onwards you get a 15% cut”) and others referred to a single bivariate value but with comparative language (“If you call 26 and over you get cheaper rates”).

Two hundred and sixteen responses described covariation in the wrong direction or an unspecified direction. Some examples of the wrong direction of the covariation were technically correct about the situation of increasing cost, such as “The longer the call the more you pay,” but failed to specify the information in the graph concerning decreasing rate. Those that did not specify direction mentioned both variables, either without any notion of dependency, such as “Shows the length of a call and the discount,” or with a notion of dependency that did not follow a trend, such as “it tells us how much we get off after spending a certain amount of time on the phone.”

6.03.04.04 Level 3. Appropriate Covariation

Appropriate Covariation was evident in 523 responses referring to both variables and the negative covariation between them. Of the 523, 166 were significant paraphrases of the title that provided evidence the student understood the correspondence of the variation of values for phone call duration and rate, such as “For each minute you’re on the phone calling overseas the cheaper the rate or % off.” A further 284 described a cheaper rate as a greater discount, thus making the direction of covariation positive and yet correct, such as “The longer your call the greater discount rate you get.” Finally 73 provided various appropriate expressions, some responses listing a series of values, and others elaborately describing the covariation in an appropriate manner, such as the distinction between decreasing rate and increasing cumulative cost. Examples include the following.

- R1. The longer your call is overseas the more money you get off, but it doesn’t mean a short call is more expensive than a long call.

- R2. If you speak for under 4 minutes and if you speak for 4 to 11 minutes you get 3% off, from 11-16 minutes you get 5% off, from 16 to 26 minutes you get 10%, and 26 minutes and over you get 15% off.
- R3. As you talk, the cost will increase more and more slowly.
- R4. It's telling or showing us that the longer we talk at one time is cheaper than a heap of short calls.

6.03.05 Results – Numerical Graph Interpretation (Q2)

Q2 asked students if there was anything unusual about the graph. Table 6.03 shows the distribution of categories of response by grade level. High proportions of students offered no response, or a response that nothing was unusual about the graph, and less than 10% of students commented on the unusual scale or misrepresentation of the graph.

Table 6.03.

Response Counts and Percentages of Responses by Grade at each Level of Numerical Graph Interpretation to Q2

Numerical Graph Interpretation Response Level	Response	Grade				
	Count N=2190	6 n=644	8 n=424	9 n=674	10 n=236	11 n=212
0. Nonstatistical	1602	86	78	66	59	63
1. Single Statistical Aspect	401	11	18	22	27	20
2. Inadequate Covariation	101	2	3	7	4	8
3. Appropriate Covariation	86	1	2	5	10	8

6.03.05.01 Level 0. Nonstatistical

Most students at Level 0 gave a minimal response. Of 1602 responses, 992 commented nothing was unusual, such as “No,” or “No, not that I can see,” 273 offered no response or an irrelevant one, such as “I don’t know,” and 86 commented “Yes” without further elaboration. The remaining 251 responses considered the

graph unusual based on surface features of the visual presentation, such as what students thought a phone should look like.

- R5. There's no way you can do that to a phone. In the last phone, the ear part of the phone wouldn't reach when you're talking in it.
- R6. They used phone instead of a line graph.

6.03.05.02 Level 1. Single Statistical Aspect

A total of 401 responses were classified at Level 1. Of these 401, 260 responses had difficulty comprehending that decreasing rate and increasing cost are compatible, similar to the rate/value confusion for students interpreting physics graphs found by other studies (e.g., Beichner, 1994). The following examples indicate how students found this claim unusual, based on the students' misunderstanding that it referred to reduced cost rather than reduced rate.

- R7. Yes, it's usually dearer.
- R8. Yes the phone call should be dearer.
- R9. Yes, the longer you call it costs extra for lines etc. Why does it get cheaper?
- R10. It should cost more the longer you talk but it doesn't. And if you talk for a great amount of time it might end up being a 100% off.

6.03.05.03 Level 2. Inadequate Covariation

There were 101 responses that considered the graph unusual in relation to the direction. Of these 101, 61 considered the graph misleading, many indicating an awareness that the costs would be higher for longer calls despite the rate decreasing. The remaining 40 responses commented on the unusual nature of the representation with respect to the reversed scale or direction in contrast either to their expectations of the covariation of the context or to their expectations of usual graphs.

- R11. Yes, at first glance you'd think you could stay on there all day and not pay anything, but you can't.

- R12. The presentation is a bit silly (the phones). The untrained eye might think their call actually got cheaper!
- R13. The highest rate of discount is at the bottom - the start is higher and it declines when graphs usually incline.

6.03.05.04 Level 3. Appropriate Covariation

Of 2190 responses, only 86 responses commented on the non-linear aspect of the scales, acknowledging that this gives a visual misrepresentation that might mislead people about the magnitude of the discount. Some commented the times or percentages were not evenly spaced, and a few commented directly that the visual height of the phone did not represent the magnitude of the discount.

- R14. It's not drawn to scale. It's kind of back to front. You'd think the big phone would represent big saving.
- R15. The space from 3% - 5% is the same as the space from 10% - 15%.
- R16. On the graph, the 15% mark is around $\frac{1}{4}$ of the original price. Once you reach 26 minutes the charge stays fixed.
- R17. The size of the phone handle is not representative of what percent you get off.
- R18. The prices only go down a fraction, and not alot, like most people would think it was.

6.03.06 Results – Numerical Graph Interpretation (Q3 and Q4)

Q3 and Q4 asked students to determine call costs from the graph. Q3 required students to read the correct time value and the corresponding discount, and use this reading to determine the cost. Q4 depended on the skill required in Q3, but required a greater appreciation of the covariation, notably that the cost rate per minute for each part of the call reduced for each successive segment of time as time passed, as detailed in the time intervals (e.g., “4th-10th” minutes of the call) on the horizontal axis. Due to this relationship between the questions, they were considered a single

question, with Q3 targeting a response at Level 2 in reading a value and interpreting it in a rate context, and Q4 targeting Level 3 in requiring complete awareness of Appropriate Covariation presented in the graph. Calculations errors were not penalised where there was evidence of appropriate methods.

The percentages of students responding at each level are shown in Table 6.04. Students in higher grades tended to respond at higher levels of response, with 20% of eleventh-grade students undertaking a multi-stage calculation appropriate to the multi-stage rate changes.

Table 6.04.

Response Counts and Percentages of Responses by Grade at each Level of Numerical Graph Interpretation to Q3 and Q4

Numerical Graph Interpretation Response Level	Response	Grade				
	Count N=1928	6 n=608	8 n=361	9 n=561	10 n=204	11 n=194
0. Nonstatistical	1287	89	76	53	46	39
1. Single Statistical Aspect	218	5	11	17	6	20
2. Inadequate Covariation	274	4	9	19	33	22
3. Appropriate Covariation	149	2	3	11	15	20

6.03.06.01 Level 0. Nonstatistical

Of 1928 responses, 1287 were classified as Nonstatistical across Q3 and Q4. Approximately 500 students offered no responses. Approximately 600 responses ignored the graph using the cues “10 minutes” and “rate of \$1/min” to offer responses of “\$10, \$30” or “\$40, \$30.” The most significant feature of these responses was that they show no evidence that the graph played any role in the students’ reasoning for the task at hand. The remaining Nonstatistical responses included a variety of values such as “\$5,” “\$20,” or “a lot.”

6.03.06.02 Level 1. Single Statistical Aspect

Of those who did employ the graph in their response, 218 students selected a value from the graph. Those who used the cue “10 minutes” responded with an inappropriate response of “3% off.” Other students responded “15% off” by using the cue “30 minutes” to select the appropriate value to read off the graph. These students, however, did not go on to use this to calculate a correct response of the cost for another 10 minutes. Some responses used “15% off” to give responses of “15c” or “85c,” without evidence of how the percent was applied with rate information and duration to arrive at a cost.

6.03.06.03 Level 2. Inadequate Covariation

A total of 274 students showed evidence of reading the 15% discount off the rate and combining it with the standard rate and the duration. Most of these students offered an appropriate response to Q3, but failed to attempt the appropriate multi-stage calculation for Q4. Common responses were \$8.50, \$25.50 or \$34.00, related to 10 minutes, 30 minutes or 40 minutes, respectively. A few students offered responses of \$1.50 or \$4.50, which were considered to have combined rate information with the appropriate percentage from the graph, but forgotten to calculate the cost from this discount amount.

6.03.06.04 Level 3. Appropriate Covariation

For Q4, the appropriate calculation of interval stages of discounts yielding a result of \$27.79 was achieved by only 32 students. A further 117 students showed clear evidence of the appropriate method but introduced calculation errors. For the purpose of this investigation, the interest was how students interpreted the covariation of reducing rate with increasing duration to arrive at a cost, hence the most significant features in responses were evidence of recognising that the discount

must be calculated in five stages and evidence of using the percentage discount from the graph to determine costs for each stage. For example, a number of students responded \$23.12. This value is \$4.67 less than \$27.79 and would result if a systematic error were made in calculating the time interval by subtraction, yielding all time intervals one minute less than the correct value.

6.03.07 Discussion of Investigation 7

Four levels of Verbal Graph Interpretation and Numerical Graph

Interpretation were used to describe the degree to which responses represented the covariation between cost rate and duration presented in the graph in Figure 6.01. Many students verbally interpreted the graph with a simple paraphrase of the title, involving both variables at Level 2, but few students explained further the covariation at Level 3. About one third of students noted aspects of the representation that may be unusual, however few commented on the uneven scaling of both axes. When asked to calculate a cost, supplied with the graph and rate information, many students ignored the graph altogether. Asking for “cost” may trigger a calculation mode, which excludes consideration of the graph. Further, some students who did read the graph stated individual facts based on extracting one aspect or data value from the graph. To go further, it is necessary to understand rate in context, to understand discount, to calculate unusual interval lengths where subtraction of “endpoints” is invalid, and it is necessary to have an overview of a multi-step procedure.

The differences among these results from different questions might be considered in relation to limitations of the task wording and purpose. Q1 and Q2 were worded to invite explanation and identification of unusual features, but did not specify the need to clarify aspects of the graph that might be misleading in relation to

the covariation between rate and duration. From a consumers' reading of the graph, the purpose might be limited to knowing the title of the graph, "The longer you call the cheaper the rate," and details of the percentage discount and cost calculations may simply be implicitly agreed to by the consumer in paying the phone bill.

Another consideration for the differences in results among the questions comes from a three-tiered hierarchy of statistical literacy (Watson, 1997). The first tier, involving statistical terminology and concepts, is closely aligned with the verbal interpretation of the graph (Q1). The second tier, appreciating the statistical concept in a real context, is similar to the involvement of cost rates with the context of the graph (Q3 and Q4). The third tier, critical interpretation, is similar to the identification of unusual or misleading aspects of the graph (Q2). Links between motivation and purposes for representing and critiquing claims of statistical covariation are explored further in Chapter 7.

This investigation, as for Investigation 1, found many students offered low level responses for a task embedded in a media context with complex measures for variables. The highest responses levels were to Q1, a result likely to have been influenced by the support of the title supplied, which permitted students to produce a paraphrase involving both variables without reference to the graph.

The following investigation removed the support of a title, the complexities of a misleading media graph, and the complex unfamiliar measures. It explored further students' abilities and challenges when interpreting statistical covariation in graphs.

6.04 INVESTIGATION 8A: NOISE VERSUS NUMBER OF PEOPLE (SURVEYS)

6.04.01 Introduction and Aims

This investigation aimed to consider student approaches to Verbal Graph Interpretation and Numerical Graph Interpretation in a task with familiar variables and simple graphical presentation, aligned with Research Aim 1 of the wider study (Section 3.02). The familiar variables of this task were in contrast to Investigation 7 where the media context, unusual graph form, and rate measures may have made the task more challenging. Hence it was expected that students would tend to offer higher level interpretations in this simpler context, just as in Investigations 2, 3, and 4, higher level responses were more often observed than in Investigation 1. The task was administered to the same sample of students as for tasks involving Speculative Data Generation and Coordinate Graph Production (Task 3 – Heights, Task 4 – Test scores, Task 1 – Heart deaths, and Task 5 – Temperatures). This common sample was important in exploring conceptual development across the various tasks (Chapter 7), aligned with Research Aim 4 of the wider study (Section 3.02).

6.04.02 Task

The task in this investigation is shown in Figure 6.02. The context was chosen such that students would be familiar with the variables; noise level and number of people in a classroom, though rarely measured, are at least intuitively experienced by students in schools. The context was also chosen so that students might expect a positive covariation between the variables, but the task described a negative covariation forcing students to rely on the data rather than prior beliefs. The task was worded with contextual information about the data gathered in order to

support a statistical context for covariation, such as awareness of the data collected and of possible variability from a perfect linear fit. The data were six cases, and included repeated values of each variable.

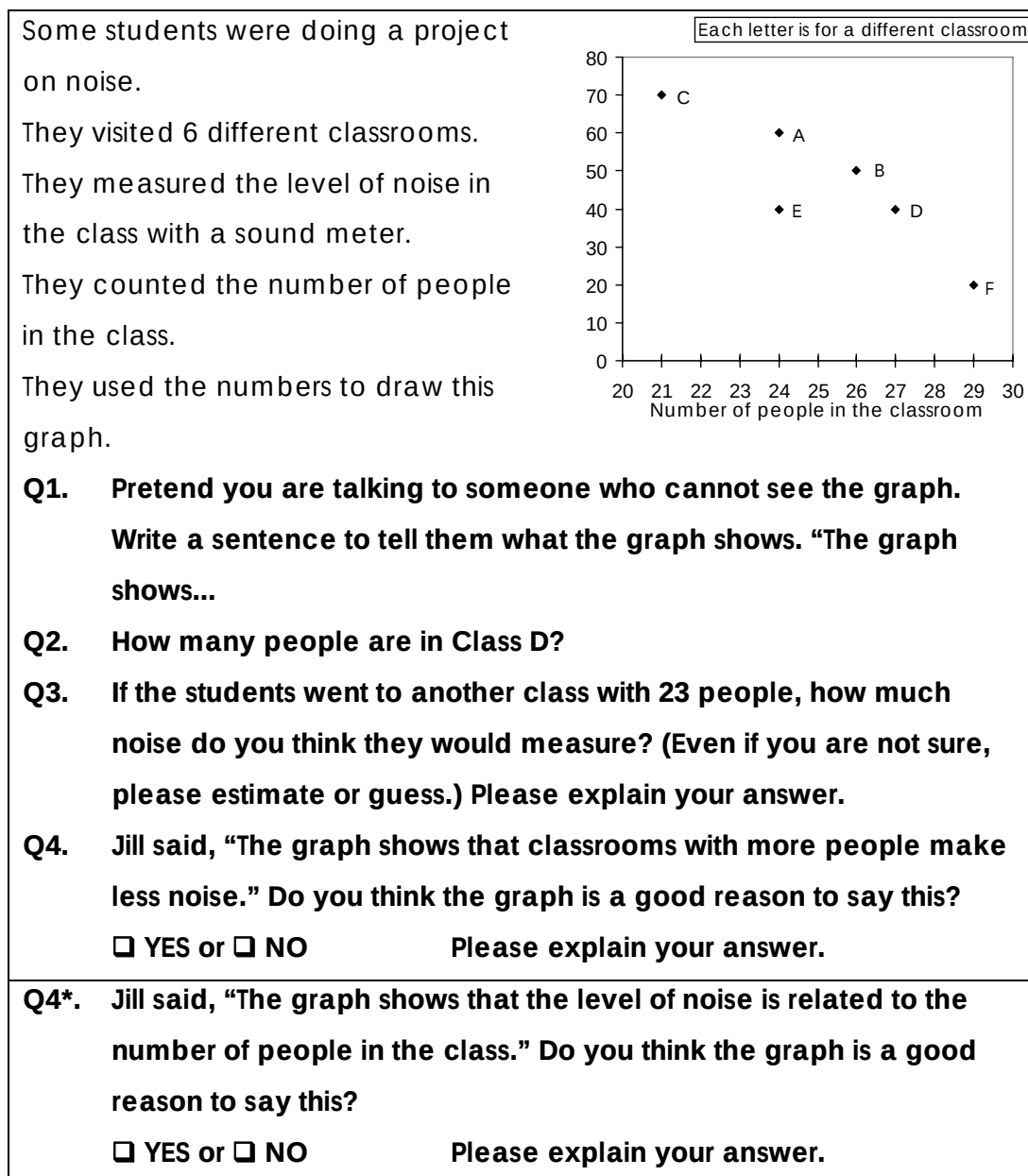


Figure 6.02. Task 6 to assess Verbal Graph Interpretation (Q1 and Q4) and Numerical Graph Interpretation (Q2 and Q3).

Q1 was worded in an open manner to avoid the assumption that an association exists (Donnelly & Welford, 1989). Because students may have avoided comment on covariation in Q1, Q4* was included and then revised to Q4 to provide

a more specific cue about covariation. Q4* was administered to third- and fifth-grade males, then revised to Q4 for remaining students. Q2 involved reading a value, and Q3 was designed to identify whether students based interpolation on proximity to one or more of Classes A, C, and E.

6.04.03 Participants and Method

The data collection followed the method specified for Data Collection 2 in Chapter 3. Responses were coded by successive refinement into the four levels as described in Table 6.01. Verbal Graph Interpretation was assessed using Q1 and Q4. Numerical Graph Interpretation was assessed combining Q2 and Q3. Q3 permitted responses at Levels 0, 1, 2, or 3, whereas Q2 contributed to assessment of Levels 0 or 1 only. The level of Numerical Graph Interpretation was based on the level of the response to Q3, except in cases where students responded at Level 0 to Q3 and Level 1 to Q2, in which case Numerical Graph Interpretation was assessed as Level 1.

6.04.04 Results – Verbal Graph Interpretation (Q1 and Q4)

Questions Q1 and Q4 (and Q4*) in Task 6 involved verbal responses in interpreting a scattergraph. To express the dual notions of appropriate variation and correspondence, responses to Q1 needed (a) to identify “noise” and “number of people” or paraphrases, and (b) to make appropriate use of comparative values such as “less” or “more.” Q4 involved the statement “more people make less noise,” which may refer to a statistical association of the six classrooms surveyed, or to a causal association of classrooms in general, depending particularly on interpretation of “make.” Asking students whether they agreed with this statement thus created an opportunity for students to clarify this ambiguity and thereby to demonstrate an

understanding of the distinction between statistical association and of causal inference.

The characteristics of the four levels of responses are shown in Table 6.01. In most cases coding was based on response to Q1; however, in some cases Q4 (or Q4*) served to demonstrate a student's ability to interpret verbally at a higher level than demonstrated in Q1. Further details are provided in the following sections for each level of response. Students in higher grades tended to respond at higher levels, as seen in Table 6.05; in particular, all students in grades 7 and 9 were able to identify at least a Single Statistical Aspect at Level 1. Seventh- and ninth-grade males performed better than their female counterparts, although this is likely due to the classes sampled rather than the students' sex.

Table 6.05.

Response Counts and Percentages of Responses by Gender and Grade at each Level of Verbal Graph Interpretation to Q1 and Q4

Verbal Graph Interpretation	Response Count	Female Grade				Male Grade			
		3	4/5	7	9	3	5	7	9
Response Level	N=121	n=13	n=23	n=20	n=12	n=13	n=7	n=12	n=21
0. Nonstatistical	13	31	13	0	0	31	29	0	0
1. Single Statistical Aspect	25	38	22	20	8	46	14	8	10
2. Inadequate Covariation	43	23	57	45	67	15	43	33	5
3. Appropriate Covariation	40	8	9	35	25	8	14	58	86

Note: Third- and fifth-grade males were administered Q4* rather than Q4.

6.04.04.01 Level 0. Nonstatistical

Thirteen students offered responses that described no covariation. These included non-responses, responses generically about the topic, such as, “that there is 6 classrooms and each dot shows that that is each classroom” (G3f) or “the graph shows class C, class A, class B, class D, class F, class E and numbers” (G5f).

6.04.04.02 Level 1. Single Statistical Aspect

Twenty-five students referred to a Single Statistical Aspect, either a single data point or a single variable. One student commented on a single data point, “it shows that class C had 21 children in there and sound level is 70” (G3m). Twenty-four students referred to one variable, the level of noise, without reference to number of people in the classroom, although some mentioned that classrooms were involved. Some of these mentioned no values, with responses such as “noise” (G3m) or “the level of noise in a classroom” (G7f). Some commented that noise values varied, such as “it shows that some classes are noisier” (G3f). Others referred to specific values of noise, such as “80 is the most loud and zero is the most soft” (G3f) or “letters of different classes, if match it up you will get their noise level, A.60 B.50, C.70, D.40, E.40, F.20” (G5f).

6.04.04.03 Level 2. Inadequate Covariation

Forty-three responses were classified as representing Inadequate Covariation, inadequate with respect to either correspondence of variables or variation of corresponding values. Some students referred to both variables but did not describe any covariation in the data, such as “the number of people in each class and the noise level” (G5f), or “level of noise goes up in 10’s and going across is the number of people in the class room which is going up from 20, 21, 22, to 30” (G9f). Possibly these students read the axis labels but not the data series. Others mentioned both variables and gave some evidence of generalizing covariation between the two variables but in the wrong direction, for example, “if you have more students it makes more noise” (G3f). Others referred to points, such as stating both extremes, for example “that the classroom with the least people is the noisiest and the classroom with the most is the quietest” (G7f), or an extreme value in a comparative

sense with the rest of the data, such as “that the class with the least people in it is making the most noise” (G5m).

6.04.04.04 Level 3. Appropriate Covariation

Forty students generalized the graphs into a pattern statement, namely a description of the negative covariation, representing Appropriate Covariation. Some responses were simply stated, such as “that less people make more sound” (G7m), and some built up to the idea, for example, “Room C is the noisiest then A followed by B, E and D are each forty, then F brings up the rear, so the more people the less noise” (G7m). Some emphasized both ends of the generalization, similar to those at Level 2 but describing “classes” in the plural to generalize either to the set of six classes or to classes in general: “The classes with less people are the loudest. The rooms with more people are the quietest” (G9m). Other students mentioned the imperfect nature of the covariation: “In most cases the higher the amount of noise the lower the amount of people with the exception of E” (G9m). Responses included statements that emphasized variation by comparison across cases such as “the more X , the less Y ,” “cases with more X have less Y ,” and “as X increases, Y decreases.” No students gave responses that objectified the correspondence or relationship at the expense of variation, such as “ X and Y are negatively/inversely related.” All preferred the language of incremental change across cases, implied by ordering each variable.

6.04.05 Results – Numerical Graph Interpretation (Q2 and Q3)

Numerical Graph Interpretation was assessed combining responses to two questions, one involving reading a value (Q2), and the other involving interpolation (Q3). The coding of the levels of response is shown in Table 6.01. Q3 permitted

responses from Levels 0, 1, 2 or 3, whereas Q2 contributed to assessment of Level 0 or Level 1 only. The highest level evident from Q2 and Q3 was taken as the level of Numerical Graph Interpretation. The percentages of students who responded at each level are shown in Table 6.06. Notably, no third- or fifth-grade students responded at Level 3, whereas no seventh- or ninth-grade students responded at Level 0.

Table 6.06.

