Children's understanding of randomness as a model

Peter Johnston-Wilder

University of Warwick

This paper presents two views of randomness from literature on philosophy of science. Each of these is discussed in relation to learning randomness, and they are related to a common restricted understanding of randomness. These views are used to analyse extracts from interviews with secondary school pupils in which the pupils were invited to participate in simple experiments involving randomness, and to discuss situations using the idea of randomness. Finally, the paper presents some conclusions and questions that follow from this discussion.

Keywords: Randomness; pupil's understanding.

Two philosophical views

Much probabilistic work in the secondary curriculum is founded on the standard frequentist approach of Von Mises (1939). The view of randomness underpinning this approach is incompatible with applications to situations involving finite sequences of outcomes, or unrepeatable single events. Recent accounts of randomness by philosophers have re-examined the concept to clarify it, and to extend the rigour of the idea to contexts in which it is naturally applied. In this paper, I refer to two writers, Kyburg and Eagle, whose ideas enlightened my thinking. The first writer offers an interpretation whereby randomness provides a model for thinking about situations that are uncertain or unpredictable. The second extends this by defining randomness as unpredictability, and in so doing seems to me to extend the applicability of the model of randomness to the real world.

Randomness as a model for incomplete knowledge

Kyburg offers an interpretation of randomness which draws attention to the role of the state of knowledge of the person making the judgement that the situation is 'random'. He says randomness is "a concept which is relative to our body of knowledge, which will somehow reflect what we know and what we don't know." (1974, 217)

This interpretation of randomness accords with the idea of randomness as a model for situations of uncertainty. It allows the use of pseudo-random generators without difficulty, provided the user does not have knowledge of the underlying algorithm to enable prediction of outcomes. For example, the sequence of integers in the decimal expansion of π may appear to be the outcome of a random process to a person who did not know their source, even though they are a determined sequence.

Randomness as Unpredictability

The concept of randomness is used in a variety of ways. Some writers (Howson and Urbach 1993) suggest that it may be impossible adequately to account for all of the various applications through a single united concept. Anthony Eagle has recently attempted such unification, proposing that "randomness is to be understood as a special case of the epistemic concept of the unpredictability of a process" (2005, 1).

There are interesting resonances in Eagle's ideas with the National Curriculum for Mathematics in England, which requires that "pupils should be taught to see that random processes are unpredictable". While the meaning ascribed by Eagle to the idea of 'unpredictability' is more rigorous and precise than that intended in the National Curriculum, the resonance of ideas is worth pursuing. Eagle identifies four competing demands to be satisfied by rigorous analysis of randomness as a single concept. These demands, required for an intuitive use of the concept, offer interesting resonance with pupils' experiences of randomness. I briefly review each of these.

Statistical Testing

The idea of random sampling is introduced in Key Stage 3 of the English National Curriculum for Mathematics; pupils are expected to select and justify random samples. They are expected to work with computer-generated pseudo-random numbers, and possibly also with random number tables. This requires at least an informal idea of what might be expected of a random sequence and some idea of how to recognise a random sequence. Although the treatment of these ideas for secondary school pupils is informal, their discussion still raises some of the difficult issues underpinning the concept of randomness. Seeing randomness as unpredictability emphasises the idea that an individual's state of knowledge about the process is a key factor in attributing randomness. Pseudo-random numbers are self-evidently unpredictable for a pupil using computer software.

Finite randomness

The National Curriculum requires that "pupils should be taught to see that random processes are unpredictable". Thus, inevitably, pupils consider finite processes and single events. They often speak of a single event as random, even though the standard frequentist explanation of probability, often presented in secondary texts, uses a view of randomness that is incompatible with such usage. However, if an event is clearly 'unpredictable', then randomness may intuitively be an appropriate model to consider.

