Diagnosing Children's Probabilistic Understanding

John Truran

University of Adelaide

Abstract

Research into children's understanding of probability has revealed a number of significant misconceptions. But much research has been eclectic, with piecemeal influence on classroom practice in both teaching and assessment. A good test of probabilistic understanding is sorely lacking.

This paper uses the results of previous research to propose a multi-dimensional classification of the very wide range of situations and appropriate responses which together make up what we refer to as "an understanding of the idea of probability". Ways in which the classification may be applied to classifying assessment questions and complete tests are discussed briefly.

Introduction

This paper argues that an important requirement for the development of a test of probabilistic understanding is a sound theoretical understanding of the concept being measured. It presents a provisional theoretical framework which can help to describe what any given test is measuring.

Several tests of probabilistic understanding currently exist and others are being developed at the moment. These tend to be *ad hoc* collections of questions many of which have been taken from the research literature. In general the questions do seem to be testing aspects of probabilistic understanding, and do present situations which research has shown to be good discriminators of depth of understanding. But it is also known that subjects respond quite differently to slightly different probabilistic situations.[1] Furthermore, the tests seem to lack any overall structure which might ensure that they really are measuring what they claim to be measuring. This difficulty is most apparent in Green's (1982) test which originally contained some 50 questions. However, in order to work within the theoretical framework of Guttman scalogram analysis Green was forced to reduce the number of questions to 19, with a consequent loss of coverage of ideas.

In this paper I present a structure for evaluating the coverage of tests designed to assess the understanding by a person (P)[2] of probabilities which might be attached to single outcomes of a random generator (RG). At the end examples are presented briefly showing how the structure might be applied to individual questions, and complete tests.

---

1 For a summary of this see Truran (1992, pp. 50 - 97)

2 "Person" is deliberately used, because it is more comprehensive that either "child" or "subject".

Clearly the classification proposed here is a tentative one. There is not space to present literature support for all of its categories. But those which seem to me to be less commonly heeded by researchers and teachers than others are discussed in more detail. A different coding is provided for elements of each part of the structure as a convenient abbreviation where appropriate and also to assist in randomly selecting a sample of situations for testing or teaching purposes.

The structure has four major parts which categorise:

- types of random generators (RGs) (bold numeral code - **5**);
- the general environment in which an RG is encountered (alpha-numeric code – B3);
- questions which might be asked about RGs (Greek letter code – σ);
- relationships between a specific experience with a RG and outcomes from the same or other RGs (Roman numeral code - VII).

Classification of Random Generators

1	Disc	**2**	Coin	**3**	Die
4	Urn	**5**	Cards	**6**	Spinning Pointer over Contiguous Sections
7	Spinning Pointer over Non-contiguous Sections	**8**	Spinning Disc with Contiguous Sections	**9**	Spinning Disc with Non-contiguous Sections
10	Roulette Wheel	**11**	Electronic	**12**	Human
13	Asymmetric Solids				

Table 1
Types of Random Generators

The list in Table 1 is clearly not exhaustive, but it does seem to cover the main sorts of RGs which are likely to be encountered by P. Discs and coins are distinguished because Ps usually have a lot of experience of, and strong subjective views about, coins as RGs, whereas they do not always see discs as being isomorphic with coins. The term "contiguous spinner" means a spinner where all the elementary events which form a specific event are next to each other on the spinner. If they are scattered around the spinner then the spinner is referred to as "non-contiguous". A roulette wheel is distinguished because it is a disc where both the pointer and the disc are free to move. The example of human RGs has been included to cover situations like the fact that when a large group of people are asked to think of a number at random within a given range their selections are asymmetrically distributed across the range. The term "asymmetric solids" refers to a wide variety of objects such as bones, drawing pins and polystyrene cups.

Classification of Environment

Table 2 classifies different types of RGs and the general environment in which they are encountered. Each column lists the possible forms which each of the independent parameters may take. This classification is almost certainly not yet complete, but I believe it represents a useful compact way of summarising critical issues which have been identified in the research literature.