Response Counts and Percentages of Responses by Gender and Grade at each Level of Numerical Graph Interpretation to Q2 and Q3

Numerical Graph Interpretation	Response Count	Female Grade				Male Grade			
		3	4/5	7	9	3	5	7	9
Response Level	N=121	n=13	n=23	n=20	n=12	n=13	n=7	n=12	n=21
0. Nonstatistical	12	46	13	0	0	15	14	0	0
1. Single Statistical Aspect	39	23	52	25	33	62	29	25	10
2. Inadequate Covariation	40	31	35	35	50	23	57	50	10
3. Appropriate Covariation	30	0	0	40	17	0	0	25	81

6.04.05.01 Level 0. Nonstatistical

Of 12 Nonstatistical (Level 0) responses, four did not respond to Q2 and five responded “23,” probably because of it appearing on the next line for Q3. Other responses to Q2 included “30,” “45,” and “91 children.” Nonstatistical responses to Q3 were idiosyncratic, such as “50, because some talk and some don’t” (G3f) and “100 because they’re the noisy people and all together they would be really noisy” (G3f).

6.04.05.02 Level 1. Single Statistical Aspect

Thirty-nine students read a single value from the graph for Q2. Some wrote 26 or 28, possibly due to inaccurate visual alignment, and some read a value of 40, reading from the correct point but from the incorrect axis. For Q3, students either

acknowledged they did not know or gave responses that used single points in an idiosyncratic argument such as “30, under E” (G7m). One response, “80, because there would have been 50 people in the room” (G5f), showed a misinterpretation of “another classroom” and assuming for Q3 that an additional 23 students were added to the 27 in Class D referred to in Q2.

6.04.05.03 Level 2. Inadequate Covariation

Forty students responding to Q3 at Level 2 interpolated values in the ranges 39–54 or 71–80, and/or provided reasons related to adjacent data points. Thirteen of these responses predicted 50, with varying evidence of reasoning, for example, “50, it was a guess” (G9m) and “If 23 people were in the class I would estimate 50 because in the classes of 24 they’re 40 and 60 and 50 was in the middle” (G9f). A few students misaligned 23 people to Class E (with 24 people) such as, “40, because it says 23 and then you go up and it should say E then you go across and then it says ‘40’” (G3f). Some gave surprising values such as, “The noise would be at about level 75. This would fit the trend in the graph” (G9m).

Some students showed evidence drawn on the graph for arriving at a response. One student responded “50 db” by interpolation between Classes C and E as shown in Figure 6.03.

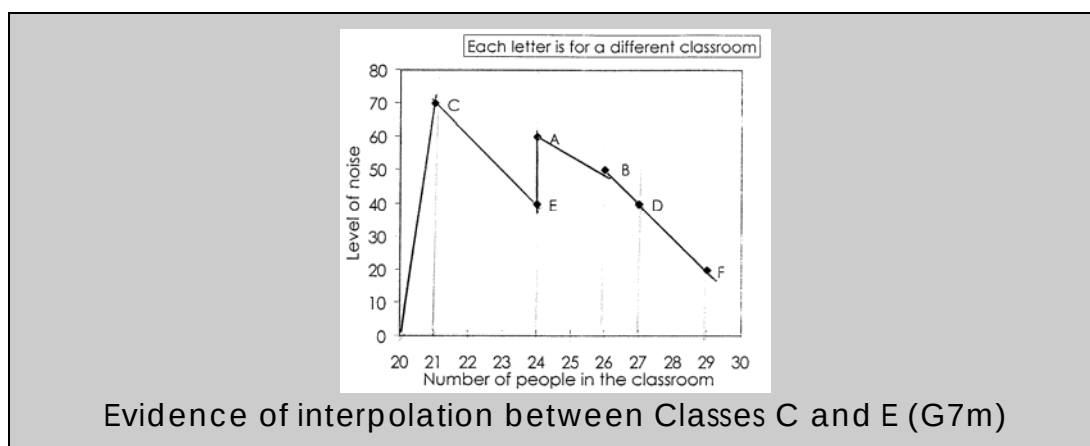


Figure 6.03. Student response – Level 2 – Inadequate Covariation.

6.04.05.04 Level 3. Appropriate Covariation

Thirty responses coded at Level 3 showed evidence of interpolation using the trend of the data to predict a value in the range 55–70. Of the 30, 17 predicted a value of 65, often with reasoning such as, “about 65 because in the class of 24 it is 60 and in the class of 21 it is 70” (G7f). Some predicted other values in the range 55-70, such as, “60 because that is the trend of the graph” (G9m).

A few students drew on the graph providing further evidence of their reasoning. One student responded with reference to the graph shown in Figure 6.04a, that a trend was evident, with an outlying point that could be explained, “65, Class E must have had a strict teacher. The curve signals the amount of noise in a class of 23 would be 65.” Another student also drew on a curved line of best-fit, excluding Class E, as shown in Figure 6.04b, and responded, “65, because the noise gets lower as the people numbers go down, and there is a noticeable wave through this graph that you can follow.”

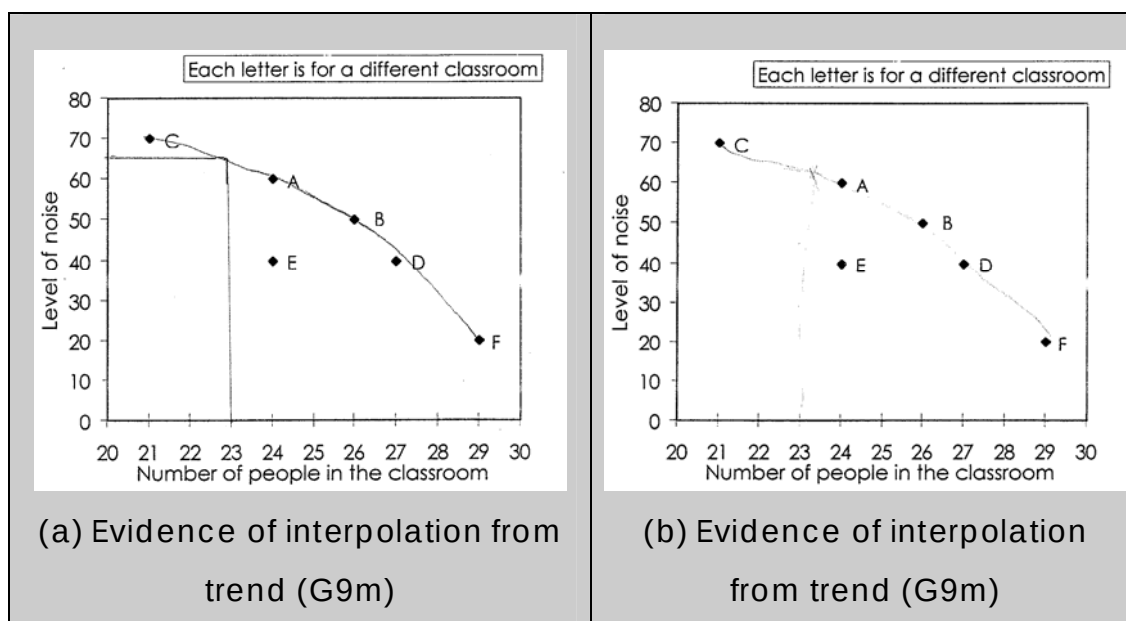


Figure 6.04. Student response – Level 3 – Appropriate Covariation.

6.04.06 Discussion of Investigation 8A

Using the four levels of the framework described in Table 6.01, students in higher grades tended to respond at higher levels, and most students in grades 7 and 9 offered Level 2 or 3 responses for both Verbal Graph Interpretation and Numerical Graph Interpretation. In Investigation 7, the task to explain a graph that was titled yielded many bivariate paraphrases at Level 2. In Investigation 8A, however, the absence of a graph title and the wording to produce a verbal statement elicited students' verbal statements providing greater evidence of understandings at various levels. For Numerical Graph Interpretation, in contrast to Investigation 7, most students, even third-graders, offered responses that identified at least a Single Statistical Aspect related to the data, such as reading a value from a scatterplot, which demonstrated they could engage the task. This contrast also might be related both to the task, which presented a graph with a simple design rather than a graph with complex measures and uneven scales, and to the questions, which clearly focused on the data in the graph presented rather than the involvement of other contextual stimulus (rate information) and calculation of rates. Some students interpreted the data set as a positive covariation based on prior beliefs despite a negative covariation presented in the data. The counterintuitive nature of the tasks was important for assessment purposes in clarifying that these responses were not based on the data presented.

When describing or interpolating, many students referred to one or two bivariate data points. Students' reasoning about isolated data points emphasized correspondence of two measures but did not describe variation to indicate covariation adequately. Development of the pointwise approach in verbal interpretations may be considered as a progression of comparisons within variables, from single-point values ("Class C had 21 children . . .") to comparison of points

(“the classroom with the least people is the noisiest . . .”) to generalizing beyond the available points (“the more people the less noise . . .”). This follows the levels of “reading the data,” “reading between the data,” and “reading beyond the data” described by Curcio (2001).

Many students described only the variable noise in verbal descriptions of the scattergraph. These students emphasized variation but did not describe correspondence of two measures to indicate covariation adequately. All of those who had success in describing the covariation verbally used the language of incremental change across cases.

6.05 INVESTIGATION 8B: NOISE VERSUS NUMBER OF PEOPLE (INTERVIEWS)

6.05.01 Introduction and Aims

This interview-based investigation aimed to explore further the findings of Investigation 8A. Interviews were employed to gain evidence from dialogue of the conceptual challenges for students and the language they used to reason and formulate interpretations. The investigation also aimed to provide more detailed response data to check the validity of the four-level framework for both Verbal Graph Interpretation and Numerical Graph Interpretation in describing the diversity of conceptions evident in student responses, aligned with Research Aim 1 of the wider study (Section 3.02).

6.05.02 Participants and Method

The interview-based data collection followed the method specified for Data Collection 4 in Chapter 3. Interviews were conducted in relation to Task 6, shown in Figure 6.01. The 13 students interviewed were assigned unique interview codes, as used in other investigations, to permit reference in the description of interview dialogue. The students included three third-grade males (G3m1, G3m2, G3m3), three third-grade females (G3f2, G3f4, G3f5), three fifth-grade males (G5m1, G5m2, G5m3), two fifth-grade females (G5f2, G5f6), one seventh-grade male (G7m2), and one ninth-grade male (G9m3). Four interviewees (G5m2, G5m3, G7m2, G9m3) had not attempted this survey task due to limited time when surveyed. The small number of students interviewed was due to this task appearing after a number of graphing tasks that used much (and in some cases all) of the interview time available. The interviewer sometimes omitted some of these four questions, for example questions

of reading the data in cases where student interpretations already demonstrated such skills. Interview dialogue was coded in relation to the levels in Table 6.01 using iterative qualitative methods (Miles & Huberman, 1994). Evidence for levels of Verbal Graph Interpretation and Numerical Graph Interpretation was taken from any dialogue during consideration of Task 6; for example, although Q1 assessed Verbal Graph Interpretation, in interviews, discussion of Q1 sometimes yielded evidence of reading values from the scale, that is, Level 1 Numerical Graph Interpretation, rendering response to Q2 unnecessary.

6.05.03 Results

The results for the 13 students are shown in Table 6.07, according to both Numerical Graph Interpretation and Verbal Graph Interpretation. Of the 13 students interviewed, 10 demonstrated the ability to read values, that is Numerical Graph Interpretation at Level 1 or greater, and 10 referred to both variables with Verbal Graph Interpretation at Level 2 or 3. Only 2, however, interpolated numerically based on a trend, and only 2 supplied Verbal Graph Interpretations of Appropriate Covariation (Level 3). For these 13 students, there was an apparent association between levels of Verbal Graph Interpretation and Numerical Graph Interpretation. In the results that follow, both Verbal Graph Interpretation and Numerical Graph Interpretation are discussed together as they were intertwined in the interview dialogue, however the results are presented in the order of increasing level of Numerical Graph Interpretation.

Table 6.07.

Students' Level of Verbal Graph Interpretation by Level of Numerical Graph Interpretation

Level of Verbal Graph Interpretation	Level of Numerical Graph Interpretation				Total
	0. Non-statistical	1. Single Statistical Aspect	2. Inadequate Covariation	3. Appropriate Covariation	
0. Nonstatistical	G3m1, G5f6	-	-	-	2
1. Single Statistical Aspect	-	G3f5	-	-	1
2. Inadequate Covariation	G3f4	G3m2, G3f2, G5m1, G5f2, G7m2	G3m3	G5m3	8
3. Appropriate Covariation	-	-	G5m2	G9m3	2
Total	3	6	2	2	13

6.05.03.01 Level 0. Nonstatistical (Numerical Graph Interpretation)

Two students (G3m1, G5f6) responded with Nonstatistical responses (Level 0) for both verbal and numerical interpretations, primarily concerned with the letters used for classroom case labels. G5f6 considered the classroom labels to form an alphabetical order, and gave minimal responses to most questions. G3m1's survey response "the classrooms" for Q1 lacked identification of a statistical variable, and his numerical response of 27 for Q2 may have been copied given that he exhibited a lack of understanding in interview: he was initially concerned with letter labels to spell his name, he ignored data points by referring to the extremes of the graph scale for his numerical interpretations.

I: [Q2] ... if you look at that, can you tell me how many people are in Class F?

G3m1: I would say there would be... 25.

- I: And how did you decide that?
- G3m1: I decided that because F would be A, B, C, D, E, F [pointing to each point in turn] and that would be how many people are in Class D... hang on, I would give it about 30, because it is the last letter and I would say the last letter is the most of the people in the class [referring to horizontal scale].
- I: What about A?
- G3m1: I would say it would have 20 people, because it is the least number. [referring to horizontal scale] [...]
- I: And why did you decide 20, and not 15 or not 25, or...?
- G3m1: Because there is not 15 there [looking at graph], and that's how many people were in all of those classes.

He relied on intuitive knowledge of the covariation for his verbal interpretations for

Q4.

- G3m1: [Q4] No. [...]
- I: Do you think that as there's more people in the classroom, it might, the graph, be showing that they get noisier or that they get quieter?
- G3m1: They get, the more people in the room I would say that they get noisier. So that I think that shows that.
- I: Mmm, okay. And if you were to look at the graph there, does it show you how much noise that Classroom C is making?
- G3m1: I would say they would about have 30 is their level noise [sic], because it's the third one [letter] on the graph and that 30 is the third one on the graph [vertical axis scale].

One third-grade female (G3f4) verbalised both variables (Level 2) but struggled to read numerical values (Level 0). She wrote in the survey that the graph showed a single variable, "that the students are measuring the level of the noise." In interview she referred to both number of people in classrooms and noise in classrooms, possibly because the interviewer had just read the question, but she did not appear to attempt to relate them as bivariate data.

- I: [Q1] And I asked you to tell me a bit about what it shows. And you said it was about the level of noise in there. Can you tell me what else you noticed in the graph?
- G3f4: How many classes there is. And it shows you how many people are in the classroom. And the level of noise they're making in the classroom for each one.
- I: Do you think there is a pattern in the graph? Or not really?

S: It sort of goes, sort of like A, B, C, D, E, F, it's alphabet. That is all. I don't know.

She appeared to be learning to read numerical values (Levels 0 to 1): in the survey, for Q2 she wrote "23" (Level 0), possibly lifted from Q3, but in the interview, amid difficulty reading values from the scale, she noticed the noise level for C of 70 (Level 1), possibly due to the proximity of the scale to the data point.

I: [Q2] And do you think it tells you how many people are in class E?

G3f4: [examines graph] Not really.

I: No? It's not a very good way for showing that?

G3f4: No, because it could be like A could be here [tracing from A on diagonal to horizontal axis at 21] and then I would have no idea.

I: Does it tell you anything about how much noise that Class C was making?

G3f4: I think it was 70. I'm not quite sure. Because it had that across there. I'm not sure. [...]

I: So what would it mean for like F then, would you say?

G3f4: About 20. Because it goes straight along the line [traces from 20 on vertical axis across to point F].

I: Right, so that might be how much noise class F was making. And can you have a guess at how many people would be in class F?

G3f4: Umm... [examining graph] About 19 or 16 or something.

I: Yes okay. Do you think that Class C would have more people in it or less people?

G3f4: Umm... [examining graph, 10 second pause] I'm not quite sure.

6.05.03.02 Level 1. Single Statistical Aspect (Numerical Graph Interpretation)

One third-grade female (G3f5) gave Single Statistical Aspect responses (Level 1) for verbal and numerical interpretations. For Q3, she responded "30... maybe 35" by proximity to Classroom E, and she agreed with the interviewer's probe that the graph showed more people make more noise, although when asked, was not sure where on the graph showed it.

Five students (G3m2, G3f2, G5m1, G5f2, G7m2) gave Single Statistical Aspect numerical interpretations (Level 1) and verbal interpretations that were

bivariate (Level 2) but unconvincing with respect to the appropriate covariation.

Dialogue often suggested uncertainty or ambivalence in interpreting covariation, referring to selected data classes or indicating the direction of the covariation being counter to students' intuition of causality. G3m2 struggled to read values, responded to Q3 by suggesting the new value should be unique ("30, because there's none on 30"), and disagreed that the graph showed "level of noise is related to number of people," commenting, "Well Class C is on 70 and it only has about 21 people in it, so the less people probably the more noise, and the more people the more noise" [sic]. G5m1 agreed in the survey that noise and number of people were related ("Yes, more people to make noise"), but in interview commented, "I kind of disagree with that now, because it says C has one of the loudest noise levels but only 21, and like F has 29 people and that has only 20." He was hesitant in describing the graph overall ("I don't really know, really") but when prompted, hesitantly offered, "Well, kind of... [10 second pause] the more people there is, the quieter it seems to be on this graph."

Two females (G3f2, G5f2) demonstrated the ability to read points, but not interpolate (Q3 responses "15" and "24, because one more voice," respectively). G3f2 gave a verbal summary "as it gets bigger, it starts to get lower," but was hesitant about the direction, twice disagreeing the graph showed "more people make less noise," and yet reaffirming "when it got big, the level of noise would get lower": the verbal formulation appeared rigid and notably class size is implicit in "it." She needed some nudging to voice her interpretations, which she offered with few words but which conveyed Inadequate Covariation (Level 2) for Verbal Graph Interpretation.

G3f2: [Q1] 21 people, and it was 70.

I: 70? What was 70 to do with? [pause] What was it that they were measuring with 70?

G3f2: Noise. [...]

I: So do you think overall, that the graph has any pattern to it or not really?

G3f2: As it gets bigger, it starts to get lower.

G3f2: [...] [Q4] No, not really.

I: Not really. But you were saying that it would show that... Tell me again what it was showing overall?

G3f2: When it got big [looking at graph], the level of noise would get lower.

I: Is that what you would expect?

G3f2: Not really. [...]

I: What would you normally expect?

G3f2: Probably when it gets bigger it gets louder.

G5f2 twice disagreed the graph showed more people make less noise, once in the survey (“No, because more noise if there is more people”), and once in the interview. In the interview she also disagreed with her own survey response (“It shows the less people were in the class that have the most noise. I don’t know, it might have really loud people in there, I don’t know. Yeah, it is loud.”).

One seventh-grade male (G7m2) considered the bivariate nature of the data set, but was not clear about the direction of the covariation, preferring to refer to other causal variables that might confound interpretation. For Q4, he initially agreed it showed a pattern, but rejected the causal connection, and later disagreed that the graph had an overall pattern, noting the repetition of data points with the same number of people as a justification, since the data were not a one-one monotonic trend.

G7m2: [Q1] That class [pointing to C] is very noisy and the teacher would get a headache very easily. [...] This graph is showing how many people are in the classroom, and how much noise they make.

G7m2: [...] [Q4] It shows that, but it’s not a good reason to say that.

I: How does it show it, firstly?

G7m2: Well Classroom F [...] The lowest classroom, they had 29 people.

I: So that shows that classrooms with more people make less noise, does it?

G7m2: Yeah, but they probably had a teacher in there at the time, yeah so they didn’t like do that in between the periods.

I: So as you look at the graph there, do you think it does have an overall pattern to it that it might be showing that?

G7m2: No...

I: ... that classrooms with more people make less noise?

G7m2: Well these two classrooms have exactly the same amount of people except this class is making more noise than this class.

The examples of students who responded at Level 1 for Numerical Graph

Interpretation provided indicated frequent uncertainty about general verbal statements concerning covariation in the graph.

6.05.03.03 Level 2. Inadequate Covariation (Numerical Graph Interpretation)

Two students (G3m3, G5m2) offered numerical interpretations at Level 2.

Both offered hints of appreciating the covariation, but were influenced by certain data cases, particularly Class E, when formulating verbal statements for Q1 or interpolating for Q3. G3m3 interpolated “50, I had a guess,” presumably the midpoint of Classes A and E, however the interview concluded before this could be confirmed. He began to compose a verbal statement of covariation by referring to an extreme point, but then began to cloud the interpretation with reference to repeated values and causal reasoning.

G3m3: [Q1] It shows that Class F had the most people and they had the least noise. And then D is probably the second best class, because they had more people but they were about the same as E, but E had less people [mumbles].

I: [Q4] [...] more people in the classroom means there will be less noise...?

G3m3: Well, if the classes knew they were coming, they would probably be quiet that morning, but if they didn't know, they might not.

G5m2 gave an appropriate verbal statement (“mostly the more people, the quieter it is, but these two classes in, around that number of people, can be noisier”), but he suggested between 30 and 40 for Q3, referring to the trend as well as Class E

(“Because they’ve got 24, and most of the graph shows that the less people the noisier. But I reckon, so, around that number.”).

6.05.03.04 Level 3. Appropriate Covariation (Numerical Graph Interpretation)

Two students (G5m3, G9m3) interpolated appropriately based on the trend. In interview, G5m3 referred to extreme points but did not generalize or agree to a general verbal statement, yet he interpolated 65 “because on this graph, it has C with the least amount of people with the most noise, so if you just keep going in that order....” The sole ninth-grade student (G9m3) interviewed did not attempt these questions in the survey, but he fluently answered the two questions posed by the interviewer, intertwining verbal statements with numerical values about the extremes or intermediate values.

G9m3: [Q1] The graph there is showing that the more people in the class, the less noise there was. Yeah, it just shows, one of them, 29 people in the class, the level of noise is only at 20, while when you have 21 people in the class, the level of noise is at 70. And so, sort of, as the... as more students in the class, it’s sort of declining noise.

G9m3: [Q3] Using the graph, probably about 63 or 4. [...] I looked at sort of the decline, as more people, or the increase [motioning diagonal in data] as people got less [horizontal scale right to left] and if you put 23 in order to draw that, it would make it, the full incline, it would probably, you’d put it somewhere around that area there, just to join the line.

6.05.04 Discussion of Investigation 8B

Responses were observed at each of the four levels of both verbal and numerical interpretations, despite the small sample of students. This range of response provided additional support to the four-level framework for assessing Verbal Graph Interpretation and Numerical Graph Interpretation. Dialogue illustrated, for example, how students often referred to the extreme points when interpreting the graph.

The interview dialogue often revealed challenges for students in interpreting the graph. Some students gave responses at two adjacent levels in different segments of dialogue, for example, conflict about the direction of the covariation. Many students referred to a belief that more people would make more noise; some allowed this to override careful interpretation of the graph. A data point that involved repeated values, Class E, was also a significant point in students' reasoning as a case contrary to a deterministic or monotonic trend. This feature of Task 6 paralleled the feature of constant height with changing age employed in Task 3 (Investigation 2). The counterintuitive direction, and the cases that do not conform to a deterministic, monotonic covariation, were issues of limited conceptions of covariation identified in previous research (Batanero et al., 1996, 1997; Crocker, 1981).

