EXPRESSIONS OF UNCERTAINTY WHEN VARIATION IS PARTIALLY-DETERMINED

Janet Ainley[1] and Dave Pratt[2]
[1]School of Education, University of Leicester, UK
[2]University of London Institute of Education, UK
janet.ainley@le.ac.uk

Research into students' knowledge about uncertainty has tended to focus on contexts where the variation observed might easily be attributed to randomness. Yet, the variation observable in many everyday phenomena might be seen as partially-determined, in the sense that one main factor might explain the signal in the variation but additional noise is often inexplicable and might be accounted for as random error. We set out to research how students (age 11) accounted for variation that is in this sense partially-determined. In this paper, we describe how students' expressions of uncertainty were shaped by particular features within the activity structure in which children recorded and represented the results of their experiments and then modelled the variation.

PARTIAL DETERMINATION

Statisticians often set out to make sense of data using models that predict a *signal* that accounts for variation of a dependent variable in terms of the variation in one or more independent variables. Inevitably there is more variation in the data than can be explained through the signal. The remaining variation is often described as *noise*. A full description of the many ways that statisticians model phenomena involving signal and noise, and the range of terms used to describe them, is given by Brovocnik (2005). In this paper, we refer to phenomena that might be modeled in this way as *partially-determined*.

A very wide range of phenomena can in this sense be regarded as partially-determined, and yet, educational research on probabilistic thinking has largely focused on phenomena that are generally modelled as undetermined, or totally random, with a focus on familiar random generators such as coins, dice and spinners. (see, for example, Konold, 1995; Pratt, 2000; Ireland & Watson, 2009.) Indeed, there is very limited research on students' understanding of partially-determined phenomena. Grotzer and Perkins (2000) established a taxonomy of how causality might be modeled and how students reflect on those models. One element of their taxonomy was *probabilistic causality,* where a level of uncertainty is introduced into the causal relationship. The notion of probabilistic causality lays bare the idea that even coins, spinners and dice are not in themselves perfectly random but that, in using a model that is totally undetermined, the unmanageable layers of causality that determine the outcome are intentionally masked. Partially-determined phenomena are more evident when there are easily identifiable signals but when it is also apparent that the signal is insufficient to account for variation in outcomes. Prodromou and Pratt (2012) have reported on older students making sense of simulated basketball throws, where the success or failure of a throw is not entirely determined by the kinetics of the throw itself. We report here how younger students made sense of partially-determined phenomena.

ACTIVE GRAPHING AND EXPLORATORY DATA ANALYSIS

Ainley et al (2001) explored young children's understanding of scatter graphs by engaging them in experiments that generated bivariate data. In fact, the scatter graphs demonstrated variation resulting from partially-determined phenomena. In one task, children explored the amount of mass that could be supported by a paper bridge, where the independent variable took the form of the number of folds in the paper, in order to design a bridge that could support a precious china egg. In another example, the focus was on the time of flight of a paper flying machine, where the independent variable took the form of the length of its wings, the aim being to make a champion flyer. The approach in this research, which they termed *Active Graphing,* allowed students to gain a sense of the relationship between the signal and the dependent variable by gradually making sense of the scatter graph. The noise in the data was treated as something to discount rather than explain; indeed, the focus was very much on the signal. Active Graphing requires students to anticipate how they might change the independent variable in their experiment to clarify that relationship, for

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

example by filling in apparent gaps in the data. One limitation of Active Graphing is that it requires an independent variable and so is not a valid pedagogic approach when teaching about observational data. For example, a typical classroom project is to ask the class to measure themselves and look for relationships within the data, such as a connection between arm length and height. In such an activity, the children select an individual child and measure both variables; there is no independent variable and so Active Graphing does not apply.

Pedagogic approaches developed to handle observational data mostly exploit the power of digital technology to manipulate, compute and dynamically represent data using Exploratory Data Analysis (EDA) (Tukey, 1977). Probability is generally seen as a difficult topic to teach and learn (for example, Falk & Konold, 1997; Shaughnessy, 2003; Jones, 2005) but EDA was seen as a technique that could bypass those difficulties and thus facilitate the identification of ‘stories’ in data (Cobb, 2005). However, EDA relies heavily upon graphical interpretation, which is itself known to be challenging for students (Curcio, 1987; Monteiro & Ainley, 2004).