	A	B	C	D	E	F	G	H	I	J	K
	Previous Exper-ience with RG	*Formal Practical Experi-ence with RG*	*Form-al Theor-etical Exper-ience with RG*	*Operator of RG*	*Present-ation of Situat-ion*	*Style of Res-ponse*	*Number of Ele-mentary Events*	*Num-ber of Events*	*Structure of RG*	*Know-ledge of Struct-ure of RG*	*Reward*
1	Own Culture	> 7 days	> 7 days	Self	Individ-ual	Oral	2	2	Sym-metric	Known	None
2	Unusual	2 - 7 days	2 - 7 days	Other - Present	Group with discuss-ion	Written	Small	3 - 6	Slightly Non-Sym-metric	Un-known	Hypo-thetical
3	Different Culture	<1 day	<1 day	Other - Absent	Group without discuss-ion	Mult-iple Choice	Large	6 - 12	Very Non-Sym-metric		Actual
4						Non-Lingu-istic	Very Large	>12	Deceptive		

Table 2

Classification of Environment

It will not always be possible to determine all of these parameters as can be seen from the comments on them in Table 3.

A	*Previous Experience with RG*	These will clearly vary from country to country, from school to school and within the same class. For example, not all Ps have experience of playing cards. An unusual RG within the dominant Australian culture might be a polystyrene cup with which Ps are familiar as an object but not as a RG. An RG from a different culture might be the use of a traditional Australian two-up board by a recent immigrant from South-East Asia.
B	*Formal Practical Experience with RG*	This classification is fairly rough, and not always controllable. It is included to take account of P's structured experiences. and possible introduction to non-naïve strategies for analysing probabilistic situations.
C	*Formal Theoretical Experience with RG*	This classification is fairly rough, and not always controllable. It is included to take account of P's possible introduction to non-naïve strategies for analysing probabilistic situations.
D	*Operator of Random Generator*	This often neglected variable has been shown to have an important influence on the responses of both children and adults.

E	*Presentation of Situation*	Social influence on P's responses is an area which has been little studied in probability but which may be a significant variable.
F	*Style of Response*	Non-linguistic responses might be selecting which of two RGs would be the better to choose to obtain a desired outcome or assessing an RG by placing a mark on a line segment representing a probability scale.
G	*Number of Elementary Events*	"Small" means that P can rapidly decide how many elementary events there are—normally less than ten. "Large" means that P can count the number without losing interest—roughly between 10 and 30. "Very large" means that counting is impracticable.
H	*Number of Events*	These are different from G and are an estimate of the level of difficulty involved in focussing on outcomes. "Two events" is distinguished because some Ps have a special tendency to regard such a case as symmetric.
I	*Structure of RG*	"Slightly non-symmetric" is distinguished to allow for situations like gender balance at birth. "Very non-symmetric" refers to generators like polystyrene cups whose shape discourages a symmetric argument. Deceptive RGs are RGs like loaded dice.
J	*Knowledge of Structure of RG*	Many real-life probabilistic decisions need to be made without understanding the structure of potential non-deterministic influences.
K	*Reward*	Ethical reasons may inhibit evaluating this parameter in practice, but it should not be neglected in theory.

Table 3

Comments on Classification of Environment

Classification of Questions

Table 4 presents five basic ways in which questions about RGs might be used. There may well be others. No matter which question is asked P needs to be clear whether or not a "don't know" response is acceptable. It is fairly easy to make this clear with multiple choice questions, but in other cases it would be easy for P to feel that any response was preferable to an admission of ignorance.

α	*Prediction of Out-come*	"If I draw a ball from this urn which colour do you expect to come out?"	This has a number of methodological difficulties; it does not determine what of several possible strategies are being used.
β	*Compar–ison of RGs*	Presenting P with two RGs and asking which would be the better one to choose in order to achieve a desired outcome.	This approach has the value of being non-linguistic, and of enabling comparison between quite different types of RGs. It avoids most of the difficulties of α.
γ	*Compar-ison of Outcomes*	Asking P which of 2 outcomes of an RG is more likely.	This may be answered on other than probabilistic grounds.
δ	*Fair Allocation of Payout for Bets*	The meaning of "fair" may be confused with that of "symmetrical"	

ε	*Sequences of Outcomes*	This may be one-dimensional or two dimensional, e.g. the snow-flake pattern.	P may be asked either to present a possible sequence or which RG is most likely to have produced a specified sequence.
ξ	*Questions of Technical Knowledge*	"If I draw out one ball at random from this urn, what is the probability that I will draw out an red ball?"	