Explanation and Confirmation

The problem of how to recognise, or test whether a process is random is a particularly difficult one for pupils with standard views of randomness. Typically pupils might use the idea of pattern breaking; this can lead to difficulties. If the pupil is simply looking to see if outcomes are unpredictable, some of these issues can be avoided.

Determinism

Some pupils may consider that a process that is essentially deterministic cannot be seen as random, and vice versa. Such a belief would be difficult to reconcile with approaches whereby some phenomena are seen as both deterministic and random.

I now consider the relevance of speaking about randomness as unpredictability for children in secondary school.

Degrees of randomness

Researchers have reported pupils' confusion between the related ideas of randomness and fairness (for example, Batanero and Sanchez 1999). Fairness applied to a die usually describes 'equiprobability', and yet it is sometimes used as a synonym for 'random'. A 'fair' die is sometimes thought of as being 'more random' than a biased die. This leads to the idea that some pupils think of randomness as a matter of 'degree': some processes are seen as random, but less random than others. The term 'degree(s) of randomness' is used in various contexts on websites, including many academic sites in fields such as geography, mathematics and linguistics. This may indicate that the underlying idea is emerging into wider use, and it might not be surprising to find that children are using the idea in their discourse about randomness.

The idea of 'degrees of randomness' relates to a fundamental philosophical issue about the nature of randomness. If randomness is 'unpredictability', then a probability distribution may be considered to be a measure of the degree to which an outcome is predictable. A situation with a finite number of possible outcomes has maximal unpredictability if the outcomes are equally likely, and might be considered the most random possible. The a biased die is to some extent predictable; in that sense bias is the antithesis of randomness. Alternatively, if randomness is considered absolutely the opposite of deterministic, then a situation in which the outcome is not completely determined can be said to have a random component.

Pupils' ideas about randomness

In this section, I present some ideas about randomness expressed by secondary school pupils in interviews, and I relate these to the philosophical positions discussed above. The interviews, each lasting at least 40 minutes, were conducted with 18 pupils aged 13 to 17 as part of a study of learners' conceptions of randomness. Interviewees discussed a series of experiments involving rolling three different dice (a biased die, a spherical die and a cracked die). At the end of each interview, I invited the pupil to describe other situations that they might think of as being random in nature. Several of the examples discussed below are taken from that final stage of the interviews.

Model of incomplete knowledge

When I asked David (age 14.1) to think of examples, other than dice and coins, where things happen 'by chance', his first suggestion was 'football matches', although he went on to qualify this example by excluding factors such the skill of the teams.

> … someone like Bradford playing Manchester United, everyone would say Manchester United are going to win, cos they're the best team. But although it is sometimes in football down to skill, a lot of it can be chance and good luck, so I'd say they could still have a chance to win. It's still just two teams and one of them has to win. So you couldn't ever know definitely which one's going to win.

When considering the result of a match, David did not see chance as an external agency, like luck, but rather he used chance as a model to account for the remaining uncertainty after taking account of what he knew about the skill of the teams. This informal way of thinking is reminiscent of a statistician using analysis of variance to account for a proportion of the variability and looking on the residual variability as due to 'chance'. David's approach also resonates with Kyburg's (1974) suggestion that a judgement of randomness is relative to the knowledge of the judge

Ben (age 15.7) thought sporting results (football, or horse-racing) were essentially deterministic and, to some extent, predictable by people in the know. However, he still insisted that these events had an element of chance.

> They're not completely random, but there will be some anomalies that just… turn up. So, I'd say that they're almost not random, but they sort of are as well.

Ben mentioned very rare events, such as "getting knocked over by a bus", and commented "it's such a small chance, hopefully, that … you just don't treat it as anything random at all – it's just not going to happen". This suggests that he may have been thinking of probability as a measure of the extent to which something is not determinate, and hence 'random'.

