A Study of Children's Probability Judgements

Jenni Way
University of Western Sydney, Nepean

This paper reports some preliminary results of an ongoing study of children's intuitive understanding of basic probability concepts. The conceptual framework of the study is outlined. Task-based interviews, conducted with fifty K-6 children in N.S.W., have revealed some developmental trends as well as some interesting insight into the types of strategies applied by children when attempting to quantify their perceptions of random devices and to compare probabilities.

In the words of Kapadia and Borovcnik (1991, p2). "Probability can be thought of as the mathematical approach to the quantification of chance, just as rulers measure distance." The concept of chance is an awkward one; it does not sit well with the causal, logical and deterministic thinking commonly applied in the world of mathematics. Nevertheless, probabilistic thinking provides a significant means of applying mathematical ideas in realistic situations (Kapadia and Borovcnik, 1991). Just as with learning to measure distance, children first need to understand the essential concepts of probability and develop the accompanying skills before they are able to make meaningful quantification.

Through life experiences and the use of language children develop intuitions about such things as chance, luck, fairness, faith, certainty, possibility and impossibility. However most social and cultural situations are well beyond the average person's ability to quantify or explain mathematically. Instruction in probability usually begins with more tangible situations, through the use of random generators, like dice, spinners, raffles and coin flipping. This type of activity allows the crucial concept of sample space to be modelled in a concrete way. Random generators not only include the sample space but also the random action that, by chance, produces an outcome. Consider, for example, a box containing three blue blocks and two red blocks (the sample space), and the random action of drawing a block without looking. Knowledge of the structure of the sample space allows the observation that it is more likely that a blue block will be drawn than a red block, but an understanding of chance adds the realisation that a red block could be drawn. It is this type of thinking that forms the foundation of probabilistic reasoning.

The quantification of chance, the actual probability of an event, requires comparison of a different kind and usually the construction of a fraction. The sample space must be considered in terms of the subset of the items representing the desired outcome (say a blue block) in comparison to the total set of possible outcomes in the sample space, which results in placing 3 over 5.

This creation of a numerator and denominator by examining the structure of the sample space becomes crucial when two different sample spaces are being compared with the purpose of determining which sample space has the greater potential for producing a specified outcome. This task is not difficult when the denominators are the same, but becomes quite complex when the denominators are different (ie. the sample spaces contain a different total number of elements). It becomes necessary to construct either fractions or ratios and apply proportional thinking to compare the two relations.

One additional factor that should be considered is that each successive use of the random generator (eg. drawing of a block) produces an outcome. Repeated draws (with replacement) create a set of frequency data which, due to the element of chance, may or may not be representative of the structure of the sample space (eg. 10 draws might result in all blue blocks, which is not representative of the 3 blue, 2 red structure of the sample space). This can be a powerful distracter in the development of mathematically appropriate thinking (Shaunessy, 1992). The understanding that each random action leading to an outcome (eg. drawing of a block) is an independent event is therefore another factor in probabilistic reasoning.

A Conceptual Framework

When designing interview tasks with the purpose of investigating children's probabilistic thinking it can be helpful to consider some factors of the situation as being external to the child (interviewee) and some cognitive and affective factors being internal (Goldin, 1993). The researcher's goal is to access some of these internal factors. Random generators, and the sample spaces imbedded in their structure, presented in the context of a particular task are physical in nature and therefore external factors in probabilistic reasoning. The way in which an individual perceives the physical structure is an internal factor determined by a range of influences. These influences can include belief in the fairness of the situation, an understanding of chance and randomness, previous experiences associated with the situation, and social/cultural background including religion. Age (or cognitive development) is also an influence in terms of the development of thought processes that allow consideration of multiple outcomes and all possibilities, a functional understanding of proportions and ratios, and the consistent application of logic (Piaget & Inhelder, 1951). All of these influences can determine what mathematics, if any, can be extrapolated from the situation by the individual. Thus, when an individual is called on to make inferences from a random generator's structure (eg. predict most likely outcome) or to make comparisons between sample spaces, both external and internal factors are involved.