Table 4

Classification of Questions

P's responses when making comparisons may vary according to the relative probabilities of the desired outcome. For this reason several different comparisons need to be provided in order to obtain a full assessment of P's understanding. Appropriate categories of probabilities are zero, between zero and one-half, one half, between one-half and one, and one. Each of these situations might be compared with every other situation, in cases when the two probabilities are either equal or unequal.

Classification of Probabilistic situations

There is considerable research evidence that the environment in which a specific probabilistic situation is presented may well influence the way in which P perceives that particular situation. There seem to be several important forms, some under the control of the teacher, some under the control of P. Table 5 lists some of these situations.

I	*First Experience with RG*	This may well indicate P's naïve ideas, especially is the structure of the generator is known.
II	*Previous Results with the Same RG*	Previous results can generate a number of strategies such as that of "negative recency" where a prediction is different from the previous outcome or from the majority of a set of recent outcomes.
III	*Previous Predictions of Results from the Same RG*	While it is often argued that heuristics like negative recency operate with respect to outcomes, it has also been found that they may also operate with respect to predictions, quite independently of outcomes (Truran (1992, pp. 181-188).
IV	*Whether a Similar RG is Operating at the Same Time*	Mathematically the situation of tossing three dice at once and tossing one die three times in a row are equivalent. But it has been shown that the responses of Italian children make it clear that many of them do not view these situations as equivalent (Fischbein Nello and Marino , 1991)
V	*Previous Experience with Other Similar Types of RG*	It has been found that P's responses to less familiar embodiments are more likely to be correct when preceded by questioning about familiar embodiments. (Zaleska, 1974; Zaleska and Askévis-Leherpeux, 1976)
VI	*Changes in the RG from Trial to Trial*	When balls are drawn from an urn without replacement then the independence of the trials is not affected but the probability distribution is. This may well affect children's responses.

VII	*Whether Students are Trying to Predict an Outcome or to Maximise Their "Rewards" in the Long Term*	Some children, even quite young children, consistently adopt a Pascalian strategy of predicting the most likely outcome, regardless of previous outcomes or previous predictions.

Table 5

Classification of Probabilistic Situations

Applications of the Classification

Space makes it impossible to do more than show how this structure might be applied to individual questions or to a complete test or to classroom practice. One example of each will have to suffice.

Application to a Question

This interesting question by Tobin (1982) requires not only an answer, but also a reason for the answer.

> A gardener bought a package containing 3 squash seeds and 3 bean seeds. If just one seed is selected from the package what are the chances that it is a bean seed?
>
> a. 1 out of 2
> b. 1 out of 3
> c. 1 out of 4
> d. 1 out of 5
> e. 4 out of 6
>
> Reasons
>
> 1. Four selections are needed because the three squash seeds could have been chosen in a row.
> 2. There are six seeds from which one bean seed must be chosen.
> 3. One bean seed has to be selected from a total of three.
> 4. One half of the seeds are bean seeds.
> 5. In addition to a bean seed, three squash seeds could be selected from a total of six.

Table 6 presents a classification of Tobin's question according to the structure proposed here. The value of such a classification is of limited value for an individual question. However, when groups of questions are taken together, the classification makes it easy to establish the spread of the whole group. This will be considered in the next example.

RG	8	Urn
A	2	Unusual Culture
D	3	Other - Absent
E	3	Group without discussion
F	3	Multiple Choice
G	2	Small
H	1	2
I	1	Symmetric
J	1	Known
K	1	No Reward
	ξ	Technical Knowledge
	I	No Previous Experience

Table 6

Analysis of Sample Question

Application to a Test

Green (1982) proposed a test of probabilistic understanding. Most questions had to be omitted to give the test the required statistical validity. In Table 7 are listed those questions in the test's final form which measure understanding of situations leading to single outcomes of a random generator and classify each of these questions with the form proposed in this paper.[3]