6.06 GENERAL DISCUSSION – GRAPH INTERPRETATION

Four levels of response were detailed for tasks concerning Verbal Graph Interpretation and Numerical Graph Interpretation. These levels related closely to levels described in previous research of correlational reasoning (Ross & Cousins, 1993; Swatton, 1994) and graph comprehension (Curcio, 2001). Themes evident in responses at various levels were consistent with those observed for Speculative Data Generation. Some students did not refer to the data but based responses on topic knowledge (Level 0), and some verbal interpretations were limited to one variable (Level 1). At Level 2, students referred to difference or comparison of selected values or found difficulty clarifying the nature of the correspondence between two variables. At Level 3, students made use of incremental language and global trends to describe covariation.

In Investigation 8A, most students, even third-graders, offered responses that identified at least a Single Statistical Aspect related to the data, such as reading a value from a scatterplot, which demonstrated they could engage with the task. Although acknowledging the small sample of students was likely to be above average ability, this result indicates tasks of handling bivariate data and a Cartesian system are accessible for primary school students given appropriate support from the task description or discussion. This finding follows recent curriculum suggestions that bivariate data need not be reserved for senior students (NCTM, 2000). In contrast, in Investigation 7, many students offered Nonstatistical responses. This contrast to Investigation 8 was likely due to the differences in tasks. The task in Investigation 7 involved rates that were unfamiliar variables, set in a complex newspaper context, whereas in Investigation 8, use of familiar variables supported students to appreciate the meaning of individual data values.

Further elements of task design – covariation that was negative (Investigations 7, 8), counterintuitive (Investigation 8), and involving repeated values on one variable (Investigation 8) – were important to assess the degree to which students appropriately interpreted the correspondence and variation presented in the data.

The coding of responses supported the use of assessment frameworks of four levels of response for Verbal Graph Interpretation and Numerical Graph Interpretation. The names of these levels were identical to those employed for other covariation-based tasks involving Speculative Data Generation and Coordinate Graph Production. The associations among these assessment frameworks are discussed in the following chapter.

CHAPTER 7. ASSOCIATIONS AMONG SKILLS

7.01 INTRODUCTION

This chapter explores the associations among the skills of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation, as seen in Figure 1.01. Each of these skills was described according to four levels: Nonstatistical, Single Statistical Aspect, Inadequate Covariation, and Appropriate Covariation. Some tasks involving various of these skills were administered to common students. Associations among these skills were explored by correlations of response levels, cross-tabulations of response counts, and examples of responses of individuals to tasks for each of these skills.

7.02 SIMILARITIES AND DIFFERENCES AMONG ASSESSMENT FRAMEWORKS

The assessment frameworks for Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation each involved four levels of the same name with similar features. For Speculative Data Generation, within these response levels, similar response categories were observed across three tasks (Task 1 – Heart deaths, Task 3 – Heights, and Task 4 – Test scores). Stability of most categories across the tasks provided broader evidence of the usefulness of the assessment frameworks for exploring Speculative Data Generation.

A number of categories within the assessment framework for Coordinate Graph Production were notable in contrast to frameworks for tasks of Speculative Data Generation (Figure 4.02, Figure 4.27). A Table (Category 1A) response for

Coordinate Graph Production contrasted with the level assigned for a similar response for Speculative Data Generation (Category 3B, Figure 4.27). The reason for this difference was that for the task in relation to Coordinate Graph Production, a table was a reproduction of that given, whereas the task was to represent these data graphically, that is, clearly involving a translation or transnumeration (Pfannkuch & Wild, 2004). For the tasks of Speculative Data Generation, the focus was not upon the form of the representation, but the structure of the speculative data provided in the response.

A second contrast of the frameworks was in relation to Series Comparison graphs (Level 3) for tasks of Speculative Data Generation, for which similar responses were considered as Spatial Variation graphs (Category 2B) for the task of Coordinate Graph Production. For Speculative Data Generation, often a case label in the response, such as calendar years or student names, presented the correspondence demanded to show covariation. In some cases, time as a covariate was implied as the common baseline. In contrast, the task of Coordinate Graph Production included the data with correspondence provided, and Category 2B responses represented each data variable with no clear evidence of a case-label to show the correspondence of the data values.

A final contrast between frameworks was the differentiation of categories at Level 3. For Speculative Data Generation, Coordinate Variation graphs formed a category, whereas for Task 5 (Temperatures) with data values supplied, many students responded with coordinates, and the differentiation of coordinate Bar Graphs and Line Graphs was viable to differentiate those emphasising discrete values or continuity.

For Verbal Graph Interpretation and Numerical Graph Interpretation, four levels were described, however no categories within levels were discerned. Graph interpretation tasks generated student responses that were less specific in structure of correspondence and variation than graphic responses for Speculative Data Generation and for Graph Production. In this respect, the tasks and assessment frameworks for Speculative Data Generation and Coordinate Graph Production provided a detailed window on students' understanding of covariation, with greater richness than could be discerned using only tasks for Verbal Graph Interpretation and Numerical Graph Interpretation.

7.03 ASSOCIATION OF TASK 1 (HEART DEATHS) AND TASK 2 (TELEPHONE CALL RATES)

As part of Data Collection 1, Task 1 (Heart deaths), as presented in Investigation 1, and Task 2 (Telephone call rates), as presented in Investigation 7, were administered to 1601 common students. Table 7.01 shows the counts of student responses at each level of Speculative Data Generation for Task 1 (Heart deaths) by the level of Verbal Graph Interpretation and Numerical Graph Interpretation for Task 2 (Telephone call rates). Speculative Data Generation for Task 1 (Heart deaths) was moderately correlated with the level of Verbal Graph Interpretation ($r = 0.39$) and slightly correlated with Numerical Graph Interpretation ($r = 0.31$).

Table 7.01.

Frequency of Student Responses at each Level of Speculative Data Generation for Task 1 by Level of Verbal Graph Interpretation and Numerical Graph Interpretation for Task 2

Level of Speculative Data Generation: Investigation 1:	Investigation 7: Task 2 - Telephone call rates							
	Level of Verbal Graph Interpretation (Q1)				Level of Numerical Graph Interpretation (Q3 and Q4)			
	0	1	2	3	0	1	2	3
Task 1 - Heart deaths								
0. Nonstatistical	121	86	121	51	261	21	19	6
1. Single Statistical Aspect	73	89	216	65	311	38	37	10
2. Inadequate Covariation	24	43	144	78	152	40	48	21
3. Appropriate Covariation	25	34	206	225	148	70	122	80
Total students (N)	243	252	687	419	872	169	226	117

Response levels for Verbal Graph Interpretation were in general slightly higher than for Speculative Data Generation. In particular, higher frequencies of students responded at Levels 0 and 1 for Speculative Data Generation than for Verbal Graph Interpretation. The high frequency for Level 2 Verbal Graph Interpretation, however, was in part due to responses that were paraphrases of the title provided by the news advertisement in the task. A high number of students (206) achieved Level 3 Speculative Data Generation with only Level 2 Verbal Graph Interpretation. This may suggest that, for Task 2 (Telephone call rates), it was easy for students to provide a verbal statement involving both variables, but very challenging to articulate the relationship appropriately, whereas for Task 1 (Heart deaths), it was appropriately challenging to present speculative data showing covariation.

Response levels for Numerical Graph Interpretation, assessed in Task 2, were lower than for Speculative Data Generation, assessed in Task 1 (Heart deaths). There was a high frequency for Level 0 Numerical Graph Interpretation (872 responses), in some cases non-response, possibly due to lack of motivation. Even excluding Level 0, however, many students responded at higher levels for Speculative Data Generation than for Numerical Graph Interpretation. This may suggest that it was easier for students to translate aspects of a verbal statement of covariation to construct a graphical form of their own design, rather than to read values from a conventional coordinate scattergraph and interpolate based on discernment of a global trend. The activity of constructing the graph, step by step, and the freedom of design, appeared to engage and challenge students.

There was a more uniform distribution of students across levels for Task 1 (Speculative Data Generation) than for Task 2 (both Verbal Graph Interpretation and Numerical Graph Interpretation). Thus assessment based on Speculative Data Generation framework used for Task 1 discriminated among students more than assessment based on the frameworks for Task 2. This empirical evidence supports the suggestion in Section 7.02 that Task 1 for Speculative Data Generation provided a more informative window on students' understanding of covariation than Task 2 for Verbal Graph Interpretation and Numerical Graph Interpretation.

7.04 RESPONSE COUNTS AMONG TASKS FOR SPECULATIVE DATA GENERATION

Investigations 3A, 4A, and 5A concerned Speculative Data Generation with common students responding to different tasks. Further, Investigation 3A involved two assessments of Speculative Data Generation, one from Q1&2, and the other from Q3. Correlations were calculated among scores for pairs of tasks, where the score is the numerical value of the response level. It is acknowledged that this method does not take into account the conceptual differences among response levels, that response levels of 0, 1, 2, and 3 do not constitute a continuous measure, and that the response levels may be subject to ceiling effects. Hence interpretation of the correlations must be complemented by consideration of the cross-tabulations of response levels to each pair of tasks. Correlations among pairs of responses levels are shown in Table 7.02, with the numbers of responses in common across both tasks shown in parentheses. In general, moderate correlations were observed between all pairs. Although correlations were statistically significant, the percentage of variation explained between tasks varied from just 13% to 27%. In general the correlations involving Task 1 (Heart deaths) were slightly weaker than other correlations.

Table 7.02.

Response Counts and Correlations Among Pairs of Response Levels

Task	Task		
	3 (Q1&2)	3 (Q3)	4
Task 3 (Q3) (Height and sex)	0.52 (83)	-	-
Task 4 (Test scores)	0.46 (167)	0.45 (76)	-
Task 1 (revised) (Heart deaths)	0.44 (91)	0.43 (70)	0.36 (87)

Tables 7.03 and 7.04 show associations among levels of Speculative Data Generation for the different tasks. Task 4 and Task 3 (Q1&2) were administered to third- and fifth-grade students, whereas other tasks were not, which resulted in higher numbers of students for the combination of Task 4 and Task 3 (Q1&2) than for other combinations of tasks. The dominant feature of the six cross-tabulations shown was the high proportions of students who responded at Level 3, Appropriate Covariation, for pairs of tasks. This feature contributed to the observed correlations. Another feature was the high proportion of students who responded at the same level across two tasks. These numbers are summarised in Table 7.05. In general, about 50% or more of responses across two tasks were at the same level for both tasks, and most remaining responses tended to differ by one level of Speculative Data Generation, as shown in Tables 7.03 and 7.04. Table 7.05 indicates that overall, students responded with similar levels of Speculative Data Generation for Task 3 Q1&2, Task 3 Q3, and Task 4, whereas they responded with lower levels to Task 1 (Heart deaths). Possible reasons for this, already described, include familiarity of the measures, complexity of the phrases involving “increase,” that is change over time, and difficulties conceptualising a bivariate data case.

Table 7.03.

Frequency of Student Responses at each Level of Speculative Data Generation for Tasks 3, 4 and 1 (revised) by Level of Speculative Data Generation for Task 3

Level of Speculative Data Generation	Level of Speculative Data Generation					
	Task 3 Q1&2 Height vs Age			Task 3 Q3 Height vs Age by Sex		
	1	2	3	1	2	3
Task 3 Q3: Height vs Age by Sex						
1. Single Statistical Aspect	1	2	0	-	-	-
2. Inadequate Covariation	0	10	6	-	-	-
3. Appropriate Covariation	1	9	54	-	-	-
Total students (<i>N</i>)	2	21	60	-	-	-
Task 4: Test scores vs Study times						
0. Nonstatistical	5	10	3	1	1	2
1. Single Statistical Aspect	5	12	1	1	2	1
2. Inadequate Covariation	0	21	14	1	3	9
3. Appropriate Covariation	3	32	61	0	6	49
Total students (<i>N</i>)	13	75	79	3	12	61
Task 1 (revised): Heart deaths vs Use of motor vehicles						
0. Nonstatistical	2	12	8	1	2	7
1. Single Statistical Aspect	0	6	6	1	4	3
2. Inadequate Covariation	0	2	4	0	1	4
3. Appropriate Covariation	0	9	42	0	2	45
Total students (<i>N</i>)	2	29	60	2	9	59

Table 7.04.

Frequency of Student Responses at each Level of Speculative Data Generation for Task 1 (revised) by Level of Speculative Data Generation for Task 4

Level of Speculative Data Generation	Level of Speculative Data Generation			
	Task 4: Test scores vs Study time			
	0	1	2	3
Task 1 (revised): Heart deaths vs Use of motor vehicles				
0. Nonstatistical	3	2	2	12
1. Single Statistical Aspect	2	2	3	5
2. Inadequate Covariation	0	0	0	6
3. Appropriate Covariation	0	2	5	43
Total students (N)	5	6	10	66

Table 7.05.

Frequency of Comparison of Response Levels of Speculative Data Generation for Four Tasks

Level of Speculative Data Generation for Task in row compared to Level for Task in column	Speculative Data Generation Task		
	3 (Q1&2)	3 (Q3)	4
Task 3 (Q3): Height vs Age by Sex			
Higher	10	-	-
Same	65	-	-
Lower	8	-	-
Task 4: Test scores vs Study times			
Higher	35	7	-
Same	87	53	-
Lower	45	16	-
Task 1 (revised): Heart deaths vs Use of motor vehicles			
Higher	9	2	9
Same	44	47	48
Lower	38	21	30

7.05 RESPONSE COUNTS AMONG SKILLS

Associations among Speculative Data Generation, Coordinate Graph

Production, Verbal Graph Interpretation, and Numerical Graph Interpretation were considered for various tasks. Correlations among pairs of responses levels are shown in Table 7.06, with the numbers of responses in common across both tasks shown in parentheses. The strongest correlations were between response levels for

- Coordinate Graph Production and all other responses, and
- Verbal Graph Interpretation and Numerical Graph Interpretation.

Weaker correlations were observed for response levels for Speculative Data Generation tasks and those for graph interpretation tasks. The strong correlation of Coordinate Graph Production with other skills may indicate the common element of graph design in responding to these tasks. Notably, the Speculative Data Generation Task 4 (Test scores) is less strongly correlated with Coordinate Graph Production, and more strongly correlated with Numerical Graph Interpretation, than other tasks for Speculative Data Generation. This task clearly specified six discrete data cases in a bivariate context. These correlations might indicate that the task assessed speculation of numerical values of data to represent covariation with less emphasis on graph design.

Table 7.06.

Response Counts and Correlations Among Pairs of Response Levels

Skill and Task	Skill and Task		
	CGP: Task 5	VGI: Task 6	NGI: Task 6
Speculative Data Generation: Task 3 (Q1&2) (Height):	0.58 (133)	0.43 (120)	0.41 (120)
Speculative Data Generation: Task 3 (Q3) (Height and sex):	0.65 (58)	0.37 (59)	0.32 (59)
Speculative Data Generation: Task 4 (Test scores)	0.46 (133)	0.29 (111)	0.47 (111)
Speculative Data Generation: Task 1 (revised) (Heart deaths)	0.67 (66)	0.36 (55)	0.25 (55)
Coordinate Graph Production: Task 5 (Temperature)	-	0.54 (98)	0.54 (98)
Verbal Graph Interpretation: Task 6 (Noise)	-	-	0.53 (121)

Tables 7.07 and 7.08 show that there were many students who responded with different response levels to two tasks, in some cases as diverse as Level 3 (Appropriate Covariation) for one task and Level 0 (Nonstatistical) for the other. Speculative Data Generation is shown for Task 4 (Test scores) in preference to Task 3 (Heights), which involved graph production features such as axis allocation, and in preference to Task 1 (Heart deaths), which involved a media context and one for which fewer students were administered the task.

Comparable numbers of responses for Coordinate Graph Production were higher (20) or lower (25) than those for Verbal Graph Interpretation (Table 7.09), however Table 7.07 reveals that students with higher levels of Coordinate Graph Production than Verbal Graph Interpretation were rarely higher by more than one response level (5 of 20 responses), whereas students with higher levels of Verbal Graph Interpretation than Coordinate Graph Production were more commonly higher

by two or three response levels (11 of 25 responses). This provided some evidence that for these students, understandings that support Verbal Graph Interpretation developed prior to those supporting Coordinate Graph Production.

Table 7.07.

Frequency of Student Responses at each Response Level for Tasks 5 and 6 by

Response Level for Tasks 4 and 5

Response Level	Task 4 Test scores vs Study time				Task 5 Temperature vs Time			
	Level of Speculative Data Generation				Level of Coordinate Graph Production			
	0	1	2	3	0	1	2	3
Coordinate Graph Production: Task 5 Temperature vs Time								
0. Nonstatistical	8	5	10	10	-	-	-	-
1. Single Statistical Aspect	3	6	8	9	-	-	-	-
2. Inadequate Covariation	2	3	3	13	-	-	-	-
3. Appropriate Covariation	1	1	6	45	-	-	-	-
Total students (<i>N</i>)	14	15	27	77	-	-	-	-
Verbal Graph Interpretation: Task 6 Noise vs Number of people								
0. Nonstatistical	0	3	4	6	5	4	1	1
1. Single Statistical Aspect	5	6	5	8	7	7	4	3
2. Inadequate Covariation	5	3	10	17	6	5	12	7
3. Appropriate Covariation	2	1	2	34	3	2	2	29
Total students (<i>N</i>)	12	13	21	65	21	18	19	40
Numerical Graph Interpretation: Task 6 Noise vs Number of people								
0. Nonstatistical	4	3	5	0	5	4	2	0
1. Single Statistical Aspect	2	8	10	16	9	7	11	6
2. Inadequate Covariation	6	1	5	23	6	7	6	12
3. Appropriate Covariation	0	1	0	26	1	0	0	22
Total students (<i>N</i>)	12	13	21	65	21	18	19	40

Table 7.08.

Frequency of Student Responses at each Level of Numerical Graph Interpretation for Task 6 by Level of Verbal Graph Interpretation for Task 6

Response level	Level of Verbal Graph Interpretation			
	Task 6 Noise vs Number of people			
	0	1	2	3
Numerical Graph Interpretation: Task 6 Noise vs Number of people				
0. Nonstatistical	2	6	4	0
1. Single Statistical Aspect	7	10	19	3
2. Inadequate Covariation	4	7	15	14
3. Appropriate Covariation	0	2	5	23
Total students (<i>N</i>)	13	25	43	40

Table 7.09 shows a summary of these comparisons in relation to higher, lower, or the same levels for pairs of response levels to different tasks. Of 133 students, Speculative Data Generation response levels for Task 4 (Test scores) were higher, the same, and lower than Coordinate Graph Production levels for 55, 62, and 16 students respectively, providing some evidence that for these students, understandings to support levels of Speculative Data Generation were developed prior to those for Coordinate Graph Production. Similar observations can be made suggesting Speculative Data Generation levels were higher than those for Verbal Graph Interpretation and Numerical Graph Interpretation.

Table 7.09.

Frequency of Comparison of Response Levels of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation

Level of Speculative Data Generation for Task in row compared to Level for Task in column	Skill and Task		
	Speculative Data Generation Task 4	Coordinate Graph Production Task 5	Verbal Graph Interpretation Task 6
Coordinate Graph Production Task 5 (Temperatures)			
Higher	16	-	-
Same	62	-	-
Lower	55	-	-
Verbal Graph Interpretation Task 6 (Noise)			
Higher	18	25	-
Same	50	53	-
Lower	43	20	-
Numerical Graph Interpretation Task 6 (Noise)			
Higher	10	23	25
Same	43	40	50
Lower	57	35	46

Students tended to offer lower levels of response to tasks involving Numerical Graph Interpretation than those involving Verbal Graph Interpretation and Coordinate Graph Production (cf. Table 7.07). In combination, these findings are shown in Figure 7.01, with arrows suggesting a developmental order (1) Speculative Data Generation, (2) Verbal Graph Interpretation, (3) Coordinate Graph Production, and then (4) Numerical Graph Interpretation. In this sense, the skill to produce coordinate graphs appears often to be developed prior to (and possibly enables) the skill to interpret numerical trend on a graph, likely due to the common need to understand the coordinate system. An obvious limitation of evidence of this model is the assessment of skills using only a single task.

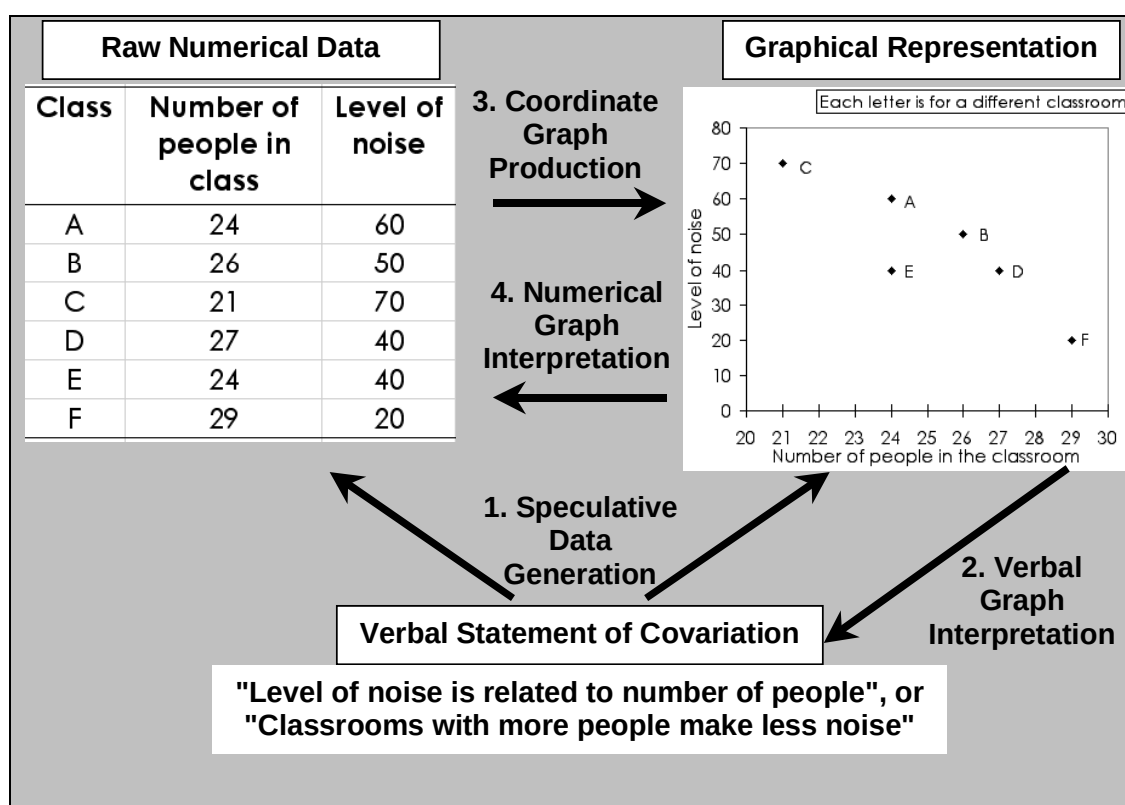


Figure 7.01. Proposed development order of four skills for understanding and representing statistical covariation.