Examples from sport are complex situations in which there are clear causal factors, but Ben was able to see that, in spite of these, a random model could be applied appropriately. When discussing situations in which he could consider the outcomes to be random, the very large number of factors that affect the outcomes was the critical element in making something random. However, when he tried to apply this reasoning to an imagined person involved in a recent rail crash, for which no cause was yet known, he stated that the outcome was strictly "not completely random" because there are contributing factors that affect the outcome, and these factors could be predicted. Nonetheless, the number of contributing factors appeared so large that it became difficult or impossible to predict the outcome, and so the situation could be considered random. Ben suggested that randomness could be a useful model for events where the causes are unpredictable or unknown.

> [something random] occurs because of … millions and millions of variables that just can't be controlled… just sort of controls themselves… they just change for no reason, or they have sort of a reason, but can be affected… The reasons could be other variables, which would make it even more complicated…

Ben naturally linked the notion of randomness as unpredictability with his use of randomness as a model to account for incomplete knowledge of a complex system.

Expressing degrees of randomness

Several interviewees compared the degree of randomness shown by the three different dice. David used this idea when he finally recognised bias in the die (after 23 throws); he still thought the behaviour of the die was 'chance', "but not as well as a normal dice". This seemed a reasonable comment in view of the fact that the outcomes from the biased die were not certain. Later, when considering the spherical die, he commented "this dice is more chance than the other one, because… it was high numbers then low numbers". Ben said something similar after ten outcomes from the spherical die, when he commented that this die appeared to be "a lot more random than the other one… although we're getting lots of 1s and 5s", referring to the biased die. He meant by this that it didn't "seem to be so biased". Note that neither Ben nor David were saying that the spherical die appeared to be completely random; both were still concerned about aspects of the outcomes that they had observed so far.

Later, when discussing the cracked die, Ben seemed to suggest that he could control the degree of randomness in the outcomes by the way he rolled the die.

> If you pick it up like that, and then roll it…it's more likely to get… a number on one of these faces, so long as it'll land straight, and doesn't bounce. And then if you just… pick it up and chuck it, then… that's more random, rather than just keep rolling it all the time, which is what I'm doing on this table. Because if I just… chuck it, it would bounce around on the carpet a bit and go away. So you're more likely to get… the same type of numbers recurring over again. Well… it might not be the same numbers, but it will be less random because, if we're always picking it up in the same way and rolling it in the same way, then it should start occurring in a pattern rather than… just completely random.

Ben suggested that, by rolling the die in controlled manner along the table, he could restrict the outcomes, whereas if he were to "chuck it, it would bounce around on the carpet a bit…".

The expression of degrees of randomness involves understanding beyond seeing randomness as only equivalent to fairness. It also requires a facility with language related to probability and randomness. Ben and David expressed this idea particularly clearly; both also showed that they were able to think of randomness as a model for incomplete knowledge in various real world contexts.

In another interview, Abby (age 17.7) came to the interesting conclusion that "randomness is personal"! She had been discussing the idea that the outcome of rolling a die was unpredictable, but that if she knew enough about how it was rolled, how much it weighed, and about the physics, then it could become predictable. She went on to consider whether this would still be random.

> [It] depends on what randomness means. Because if it does mean something that is unpredictable, then… that depends what you mean by unpredictable, because I can't predict the weather, but the Met office can, to a point. ... So that all depends on what is predictable and what is random. I see random as something that's unpredictable, but … I've all of a sudden got a bit stumped as to what predictable and unpredictable mean. Perhaps random is a personal thing.

Abby seemed to be creating for herself a new way of understanding 'randomness' as she spoke. The degree of randomness was related to the knowledge of the person making the judgement. She returned to this idea again later in the interview and gave further examples. This example shows how the idea of degrees of randomness can be extended to form the more complex idea that randomness can be a model that depends on the knowledge of the individual. Abby has expressed a sophisticated understanding of randomness that is close to Kyburg's view.