Question	Type of RG	Environment	Question Form	Probabilistic Situation
2	urn	A1, D2, E3, F3, G3, H1, I1, J1, K1	γ	I
3	6	A1, D3, E3, F3, G2, H2, I1, J1, K1	β	I
4	die	A1, D3, E3, F2, G2, H2, I1, J1, K1	γ	I
5	coin	A1, D3, E3, F3, G1, H1, I1, J1, K1	ε	II
6b (answer)	urn	A1, D1, E3, F3, G2, H1, I1, J1, K1	β	I
6b (reason)	urn	A1, D1, E3, F2, G2, H1, I1, J1, K1	β	I
6c (answer)	urn	A1, D1, E3, F3, G2, H1, I1, J1, K1	β	I
6c (reason)	urn	A1, D1, E3, F2, G2, H1, I1, J1, K1	β	I
6d (answer)	urn	A1, D1, E3, F3, G2, H1, I1, J1, K1	β	I
6d (reason)	urn	A1, D1, E3, F2, G2, H1, I1, J1, K1	β	I
6e (answer)	urn	A1, D1, E3, F3, G2, H1, I1, J1, K1	β	I
6e (reason)	urn	A1, D1, E3, F2, G2, H1, I1, J1, K1	β	I
9	die	A2, D1, E3, F2, G2, H2, I1, J1, K1	δ	I
18	urn	A1, D3, E3, F3, G2, H2, I1, J1, K1	ε	II
19 (answer)	7	A1, D1, E3, F3, G2, H1, I3, J1, K1	β	I
19 (reason)	7	A1, D1, E3, F2, G2, H1, I3, J1, K1	β	I

Table 7

Analysis of Test of Green (1982)

3 Qq. 10 and 26a have been omitted from Green's test because they are not concerned with single outcomes of a random generator.

It is easy to see the limited range of matters which have been examined in this test by glancing down the columns. Some of the lack of variation is caused by the limitations of the testing situation, but certainly not all. A similar analysis could easily be applied to classroom practice, either by examining examples presented by a teacher or printed in a text. Here the standards of comprehensiveness might well be set higher because of the greater freedom available in a classroom situation.

Discussion & Summary

There is an inevitable tension between reductionist and holistic approaches to teaching, learning and assessment. The Uncertainty Principle found in Physics and Theology[4] has its analogue in Education. This paper has been prepared on the premise that when researchers are investigating a new field the reductionist approach provides a valuable way of assessing whether more holistic approaches really are holistic. As with all reductionist approaches it is capable of infinite refinement, so it is not presented as an absolute criterion. If it is used, then experience will help to determine the level of refinement which is of most practical value. But it is argued that unless a test is reasonably comprehensive by the standards of this structure, then it cannot reasonably be claiming to assess an understanding of the whole concept of probability.

References

Fischbein, E., Nello, M. S. & Marino, M. S. (1991). Factors Affecting Probabilistic Judgements in Children and Adolescents. Educational Studies in Mathematics 22: 523-549.

Green, D. R. (1982). Probability Concepts in 11 - 16 Year Old Pupils. Report of Research Sponsored by the Social Science Research Council 1978-198. Loughborough, U.K.: Centre for Advancement of Mathematical Education in Technology, University of Technology, Loughborough.

Tobin, K. G. (1982). Patterns of Reasoning: Probability in L. Rattray-Wood, & P. Ferguson (Eds) (1982) Research in Science Education Volume 12 Papers from 12th Annual Conference Macquarie University May 1982. pp. 42-49 Melbourne, Victoria: Australian Science Education Research Association.

Truran, J. M. (1992). The Development of Children's Understanding of Probability and the Application of Research Findings to Classroom Practice. Unpublished master's thesis, University of Adelaide.

Zaleska, M. (1974). Influence du mode de présentation de séries aléatoires sur le choix d'une stratégie: I.—Situations hypothétiques L'Année psycholigique 74: 125-144.

Zaleska, M. & Askévis-Leherpeux, F. (1976) Influence du mode de présentation de séries aléatoires sur le choix d'une stratégie: II.—Situations réelles L'Année psycholigique 76: 501-514.

4 "He who would seek to gain his life ... must lose it"