Language must also be considered because it is usually the vehicle for communicating ideas between the people involved. The use of the game context assists in aligning the researcher's goals with the interviewee's goals and so reduces the dependence on words to convey meaning.

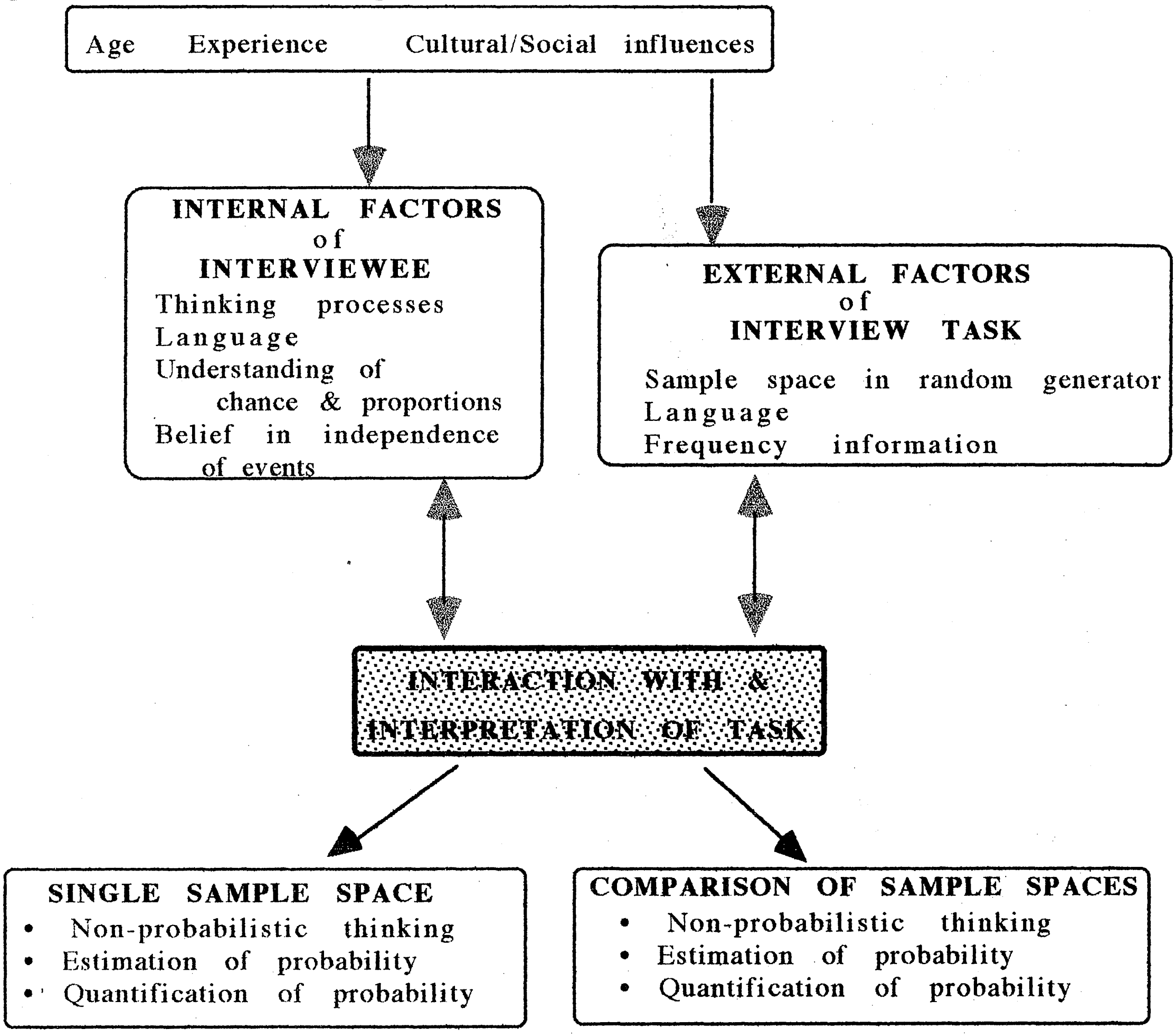
When the interviewee indicates a choice of outcome or a choice of sample space the researcher is witnessing an external action, but can only guess at the internal thought processes that have influenced this choice. Asking the interviewee to give reasons for the choice permits the researcher to gain some insight into the otherwise hidden internal influences. Of course, there is no guarantee that the interviewee will truthfully explain their thinking, or will have the appropriate language available to accurately communicate their thinking, nor that the researcher will correctly interpret what the subject has said. However, cautious interpretation of the information is definitely more useful than no information at all.

