

'Reasoning from Evidence' Tasks

Malcolm Swan
Mathematics Education
University of Nottingham
Malcolm.Swan@nottingham.ac.uk

Jim Ridgway
School of Education
University of Durham
Jim.Ridgway@durham.ac.uk

WHY USE REASONING FROM EVIDENCE TASKS?

Newspapers, television and the web present citizens with assertions, and arguments often based on plausible 'mathematical' reasoning. Much of this is intended to persuade, impress, and affect behavior. It is therefore an important life skill to be able to analyze data and interpretations of data, argue critically and make informed decisions based on sound reasoning.

For students who choose a career in mathematics, science, and/or science, and/or engineering, the development of this skill in analyzing data is central to the discipline.

Usually data are analyzed via computer. Some teachers use these tasks to assess students' abilities to use computer packages such as Excel to support their mathematical thinking.

WHAT ARE REASONING FROM EVIDENCE TASKS?

The tasks require students to analyze unsorted data. This will assess students' abilities to organize information, represent it in a meaningful way, and draw sensible conclusions.

WHAT IS INVOLVED?

Instructor Preparation Time:	Minimal if use existing tasks.
Preparing Your Students:	Students will need some coaching on their first task.
Class Time:	45 minutes.
Disciplines:	Appropriate for all, requires proportional reasoning and graphical skills. Superior solutions can often be created using a spreadsheet.
Class Size:	Any.
Special Classroom/Technical Requirements:	None, unless the data analysis is done via computer.
Individual or Group Involvement:	Either.
Analyzing Results:	Intensive for formal scoring for large classes. Best used as an informal way to get your students thinking mathematically.
Other Things to Consider:	Fairly demanding task for students who are unfamiliar with open-ended problems.

Description

'Reasoning from Evidence' tasks consist of questions asking students to organize and represent a collection of unsorted data, and to draw sensible conclusions.

For example, students are given a collection of data concerning male and female opinions of two deodorants. The experiment has been designed to test two variables; the deodorant name/ packaging and the fragrance. Both forms of packaging are tested with both forms of fragrance. Data consist of responses from males and females on a five point scale (from 'Love it' to 'Hate it') to each combination of packaging and fragrance. The data are presented to students as an unsorted collection of responses from 40 people that they have to organize. They may begin, for example, by dividing the data into two piles, male and female. They may then allocate numerical values to the data and calculate mean ratings, draw graphs and so on. This should enable them to draw simple conclusions. The demand here is not so much in the performance of technical skills as in the mathematical processes of organization, representation and interpretation.

Assessment purposes

- To see if students are able to draw sensible conclusions from unsorted data;
- (Optional) To see if students can analyze data using a computer spreadsheet.

Limitations

As these tasks involve a considerable amount of organization and reasoning, the output of students will be difficult to assess rapidly. Little mathematical knowledge is assumed apart from fundamental statistical ideas of chance and proportion. Of course, more advanced students may be able to implement more sophisticated ideas, such as tests of significance, but these are not essential.

Teaching Goals

- Students learn how to organize information using tables, graphs and charts
- Students learn how to test hypotheses and draw conclusions from data which are not always clear cut
- Students have the opportunity to compare and contrast different approaches to handling data

Suggestions for Use

Introducing 'Reasoning from Data' tasks for the first time.

Many students will find the tasks unfamiliar. In statistical work, students are usually presented with "clean" data (e.g. data already aggregated into tables) and are told which methods to use for its interpretation. In situations where students have conducted experimental work to gather data, often their teacher has told them which representations or methods to use. In our experience, students rarely have the opportunity to make such decisions for themselves.

Thus, the first time they see the "Reasoning from Data" tasks, it is helpful for them to spend some time discussing possible approaches to the task in pairs or small groups. They may even be able to share out some of the workload of aggregating results. Sometimes, one member of a group will suggest drawing a bar chart, while another will suggest a line graph or a table. They should be encouraged to do all of these and compare the relative advantages of each representation.

Occasionally, we find that students decide to use software (such as a spreadsheet program like Microsoft Excel) to analyze the data. This is not always as straightforward as it sounds, since it adds some complexity. Students will need to think about the format of the data as it is entered (e.g., decimals, times and dates, text) and choose sensible graphical representations to use when analyzing the output. It is easy to select meaningless graphical outputs!

Data are presented here both in printed form and as Excel spreadsheets. Your choice of how to present the data to students will be determined by your teaching goals.

Providing guidance as students work on 'Reasoning from data' tasks.

Whether your students work in groups or individually, many will ask for guidance while doing the tasks. The amount of guidance that students need should decline as they become familiar with these types of problems. The amount and type of help you provide the students depends upon your goals for the task. For instance, if your primary goal is that the students struggle with solving the problems on their own (and learn that they can "do math"), you may choose to provide very little assistance; however, if your goal is that the students can understand and evaluate their reasoning process, then you may provide them with more assistance in that direction.