7.06 EXAMPLES OF ASSOCIATIONS AMONG RESPONSES

Responses across various tasks for each of seven students are presented in Figures 7.02 to 7.08. The third-grade female who offered responses shown in Figure 7.02 illustrated developed skills for Speculative Data Generation and Verbal Graph Interpretation in advance of Coordinate Graph Production and Numerical Graph Interpretation, consistent with the developmental order suggested in Figure 7.01. Her responses for Speculative Data Generation showed data but not using conventional graphing methods. Her responses for Coordinate Graph Production and Numerical Graph Interpretation also indicated lack of graphing skills for production or interpretation of coordinate graphs. Her response for Verbal Graph Interpretation ("If

you have more students, it makes more noise”) was coded as Level 2, being bivariate but in the wrong direction. When considered in relation to her apparent lack of graphing skills, this response, and others like it, could arguably have been considered at Level 0 (Nonstatistical) based on contextual understanding without engaging the graph. This set of responses illustrates in part the developmental model proposed in Figure 7.01, namely that Speculative Data Generation may develop prior to Verbal Graph Interpretation and the adoption of graphing conventions to support Coordinate Graph Production and Numerical Graph Interpretation.

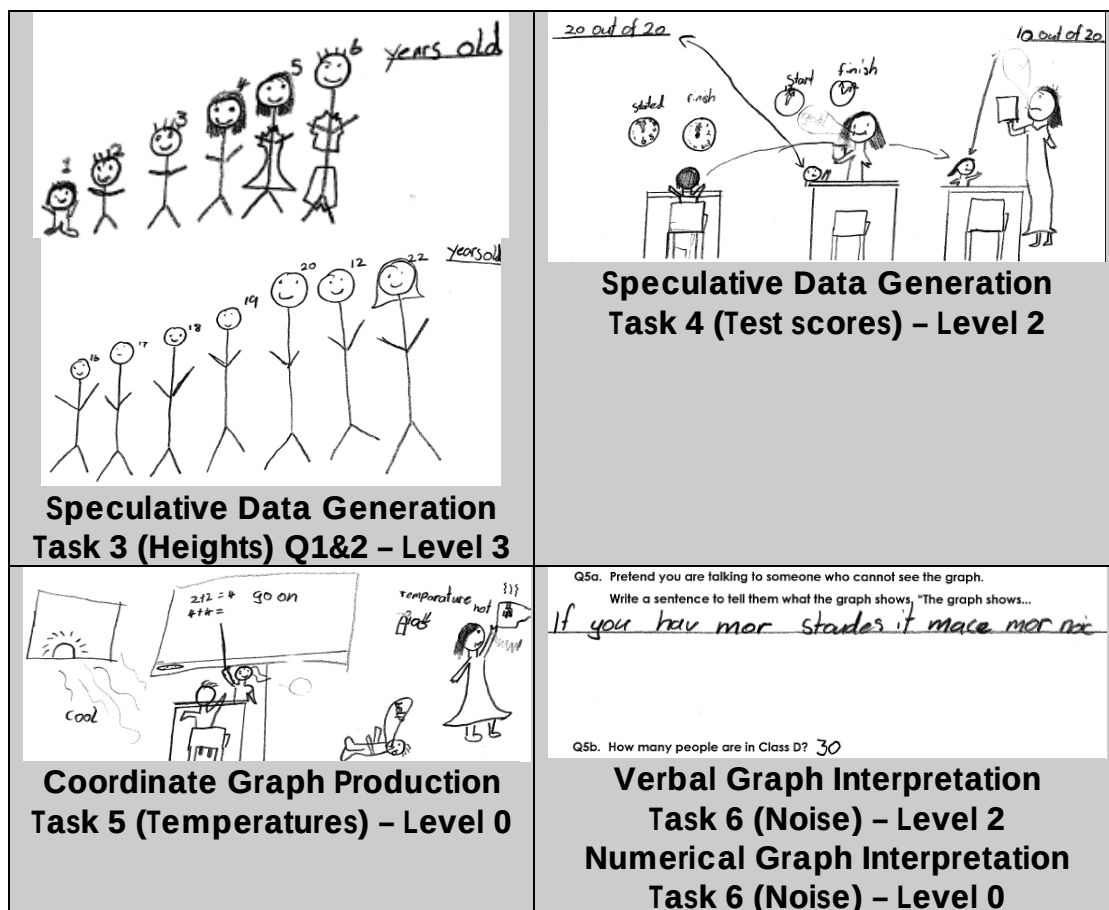


Figure 7.02. Responses from a Grade 3 female.

Another third-grade female offered responses shown in Figure 7.03. Her responses for Speculative Data Generation to Tasks 3 and 4 were similar to those in Figure 7.02, but with the addition of axis frameworks. Throughout her responses, she evidenced thinking in relation to the context of each task, which found expression in tasks for Speculative Data Generation, but not for Coordinate Graph Production or tasks involving Verbal Graph Interpretation and Numerical Graph Interpretation that involved graphing conventions. These responses were consistent with the developmental order suggested in Figure 7.01, highlighting that limited understanding of graphing conventions did not limit the student in generating speculative data involving covariation. For Coordinate Graph Production, she did not effectively use the axis framework, and she listed the six time values but made up values for temperature without regard for the statistical data provided. For Verbal Graph Interpretation, she read a series of noise values, as if the graph represented a single variable, and did not offer evidence of reading the coordinate system and scales in relation to Numerical Graph Interpretation.

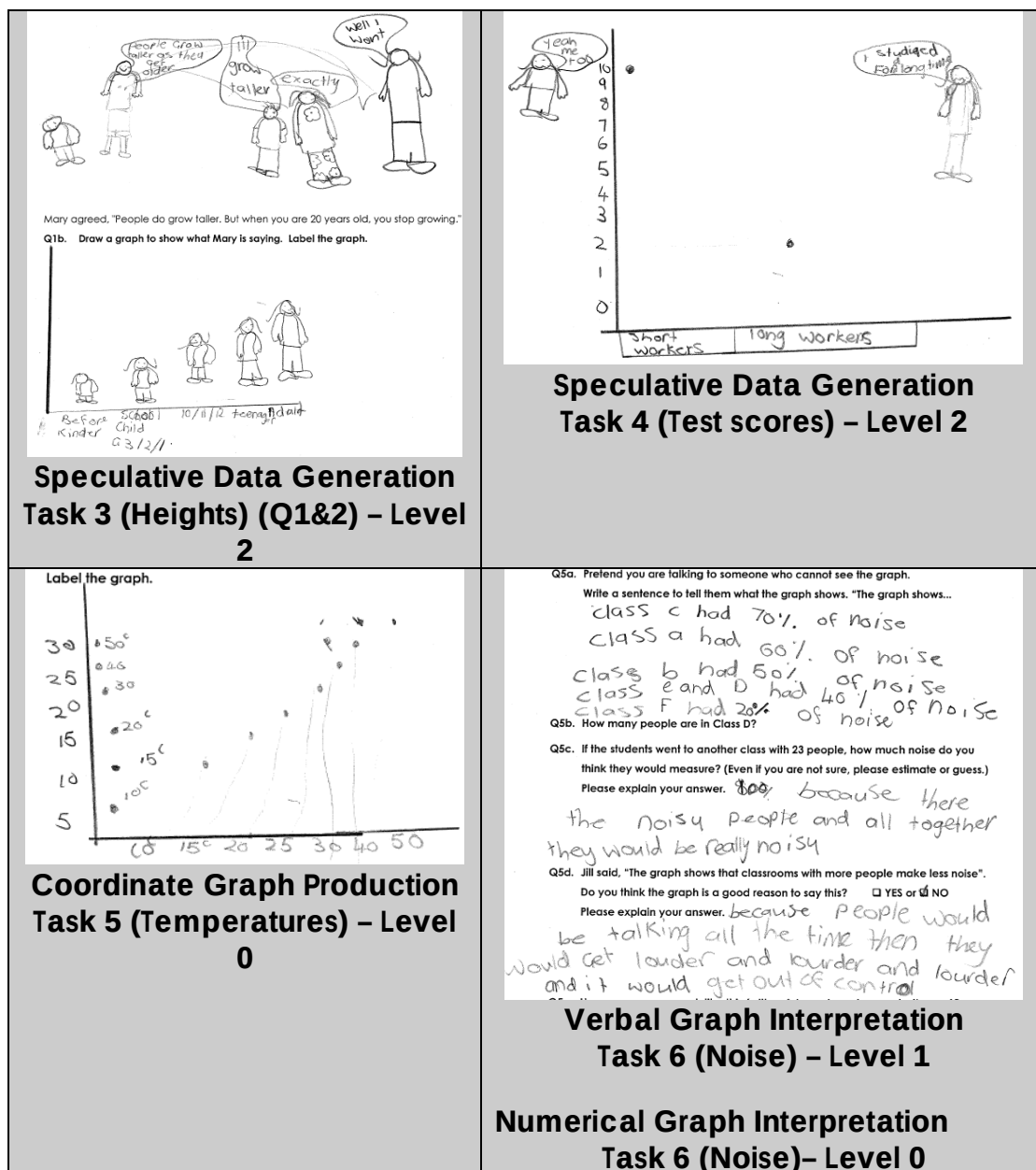


Figure 7.03. Responses from a Grade 3 female.

The fifth-grade male whose responses are shown in Figure 7.04 drew bar graphs for all graphical responses. He was one of many students who responded at similar levels across different tasks. He appeared to approach tasks for Speculative Data Generation drawing bars for comparison case by case. For Task 3 (Heights), he showed a pair of bars, for Task 1 (Heart deaths), he showed a pair of bars labelled with case years, and for Task 4 (Test scores), he showed six bars numbered as cases and separated from each other by vertical lines. His response for Verbal Graph

Interpretation concerned a single bivariate data point, but this case was selected as a superlative “least people, most noise” to provide some comparison to other values. His response for interpolating in relation to Numerical Graph Interpretation referred to two nearby points as references. Hence by bars or verbal references, these responses illustrated the case-by-case construction of covariation, which supported similar response levels across tasks for different skills.

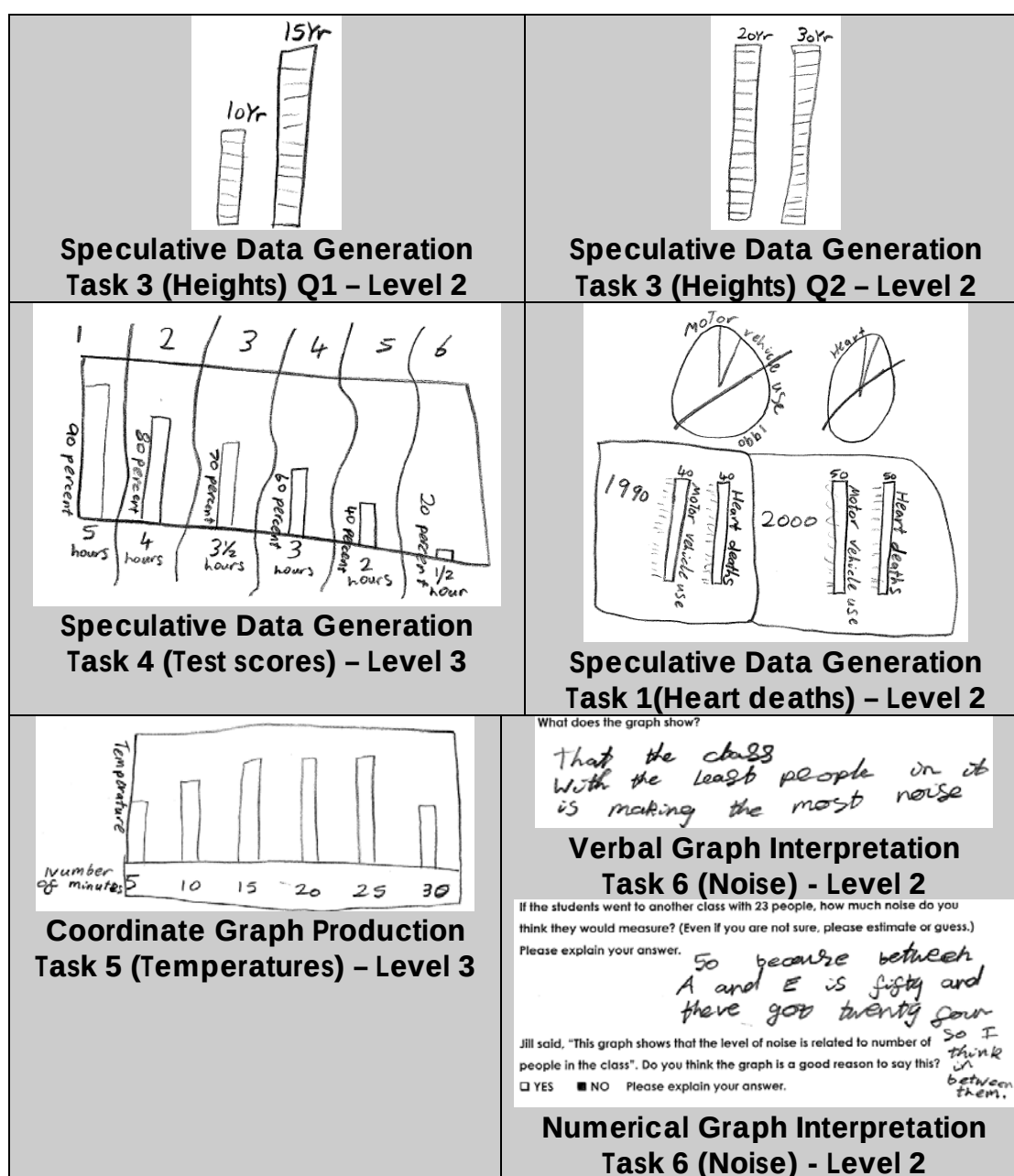


Figure 7.04.

Responses from a Grade 5 male.

Responses from a seventh-grade female in Figure 7.05 illustrated ways that quite different response levels were sometimes achieved for different tasks due perhaps to the presence of elements in the tasks to support the structure of the response. She drew Series Comparison responses in two bar graphs for each of Task 3 (Heights) (Q3) and Task 1 (Heart deaths), but only offered a Single Variable response showing test scores for Task 4 (Test scores); had she chosen to draw another bar graph showing corresponding study times, she would have presented another Series Comparison response. The response for Task 4 (Test scores) showed nominal case labels of people's names. These labels involved thought about each case to create names, and did not encourage ordering of names, nor of test scores. Hence the graph data appeared as six discrete cases. In contrast, for Task 3 (Heights) and Task 1 (Heart deaths), data were ordered by case label and by the variable on the vertical axis, hence supporting conceptualisation of a variable.

Series Comparison graphs also featured in responses of a seventh-grade male, as shown in Figure 7.06, who responded at consistently high levels across the tasks. He demonstrated Coordinate Graph Production skills for Task 5 (Temperatures), and also in producing the curvilinear relations with graphing conventions for Task 3 (Heights) (Q3). His response to Task 4 (Test scores) is noteworthy in two respects. The first is the level of response: he drew the graph in the wrong direction for the task, but was aware of this, stating in a subsequent question (not analysed but shown in Figure 7.06) that his graph showed “students who studied longer got better marks.” Hence his response did not indicate limited skills for Speculative Data Generation or Coordinate Graph Production, but rather his unwillingness to accept the given verbal statement as possibly true. The second noteworthy aspect of this response was the bar graph form and case labels (S1, S2, ...). As observed for the

female response in Figure 7.05, it is possible that the time basis, in years, for Task 3 (Heights) (Q3) and Task1 (Heart deaths) was what supported the Series Comparison responses.

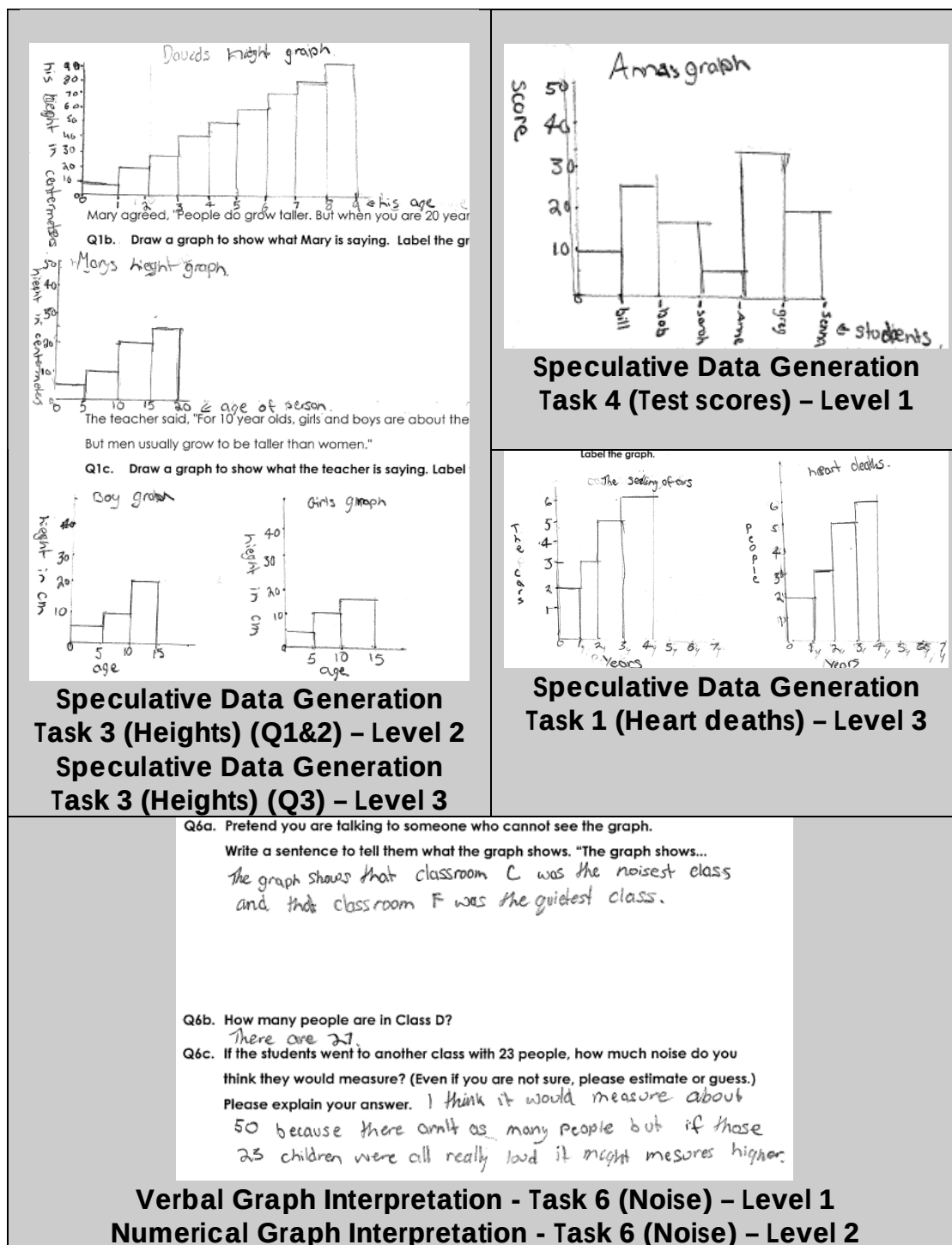


Figure 7.05.

Responses from a Grade 7 female.

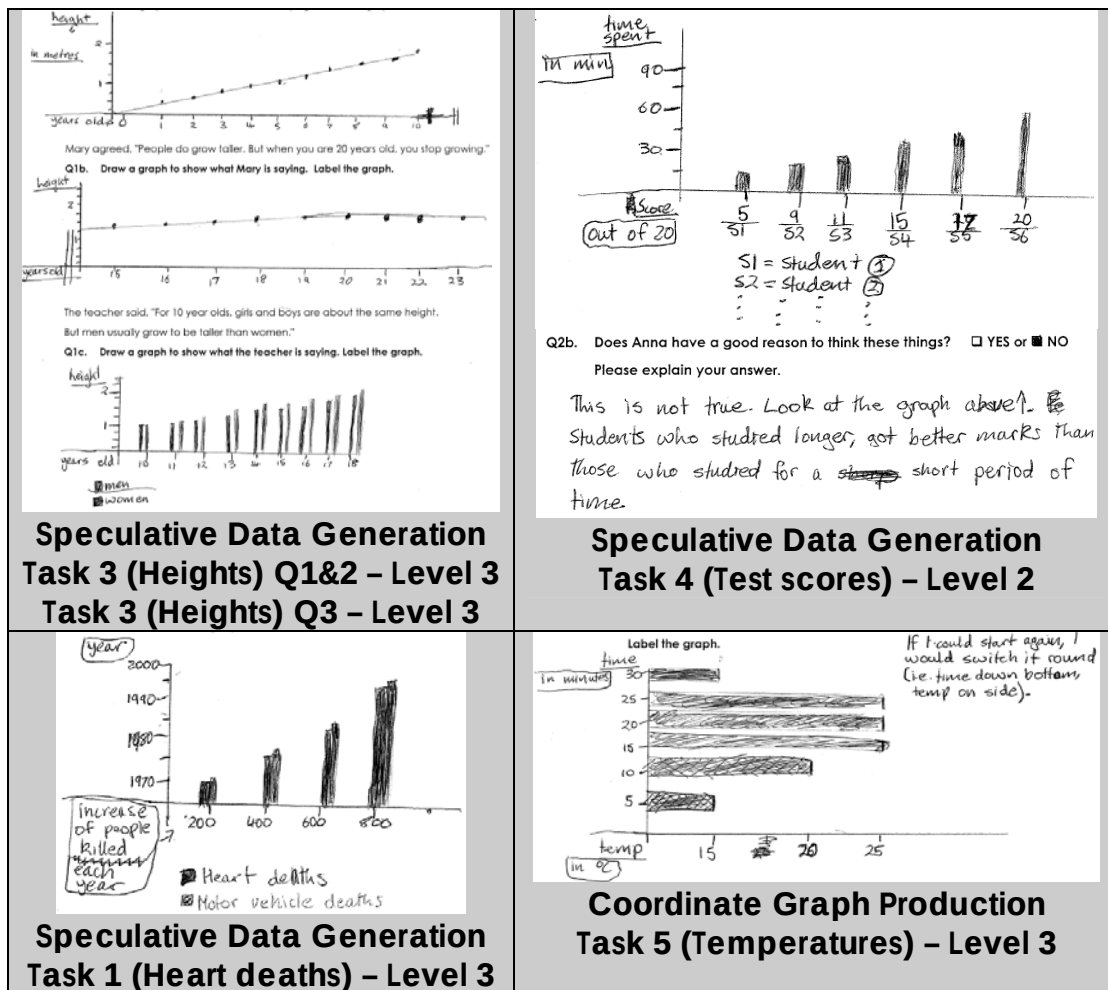


Figure 7.06. Responses from a Grade 7 male.

The responses of the seventh-grade male shown in Figure 7.07 indicated a consistent use of the coordinate system. He offered Coordinate Variable responses to Tasks 4 (Test scores) and 1 (Heart deaths), and he showed a Series Comparison to introduce a third variable when required for Task 3 (Heights) Q3. His use of line graphs suggested a view of continuous change rather than casewise consideration, indicated by the drawing of line segments on the scattergraph for Task 6 (Noise) to support interpolation, and even in Task 4 (Test scores) by the representation of more than the expected six data cases.