There is as yet no one comprehensive theory on the cognitive and psychological components of the acquisition and application of probabilistic reasoning. Therefore the theoretical parameters of this study are based on several established theories that may, in combination, provide a framework for understanding this field. Firstly, there is an underlying constructivist perspective which permeates the whole research design. Secondly, there is an assumption that there will be a change in performance with age, which is consistent with Piaget's developmental theory. This does not exclude the interpretation of result via the more recently developed model of the SOLO Taxonomy (Biggs & Collis, 1982). Thirdly, Fischbein's (1975) ideas about an individual's progress from primary intuitions to secondary intuitions as a result of experience and instruction might provide explanation of some results where Piaget's theories cannot. Finally, the interpretation of the results of some tasks, particularly with the youngest subjects, may be assisted by considering information processing theories such as Brainerd's (1981) working memory model.

As explained in the introduction to this paper, comparing sample spaces can require higher level thinking than considering just one sample space, and therefore these two have been separated in the conceptual diagram. Within these two categories there can be different levels of responses; where probability is not actually considered, where a non-numerical estimate is made, and where quantification actually occurs.

Two major points of contention arise from the research literature. One centres on the age at which a basic understanding of probability arises (Carlson 1970; Falk 1980, Fischbein 1975, Green 1983; Perner 1979; Pumphrey 1968). The other involves a debate

Figure 1. A model of the conceptual framework



about the validity of certain research tasks in determining the presence or level of probabilistic understanding (Acredolo et. al.(1989), Hoemann & Ross 1982, Falk et al, 1980; Perner 1979). The two issues are linked because the latter has been used to explain discrepant results related to age.

Method

The research project being reported here aims to take advantage of the situation that probability is not yet part of the NSW K-6 Mathematics Syllabus and hence has not been formally taught in the schools. Therefore it can be assumed that the understandings that the children hold will be intuitive in nature. The study is exploratory and seeks information about the basic notions of probability discussed above. Of particular interest is the children's perception of the relationship between sample space and the probability (likelihood) of certain events; and any differences in the children's performance in one-sample-space tasks and two-sample-space tasks.

Interview Protocol

A structured interview protocol was developed consisting of four tasks with accompanying questions presented within game contexts. A fifth task was added for half the sample. The tasks utilised two different random generators and required children to consider the structure of sample spaces, consider probabilities of events (without numerical probabilities) and to compare sample spaces (with and without numerical thinking encouraged). The tasks were always presented in a single session (about half an hour) and in the same order The questions were asked in the same sequence each time. Some variation arose due to the random nature of the outcomes in the games, and also due to extra probe questions being used when necessary. All interviews were conducted by the same researcher and were tape recorded. The recorded interviews were transcribed directly into a data-base matrix, where information could be retrieved for individual children, particular tasks, specific questions or for various groupings (such as age groups).