In this research project we asked whether there might be a pedagogic approach that somehow marries the strengths of Active Graphing to support the appreciation of graphs and EDA to unearth the signal by managing the noise. We noted the new functionality in TinkerPlots 2 (TP2) (https://www.keycurriculum.com/products/tinkerplots) to create statistical models of phenomena. Consider again the task for children to measure their arm lengths and heights, which is accessible to EDA but not to Active Graphing because there is no independent variable. Suppose that the focus of the task becomes to model the relationship between arm length and height in TP2. Such a model might be understood as a machine to create heights from arm lengths (as in the ‘cat factory’, Konold et al, 2007). A successful machine (i.e. the model) would need to reflect the signal, in the form of a relationship between arm length and height, since arm lengths and heights are not independent, and the noise, as not all children with a specific arm length have the same height. Such a modeling approach creates an artificial independent variable, arm length in this example, much as statisticians model phenomena by conjecturing signal variables to see if they do in fact account for significant amounts of variation in the dependent variable. The creation of an artificial independent variable opens up the possibility that Active Graphing might be a valid pedagogic approach in tasks based on observational data.

In this study, we set out to design tasks that incorporate this modelling approach, aiming to draw on the following attributes from Active Graphing and EDA: (i) it is possible to design tasks that are purposeful from the student’s point of view; (ii) such tasks might generate meaningful data, familiar to the students because they are collecting it themselves; (iii) there is a real need for the children to use the graph in order to complete the activity. We aimed to trace the relationship between the design (of the tasks and the modeling tools in TP2) and how the children expressed uncertainty when the variation was partially-determined.

METHOD AND SETTING

We worked with four different groups of three 11 year-olds from state primary schools in England. These children had little or no experience of using TP2, but were familiar with conducting experiments in science in which they had to design ‘fair tests’, take measurements and record data. Two tasks were developed, and each was trialled with two groups of children.

The Angry Emu Task

The children were asked to help Rovio (developer of the popular *Angry Birds* game) to plan a game with a new bird: *Angry Emu*. As in the Angry Birds game, players use a slingshot to launch the bird but Angry Emu cannot fly in the air, and moves only horizontally. The children were asked to prepare a data-based recommendation to help develop the Angry Emus computer game in a way that will resemble the real movement of a toy bird (seated on a toy car) launched from a sling made from a length of elastic. The activity structure included four main stages:

1. The researcher introduces the task. The children engage in free play with the equipment. They experience how the bird physically moves, and begin to measure the stretch of the sling, and the distance travelled. They are encouraged to test each stretch a number of times, and to begin to predict outcomes.

2. The researcher shows how to record and graph data in TP2. The children repeatedly generate graphs, make predictions and gather more data as necessary. They experience variation in outcomes, and begin to express uncertainty and deterministic explanations.
3. The researcher demonstrates how to create a machine. The children express their ideas about signal and noise by making a machine, which links two devices representing the stretch and the distance. They use this to generate data and produce graphs.
4. The children compare the data generated by their model with their previous experimental data. They predict outcomes for other values of the stretch which not yet explored.

The Angry Emu task is similar to the original Active Graphing tasks insofar as the children engaged with a partially-determined phenomenon (distance travelled by the Angry Emu) by collecting data, but, in comparison with the tasks used in earlier research, where the purpose was to make a particular product, signal was not prioritised over noise.

The 101 Dalmatians Task

The children were asked to imagine that they were creating a scene as part of a theme park, which will show 101 Dalmatians. The dalmatians needed to be different sizes (as in the original story), and look reasonably realistic. They were given some data about five dalmatians:

Table 1: Initial data: five Dalmatians

spots	height	tail	body	leg
brown	41	23	40	22
black	37	23	37	18
black	26	13	27	14
black	30	19	30	16
black	30	15	31	17

The data focus on quantitative variables but the colour of the dalmatians' spots was included as a simple access point to TP2 for the children. The data is carefully designed to incorporate the following approximate relationships: (i) body length = height at shoulder; (ii) leg length is between half and two thirds of height at shoulder; (iii) tail length is marginally more than half of the body length. We also included examples where the same height did not relate to the same lengths in other variables and similarly the same tail length failed to correlate with the same values in other variables. The planned activity structure had the following inquiry stages:

1. The researcher introduces the task and leads a discussion of the data with the children, using plots to compare variables, and encourages the articulation of relationships.
2. The researcher uses the data on the colours of spots to show how to create a machine. The children add devices for height and one other variable (body, leg or tail length) to the machine. When data are generated and plotted the researcher encourages discussion about whether the data are realistic, with the aim that the children will notice that the machine generates nonsense dogs because no relationship between variables has been built into the machine.
3. The researcher shows how to build dependency links. The children use this method to express their ideas about the relationship between the variables. In doing so the children are likely to recognize the need to include some variation, so the researcher demonstrates how to input a range of values. The children use this facility to express their ideas about noise alongside signal.
4. The children generate graphs and discuss the appropriateness of the data, both in terms of producing realistic dogs and in relation to the original data. They may then go on to modify their machine accordingly.

Below we discuss the activity that emerged for these tasks. The key issues that we report were those identified from our analysis based on a progressive focusing of the data.