Responses from a ninth-grade male showed further sophistication using the notion of continuous change. He employed line graphs without featured data points for Task 1 (Heart deaths) and Task 5 (Temperatures), and produced a curved line of

fit for the scattergraph in Task 6 (Noise). His responses in relation to Verbal Graph Interpretation were somewhat surprising in relation to other responses: he only referred to a single variable of the noise range of values, and for Task 6 (Noise) Q4, he offered contextual ideas to override the trend, which he even misquoted for Q3, "... the noise gets lower as the people numbers go down." It is speculated that such inconsistencies for this student do not reflect limited understandings, but random variation in assessment measurements due to casual or careless responses.

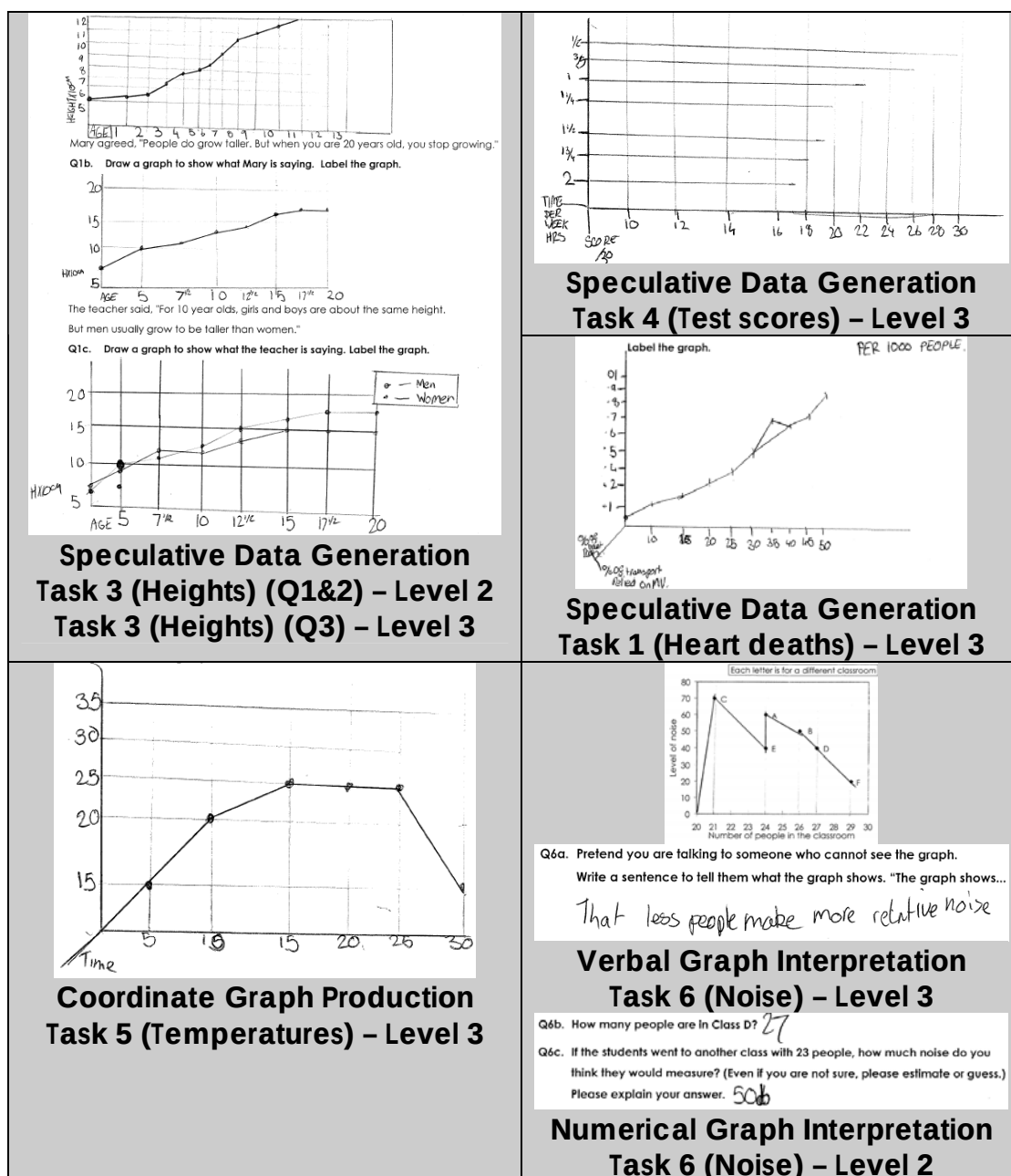


Figure 7.07. Responses from a Grade 7 male.

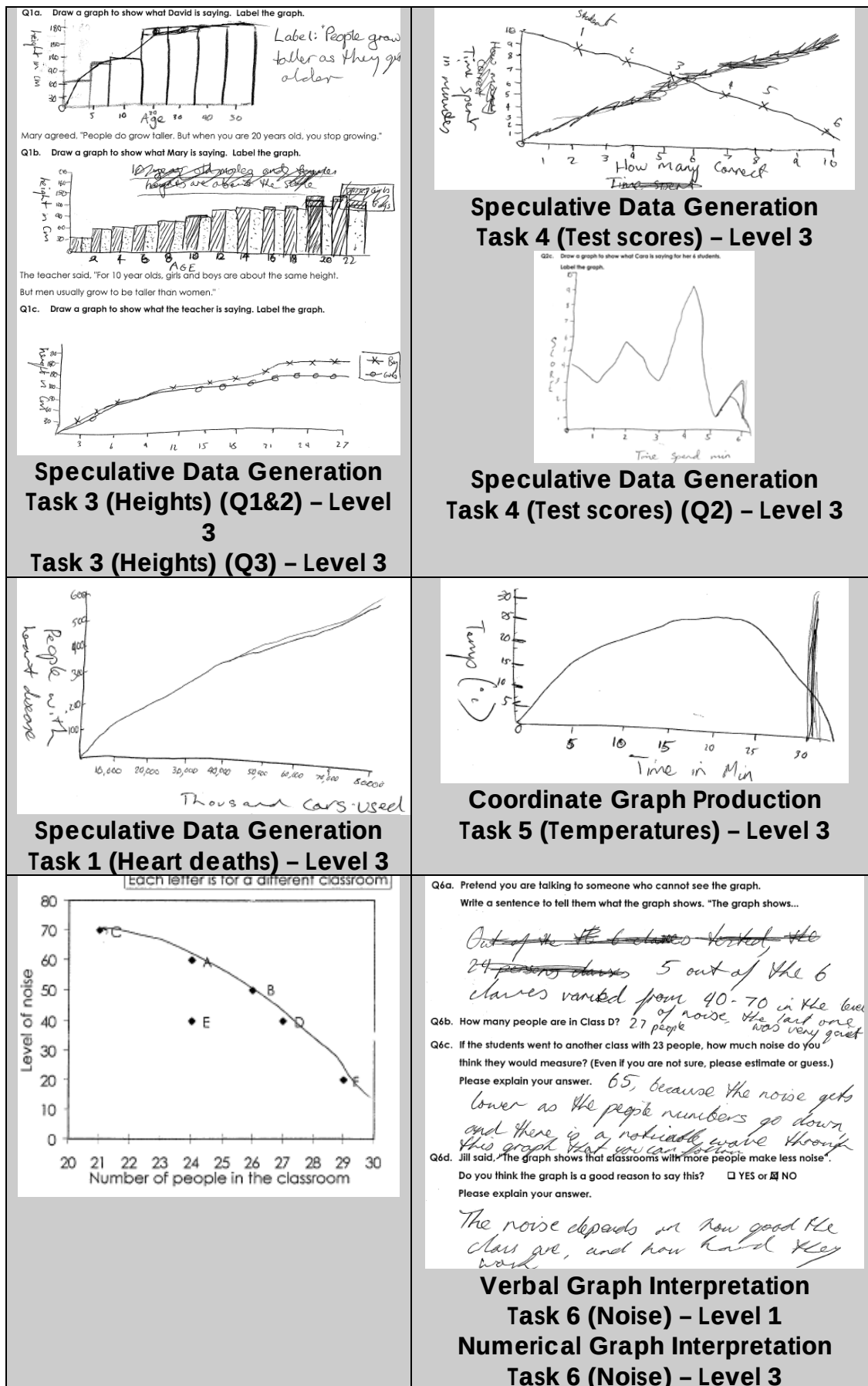


Figure 7.08.

Responses from a Grade 9 male.

7.07 DISCUSSION

Response levels for various skills were related to each other according to the assessment frameworks related to understanding and representing statistical covariation. Moderate correlations were observed among responses to different tasks for Speculative Data Generation, likely due to some differences among tasks of the same skills, such as task features of complexity of data and context. Moderate correlations were also observed among responses to tasks for different skills of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation, likely due to commonalities of these skills, such as graph design.

There was some indication in these data that the skill of Speculative Data Generation—involving data handling with a sense of global covariation—may commonly develop prior to the skill to produce a coordinate graph to represent such covariation. This finding follows Chick and Watson (2001) in suggesting that the ways in which students structure data are the bases for classifying levels of representations.

There was also indication that Verbal Graph Interpretation may develop prior to Coordinate Graph Production, suggesting that the global sense underlying global interpretation in verbal statements may be important to support conceptualisation to produce coordinate graphs. Numerical Graph Interpretation at Level 3, to interpolate values based on a trend shown in a coordinate graph with reference to the numerical scales, was observed to be the most challenging level to achieve. This finding was possibly due to the demands of understanding the design and scales of the coordinate system, and also possibly due to the task involving covariation that was counterintuitive.

The examples provided also illustrated students' use of similar graph forms and structures across various tasks. In some cases, however, students' responses across tasks showed evidence of re-structuring or of varying emphases, apparently a function of the task features.

CHAPTER 8. GENERAL DISCUSSION

Discussion of results in relation to the research aims of each investigation were provided in Chapters 4 to 6, and findings across various investigations were considered in Chapter 7. The following discussion explores further themes in relation to the literature review provided in Chapter 2. Discussion includes how the findings of this study have contributed to the research literature in relation to developing students' understandings and representations of statistical covariation, and implications of these findings for teachers. Limitations of the research are acknowledged, and suggestions for future research are offered.

8.01 ASSESSMENT FRAMEWORKS

Four levels for each of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation were described. Levels indicated degrees of success in representing correspondence of variation, whether in a speculative data set, a coordinate graph, a verbal statement, or a numerical interpretation. The four levels – Nonstatistical, Single Statistical Aspect, Inadequate Covariation, and Appropriate Covariation – related closely to a general model of cognitive development (Biggs & Collis, 1982) concerning the structure of the observed learning outcome, involving prestructural, unistuctural, multistuctural, and relational levels. They also related to (a) levels of graph interpretation tasks, from reading values, reading between values, to interpreting beyond values to a trend (Curcio, 1987), and (b) the development of data interpretation suggested by mathematics and science educators, from local values to global trends (Ben-Zvi & Arcavi, 2001; Gerber et al., 1995), and more specifically, from no variables, to a single variable or data case, to two variables or data cases, and finally to multiple

variables or two variables appropriately controlled (Donnelly & Welford, 1989; Ross & Cousins, 1993b; Swatton, 1994; Swatton & Taylor, 1994). The frameworks were similar for each of Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation and Numerical Graph Interpretation, drawing common themes from a diverse literature including mathematics and statistics education, science education, and psychology. The frameworks suggested levels for previously isolated categories observed for Speculative Data Generation (Mevarech & Kramarsky, 1997) and graph production (Brasell & Rowe, 1993; Chick & Watson, 1998, 2001; Wavering, 1989).

Within each level, categories of response were evident, often emphasizing either (a) correspondence or comparison of measures in bivariate data cases, or (b) variation of values of measures. These aspects of covariation have been previously acknowledged (Clement, 1989; Nemirovsky, 1996a; Wavering, 1989). At Levels 1 and 2, categories differed in the presence and structuring of the relevant elements of the task, which included identifying appropriate variables, showing correspondence of values, and showing variation. Ordered Graphs and Trend Graphs illustrated corresponding variation of values but not labeling of appropriate variables for which the covariation was intended. Single Comparison and Double/Group Comparison responses illustrated correspondence of values but not appropriate variation for each variable. Single Variable and Double Variable responses illustrated variation of values, that is, variables, but not appropriate correspondence between the variables. Implications for teaching thus depend not only upon taking into account students' levels of response but also upon which of these relevant aspects students demonstrate or fail to demonstrate. Some categories, such as Single, Double, and Series Comparison responses, were similar to types of interpretation tasks (Meyer et al.,

1997) and responses to graphing tasks (Brasell & Rowe, 1993; Mevarech & Kramarsky, 1997). Evidence from various tasks of the importance of case labels for conceptualising graphing and for providing a language to refer to data cases strengthened findings from classroom learning about statistical covariation (Cobb et al., 2003), and evidence of links between ordering and comparative language supported research about functions (Yerushalmy, 1997).

As was stated as Research Aim 1 (Section 3.02.01), the assessment frameworks were designed to provide support for teachers when assessing levels of student responses. They are also considered to be useful to inform teaching aspects that might promote development of students' understandings and representations of statistical covariation.

8.02 TASKS TO ASSESS THE SKILLS

The tasks elicited responses from students in grades 3 to 9 that informed the levels and categories of the assessment frameworks for Speculative Data Generation, Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation. In particular, tasks were designed that assessed the skill of Speculative Data Generation, a skill not previously named within the research literature. Although some of these tasks were strongly influenced by Graph Production skills (e.g., Task 3 – Heights), others (e.g., Task 4 – Test scores) elicited tabular responses illustrating Appropriate Covariation (Level 3), indicating the possibility of assessing Speculative Data Generation without demonstration of Graph Production skills.

Some issues of task design were similar to those faced in broader education settings, such as supporting accessibility of the task by using familiar contexts and language, facilitating student expression of ideas through open-ended task wording,

and scaffolding parts of tasks to assess various levels of understanding. Other issues of task design were more specific to covariation.

- Tasks differed on *variables involved*, from realistic multivariate contexts in which the relevant variables had to be discerned (e.g., Task 1 – Heart deaths), contexts that were clearly bivariate (e.g., Task 6 – Noise), contexts that indicated differences either among or within cases (e.g., Task 3 – Heights), and contexts for which variables were familiar but measurement required consideration (e.g., Task 4 – Test scores). Hence a variety of tasks can be successively employed with students to explore these different issues.
- *Counterintuitive covariation* was considered important to distinguish whether students were making knowledge-based or data-based judgements.
- *Constant or repeated values* were considered important to assess ability to represent independent variables in a coordinate system and to explore deterministic conceptions of covariation.
- In some cases, *data sets with particular features* supported tasks. For example, Numerical Graph Interpretation could be assessed at Level 3 in a survey task, where explanations are often limited, only by employing a task with a data set structured so that the value interpolated based on a trend (Level 3) would be clearly distinguishable from a value based on reading of local data points (Level 2).

8.03 PERFORMANCE BY GRADE LEVEL AND ACROSS TASKS

Students from grade 3 and above engaged the tasks, and in general they were appropriately challenging for students in grades 3 to 9. Some students from grade 3

responded at the highest level (Appropriate Covariation, Level 3) for each of the tasks. This demonstrated performance to types of tasks that teachers may not be using with their students. Coupled with tasks of Speculative Data Generation, which are novel to many, this study can inform teachers of ways students might be able to engage tasks above their teachers' expectations.

Students from higher grades were more likely to respond at higher levels, however, for some tasks, less than half of the students in any grade represented covariation appropriately. Across most tasks of Speculative Data Generation, levels were generally higher than observed by Swan (1988), where 37% of 192 thirteen- to fourteen-year-olds appropriately drew a graph that was decreasing, and by Mevarech and Kramarsky (1997), where 55% of eighth-grade students successfully represented claims of bivariate association. Performance comparisons to other studies must include acknowledgement of the criteria for judging representational adequacy: when assessing some responses in this study with respect to Speculative Data Generation, higher levels were assessed for responses with little or no evidence of traditional conventions of graph production, or for responses such as those based on Series Comparisons, which others (e.g., Brasell & Rowe, 1993) assessed at lower levels.

The findings across tasks indicated that task features, particularly familiarity of variables and involvement of time as a variable, may affect performance for Speculative Data Generation. Response levels for Speculative Data Generation in this study were highest for Task 3 (Heights). This task was supported by intuitive contexts and representations, such as height and age, variables previously observed to support high level responses (Ainley, 1995), and similar to the natural mappings and times-series graphs that featured early in the historical development of graphing (Biderman, 1990; Tilling, 1975). Speculative Data Generation skills appeared to

develop for many primary school students prior to the appreciation of coordinate system conventions that influence Coordinate Graph Production and Graph Interpretation of coordinate graphs. These results echo some previous research (e.g., diSessa, 2001; Krabbendam, 1982; Nemirovsky & Tierney, 2001), and support curriculum expectations from Australia, the United States, and New Zealand:

- that primary students should have experiences with functions to “represent (verbally, graphically, in writing and physically) and interpret relationships between quantities” (AEC, 1991, p. 193);
- that in grades 3-5, as part of the algebra standard of “analyze change,” students should “represent and analyze patterns and functions, using words, tables, and graphs” (NCTM, 2000, p. 158) and “investigate how a change in one variable relates to a change in a second variable” (p. 158); and
- in upper primary grades, students should be “collecting and graphing simple time-series data such as the height of a classroom-grown bean plant at midday each day” (ME, 1992, p. 179).

In particular, this support is for active involvement in analysing data (e.g., NCTM, 1989), rather than prescriptive graphing outcomes that might imply, for example, that students in primary school draw univariate graphs and only engage bivariate data in secondary school (AEC, 1994). This suggests it may be most effective to begin teaching of coordinate graphing with hypothesizing relationships between variables, which introduces a clear purpose to undertake data collection and exploratory data analysis. Experiences could also include production and interpretation of qualitative graphs without specific numbers (e.g., Leinhardt, et al., 1990; Mevarech & Kramarsky, 1997). Students should be encouraged to engage and represent data in ways meaningful to them rather than, or at least prior to, emphasizing the mechanics

of conventional graph construction, such as scaling axes. When graphs are to be interpreted by others, however, there may be good reason to adopt (or adapt) conventions, such as coordinates.

8.04 COGNITIVE CONFLICT

Some cognitive conflict was observed in survey responses to some tasks, such as the introduction of the constant function into a task that prompted re-structuring graphs towards the coordinate system. Careful task design, such as “garden-path tasks” (Posner & Gertzog, 1982) leading students along a path to see the consequences of their conceptions, not only may assess student misconceptions, but also may act as a trigger to help students become aware of inconsistencies in their reasoning.

Students responding to graphs of others in interviews commented both on features of graphical presentation and on the whether the speculative data were appropriate to the verbal statement given. In some cases, these two aspects – presentation and data – resulted in ambivalence. As a research technique, the various investigations in this study involving cognitive conflict (Investigations 3B, 4B, 5B) did not always yield conclusive evidence whether students had changed their minds. Other studies have employed the cognitive conflict method in relation to representing pictographs (Watson & Moritz, 2001b), chance measurement (Watson & Moritz, 2001a), and comparing univariate distributions (Watson, 2002). In all of the studies, prompts differed on a single critical feature that involved clear conceptual differences: configuration in clumps or rows for pictographs (Watson & Moritz, 2001b), and frequency versus proportional comparisons for chance measurement (Watson & Moritz, 2001a) and for comparing univariate distributions (Watson, 2002). The ambivalence evident in the current study highlights the complex

intertwining of Speculative Data Generation and Coordinate Graph Production, and the need to future research to isolate aspects for analysis purposes.

Disagreement with new prompts was more common than agreement, indicating a generally conservative approach to changing ideas. Similar findings have arisen from previous studies (Watson, 2002; Watson & Moritz, 2001a, 2001b). Watson (2002) hypothesized “greater success rates for the use of cognitive conflict, if the tasks are easier and not as much improvement is required to reach an optimal response” (p. 250). In the current study, agreement was slightly more likely for prompts at levels close to the level of the interviewee’s initial response. This finding might be expected based on Vygotsky’s notion of the *Zone of Proximal Development* (Forman & McPhail, 1993; Goos, 2000) in which the range of intelligible ideas includes those at a slightly higher level than the student alone could supply. This indicated, in some cases, promise for the cognitive conflict technique for developing student reasoning. As a teaching technique, critique of other responses differing on various aspects, including the data itself and the representation, allows teachers to lead discussions concerning how specific features such as the variables, data structure, data measurement units, graph form, and scales contribute to supporting understanding of statistical data and covariation.

8.05 EDUCATIONAL IMPLICATIONS

The following discussion addresses the levels observed and certain student conceptions evident within response categories at each level. Consideration is given throughout to identifying teaching strategies that may assist in developing student understanding to respond at higher levels.

8.05.01 From Single Data Points to Global Trends

Many students referred to one or two bivariate data points or drew Single or Double Comparison graphs, that is, comparing one or two bivariate data points. Students' reasoning about isolated data points reduced the complexity of the task by focusing on local, individual values rather than on global trends (Bell et al., 1987a; Leinhardt et al., 1990). Responses emphasized correspondence of two measures but indicated no variation in either measure, which made it impossible to indicate how the two measures vary together. Adding bivariate points to build variation of values has been described previously as the pointwise approach to teaching covariation (Nemirovsky, 1996a).

Development of the pointwise approach in verbal interpretations may be considered as a progression of comparisons within variables, from single-point values ("Class C had 21 children...") to comparison of points ("the classroom with the least people is the noisiest...") to generalizing beyond the available points ("the more people the less noise..."). This follows the levels of "reading the data," "reading between the data," and "reading beyond the data" described by Curcio (2001). For Speculative Data Generation, a pointwise approach was the building block used by some young students who added more data points to initial values showing correspondence. In generating more points, students appeared to find it easy to maintain the appropriate correspondence between the measures. The difficulty in generating more data points appeared to be generating appropriate variation that ensured both numerical variables did vary; some resorted to group comparisons or repeated values.

Teachers can encourage students to compare one bivariate data point with another, as a prerequisite to asking questions of the whole data set such as, "as one

measure increases, does the other measure increase or decrease?” One step toward this goal was observed in Double Comparison graphs, which represented two values for each of two variables. This is similar to a partially correct strategy identified by Estepa and Batanero (1996) titled “correct interpretation of isolated points.” The addition of values makes the points less isolated and permits comparisons for two different values. Further addition of values appeared in Series Comparison graphs that involved a comparison of two data series to show covariation.

Pointwise approaches may provide an important way into many statistical issues—such as repeated values in either variable and the contextual understanding of data elements involving measurement and sampling issues—that do not occur in algebraic studies of continuous functions. In this respect, Tables and Series Comparison graphs may be significant representations for reasoning about covariation, since they devote a feature (column or axis) to retain case information, such as the name of a person, and can represent two cases with identical bivariate values, which are slightly problematic to display in Cartesian coordinates.

In Investigation 6A, many students drew tables of data, and when interviewed, these students often read corresponding values from tables or graphs. A preference for tabular data to show values has a strong tradition in the early history of graphing (Biderman, 1990) and persists today in academic circles – for example, the presentation of results in this study, as seen in Table 5.01. Nevertheless, students should be encouraged to use the value of graphs for showing change, or as Playfair put it, “gradual progress and comparative amounts, at different periods” (Funkhouser, 1937b, p. 281). It may be, however, that the notion of “gradual progress” of trends is not simple to master; students may be able to read discrete data points in a coordinate scatterplot, but fail to identify a trend when asked to verbalise

it or interpolate (Investigation 8A). The idea of “comparative amounts,” however, may be accessible from iconic forms such as two thermometers side-by-side, used by some students. Notably, in his youth, Playfair also represented temperatures on a divided scale according to the literal height on a thermometer, forming the basis for his “lineal arithmetic” and development of graphing (Biderman, 1990). This strategy can build to comparison of data slices in scatterplots, as advocated by Cobb et al. (2003).