Reporting out of individual or group work

If you decide to come together as a large group to discuss what students came up with (or report out), it is again helpful to decide the degree to which you will participate in these discussions, which will depend upon your goals for the session. For instance, you can facilitate the students' discussion, having them defend their ideas and write their ideas on the board, while adding almost none of your own. This approach can direct students away from viewing you as the authority of the information. Or, you can lead the discussion, soliciting student comments and organizing them in a useful manner and adding comments to guide them into an understanding of the problem.

Formal and informal use of 'Reasoning from data' tasks

These tasks can be used formally or informally. In formal assessment (where you grade the assignment as an examination), do not intervene except where specified. Even modest interventions – reinterpreting instructions, suggesting ways to begin, offering prompts when students appear to be stuck – have the potential to alter the task for the student significantly.

In informal assessment (an exercise, graded or non-graded), you may want to be less rigid in giving the students help. Under these circumstances, you may reasonably decide to do some coaching, talk with students as they work on the task, or pose questions when they seem to get stuck. In these instances you may be using the tasks for informal assessments—observing what strategies students favor, what kinds of questions they ask, what they seem to understand and what they are struggling with, and what kinds of prompts get them unstuck. This can be extremely useful information in helping you make ongoing instructional and assessment decisions. However, as students have more experiences with these kinds of tasks, the amount of coaching you do should decline and students should rely less on this kind of assistance.

Group work versus individual work

The open-ended nature of 'Reasoning from Data' tasks makes for great group work problems. Students can discuss various measures and their merit and are likely to come up with many more ideas than if they worked alone. The CL1 Collaborative Learning web site can provide instructions on how to use group work effectively within the classroom. However, individual work may give you more clues as to each student's sophistication with this type of problem.

Presumed background knowledge

One nice attribute of 'Reasoning from Data' tasks is that they require very little mathematical knowledge, yet they allow students to use advanced mathematical knowledge where they do possess it.

Students do need to have a basic understanding of chance and proportional reasoning. They will also need to use a calculator and draw graphs when analyzing the data. Of course, as the choice of analysis tools are left to the student, they may decide to use more sophisticated methods, such as graphing facilities in spreadsheets, tests of significance and so on. However, the value of these tasks lies not in the sophistication of the methods used, but with students' ability to draw sensible conclusions from the data.

Step-by-Step instructions

1. Prepare by reading through the 'Reasoning from Data' task on your own and coming up with your own solutions
2. Hand out copies of the task to students, either working individually or in groups.
3. State the your goals for the 'Reasoning from Data' task, emphasizing that they should be able to defend both their choice of method and the reasoning which leads to their answer.
4. Walk around and listen to students as they discuss and work through the problems, providing guidance as necessary.
5. Have students present their solutions, either in written or verbal form.

Variations

The tasks included in this site can be downloaded and used without modification. If you choose to develop your own 'Reasoning from Data' task, you can follow the pattern used in these tools. Where unsorted data is to be analyzed, simply collect together the raw results from an experimental study in a relevant field, and present these to your students with some background discussion of the reasons why the data were collected and the method of collection. Then allow the students time to develop their own ways of analyzing the data.

Analysis

Student work can be measured against the following criteria:

- can students select appropriate variables for sorting a data set?
- can students select appropriate methods for analyzing a data set?
- can students construct, read and interpret graphical representations of a data set? and,
- can students draw sensible conclusions from a data set?

This generic scoring rubric may be modified and adapted for specific tasks.

Category of performance	Typical response
The student needs significant instruction	Student can begin to organize the data and makes a limited analysis using a single statistic. The student may not have attempted to represent the data in tables or graphs. Only one variable is typically considered.
The student needs some instruction	Student has made an attempt to organize the data and has attempted to represent it and draw conclusions from it. Again, the response may show that only one variable has been considered. The representation used may be inappropriate and the conclusions invalid.

The student's work needs to be revised	Student has selected appropriate variables and methods for sorting, analyzing and representing the data. There may be errors in the calculations and graphs. The student attempts to draw conclusions from the data but these may be flawed.
The student's work meets the essential demands of the task	Student has selected appropriate variables and methods for sorting, analyzing and representing the data. The student has used a variety of analytic tools to interrogate the data set. The conclusions/recommendations follow from and are supported by their analysis of the data

The example below shows how the generic rubric can be modified to fit the 'Emergency 911! Bay City' task:

Category of performance	Typical response
The student needs significant instruction	Students calculate a single statistic (e.g., mean or median response time). They recommend one ambulance service over the other on the basis of a comparison of this single statistic even though the mean difference is only .2 minute, not significant for making a policy recommendation. The analysis of the data ignores all other variables except response time.
The student needs some instruction	Students may calculate measures of center and explore the data with other kinds of analysis (e.g., box plots, stem and leaf plots) but they consider only a single variable – the response times of the two ambulance services. They demonstrate some ability to use their statistical "toolkit" but the analysis is not connected to the real-world context of the problem and the argument is weak.
The student's work needs to be revised	Students select appropriate variables for analyzing the data (e.g., response time in relation to time of call), make appropriate calculations, use appropriate graphical representations, and make a reasonable recommendation based on their analysis. There may be errors in the calculations and in the graphs. However, students do not fully interrogate the data set, thereby not ruling out other possible salient relationships (e.g., response time in relation to day of the call). The recommendations follow from the analysis but the report may lack clarity and thoroughness.
The student's work meets the essential demands of the task	Students select appropriate variables for sorting, analyzing and representing the data. Students consider a number of relationships and use a variety of analytic tools to fully interrogate the

data set. Their recommendations follow from and are supported by their analysis of the data.

Malcolm Swan

Mathematics Education
University of Nottingham
Malcolm.Swan@nottingham.ac.uk



Most assessment practices seem to emphasise the reproduction of imitative, standardised techniques. I want something different for my students. I want them to become mathematicians - not rehearse and reproduce bits of mathematics.

I use the five 'mathematical thinking' tasks to stimulate discussion between students. They share solutions, argue in more logical, reasoned ways and begin to see mathematics as a powerful, creative subject to which they can contribute. Its much more fun to try to think and reach solutions collaboratively. Assessment doesn't have to be an isolated, threatening business.

Not just answers, but approaches.

Malcolm Swan is a lecturer in Mathematics Education at University of Nottingham and is a leading designer on the MARS team. His research interests lie in the design of teaching and assessment. He has worked for many years on research and development projects concerning diagnostic teaching (including ways of using misconceptions to promote long term learning), reflection and metacognition and the assessment of problem solving. For five years he was Chief Examiner for one of the largest examination boards in England. He is also interested in teacher development and has produced many courses and resources for the inservice training of teachers.

Jim Ridgway

School of Education
University of Durham
Jim.Ridgway@durham.ac.uk



Thinking mathematically is about developing habits of mind that are always there when you need them - not in a book you can look up later.

For me, a big part of education is about helping students develop uncommon common sense. I want students to develop ways of thinking that cross boundaries - between courses, and between mathematics and daily life.

People should be able to tackle new problems with some confidence - not with a sinking feeling 'we didn't do that yet'. I wanted to share a range of big ideas concerned with understanding complex situations, reasoning from evidence, and judging the likely success of possible solutions before they were tried out. One problem I had is that my students seemed to learn things in 'boxes' that were only opened at exam time. Thinking mathematically is about developing habits of mind that are always there when you need them - not in a book you can look up later.

You can tell the teaching is working when mathematical thinking becomes part of everyday thinking. Sometimes it is evidence that the ideas have become part of the mental toolkit used in class - 'lets do a Fermi [make a plausible estimate] on it'. Sometimes it comes out as an anecdote. On graduate told me a story of how my course got him into trouble. He was talking with a senior clinician about the incidence of a problem in child development, and the need to employ more psychologists to address it. He 'did a Fermi' on the number of cases (wildly overestimated) and the resource implications (impossible in the circumstances). He said there was a silence in the group...you just don't teach the boss how to suck eggs, even when he isn't very good at it. He laughed.

Jim Ridgway is Professor of Education at the University of Durham, and leads the MARS team there. Jim's background is in applied cognitive psychology. As well as kindergarten to college level one assessment, his interests include the uses of computers in schools, fostering and testing higher order skills, and the study of change. His work on assessment is diverse, and includes, the selection of fast jet pilots, and cognitive analyses of the processes of task design. In MARS hhe has special responsibility for data analysis and psychometric issues, and for the CL-1 work.

About MARS

The Mathematics Assessment Resource Service, MARS, offers a range of services and materials in support of the implementation of balanced performance assessment in mathematics across the age range K to CL-1. MARS is funded by the US National Science Foundation, and builds on earlier funding which began in 1992 for the Balanced Assessment Project (BA) from which MARS grew.

MARS offers effective support in:

The Design of Assessment Systems: assessment systems are tailored to the needs of specific clients. Design ranges from the contribution of individual tasks, through to full scale collaborative work on test development, scoring and reporting. Clients include Cities, States, and groups concerned with educational effectiveness, such as curriculum projects and professional development initiatives.

Professional Development for Teachers: most teachers need help in preparing their students for the much wider range of task types that balanced performance assessment involves. MARS offers professional development workshops for district leadership and 'mentor teachers', built on materials that are effective when used later by such leaders with their colleagues in school.

Developing Design Skills: many clients have good reasons to develop their own assessment, either for individual student assessment or for system monitoring. Doing this well is a challenge. MARS works with design teams in both design consultancy and the further development of the team's own design skills.

To support its design team, MARS has developed a database, now with around 1000 interesting tasks across the age range, on which designers can draw, modify or build, to fit any particular design challenge.