DISCUSSION AROUND EXCERPTS FROM THE DATA

(R = researcher; other letters = specific children; italicized comments are added for clarification.)

The Angry Emu Task

In Stage 1, the children tended to account for variation in how far the angry emu travelled through actual or imagined deterministic causes:

R. So how would you say it moves when you pull the springy to 35? *(The children invented the name* *'springy' to stand for how far the elastic was pulled back before release.)*

C. Out of control.

L. Yeah he goes out of control sometimes it works and sometimes he does backflips and sometimes he just does spins.

C. He does spins and goes sideways and like just crashes into things.

K. He goes rapidly.

C. He goes like rapidly he goes. *(He gestures with his hands a motion where the hands veer away to one side.)*

Such deterministic explanations were apparent in the responses of all children on this task but such explanations were supplemented by a different type of expression as the activity developed.

When the children were encouraged to predict the outcomes of experiments, or to imagine reporting their conclusions to Rovio, they began to use verbal indicators of uncertainty.

R. So what are we going to say to Rovio when the springy is 20?

L. I think it'll be 50.

R. So each time we do the springy 20 it goes exactly 50.

C. No it goes just under or just over.

L. About 50.

C. About… around.

The children occasionally used a range of numbers to express uncertainty in the distance travelled, and this manner of expressing uncertainty became the norm in the next stage when they began to build a machine to generate distances. There are various ways of expressing the distribution of values in TP2 but the simpler methods that we were using require a range of values (or even more simply each separate value can be entered but this soon becomes tiresome). As a result of this requirement, the idea that the uncertainty can be expressed as a range became fixed as the favoured type of expression. During stage 3, one girl expressed this idea as a template:

L. Yeah and then in here that's where we put the something to something.

We wondered how the children were deciding what the minimum and maximum values in the range should be:

R. Ok now before you do anything I want you to explain to me why you chose those numbers. *(They have chosen 40 to 75).*

C. Because we've got 43 and 73 so…

L. Yeah 43 is like over 40 so if we put it at a 10s number and round it up that'll make it easier and 75 because 80 would be a bit too much so we can round it up too to 75.

To identify the range of distances for each stretch the children looked at the lowest and highest values of the distance travelled in their experiment, and then used a mixture of numerical rounding and visual approximation to choose the values they put into the model.

An expert might use a similar method, although they would be aware that some allowance needed to be made for the distribution. For example, an expert might want values near to the centre of that range to be chosen by the machine rather more frequently than those near to the extremes. TP2 would allow such methods based on distributional thinking but we tended to place emphasis on simpler methods, not expecting such young children to have access to this type of reasoning.

The account of the activity so far places emphasis on noise and little has been said about signal (in contrast the original Active Graphing research). Perhaps they did not articulate a relationship between the stretch and the distance travelled simply because it was too self-evident. Occasionally, there were glimpses of the children's accounting for the signal in the variation:

R. Ok so you feel you can't be sure about how far it's going to go?

D. No but..

B. But you could…

R. But is it completely unpredictable?

B. No you can have a strong … you can have a really strong guess at it.

The notion of 'having a really strong guess at it' seems to capture informally the interplay between signal and noise.

R. If it's 15 *(referring to the stretch),* could you tell what would be your prediction for how far it might go?

K. 35.

R. About 35. Can you explain to me how you worked that out?

K. Well I think it might be it because 10 is round about 25…*(K is looking at the graph where there is a data point at 'spring' (stretch) 10 and distance 25)* But, like, 5 more springy might be like 10 more distance

K. appears to have an early algebraic appreciation of the relationship whereby the signal in how far the angry emu travelled was determined by adding 10 more onto the distance for each five that is added to the independent variable, the stretch. This comment expressing a sense of the gradient in the linear relationship is unusual, although at other times children used gestures which indicated a notion of slope. Although K uses precise numbers, the relationship is clearly not seen as totally determined as, at other points in the activity, all of the children articulated ideas about noise.

The 101 Dalmatians Task

The activity developed in broad terms as predicted in the task design, though the researcher needed to intervene quite explicitly in order for the children to recognise the nonsense dogs that moved them on from Stage 2 to Stage 3. One notable difference between the activity arising out of this task, compared to the Angry Emu task, is that the children rarely accounted for variation causally. This is perhaps not surprising since the data in this task are observational. Once the children began to build a machine to generate attributes of the dalmatians, they occasionally used language that suggested causality, though this related, to this machine world rather than to reality:

M. I think we should do the leg length next because then it will make the leg length out of the height.

Verbal indicators of uncertainty also emerged in this task but in a rather different way. In the Angry Emu task, the children were able to experience directly the variation in how far the emu travelled and they could see that it did not travel the same distance even when the same stretch was used. Therefore we built a similar effect into the given data so that dalmatians with the same height had different tail lengths. Verbal expressions of uncertainty, however, emerged also by reference to the relationships that had been built into the data.