8.05.02 Ordering Variation to Establish Correspondence

An important idea for development of reasoning beyond isolated points or dichotomous extremes may be the ordering of cases on a single variable (Ross & Cousins, 1993b; Wavering, 1989). For Speculative Data Generation, it is possible to generate new cases that have incrementally more or less of one measure, often at fixed differences, and then simply increment the other variable appropriately. Such fixed differences move students away from considering isolated cases that may include repeated values, to a generation of patterns within a variable, an idea that is frequent in algebra. For real-world data variables, generating new values may be restricted by the minimum or maximum possible values. For verbal interpretation, ordering of one variable allows variation of the other variable to be observed as an increasing or decreasing feature of the data series (a trend) verbally summed up as a single phrase, thus structuring the graphic language of the data series to correspond to the verbal language of change (Yerushalmy, 1997).

Students in this study who drew Single Variable and Double Variable graphs did not establish correspondence between variables. The notion of ordering values of one variable, and then observing the corresponding order of values for another variable, has been described as the variational approach to teaching covariation (Nemirovsky,

1996a). Ordering values, described by Wavering (1989) as seriation, is an important data handling skill necessary to gain an appreciation of the variation present (Ross & Cousins, 1993b). When the variable is numerical but measured across categorical cases, it is not apparent whether the cases should be ordered by value or by case information such as alphabetically. The notion that the case information, or the variable on the horizontal axis, could be ordered to aid global interpretation is what may be needed for students to progress from Single Variable graphs to Coordinate Variable responses. Once one variable is ordered, attention can then be diverted from this variable and devoted to how the other measure correspondingly varies (Konold & Higgins, 2003; Krabbendam, 1982; Nemirovsky, 1996a). Students drawing Double Variable responses could be asked whether or not the horizontal axis denotes ordered values and if so, requested to examine whether the values on the vertical axis are consistently increasing, consistently decreasing, or alternating as one moves left to right on the graph.

Selection of appropriate variables for tasks may be important to assist students to recognize the importance of ordering values. Teachers could initially select variables that are commonly understood as ordered, such as height or time (Krabbendam, 1982). Use of height as a variable naturally aligned with the vertical axis may assist in encouraging young students to transform stylized pictures into conventional graphs (Ainley, 1995). Time-series graphs are considered important in the mathematics curriculum (e.g., ME, 1992), and for early algebra learning (e.g., NCTM, 2000). As distance-time or speed-time graphs, suggested in the Tasmanian curriculum (DEA, 1993), are known to cause confusion for some students who interpret the graphs as pictures (Bell et al., 1987a, 1987b; Brasell & Rowe, 1993), it may be helpful for teachers of secondary school students to give more attention to

time-series graphs for which there is no connotation of position. For example, in tasks involving air temperature over a 24-hour period, change over time is readily expressed in verbal terms (Yerushalmy, 1997), such as “during the morning it’s getting hotter.” Historically, times-series graphs were significant in the development of statistical graphing, building on the prevailing Cartesian graph form used in idealized settings in the 17th and 18th centuries, to get real data (economic or physical measures) into line graphs and bar graphs (Biderman, 1990). For today’s students who are often exposed to bar graphs long before Cartesian coordinates, times-series graphs may provide a conceptual bridge from bar graphs to the coordinate system in which each axis is ordered.

8.05.03 From Single Variables to Bivariate Data

Some students drew graphs or referred to Single Variables, and many described only the variable noise in verbal descriptions of a scattergraph. These students emphasized variation but did not describe correspondence of two measures to indicate covariation adequately. Those who had success in verbally describing the covariation all used the language of incremental change across cases, implied by ordering each variable, rather than objectifying the correspondence as “X is related to Y.” Interpolation tasks, though numerical and often involving reference to specific points, may in fact encourage students to discuss differences between points and lead to discussion of increments more globally.

A change-over-time approach to covariation has been recommended by algebra curricula (e.g., NCTM, 2000) and researchers (e.g., Nemirovsky, 1996b). Such an approach carries with it implicitly the understanding that time is ordered, and thus verbal phrases such as “it started to grow faster, then it slowed down again” (NCTM, 2000, p. 163) allow students to focus on change of one variable without attending to

the correspondence of the variables, as is required if the independent variable is not time. Tables and Series Comparison graphs may be significant representations not just for developing reasoning to include more cases as noted earlier but also for emphasizing both variables and the correspondence of individual data values. Both of these representations treat each variable as a measured variable (often termed dependent and, if graphed, represented on the vertical axis) across a number of cases, whereas Cartesian graphs have axes conventionally considered independent (horizontal) and dependent (vertical). Aside from the implication of dependency and possibly causation, some students do not attend to the variable on the horizontal axis, such as in the many interpretations involving only the variable noise. Tables and Series Comparison graphs may be considered as natural progressions composed of two Single Variables tables or graphs. As already noted, ordering of values is a key concept that allows not only handling of variation, but also establishing correspondence case-wise. Once cases are ordered by one variable, such as in the horizontal dimension, the foundation is set for coordinating the correspondence of two variables in Cartesian coordinates. The transformation from an ordered Table (no use of dimension), or from an ordered Series Comparison graph with both data series in an axis framework (both variables denoted by vertical dimension), to the Coordinate Variable form was observed in students' responses to the task of Coordinate Graph Production, where bivariate cases had been ordered by one variable in the horizontal dimension, and vertical height incorporated to denote variation in the other variable. In these representations, moving the written value labels from the data elements to the axes results in Cartesian coordinates.

8.05.04 From Prior Beliefs to Data-Based Judgments

Nonstatistical attempts by students did not represent the data provided or indicated in the task, offering non-response, unlabelled graphs, or contextual responses. Some cases of non-response were probably due to lack of motivation to complete a survey not forming part of the student's assessment. Other non-response may reflect students being too tentative to speculate on data or the variables involved, possibly not having experienced such a task previously. Evidence from interviews indicated that minor clarification or probing from the interview supported students to provide a response at a higher level related to the data-based context, though often still with incomplete or inappropriate features. Some contextual responses involved fictional values contrary to those supplied, and some were pictures or comments based on topic knowledge without any indication of awareness of the information supplied in the task.

Some students generated or interpreted a data set as a positive covariation, at Level 2, based on prior beliefs when a negative covariation existed in the data. Others wrote the values on one axis in reverse order, thus displaying a negative covariation, at Level 3, but appearing visually as an increasing function, in accord with an alternative conception that all covariation graphs should appear in a positive direction (Mevarech & Kramarsky, 1997). The counterintuitive nature of the tasks was important for assessment purposes in eliciting these responses. An important level for these students to achieve was appreciating covariation in context, similar to Tier 2 of Watson's (2000; Watson & Moritz, 1997) statistical literacy hierarchy, evident by representing a verbal claim in a graph or by interpreting a graph in a verbal statement. To do this, students must be encouraged to suspend prior beliefs temporarily to look at the data and examine what covariation might be indicated.

8.05.05 Selecting Variables and Structuring

Representations

For Tasks 1 (Heart deaths) and 2 (Telephone call rates), statistical covariation was embedded within real data measures, as encountered in out-of-school contexts (Gal, 2002; Ross & Cousins, 1993; Tufte, 1997). This involved multiple variables or complex measures, requiring students to identify appropriate variables or measures, such as distinguishing rates and values. For Task 1 (Heart deaths), Coordinate Variable and Series Comparison graphs both showed correspondence of variation between the correct variables but were differentiated by the inclusion of cases as a covariate and by the assignment of variables to axes. In some contexts, especially those in which one variable is systematically varied, teachers could encourage Coordinate Variable graphs, as is commonly expected of students in representing covariation (e.g., Brasell & Rowe, 1993; Cobb et al., 2003), by suggesting students identify the two relevant variables and place them on opposing axes. Teachers could choose a variable that is manipulated, and suggest this measure be systemically varied at regular intervals to examine the affect on another variable. Regular spacing of covariate values would permit students to consider the covariate as ordered categories, and thus to draw a bar graph. Bar graphs were drawn by most sixth-grade students possibly being a more familiar form and simpler to construct as univariate graphs (Chick & Watson, 2001) than a Cartesian coordinate system (Brasell & Rowe). Measuring the effect of manipulations of one variable upon another uses the notion of dependency to build an appreciation of covariation (Coulombe, 1997), which may be a natural starting point for considering covariation—historically, it was first conceived this way, and only later as the correspondence between two sets

(Kieran, 1993)—however it carries with it causal implications that may make it difficult for students to distinguish covariation from causation.

In other data contexts involving two observed variables without a controlled variable, teachers could encourage students to draw Series Comparison graphs, that is, to consider change of two variables over another covariate, such as height and age as properties of different people, or heart deaths and motor vehicle use as functions of time. Series Comparison graphs, with both significant variables on the same axis, are less often valued by educators but represent a valid form of representation (Konold, 2002) for three reasons. First, many students draw them (e.g., Brasell & Rowe, 1993; Cobb et al., 2003), and thus it is likely that many students understand this graph form. Second, they appear to be constructed either by adding data cases to Single Comparison and Double Comparison graphs, or by adding a data-series to time-based Trend Graphs. If from the latter, Series Comparison graphs could otherwise be known as Double Trend graphs. Hence they are conceptualised by scaffolding upon simpler ideas. Third, Series Comparison graphs assign both variables to one dimension, which may reduce the assumption of casual dependence that may be implied by Coordinate Variable graphs for which opposing axes are often termed “independent” and “dependent” variables.

Series Comparison graphs also have the potential to retain covariate or case information on a separate axis (e.g., years, or in other contexts, names) and illustrate how bivariate data are the measured values for two attributes of that case. Case information retains stronger links with the source of the data for examination of outliers or sampling issues (Cobb et al., 2003). Consider the data shown in Figure 8.01a. This view of the data would suggest X is not related to Y , with a correlation coefficient of $-.01$, and a straight line of best fit $y = -.01x + 5.6$. To introduce case

information of points of time, the information presented in a Series Comparison graph is shown in Figure 8.01b. In this graph of the data, it can be seen that Y at a given time is the value of X at the preceding point in time. The introduction of cases in time has introduced another explanatory variable: in this instance, the association depends on the ordered nature of cases in time: if the cases were people with no reason for ordering, then the association of X and Y via an offset in case values would not be meaningful. Such reflection indicates that the goal in developing understandings and representations of statistical covariation is not mechanically producing certain graph forms, but a practice of exploratory data analysis (Biehler, 1997). Such exploration of possible explanatory variables is the skill expected but sometimes found lacking in professionals, for example the lack of a graph of appropriate variables which contributed to the *Challenger* space-shuttle disaster (Tufte, 1997).

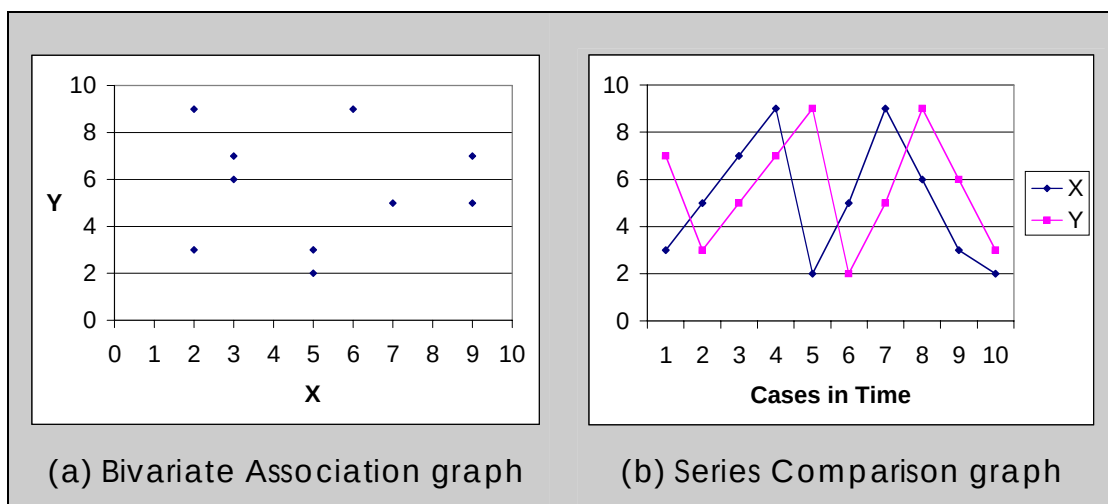


Figure 8.01. Two graphs of the same data for X and Y.

8.05.06 Graphing Conventions

As noted in Section 8.03, students should be encouraged to engage and represent data in ways meaningful to them prior to emphasising conventional graph

construction. When students were given freedom to represent in their own ways, such as in the survey tasks, students were able represent aspects of the data and the context that they considered relevant, such as case labels. It was noted, however, in interview settings that some student drawn graphs were not easily interpreted by other students. Drawing attention in the classroom to this interplay between producing a graph, possibly for oneself, and subsequent difficulties of others when interpreting the graph, provides the basis for class discussion of a variety of graphing methods and conventions to ensure the graph is an effective communication medium. In Section 2.01.05, Wainer (1992) was quoted as saying, “though there is ample evidence that the ability to understand graphically presented material is hard-wired in, there is even more evidence that the ability to draw graphs well is not” (p. 18). Classroom discussion should include the criteria for graphical adequacy and how various conventions support interpretation.

Labels were omitted for one or more variables in many graphs drawn by students, which were similar to the generalized graphs observed by Mevarech and Kramarsky (1997). Even many Series Comparison responses did not label times or cases, although they were implied covariates. These results occurred despite the specific instruction to “draw and label” the graph, and suggest that teachers need to reinforce that labels are required for graphs to communicate effectively. In algebra teaching, often the variables are abstract and assumed as “x” and “y.” In statistics tasks that include contexts of real data, students may sometimes consider that the variables given, such as in a newspaper headline, are redundant to include in a graph. Students should be encouraged to make graphs self-contained, using labeling and a key if necessary. Students could be asked to construct graphs on topics of their choice, and then to exchange their graphs, without other contextual clues, with

fellow students who are required to write verbal descriptions of what the graphs show (Chazan & Bethell, 1994). With free choice of topics, labels would be essential to convey the measures represented.

Graph forms were related to grades of students: primary school students commonly drew bar graphs, and secondary school students commonly drew line graphs, and in some cases, scattergraphs. The reasons that students drew various graph forms were not evident from their responses alone: for example, perhaps each student was familiar with only the one graph form of the response. Teachers must be alert to challenge assumptions that might be associated with graph forms, such as lines beginning at the origin. Equipped with examples of various bar graphs and line graphs as observed in this study, classroom discussion could address the assumptions and implications conveyed in each graph form. Line graphs may carry an implication of intermediate values. For contexts involving time (e.g., Task 3 – Heights, and Task 5 – Temperatures), data points were sampled from continuously changing variables, whereas in other contexts (e.g., Task 4 – Test scores, and Task 6 – Noise), data were discrete. In this sense, the graphing tasks resembled both tasks used in algebra teaching of functions and tasks of discrete data representation often used in school statistics teaching (AEC, 1991; ME, 1992; NCTM, 2000).

Axis allocations of variables were conventional for many responses, which may indicate exposure to this allocation, particularly with time on the horizontal axis as in news reports (e.g., Figure 2.02) or other experiences such as from science classes. The inversion of axes by students observed in previous studies may have been due to the variables involved not lending themselves to natural mapping; for example, in the study of Brasell and Rowe (1993), both variables concerned heights of the ball, before and after bouncing. In contrast, Ainley (1995) and Investigations 2

and 3A observed that most students showed height and age of a person on the vertical and horizontal axes respectively. In Investigation 6A, a few responses at various levels showed temperature in a thermometer form, giving some indication of one reason for assigning temperature to the vertical axis. It may also be that for students who drew coordinate graphs, the constant segment – of heights or repeated temperature values – cued students to represent the repetition horizontally rather than vertically. Horizontal repetition retains the common visual property of line graphs involving no vertical repetition, that is, a variable as a function of time such that each time had only one corresponding value. In this way, the constant function may not only help to identify student difficulties in coordinate graphing (Mevarech & Kramarsky, 1997), but also assist in drawing attention to the covariational nature of the bivariate data to be represented in coordinates. The reasons for axis allocation could be explored in future research using tasks contrasting those with and without (a) natural mappings of height or time, (b) variables with self-evident dependency, and (c) data sets with constant segments or repeated values.

8.05.07 Broader Statistical Literacy

This study has shown that graphing and verbalizing covariation, using familiar contexts, can occur before the standardization of graphing conventions. Teaching of standard graphs forms, such as Cartesian coordinates, might not eliminate alternative conceptions (Mevarech & Kramarsky, 1997), and might even inhibit reasoning about covariation, if students are able to interpret only their own representation. This study has also shown that less familiar contexts, such as media settings with complex measures for variables, provide additional challenges for students. Just as students should be encouraged to develop understandings about covariation and about graphing for effective communication, they should be encouraged to engage issues in

measurement of variables. This can be encouraged from an early age by considering how simple contexts can be coded as measured data (Lehrer & Romberg, 1996). For example, in the simple context of common pets for the class, issues that might be considered include counting numbers of people owning various animals or numbers of animals owned, and whether a tank of ten fish should count as one or ten. Evidence of the variety of performance across various tasks in this study suggests the teacher should provide focus and appropriate scaffolding on various aspects of covariation, representation, and data interpretation, to enable students to tackle complex settings such as found in the media.

Once the claim of covariation is understood in context, students must question the process of inference from statistical data to causal claim—Tier 3 of Watson’s (1997) hierarchy. At this level, awareness of prior beliefs should be encouraged, as well as their balanced integration with available data. An important feature of using tasks involving counterintuitive covariation is that they should naturally raise questions about reliability of the data set, and about generalizability to a causal inference (see Figure 1.01). The tasks involving only six data points were designed to be easy for students to break down the tasks to represent covariation as a series of corresponding cases and draw it quickly, but also importantly introduced the issue of sample size. Other questions used as part of this study, but not reported here, have elicited student responses noting that small sample size made generalization difficult (see Appendix 4B).

8.06 LIMITATIONS OF THE CURRENT STUDY

The research undertaken in this study was limited with respect to the research design, sample of students, administration, and the tasks, coding and interpretation. The research design of surveys and interviews was used to explore students’

understanding evidence in written representations and in dialogue. This method was time efficient in relation to the wealth of information gained, however did not directly gather information about effectiveness of classroom teaching and learning of statistical covariation (e.g., Cobb et al., 2003), nor did it observe student behaviour in a natural setting to assess disposition to use understandings about statistical covariation, for example to speculate about what lies behind a newspaper statement (cf. Gal, 2002).

The sample of students was taken from just a few classes, with students from above average socio-economic background, thus caution must be exercised when considering the generalizability of results. The previous and recent learning experiences of students were also unknown. A draft teacher survey was prepared to explore these possibilities, but was not viable to administer within the terms of this study for the schools involved (see Appendix 3E). The administration of tasks, though somewhat controlled, was in a classroom environment, constrained by time limits and close seating arrangements, which may have limited students' attempts and allowed copying. In addition, particularly for secondary year levels, the surveys were completed during mathematics classes, sometimes involving assumptions about the tasks, or stimulus in the classroom environment, such as simple graphs on the walls. For interviews, background noise sometimes made interpretation of video dialogue difficult.

Assessment tasks sampled the skills of students: the validity of assessing students would be strengthened by administering additional tasks. This sample was limited to a single task for Coordinate Graph Production, two tasks for Verbal Graph Interpretation and Numerical Graph Interpretation, and three tasks for Speculative Data Generation. Some other tasks were administered that did not clearly elicit

student understandings, though others administered to only a few students indicated consistent findings (see Appendix 4B). The tasks differed in context complexity and in the variables involved, hence differences in outcomes could be due to these factors. The study did, however, control for certain aspects of the task design, such as consistent use of six data points. It was also important not just broadly to sample success rates by students across various tasks, but to examine in depth for a few tasks features of students' responses and dialogue that indicated conceptual challenges.

In relation to coding and interpretation, a common issue in similar studies is whether interpretations of the data have been limited to a single view or forced to fit prior expectations. In general, this issue was mitigated by open-coding of response categories, by iterative coding methods, and by checking of other raters. It is possible that additional categories or levels may have been identified either by attending to other features or minute aspects of students' responses, or by gathering other responses related to these skills.

8.07 FUTURE RESEARCH

The conclusions from this research indicate that future related research should employ tasks with a range of contexts involving various measurement issues or multivariate complexities, and particularly counterintuitive covariation. Tasks should also avoid specification of graph axes, graph forms, or graph titles, which might limit student response options. As students sometimes need encouragement to speculate on data, future research should employ tasks involving discrete data cases, and possibly even supplying case labels to encourage specification of associated data values.

The responses gathered in this study could be considered in other ways. Students' representations of the measurement units of variables could be analysed

further. In some responses, there was rich evidence of appreciating measurement issues, whereas other responses appeared abstract and devoid of real data measurement. How this richness or abstraction might relate to understandings of statistical covariation was explored only partially in this study in relation to case labels supporting some representations. Interview probing, for example, might document further how awareness of measurement issues supports appreciating both correspondence and variation of values, but may obscure an integrated abstracted pattern of covariation. Research could also consider numerical aspects of students' responses, such as whether students considered covariation to be deterministic (e.g., Batanero et al., 1997) or to involve statistical fluctuation. Responses in this study showed evidence of differences among students, however this issue was not explored in detail.

With the use of slightly modified tasks, related issues could be investigated. Students' responses to tasks for Speculative Data Generation could be assessed at levels of Coordinate Graph Production. In this study, such analysis was not undertaken as the Speculative Data Generation tasks did not clearly require students to represent the data in a particular way. Speculative Data Generation tasks with a clear communicative purpose might be used to explore this issue further. Other tasks could be used to explore students' skills for Verbal Data Interpretation or for critiquing Causal Inference, as shown in Figure 1.01.

The potential of prompts to stimulate cognitive conflict for developing understanding and representations of statistical covariation was not clearly established in the current study. This could be explored further using more structured tasks to tease out conflict for data issues and for presentation issues. For example, tasks could be reworded, asking students to re-draw a graph by revising the data to

match a verbal statement, or revising the presentation of given data. Tasks could also be administered by computers, to examine systematically whether student reactions are based in the data shown or the graphical presentation. Other tasks to assess Verbal Graph Interpretation could involve stimulus graphs, and ask students to comment on various prompted titles.

Tasks could also be used in more authentic classroom settings, either face to face or in online environments, involving discussions among students (e.g., Cobb et al., 2003). Prompts showing only partial success on given tasks can provide a starting point for conversation and debate. Verbal comments from other students in the class models and encourages the articulation of statistical reasoning in contrast to seeing only the written product of another student. The current study considered a sample of students from selected classes. Future studies could explore whether other students with diverse learning experiences, abilities, or classroom cultures may approach these tasks differently. In particular, future research could use the tasks and frameworks presented in this study to assess students' understandings, to inform instructional intervention, and then to re-assess students' understandings in order to evaluate the effectiveness of teaching for statistical covariation.