Conclusions

When interviewees were able to think of randomness as a model for situations in which outcomes were 'caused' by many factors in a complex system, they showed subtle and delicate shifts between speaking of the outcomes as random and determined. Some interviewees spoke in terms of randomness being a matter of degree, and were able to think of some things as being more random than others. I see this idea as critical in the development of an understanding of the applicability of randomness as a model for a variety of situations in the real world.

I also see the idea of degrees of randomness as part of the solution to the paradox of seeing randomness as equivalent to fairness. Once randomness is seen as unpredictability and a probability distribution as a measure of degree of predictability, then the learner can be seen to possess the resources to construct a global perspective on randomness – that is to construct the idea of a probability distribution.

Interviewees who expressed the idea of degrees of randomness were also able to see a variety of applications of randomness to the weather, to sport and to chance events in everyday life. Other interviewees, who appeared to see randomness as an absolute concept, without a clear expression of degrees of randomness, did not appear to perceive this variety of applications (Johnston-Wilder 2006).

A more powerful view of randomness arises from the definition of randomness proposed by Kyburg (1974): randomness can provide a model for a situation in which the judge has incomplete knowledge, even though the situation might be strictly deterministic. Such a view was clearly expressed by David and Ben, and by Abby when she spoke of randomness as ‘personal’. It is interesting to note that these are the individuals who also expressed the idea of ‘degrees of randomness’, and the extracts showing this idea of randomness as a model for incomplete knowledge are from different parts of the interviews from the places where they expressed ‘degrees of randomness’. This suggests to me that these two ideas might be linked.

Ben also expressed the idea of randomness as a model for incomplete knowledge in his suggestion that randomness might be a model for situations affected by many factors. Abby is particularly interesting as her idea that randomness was a personal judgement appeared to have developed within the interview.

These results suggest that in pupils’ early experiences with random processes, it is important that learners are allowed and encouraged to inspect the generating process, to consider whether a model of randomness is appropriate, and to discuss their views. It may also be important for pupils to meet generating processes not easily modelled from a set of equally likely outcomes, since through consideration of ‘unfair’ generators learners can develop an awareness of degrees of randomness. There is a natural assumption that physical random generators will be symmetrical, but, where the process was asymmetrical, interviewees often could not take account of this. From appreciation of degrees of randomness, pupils can begin to apply their mental model of randomness to a wider variety of contexts, and to consider more carefully the relationship between deterministic contexts and randomness. Classroom discussion about whether and under what circumstances ‘scoring a winning goal in the final minute of a critical football match’ could be modelled using randomness could allow learners to consider the role of causal factors in the predictability or unpredictability of such an event. Then an appropriate modelling tool might enable learners to build a representation of their random model for this event. The random generator should be open to inspection, discussion and amendment by the learner.

There is a need to understand better how learners can develop awareness and understanding of the relationship between sample size and variability of outcomes within the sample. There is also a need to develop clearer understanding of how to structure a learner’s attention productively to develop awareness of variation and how attention is affected by the probability distribution as well as sample size.

References

Batanero, C. and L. Serrano. 1999. The Meaning of Randomness for Secondary School Children. *Journal for Research in Mathematics Education* 30(5): 558–567.

Chaitin, G. 1975. Randomness and Mathematical Proof. *Scientific American*, 232: 47–52.

Eagle, A. 2005. Randomness is Unpredictability. *British Journal for the Philosophy of Science*, available online from http://philsci-archive.pitt.edu/archive/00002134/, accessed May 2006.

Howson, C. and P. Urbach. 1993. *Scientific Reasoning: the Bayesian Approach*. 2nd Ed. Chicago: Open Court,.

Johnston-Wilder, P. 2006. Learners’ Shifting Perceptions of Randomness. *Unpublished PhD thesis*, Open University, UK.

Kyburg, H. E. 1974. *The Logical Foundations of Statistical Inference*. Dordrecht: D. Reidel.

Von Mises, R., 1939. *Probability, statistics and truth*. 2nd Ed. Trans. J. Neyman, D. Scholl and E. Rabinovitsch. New York: Macmillan.