A. It's like the same because 30 is there *(pointing to the height of one dog)* and 30's there *(pointing to the body length of the same dog)*, but it's just a bit bigger - that's why out there and then that is just, like, 30.
R. So you think that the body length and the height are about the same?
A. About the same yeah.
R. Are they exactly the same?
A. No.
M. Because, even if you look at that, that's exactly on 30 there, but there it's a bit over.
L. But it's a bit bigger - not loads bigger.

A. was also the first to notice a relationship between height and leg length.

A. I'm noticing the leg length and the height; the leg length is 16 there. If you double 16, it will get to 31, and that's close to 32. *(32 was the actual value but A made a small arithmetic mistake in doubling the leg length.)*
(After some further discussion…)
A. So the height to the leg length … height, if you half that, you'll get close to the leg length of the dog.

Almost from the outset the 101 Dalmatians task seemed to facilitate discussion of signal alongside noise. Such discussion only began to emerge towards the end of the Angry Emu task. The use of a range of values to express the variation in, say, leg length given a value of another variable (height) did not emerge in the early stages, perhaps because noise was expressed as deviation from the relationship (as signal) rather than from previously measured values (as was the case in the Angry Emu task), However, this type of expression of uncertainty did emerge once the children began to build a machine because, at that point, the software required such an input. For example, below, L. is deciding what value to give to leg lengths.

L. Between 15 and 20 really, because 10 would be too small, but then something like 22 would be too big.

The approximate relationships observed in the original data continued to be a reference point for the children when building the machine. In the next excerpt, the children were deciding what values to give the leg length when the height was specifically 37:

R. What numbers have you chosen *(for the range)*
M. 15 to 19
R. And why was that?
M. Because it's approximately half, because 15 is half of 30 and then 17 is half of 34 and there's not really a half of 37 so it's the closest halves to it.

CONCLUSIONS

In these tasks, children expressed uncertainty of outcomes with respect to deterministic causes when the task itself involved experimentation and were less likely to do so when the task involved observational data. However, when the observational data task was transformed into a task to build a machine (or model), causal explanations emerged occasionally.

Verbal indicators of uncertainty (such as ‘about’, ‘around’) emerged in both tasks. When using experimental data, these expressions of uncertainty related to previous experimental values whereas, when using observational data, they related to relationships found in the data. These verbal indicators became quantified as ranges. This happened easily when using experimental data and became fixed as a template when building a machine to generate data similar to that in the experiment. Such quantification took longer to emerge when using observational data and was only really prompted when building a machine. When using experimental data, statements about the relationship were not apparent until late in the activity when one child even expressed the relationship in early algebraic language. When using observational data, the relationships in the data quickly became a focus and were used not only to discuss noise but also as a means for deciding on the range of values in the machine.

ACKNOWLEDGEMENTS

This project was funded by the British Academy (SG112288). The tasks described were developed in collaboration with Dani Ben-Zvi, Hana Manor and Keren Aridor, University of Haifa.

REFERENCES

Ainley, J., Pratt, D., & Nardi, E. (2001). Normalising: Children’s Activity to Construct Meanings for Trend. *Educational Studies in Mathematics, 45*, 131-146.

Borovcnik, M. (2005). Probabilistic and Statistical Thinking. *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (CERME 4), 485-506.

Curcio, F. (1987). Comprehension of Mathematical Relationships Expressed in Graphs. *Journal for Research in Mathematics Education*, *18*(5), 382-393.

Falk, R., & Konold, C. (1997). Making Sense of Randomness: Implicit Encoding as a Basis for Judgment. *Psychological Review*, *104*(2), 301-318.

Grotzer, A. T., & Perkins, N. D. (2000). *A Taxonomy of Causal Models: The Conceptual Leaps between Models and Students’ Reflections on Them.* National Association of Research in Science Teaching (NARST), New Orleans.

Konold, C. (1995), Confessions of a Coin Flipper and Would-Be Instructor, *The American Statistician,* 49(2), 203-209.

Konold, C., Harradine, A. & Kazak, S. (2007). Understanding Distributions by Modeling Them, *International Journal of Computers for Mathematical Learning*, 12, 217-230.

Monteiro, C., & Ainley, J. (2004). Exploring the Complexity of the Interpretation of Media Graphs. *Research in Mathematics Education*, 6, 115-128.

Pratt, D. (2000). Making Sense of the Total of Two Dice, *Journal for Research in Mathematics Education*, 31(5), 602-625.

Shaughnessy, J. M. (2003). Research on Students’ Understandings of Probability. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A Research Companion to the Principles and Standards for School Mathematics*, (pp. 216-226). Reston, VA: NCTM.

Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.