8.08 CONCLUSION

Tasks asking students to speculate on the data behind a verbal claim require them to supply details that provide windows on how they interpret statements of covariation. Such tasks are also important because they are authentic to common experience such as reading claims in a newspaper. They were found to be accessible to some students from third-grade. Tasks for Speculative Data Generation complement more traditional tasks of Coordinate Graph Production, Verbal Graph Interpretation, and Numerical Graph Interpretation. In all of these tasks, educators

have a role to play in placing an emphasis on understanding covariation, and the ways it is represented to communicate to oneself and others. Correspondence and variation of values are key features of covariation, and educators can assist students by unpacking data sets as discrete cases, considering change, especially change over time, encouraging verbal formulations and graphing inventions, presenting and discussing differing responses to tasks, discussing the benefits of adopting conventions, and promoting global views of data through re-presentations for the purpose of appreciating how measures co-vary.

In discussing statistical literacy, Gal (1998) commented that students “have to comprehend the meaning of any messages that they are presented with, and be both willing and able to critically examine the reasonableness of such messages or claims” (p. 277). Most students in this study were willing to attempt to represent covariation in a graph, although many failed to demonstrate adequately correspondence of variation. The results of this study indicate that there is still more to do in achieving the ideal of a statistically literate community. It is hoped that the assessment frameworks, tasks, and performance of students observed in this study will support educational developers and teachers to assist in developing students’ understandings and representations of statistical covariation.

References Cited

- Adams, D. D., & Shrum, J. W. (1988). *The effects of microcomputer-based laboratory exercises on the acquisition of line graph construction and interpretation skills by high school biology students*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (61st, Lake of the Ozarks, MO, April 10-13, 1988). (ERIC Document Reproduction Service No. ED 292 652)
- Adams, D. D., & Shrum, J. W. (1990). The effects of microcomputer-based laboratory exercises on the acquisition of line graph construction and interpretation skills by high school biology students. *Journal of Research in Science Teaching*, 27(8), 777-787.
- Adi, H., Karplus, R., Lawson, A., & Pulos, S. (1978). Intellectual development beyond elementary school VI: Correlational reasoning. *School Science and Mathematics*, 78, 675-683.
- Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, 54, 299-352.
- Ainley, J. (1995). Re-viewing graphing: Traditional and intuitive approaches. *For the Learning of Mathematics*, 15(2), 10-16.
- Ainley, J., Pratt, D., & Nardi, E. (2001). Normalising: Childrens' activity to construct meanings for trend. *Educational Studies in Mathematics*, 45, 131-146.
- Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review*, 91(1), 112-149.
- Anscombe, F. (1973). Graphs in statistical analysis. *The American Statistician*, 27, 17-21.
- Arcuri, L., & Forzi, M. (1988). Prior expectations and manner of data presentation as factors influencing covariation estimates for incomplete contingency tables. *Acta Psychologica*, 67(2), 121-133.
- Australian Education Council. (1991). *A national statement on mathematics for Australian schools*. Carlton, Vic.: Author.
- Australian Education Council. (1994). *Mathematics - A curriculum profile for Australian schools*. Carlton, Vic.: Curriculum Corporation.
- Avons, S. E., Beveridge, M. C., Hickman, A. T., & Hitch, G. J. (1983). Teaching journey graphs with microcomputer animation. *Human Learning*, 2, 93-105.
- Barnes, M. (1991). *Investigating change: An introduction to calculus for Australian schools. Unit 1: Functions and modelling*. Carlton, Vic.: Curriculum Corporation.
- Batanero, C., Estepa, A., & Godino, J. D. (1997). Evolution of students' understanding of statistical association in a computer based teaching environment. In J. Garfield & G. Burrill (Eds.), *Research on the role of*

- technology in teaching and learning statistics* (pp. 191-205). Voorburg, The Netherlands: International Statistical Institute.
- Batanero, C., Estepa, A., Godino, J. D., & Green, D. R. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education*, 27, 151-169.
- Behr, M., & Harel, G. (1990). Students' errors, misconceptions, and cognitive conflict in application of procedures. *Focus on Learning Problems in Mathematics*, 12(3&4), 75-84.
- Beichner, R. J. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62, 750-762.
- Bell, A., & Janvier, C. (1981). The interpretation of graphs representing situations. *For the Learning of Mathematics*, 2(1), 34-42.
- Bell, A., Brekke, G., & Swan, M. (1987a). Diagnostic teaching: 4 Graphical interpretation. *Mathematics Teaching*, 119, 56-59.
- Bell, A., Brekke, G., & Swan, M. (1987b). Diagnostic teaching: 5 Graphical interpretation teaching styles and their effects. *Mathematics Teaching*, 120, 50-57.
- Bell, A., Brekke, G., & Swan, M. (1987c). Misconceptions, conflict and discussion in the teaching of graphical interpretation. In J. D. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. 1, pp. 46-58). Ithaca, NY: Cornell University.
- Beniger, J. R., & Robyn D. L. (1978). Quantitative graphics in statistics: A brief history, *American Statistician*, 32, 1-10.
- Ben-Zvi, D. (1999). Constructing an understanding of data graphs. In O. Zaslavsky (Ed.), *Proceedings of the 23rd conference of the International Group for the Psychology of Mathematics Education*, (Vol. 2, pp. 97-104). Haifa, Israel: Israel Institute of Technology.
- Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning*, 2(1&2), 127-155.
- Ben-Zvi, D., & Arcavi, A. (2001). Junior high school students' construction of meanings for data representations. *Educational Studies in Mathematics*, 45, 35-65.
- Berg, C. A., & Phillips, D. G. (1994). An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs. *Journal of Research in Science Teaching*, 31(4), 323-344.
- Bergeron, J. C., & Herscovics, N. (1982). Levels in the understanding of the function concept. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 39-46). Enschede, The Netherlands: Foundation for Curriculum Development.
- Beyth-Marom, R. (1982). Perception of correlation reexamined. *Memory and Cognition*, 10, 511-519.

- Biderman, A. D. (1990). The Playfair enigma: Toward understanding the development of schematic representation from origins to the present day. *Information Design Journal*, 6(1), 3-25.
- Biehler, R. (1997). Students' difficulties in practicing computer-supported data analysis: Some hypothetical generalizations from results of two exploratory studies. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 169-190). Voorburg, The Netherlands: International Statistical Institute.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Bobko, P., & Karren, R. (1979). The perception of Pearson product moment correlations from bivariate scatterplots. *Personnel Psychology*, 32, 313-325.
- Boyer, C. B. (1947). Note on an early graph of statistical data (Huygens 1669). *Isis*, 37(3-4), 148-149.
- Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching*, 24(4), 385-395.
- Brasell, H. M., & Rowe, M. B. (1993). Graphing skills among high school physics students. *School Science and Mathematics*, 93(2), 63-70.
- Broniarczyk, S. M., & Alba, J. W. (1994). Theory versus data in prediction and correlational tasks. *Organizational Behavior and Human Decision Processes*, 57, 117-139.
- Bryant, P. E., & Somerville, S. C. (1986). The spatial demands of graphs. *British Journal of Psychology*, 77, 187-197.
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamics events: A framework and a study. *Journal for Research in Mathematics Education*, 33, 352-378.
- Carswell, C. M. (1992). Choosing specifiers: An evaluation of the basic tasks model of graphical perception. *Human Factors*, 34, 535-554.
- Chazan, D., & Bethell, S. C. (1994). Sketching graphs of an independent and a dependent quantity: Difficulties in learning to make stylized, conventional "pictures". In J. P. da Ponte & J. F. Matos (Eds.), *Proceedings of the 18th Annual Conference of the International Group for the Psychology of Mathematics Education*, Vol. 2 (pp. 176-184). Lisbon: University of Lisbon.
- Chick, H. L., & Watson, J. M. (1998). Showing and telling: Primary students' outcomes in data representation and interpretation. In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times, Proceedings of the Twenty First Annual Conference of the Mathematics Education Research Group of Australasia*, Vol. 2 (pp. 153-160). Gold Coast: Mathematics Education Research Group of Australasia.
- Chick, H. L., & Watson, J. M. (2001). Data representation and interpretation by primary school students working in groups. *Mathematics Education Research Journal*, 13(2), 91-111.

- Clarke, D. J. (1998). Studying the classroom negotiation of meaning: Complementary accounts methodology. In A. R. Teppo (Ed.), *Journal for Research in Mathematics Education, Monograph No. 9, Qualitative research methods in mathematics education* (pp. 98-111). Reston, VA: National Council of Teachers of Mathematics.
- Clarke, D., & Helme, S. (1997). The resolution of uncertainty in mathematics classrooms. In F. Biddulph & K. Carr (Eds.), *People in mathematics education, Vol. 1* (pp. 116-123). Waikato: Mathematics Education Research Group of Australasia.
- Clement, J. (1989). The concept of variation and misconceptions in Cartesian graphing. *Focus on Learning Problems in Mathematics*, 11, 77-87.
- Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79, 531-554.
- Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, 229, 828-833.
- Cobb, P., McClain, K., & Gravemeijer, K. (2003). Learning about statistical covariation. *Cognition and Instruction*, 21(1), 1-78.
- Collis, K. F., & Romberg, T. A. (1992). *Collis-Romberg mathematical problem solving profiles*. Hawthorn, Vic.: Australian Council for Educational Research.
- Coulombe, W. N. (1997). *First year algebra students' thinking about covariation*. Unpublished doctoral dissertation, North Carolina State University.
- Coulombe, W. N., & Berenson, S. B. (2001). Representations of patterns and functions: tools for learning. In A. A. Cuoco & F. R. Curcio (Eds.), *The roles of representation in school mathematics (2001 Yearbook)* (pp. 166-172). Reston, VA: National Council of Teachers of Mathematics.
- Crocker, J. (1981). Judgment of covariation by social perceivers. *Psychological Bulletin*, 90, 272-292.
- Crocker, J. (1982). Biased questions in judgment of covariation studies. *Personality and Social Psychology Bulletin*, 8, 214-220.
- Culbertson, H. M., & Powers, R. D. (1959). A study of graph comprehension difficulties. *AV Communication Review*, 7(1), 97-110.
- Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18, 382-393.
- Curcio, F. R. (1989). *Developing graph comprehension*. Reston, VA: National Council of Teachers of Mathematics.
- Curcio, F. R. (2001). *Developing data-graph comprehension in grades K through 8* (2nd ed.). Reston, VA, National Council of Teachers of Mathematics.
- Department for Education and Employment. (1999). *Mathematics: The national curriculum for England*. London: Author and Qualifications and Curriculum Authority. Retrieved 21 March 2002 from <http://www.nc.uk.net/download/cMa.pdf>

- Department of Education and the Arts (Tasmania). (1993). *Chance & data mathematics guidelines K-8: Activity booklet 1*. Hobart, Australia: Curriculum Services Branch.
- Department of Health and Human Services (Tasmania). (2000). *Personal health record*. Hobart, Australia: Author.
- diSessa, A. A. (2001). *Inventing Graphing: A Video Case Study*. Unpublished CD. Berkeley, CA: The Graduate School of Education, Berkeley, Boxer Group. (Available for researchers, teachers and teacher trainers, on request.)
- diSessa, A. A., & Sherin, B. L. (2000). Meta-representation: An introduction. *Journal of Mathematical Behavior*, 19, 385-398.
- diSessa, A. A., Hammer, D., Sherin, B., & Kolpakowski, T. (1991). Inventing graphing: Representational expertise in children. *Journal of Mathematical Behavior*, 10, 117-160.
- Donnelly, J. F., & Welford, A. G. (1989). Assessing pupils' ability to generalize. *International Journal of Science Education*, 11, 161-171.
- Dreyfus, A., Jungwirth, E., & Eliovitch, R. (1990). Applying the "cognitive conflict" strategy for conceptual change - Some implications, difficulties, and problems. *Science Education*, 74(5), 555-569.
- Dugdale, S. (1993). Functions and graphs – Perspectives on student thinking. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions* (pp. 101-130). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ellerton, N. F., & Clements, M. A. (1990). Cognitive dissonance versus success as the basis for meaningful mathematical learning. In G. Booker, P. Cobb, & T. N. de Mendicuti (Eds.), *Proceedings of the 14th Annual Conference of the International Group for the Psychology of Mathematics Education*, Vol. 2 (pp. 177-184). Mexico: PME Program Committee.
- Erlick, D. E., & Mills, R. G. (1967). Perceptual quantification of conditional dependency. *Journal of Experimental Psychology*, 73(1), 9-14.
- Estepa, A., & Batanero, C. (1996). Judgements of correlation in scatterplots:: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, 4, 21-41.
- Estepa, A., Batanero, C., & Sánchez, F. T. (1999). Students' intuitive strategies in judging associations when comparing two samples. *Hiroshima Journal of Mathematics Education*, 7, 17-30.
- Family car is killing us, says Tasmanian researcher. (1991, June 11). *The Mercury*, p. 2.
- Feliciano, G. D., Powers, R. D., & Kearn, B. E. (1963). The presentation of statistical information. *AV Communication Review*, 11, 32-39.
- Fischer, K. W., & Knight, C. C. (1990). Cognitive development in real children: Levels and variations. In B. Z. Presseisen, R. J. Sternberg, K. W. Fischer, C. C. Knight, & R. Fuerstein (Eds.), *Learning and thinking styles: Classroom interaction* (pp. 43-67). National Education Association.

- Fitzpatrick, J., Galbraith, P., & Henry, B. (1991). *Reasoning and Data*, Heinemann Educational Australia, Port Melbourne, Vic
- Forman, E. A., & McPhail, J. (1993). Vygotskian perspective on children's collaborative problem-solving activities. In E. A. Forman, N. Minick, & C. A. Stone (Eds.), *Contexts of learning: Sociocultural dynamics in children's development* (pp. 213-229). New York: Oxford University Press.
- Freudenthal, H. (1982). Variables and functions. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 7-20). Enschede, The Netherlands: Foundation for Curriculum Development.
- Friel, S. N., Curcio, F. R., Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal Research in Mathematics Education*, 32, 124-158.
- Friendly, M., & Denis, D. J. (2002). *Milestones in the history of thematic cartography, statistical graphics, and data visualization*. Retrieved 15 February 2006 from <http://httpprints.yorku.ca/archive/00000124/>.
- Funkhouser, H. G. (1937a). A note on a tenth century graph. *Orisis*, 1, 260-262.
- Funkhouser, H. G. (1937b). Historical development of the graphical representation of statistical data. *Orisis*, 3, 269-404.
- Gagatsis, A., & Shiakalli, M. (2004). Ability to translate from one representation of the concept of function to another and mathematical problem solving. *Educational Psychology*, 24(5), 645-
- Gal, I. (1998). Assessing statistical knowledge as it relates to students' interpretation of data. In S. P. Lajoie (Ed.), *Reflections on statistics: Learning, teaching and assessment in grades K-12* (pp. 275-295). New Jersey: Lawrence Erlbaum Associates.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review*, 70(1), 1-25.
- Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 47-78). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Garfield, J., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. *Journal for Research in Mathematics Education*, 19, 44-63.
- Gerber, R., Boulton-Lewis, G., & Bruce, C. (1995). Children's understanding of graphic representation of quantitative data. *Learning and Instruction*, 5, 77-100.
- Goos, M. (1994). Metacognitive decision making and social interactions during paired problem solving. *Mathematics Education Research Journal*, 6(2), 144-165.
- Goos, M. (1998). "I don't know if I'm doing it right or I'm doing it wrong!" Unresolved uncertainty in the collaborative learning of mathematics. In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times, Proceedings of the 21st Annual Conference of the Mathematics*

- Education Research Group of Australasia, Vol. 1* (pp. 225-232). Gold Coast: Mathematics Education Research Group of Australasia.
- Goos, M. (2000). Collaborative problem solving in senior secondary mathematics classrooms. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000: Proceedings of the 23rd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 38-45). Perth: Mathematics Education Research Group of Australasia.
- Goos, M., & Galbraith, P. (1996). Do it this way! Metacognitive strategies in collaborative mathematical problem solving. *Educational Studies in Mathematics*, 30, 229-260.
- Green, D. (1993). Data analysis: What research do we need? In L. Pereira-Mendoza (Ed.), *Introducing data analysis: Who should teach it and how?* (pp. 219-239). The Netherlands: The International Statistical Institute.
- Guthrie, J. T., Weber, S., & Kimmerly, N. (1993). Searching documents: Cognitive processes and deficits in understanding graphs, tables, and illustrations. *Contemporary Educational Psychology*, 18(2), 186-221.
- Hankins, T. L. (1999). Blood, dirt, and nomograms: A particular history of graphs. *Isis*, 90, 50-80. Retrieved 26 March 2002 from <http://www.journals.uchicago.edu/Isis/journal/demo/v000n000/000000/000000.web.pdf>
- Hepting, D. (1999). *What's a picture really worth?* Retrieved February 8, 2004 from <http://www.cs.uregina.ca/~hepting/proverbial/>
- Hermann, J. (1973). *Test of graphing and graph interpretation skills*. Del Mod System, Dover, DE.; Delaware State Dept. of Public Instruction, Dover. (ERIC Document Reproduction Service No. ED 093 693)
- Herscovics, N. (1982). Problems related to understanding of functions. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 67-84). Enschede, The Netherlands: Foundation for Curriculum Development.
- Hewson, P. W., & A'Beckett-Hewson, M. G. (1984). The role of conceptual conflict in conceptual change and the design of science instruction. *Instructional Science*, 13, 1-13.
- Hitch, G. J., Beveridge, M. C., Avons, S. E., & Hickman, A. T. (1982). Effects of reference domain in children's comprehension of coordinate graphs. In D. Rogers, & J. A. Sloboda (Eds.), *Acquisition of symbolic skills* (pp. 551-560). New York: Plenum Press.
- Inhelder, B., & Piaget, J. (1958). Random variations and correlations. In *The growth of logical thinking from childhood to adolescence* (Translated by A. Parsons, & S. Milgram) (pp. 224-244). London: Routledge & Kegan Paul.
- Jackson, D. F., Edwards, B. J., & Berger, C. F. (1993). Teaching the design and interpretation of graphs through computer-aided graphical data analysis. *Journal of Research in Science Teaching*, 30, 483-501.
- Janvier, C. (1978). *The interpretation of complex Cartesian graphs representing situations: Studies and teaching experiments*. Unpublished doctoral dissertation, University of Nottingham.

- Janvier, C. (1981). Difficulties related to the concept of variables presented graphically. In C. Comiti (Ed.), *Proceedings of the fifth international conference of the International Group for the Psychology of Mathematics Education* (pp. 189-193). Grenoble, France: IGPME.
- Janvier, C. (1982). Approaches to the notion of function in relation to set theory. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 114-124). Enschede, The Netherlands: Foundation for Curriculum Development.
- Janvier, C. (1987). Translation processes in mathematics education. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 27-32). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Janvier, C. (1989). Representation and contextualization. In G. Vergnaud, J. Rogalski, & M. Artique (Eds.), *Proceedings of the 13th Annual Conference of the International Group for the Psychology of Mathematics Education, Vol. 2* (pp. 139-146). Paris: PME Program Committee.
- Janvier, C., Girardon, C., & Morand, J. C. (1993). Mathematical symbols and representations. In P. S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics* (pp. 79-102). New York: Macmillan.
- Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs*, 79(1).
- Jennings, D. L., Amabile, T. M., & Ross, L. (1982). Informal covariation assessment: Data-based versus theory-based judgments. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). Cambridge: Cambridge University Press.
- Jones, G. A., Thornton, C. A., Langrall, C. W., Mooney, E. S., Perry, B., & Putt, I. J. (2000). A framework for characterizing children's statistical thinking. *Mathematical Thinking and Learning*, 2, 269-307.
- Jones, G. A., Langrall, C. W., Mooney, E. S., & Thornton, C. A. (2004). Models of development in statistical reasoning. In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 201-226). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Kaput, J. (1993). The urgent need for proleptic research in the representation of quantitative relationships. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions* (pp. 279-312). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kaput, J. J. (1987). Representation systems and mathematics. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 19-26). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Karplus, R., Adi, H., & Lawson, A. E. (1980). Intellectual development beyond elementary school VIII: Proportional, probabilistic, and correlational reasoning. *School Science and Mathematics*, 80(8), 673-683.
- Kerslake, D. (1977). The understanding of graphs. *Mathematics in Schools*, 6(2), 22-25.

- Kieran, C. (1993). Functions, graphing, and technology: Integrating research on learning and instruction. In T. A. Romberg, E. Fennema, & T. P. Carpenter (Eds.), *Integrating research on the graphical representation of functions* (pp. 189-237). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kirk, S., Eggen, P. D., & Kauchak, D. P. (1978). The effect of cue specificity and locus on learning from graphical material. *Journal of Educational Research*, 72(1), 39-44.
- Konold, C. (2002). Alternatives to scatterplots. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a statistically literate society, Cape Town, South Africa*. [CD-ROM] Voorburg, The Netherlands: International Statistical Institute.
- Konold, C., & Higgins, T. L. (2003). Reasoning about data. In J. Kilpatrick, G. Martin, & D. Schifter (Eds.), *A Research Companion to NCTM's Standards* (pp. 193-215). Reston, VA: NCTM.
- Konold, C., Pollatsek, A., Well, A., & Gagnon, A. (1997). Students analyzing data: Research of critical barriers. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 151-167). Voorburg, The Netherlands: ISI.
- Kosslyn, S. M. (1985). Graphics and human information processing. *Journal of the American Statistical Association*, 80, 499-512.
- Kosslyn, S. M. (1989). Understanding charts and graphs. *Applied Cognitive Psychology*, 3, 185-226.
- Krabbendam, H. (1982). The non-qualitative way of describing relations and the role of graphs: Some experiments. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 125-146). Enschede, The Netherlands: Foundation for Curriculum Development.
- Kuhn, D., Phelps, E., & Walters, J. (1985). Correlational reasoning in an everyday context. *Journal of Applied Developmental Psychology*, 6(1), 85-97.
- Lane, D. M., Anderson, C. A., & Kellam, K. L. (1985). Judging the relatedness of variables: The psychophysics of covariation detection. *Journal of Experimental Psychology: Human Perception and Performance*, 11(5), 640-649.
- Lehrer, R., & Romberg, T. (1996). Exploring children's data modeling. *Cognition and Instruction*, 14(1), 69-108.
- Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs and graphing: Tasks, learning and teaching. *Review of Educational Research*, 60(1), 1-64.
- Lesh, R., & Lehrer, R. (2000). Iterative refinement cycles for videotape analyses of conceptual change. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 665-708). Mahwah, NJ: Lawrence Erlbaum.
- Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 591-645). Mahwah, NJ: Lawrence Erlbaum.

- Linn, M. C., Layman, J. W., & Nachmais, R. (1987). Cognitive consequences of microcomputer-based laboratories: Graphing skills development. *Contemporary Educational Psychology*, 12, 244-253.
- Lowe, I., Grace, N., Johnston, J., Kissane, B., & Willis, S. (1993). *Access to Algebra, Book 1*. Melbourne: Curriculum Corporation.
- Macbeth, D. (2000). On an actual apparatus for conceptual change. *Science Education*, 84, 228-264.
- MacDonald-Ross, M. (1977). How numbers are shown: A review of research on the presentation of quantitative data in texts. *AV Communication Review*, 25, 359-409.
- Maher, C. A., & Martino, A. M. (2000). From patterns to theories: Conditions for conceptual change. *Journal of Mathematical Behavior*, 19, 247-271.
- Malter, M. S. (1952). Studies of the effectiveness of graphic materials. *Journal of Educational Research*, 46, 263-272.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.
- McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. (1987). Student difficulties in connecting graphs and physics: Examples from kinematics. *American Journal of Physics*, 55, 503-513.
- McKenzie, D. L., & Padilla, M. J. (1986). The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching*, 23, 571-579.
- McKnight, C. C. (1990). Critical evaluation of quantitative arguments. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 169-185). Washington, DC: American Association for the Advancement of Science.
- Mevarech, Z. R., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: Stability and change in students' alternative conceptions. *Educational Studies in Mathematics*, 32, 229-263.
- Meyer, J. (1997). A new look at an old study on information display: Washburne (1927) reconsidered. *Human Factors*, 39(3), 333-340.
- Meyer, J., Shinar, D., & Leiser D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39(2), 268-286.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.
- Ministerial Council on Education, Employment, Training and Youth Affairs (2004). *National Year 6 Science Assessment Report 2003*. Downloaded 19 January 2006 from http://www.mceetya.edu.au/pdf/nat_year6_science.pdf
- Ministry of Education. (1992). *Mathematics in the New Zealand curriculum*. Wellington, NZ: Author. Retrieved February 8, 2001 from <http://www.minedu.govt.nz/Curriculum/mathematics/statement/index.htm>

- Mokros, J. R., & Tinker, R. F. (1987). The impact of microcomputer-based labs on children's ability to interpret graphs. *Journal of Research in Science Teaching*, 24(4), 369-383.
- Moritz, J. B. (1996, Dec.). *Interrogating Collaborators: Techniques for Mathematics Education Research*. Paper presented at the symposium Contemporary Approaches to Research in Mathematics, Science, Health and Environmental Education, Melbourne, 3 December 1996.
- Moritz, J. B. (1998). Observing students learning from students. In S. Groves, B. Jane, I. Robottom, & R. Tytler (Eds.), *Contemporary Approaches to Research in Mathematics, Science, Health and Environmental Education* 1997 (pp. 160-165). Geelong, Vic.: Deakin University.
- Moritz, J. B. (1999). Graphing data: Relating representation and interpretation. In K. Baldwin, & J. Roberts (Eds.), *Mathematics: the next millennium (Proceedings of the Seventeenth Biennial Conference of the Australian Association of Mathematics Teachers Inc.)* (pp. 90-99). Adelaide: Australian Association of Mathematics Teachers Inc.
- Moritz, J. B. (2000). Graphical representations of statistical associations by upper primary students. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000: Proceedings of the 23rd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 440-447). Perth: Mathematics Education Research Group of Australasia.
- Moritz, J. B. (2002). Study times and test scores: What students' graphs show. *Australian Primary Mathematics Classroom*, 7(1), 24-31.
- Moritz, J. B. (2003). Interpreting a scattergraph displaying counterintuitive covariation. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *Mathematics Education Research: Innovation, Networking, Opportunity. Proceedings of the Twenty-Sixth Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 523-530) Melbourne: Deakin University.
- Moritz, J. B. (2004a). Constructing coordinate graphs: Representing corresponding values with variation ordered in two-dimensional space. *Mathematics Education Research Journal*, 15, 226-251.
- Moritz, J. B. (2004b). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 227-255). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Moritz, J. B., & Watson, J. M. (1997). Graphs: Communication lines to students? In F. Biddulph & K. Carr (Eds.), *People in mathematics education*, Vol. 2 (pp. 344-351). Waikato: Mathematics Education Research Group of Australasia.
- National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author. Retrieved February 8, 2002 from <http://standards.nctm.org/>

- Nemirovsky, R. (1996a). A functional approach to algebra: Two issues that emerge. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 295-313). Dordrecht: Kluwer Academic Publishers.
- Nemirovsky, R. (1996b). Mathematical narratives, modeling, and algebra. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra: Perspectives for research and teaching* (pp. 197-220). Dordrecht: Kluwer Academic Publishers.
- Nemirovsky, R., & Tierney, C. (2001). Children creating ways to represent changing situations: On the development of homogeneous spaces. *Educational Studies in Mathematics*, 45, 67-102.
- Noss, R., Pozzi, S., & Hoyles, C. (1999). Touching epistemologies: Meanings of average and variation in nursing practice. *Educational Studies in Mathematics*, 40(1), 25-51.
- Padilla, M. J., McKenzie, D. L., & Shaw, E. L. (1986). An examination of the line graphing ability of students in grades seven through twelve, *School Science and Mathematics*, 86, 20-26.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods* (3rd ed.) Thousand Oaks, CA: Sage.
- Pei, M. L. (2002). *Classical Chinese quotations*. Retrieved February 8, 2004 from <http://www.chinapage.com/quote/quote.html>
- Peterson, L. V., & Schramm, W. (1954). How accurately are different kinds of graphs read? *AV Communication Review*, 2, 178-189.
- Pfannkuch, M., & Wild, C. (2004). Towards an understanding of statistical thinking. In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 17-46). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Pfannkuch, M., & Rubick, A. (2002). An exploration of students' statistical thinking with given data. *Statistics Education Research Journal*, 1(2), 4-21.
- Piaget, J. (1983). Piaget's theory. (G. Cellerier & J. Langer, Trans.) In P. Mussen (Ed.), *Handbook of child psychology* (4th ed., Vol. 1, pp. 103-128). New York: Wiley.
- Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), *Artificial intelligence and the future of testing* (pp. 73-126). Hillsdale, NJ: Erlbaum.
- Posner, G. J., & Gertzog, W. A. (1982). The clinical interview and the measurement of conceptual change. *Science Education*, 66(2), 195-209.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211-227.
- Pratt, D. (1995). Young children's active and passive graphing, *Journal of Computer Assisted Learning*, 11, 157-169.

- Preece, J. (1983). *A survey of graph interpretation and sketching errors*. (CAL Research Group Tech. Report, No. 34) Milton Keynes, England: Institute of Educational Technology, The Open University, Walton Hall.
- Preece, J., & Janvier, C. (1992). A study of the interpretation of trends in multiple curve graphs of ecological situations. *School Science and Mathematics*, 92(6), 299-306.
- Reading, C., & Shaughnessy, J. M. (2004). Reasoning about variation. In D. Ben-Zvi, & J. Garfield, (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 201-226). Dordrecht, Netherlands: Kluwer Academic Publishers.
- Reidhaar, J. W. (1984). An overview of nontabular methods for statistical presentation of data before this century. *Information Design Journal*, 4(1), 25-35.
- Ross, J. A., & Cousins, J. B. (1993a). Enhancing secondary school student' acquisition of correlational reasoning skills. *Research in Science and Technological Education*, 11(2), 191-205.
- Ross, J. A., & Cousins, J. B. (1993b). Patterns of student growth in reasoning about correlational problems. *Journal of Educational Psychology*, 85(1), 49-65.
- Roth, W. M., & Bowen, G. M. (1994). Mathematization of experience in a grade 8 open-inquiry environment: An introduction to the representational practices of science. *Journal of Research in Science Teaching*, 31(3), 293-318.
- Roth, W. M., & Bowen, G. M. (2001). Professionals read graphs: A semiotic analysis. *Journal Research in Mathematics Education*, 32, 159-194.
- Roth, W. M., & McGinn, M. K. (1997). Graphing: Cognitive ability or practice? *Science Education*, 81(1), 91-106.
- Schools Board of Tasmania (1993). *Syllabus for Tasmanian students*. Hobart: Author.
- Sedlak, A. J., & Kurtz, S. T. (1981). A review of children's use of causal inference principles. *Child Development*, 52, 759-784.
- Seggie, J. L., & Endersby, H. (1972). The empirical implications of Piaget's concept of correlation. *Australian Journal of Psychology*, 24, 3-8.
- Sfard, A., Neshler, P., Streefland, L., Cobb, P., & Mason, J. (1998). Learning mathematics through conversation: Is it as good as they say? *For the Learning of Mathematics*, 18(1), 41-51.
- Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. *Educational Psychology Review* 14(1), 47-69.
- Shaklee, H. (1983). Human covariation judgment: Accuracy and strategy. *Learning and Motivation*, 14, 433-448.
- Shaklee, H., & Elek, S. (1988). Cause and covariate: Development of two related concepts. *Cognitive Development*, 3, 1-13.

- Shaklee, H., & Hall, L. (1983). Methods of assessing strategies for judging covariation between events. *Journal of Educational Psychology*, 75, 583-594.
- Shaklee, H., & Mims, M. (1981). Development of rule use in judgments of covariation between events. *Child Development*, 52, 317-325.
- Shaklee, H., & Mims, M. (1982). Sources of error in judging covariations: Effects of memory demands. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 208-224.
- Shaklee, H., & Paszek, D. (1985). Covariation judgment: Systematic rule use in middle childhood. *Child Development*, 56, 1229-1240.
- Shaklee, H., Holt, P., Elek, S., & Hall, L. (1988). Covariation judgment: Improving rule use among children, adolescents, and adults. *Child Development*, 59, 755-768.
- Shaughnessy, J. (1977). Misconceptions of probability: An experiment with a small-group activity-based, model building approach to introductory probability at the college level. *Educational Studies in Mathematics*, 8, 295-316.
- Shaughnessy, J. M. (1981). Misconceptions of probability: From systematic errors to systematic experiments and decisions. In A. P. Shulte (Ed.), *Teaching Statistics and Probability 1981 Yearbook* (pp. 90-100). Reston, VA: National Council of Teachers of Mathematics.
- Shaughnessy, J. M. (1982). Misconceptions of probability; systematic and otherwise: Teaching probability and statistics so as to overcome some misconceptions. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the First International Conference on Teaching Statistics* (pp. 784-801). Sheffield: Teaching Statistics Trust.
- Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan Publishing.
- Shaughnessy, J. M., Garfield, J., & Greer, B. (1996). Data handling. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education, Part 1* (pp. 205-237). Dordrecht: Kluwer Academic Publishers.
- Sherin, B. L. (2000). How students invent representations of motion: A genetic account. *Journal of Mathematical Behavior*, 19, 399-441.
- Shultz, T., & Mendelson, R. (1975). The use of covariation as a principle of causal analysis. *Child Development*, 46, 394-399.
- Smedslund, J. (1963). The concept of correlation in adults. *Scandinavian Journal of Psychology*, 4, 165-173.
- Steffe, L. P. (1990). Inconsistencies and cognitive conflict: A constructivist's view. *Focus on Learning Problems in Mathematics*, 12(3&4), 99-109.
- Strahan, R. F., & Hansen, C. J. (1978). Underestimating correlation from scatterplots. *Applied Psychological Measurement*, 2, 543-550.
- Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of science, cognitive*

- psychology, and educational theory and practice* (pp. 147-176). Albany, NY: State University of New York Press.
- Swan, M. (1982). The teaching of functions and graphs. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 151-165). Enschede, The Netherlands: Foundation for Curriculum Development.
- Swan, M. (1985). *The language of functions and graphs*. University of Nottingham: Shell Center.
- Swan, M. (1988a). Learning the language of functions and graphs. In J. Pegg (Ed.), *Mathematics Interfaces: Proceedings of the 12th Biennial Conference of The Australian Association of Mathematics Teachers* (pp. 76-80). Newcastle, NSW: The New England Mathematical Association.
- Swan, M. (1988b). Making mathematics more meaningful and relevant. In J. Pegg (Ed.), *Mathematics Interfaces: Proceedings of the 12th Biennial Conference of The Australian Association of Mathematics Teachers* (pp. 38-50). Newcastle, NSW: The New England Mathematical Association.
- Swatton, P. (1994). Pupils' performance within the domain of data interpretation, with particular reference to pattern recognition. *Research in Science and Technological Education*, 12(2), 129-144.
- Swatton, P. (1995). Pupil performance in data manipulation and its relationship to the skill of interpretation. *Educational Review*, 47(1), 43-57.
- Swatton, P., & Taylor, R. M. (1994). Pupil performance in graphical tasks and its relationship to the ability to handle variables. *British Educational Research Journal*, 20, 227-243.
- Swedosh, P. (1996). Mathematical misconceptions commonly exhibited by entering tertiary mathematics students. In P. C. Clarkson (Ed.), *Technology in mathematics education* (pp. 534-541). Melbourne: Mathematics Education Research Group of Australasia.
- Swedosh, P. (1998). Reducing the incidence of mathematical misconceptions in 'middle band' students. In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times, Proceedings of the Twenty First Annual Conference of the Mathematics Education Research Group of Australasia, Vol. 2* (pp. 469-476). Gold Coast: Mathematics Education Research Group of Australasia.
- Swedosh, P., & Clark, J. (1997). Mathematical misconceptions – Can we eliminate them?. In F. Biddulph & K. Carr (Eds.), *People in mathematics education, Vol. 2* (pp. 492-499). Waikato: Mathematics Education Research Group of Australasia.
- Swedosh, P., & Clark, J. (1998). Mathematical misconceptions – We have an effective method for reducing their incidence but will the improvement persist? In C. Kanes, M. Goos, & E. Warren (Eds.), *Teaching mathematics in new times, Proceedings of the Twenty First Annual Conference of the Mathematics Education Research Group of Australasia, Vol. 2* (pp. 588-595). Gold Coast: Mathematics Education Research Group of Australasia.
- Swedosh, P., & Clark, J. (2000). They can run, but can they hide? In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000: Proceedings of*

- the 23rd Annual Conference of the Mathematics Education Research Group of Australasia* (Vol. 2, pp. 588-595). Perth: Mathematics Education Research Group of Australasia.
- Testa, I., Monroy, G., & Sassi, E. (2002). Students' reading images in kinematics: The case of real-time graphs. *International Journal of Science Education*, 24(3), 235-256.
- The longer your overseas call, the cheaper the rate. (1993, July 22). *The Mercury*, p. 17.
- Thomas, K. C. (1933). The ability of children to interpret graphs. In *The 32nd Yearbook of the National Society for the Study of Education* (pp. 492-494). Chicago: The National Society for the Study of Education.
- Thompson, P. W. (1994b). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 179-234). Albany, NY: SUNY Press.
- Tierney, C., Nemirovsky, R., & Shulman-Weinberg, A., (1995). *Investigations in Number, Data, and Space, Grade 4: Changes over time: Graphs*. Palo Alto, CA: Dale Seymour.
- Tierney, C., Weinberg, A., & Noble, T. (1996). *Investigations in Number, Data, and Space, Grade 5: Patterns of change: Tables and graphs*. Palo Alto, CA: Dale Seymour.
- Tilling, L. (1975). Early experimental graphs. *British Journal for the History of Science*, 8, 193-213.
- Trends in International Mathematics and Science Study (1999). *Trends in International Mathematics and Science Study*. Retrieved 15 February 2006 from <http://isc.bc.edu/timss1999.html>
- Trends in International Mathematics and Science Study (2003). *Trends in International Mathematics and Science Study*. Retrieved 15 February 2006 from <http://isc.bc.edu/timss2003.html>.
- Tirosh, D. (1990). Inconsistencies in students' mathematical constructs. *Focus on Learning Problems in Mathematics*, 12(3&4), 111-129.
- Tirosh, D., Stavy, R., & Cohen, S. (1998). Cognitive conflict and intuitive rules. *International Journal of Science Education*, 20(10), 1257-1269.
- Tobin, K. G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement*, 41(2), 413-423.
- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, Connecticut: Graphics Press.
- Tufte, E. R. (1997). *Visual explanations*. Cheshire, Connecticut: Graphics Press.
- Van Weering, B. (1982). Functions and their contribution to the curriculum. In G. Van Barnveld & H. Krabbendam (Eds.), *Conference on functions* (Report 1, pp. 194-200). Enschede, The Netherlands: Foundation for Curriculum Development.

- Van Zee, E. H., & McDermott, L. C. (1987). Investigation of student difficulties with graphical representation in physics. In J. D. Novak (Ed.), *Proceedings of the second international seminar: Misconceptions and educational strategies in science and mathematics* (Vol. 3, pp. 531-539). Ithaca, NY: Cornell University.
- Vernon, M. D. (1946). Learning from graphical material. *British Journal of Psychology*, 36, 145-158.
- Vernon, M. D. (1950). The visual presentation of factual data. *British Journal of Educational Psychology*, 20, 174-185.
- Victorian Curriculum and Assessment Authority (2002). *Mathematics - Curriculum and standards framework II*. Retrieved 16 February 2006 from <http://csf.vcaa.vic.edu.au/index.htm>
- Vinner, S. (1990). Inconsistencies: Their causes and function in learning mathematics. *Focus on Learning Problems in Mathematics*, 12(3&4), 85-98.
- Wainer, H. (1980). A test of graphicacy in children. *Applied Psychological Measurement*, 4, 331-340.
- Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, 21(1), 14-23.
- Wainer, H. (1997). *Visual revelations: Graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot*. New York: Copernicus Springer-Verlag.
- Wainer, H., & Velleman, P. F. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology*, 52, 305-335.
- Ward, W. C., & Jenkins, H. M. (1965). The display of information and judgment of contingency. *Canadian Journal of Psychology*, 19, 231-241.
- Washburne, J. N. (1927). An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Psychology*, 28, 361-376 & 465-476.
- Watson, J. M. (1997). Assessing statistical thinking using the media. In I. Gal & J. B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 107-121). The Netherlands: The International Statistical Institute.
- Watson, J. M. (2000). Statistics in context. *Mathematics Teacher*, 93(1), 54-58.
- Watson, J. M. (2002). Inferential reasoning and the influence of cognitive conflict. *Educational Studies in Mathematics*, 51, 225-256.
- Watson, J. M., & Callingham, R. A. (1997). Data cards: An introduction to higher order processes in data handling. *Teaching Statistics*, 19, 12-16.
- Watson, J. M., & Moritz, J. B. (1997). Student analysis of variables in a media context. In B. Phillips (Ed.), *Papers on Statistical Education Presented at ICME-8* (pp. 129-147). Hawthorn, Australia: Swinburne Press.
- Watson, J. M., & Moritz, J. B. (1998). Longitudinal development of chance measurement. *Mathematics Education Research Journal*, 10, 103-127.

- Watson, J. M., & Moritz, J. B. (1999a). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, 37, 145-168.
- Watson, J. M., & Moritz, J. B. (1999b). The development of concepts of average. *Focus on Learning Problems in Mathematics*, 21(4), 15-39.
- Watson, J. M., & Moritz, J. B. (2000a). Developing concepts of sampling. *Journal for Research in Mathematics Education*, 31(1), 44-70.
- Watson, J. M., & Moritz, J. B. (2000b). The longitudinal development of understanding of average. *Mathematical Thinking and Learning*, 2(1&2), 11-50.
- Watson, J. M., & Moritz, J. B. (2001a). Development of reasoning associated with pictographs: Representing, interpreting, and predicting. *Educational Studies in Mathematics*, 48(1) 47-81.
- Watson, J. M., & Moritz, J. B. (2001b). The role of cognitive conflict in developing students' understanding of chance measurement. In J. Bobis, B. Perry, & M. Mitchelmore (Eds.), *Numeracy and beyond (Proceedings of the 24th Annual Conference of the Mathematics Education Research Group of Australasia, (Vol. 2, pp. 523-530)*. Sydney: MERGA.
- Watson, J. M., Collis, K. F., Callingham, R. A., & Moritz, J. B. (1995). A model for assessing higher order thinking in statistics. *Educational Research and Evaluation*, 1(3), 247-275.
- Wavering, M. J. (1984). Interrelationships among Piaget's formal operational schemata: Proportions, probability, and correlation. *The Journal of Psychology*, 118, 57-64.
- Wavering, M. J. (1985). *The logical reasoning necessary to make line graphs*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching (58th, French Lick Springs, IN, April 15-18, 1985). (ERIC Document Reproduction Service No. ED 254 409)
- Wavering, M. J. (1989). Logical reasoning necessary to make line graphs. *Journal of Research in Science Teaching*, 26(5), 373-379.
- Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, 22(5), 366-389.
- Woolfolk, A. E. (1993). *Educational psychology* (5th ed.). Boston: Allyn and Bacon.
- Wright, T., Nemirovsky, R., & Tierney, C. (1997). *Investigations in Number, Data, and Space, Grade 2: Timelines and rhythm patterns*. Palo Alto, CA: Dale Seymour.
- Yerushalmy, M. (1997). Mathematizing verbal descriptions of situations: A language to support modeling. *Cognition and Instruction*, 15(2), 207-264.
- Zacks, J., Levy, E., Tversky, B., & Schiano, D. (1999, Aug.). *Graphs in print*. Draft document. Retrieved 15 February 2006, from <http://www-psych.stanford.edu/~bt/diagrams/papers/zacksgraphsurvey99.doc.pdf>

LIST OF APPENDICES ON CD-ROM

Appendix 1. CORRESPONDENCE

Appendix 1A. University

Relevant forms and progress reports.

Appendix 1B. Schools

Letters to schools, and consent from schools to participate in the study.

Appendix 1C. Copyright

Copyright permission for various figures within Chapter 2.

Appendix 2. PUBLICATIONS

Appendix 2A. Publications with content overlapping this thesis

Reproductions of publications cited in Section 2.03.08.

1. Moritz, J. B. (1999). Graphing data: Relating representation and interpretation. In K. Baldwin, & J. Roberts (Eds.), *Mathematics: the next millennium (Proceedings of the Seventeenth Biennial Conference of the Australian Association of Mathematics Teachers Inc.)* (pp. 90-99). Adelaide: Australian Association of Mathematics Teachers Inc.
2. Moritz, J. B. (2000). Graphical representations of statistical associations by upper primary students. In J. Bana & A. Chapman (Eds.), *Mathematics education beyond 2000: Proceedings of the 23rd Annual Conference of the Mathematics Education Research Group of Australasia (Vol. 2, pp. 440-447)*. Perth: Mathematics Education Research Group of Australasia.
3. Moritz, J. B. (2002). Study times and test scores: What students' graphs show. *Australian Primary Mathematics Classroom*, 7(1), 24-31.
4. Moritz, J. B. (2003). Interpreting a scattergraph displaying counterintuitive covariation. In L. Bragg, C. Campbell, G. Herbert, & J. Mousley (Eds.), *Mathematics Education Research: Innovation, Networking, Opportunity. Proceedings of the Twenty-Sixth Annual Conference of the Mathematics Education Research Group of Australasia* (pp. 523-530) Melbourne: Deakin University.
5. Moritz, J. B. (2004). Reasoning about covariation. In J. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning, and thinking* (pp. 227-255). Dordrecht: Kluwer.
6. Moritz, J. B. (2004). Constructing coordinate graphs: Representing corresponding values with variation ordered in two-dimensional space. *Mathematics Education Research Journal*, 15, 226-251.

Appendix 3. TASKS

Appendix 3A. Early Drafts version

A draft set of tasks.

Appendix 3B. Pilot Survey

Responses from one student for the pilot survey.

Appendix 3C. Final Survey Tasks

Survey tasks for Data Collection 3, for each of grades 3, 5, 7, and 9.

Appendix 3D. Interviews Tasks

Appendix 3E. Draft Teacher Survey

A draft teacher survey, not administered.

Appendix 4. SURVEY RESPONSES

Appendix 4A. Complete response set with analysis notes

PDF copies of all survey responses for Data Collection 3. Analysis notes include spreadsheets of coding of responses.

Appendix 4B. Analysis of Selected Tasks not considered in the Thesis

Draft notes.

Appendix 5. INTERVIEW RESPONSES

Appendix 5A. Transcripts

Transcripts of all interviews. Note that many transcripts are not verbatim and contain spelling and grammatical errors, however these transcripts were used to complement viewing of the video when coding interview dialogue.

Appendix 5B. Analysis summaries

Notes summarising features of transcripts.

Appendix 5C. Video samples

The examples are chosen to show a range of grades of students and the feel of the interview setting.