Task 1 'Bears in a Box': Each child selected two colours of plastic bears which were placed into a cardboard box in the ratio of 3:1 (eg. 3 blue, 1 red). The researcher asked the child to state the most likely result of a random draw and explain the reason for their choice. A draw was then made by the child and the result recorded by placing a matching bear on the table. Six draws were made, with the bear replaced each time and the box shaken before each draw. The 'prediction' question was asked before every second draw. After looking at the results of the six draws, the children were asked to say whether another six draws would result in the same, or a different sequence. Another six draws were then made. Finally the children were asked why the box was shaken each time, and whether the game reminded them of anything they had done before.

Task 2 'Non-replacement': Using the same four bears as the previous game a 'prediction' of the colour bear most likely to be drawn was sort, together with the reason for the choice, but after each draw the bear was left out. If the single-colour bear was drawn on the first or second go, the game was repeated to allow the child the opportunity to encounter an 'equally likely outcomes' situation.

Task 3 'Racing Cars': This game used four coloured cars (counters) and a four lane straight race track divided into eight squares between the start and finish lines. Four different spinners were prepared, with Red, Yellow, Blue and Green sectors structured in the following proportions: Spinner A 2:2:2:2, Spinner B 4:2:1:1, Spinner C 3:1:2:2 and Spinner D 3:1:4:0. The first game played was a 'warm-up' game using Spinner A. For the second game, the child was shown Spinner B and asked to choose the car most likely to win. For the third game the child was instructed to be the Red car, then asked to select the spinner (from all four) that would give them the best chance at winning. A series of 'hypothetical' race questions were then asked, which asked them to identify the spinner that provided equal chance for all cars, best chance for Yellow to win, make Red the least likely winner, and make it impossible for Green to win. For two of these choices, the children were questioned about their certainty of the predicted outcome. Reasons for every choice were sought.

Task 4 'Transfer': This task required the children to substitute a 'Bears in a Box' random generator for the spinners in the racing car game. First they were asked to construct a 'fair' game by placing some bears into the box. Then they were asked to construct a different 'fair' game, followed by one in which it would be impossible for Green to win. Finally, they were shown Spinner B and asked to work out how many bears to put into the box to make the game 'work the same way' that the spinner would. Each child was asked to explain their strategy for completing this task.

Task 5 'Proportions': This task was only completed by the children from the second school. In each of four games, the child was presented with two jars, each containing a different combination of Red and Yellow bears. They were asked to choose the jar that gave them the better chance of picking out a red bear, or to declare that both jars gave the same chance. To assist comparisons the bears were set out in lines of Red and Yellow before being placed into the jars. After choosing a jar and explaining their reasons, the children were asked to 'test' their choice by making five draws (with replacement). If two or more Red bears were drawn the child won a small chocolate.

Sample

The sample was drawn from one school in Sydney's south-western metropolitan area and one country school in the south of NSW. K-6 teachers were asked to nominate four children from each grade who were representative of the range of mathematical abilities and who would not be unduly uncomfortable in an interview situation. While a few children were clearly considered to be either 'above average' or 'below average', the vast majority could be considered to be of 'average' mathematical ability. For various reasons, such as malfunction of equipment, only fifty interview recordings were useable for analysis.

Table 1. Sample Sizes

Age Group (years)	Number
5/6	12
7/8	13
9/10	14
11/12	11

Total Sample = 50 48% Male 52% Female

Results

Although several layers of analysis and interpretation are planned only the preliminary results are reported in this paper. This preliminary analysis confines itself primarily to 'correct' and 'incorrect' responses from the children. In most questions the children were asked to make a choice, (for example, that blue was the most likely outcome) and were 'correct' if the mathematically correct choice was made. A reason for making the choice was also sought and this was deemed to be 'correct' if it drew appropriate information from the sample space (eg. There are more blue bears in the box).

Figure 1 shows the percentage of correct choices and reasons across all tasks except the Transfer Task where reasons were not relevant. It can be seen that there is a general trend of being able to make a correct choice but not give a successful explanation. This trend is repeated in the results for each task. The graph also reveals the clear development of probabilistic judgement with age, but with little difference in performance between the 9/10 years age group and the 11/12 years age group. Again this trend is repeated in the data for each task.

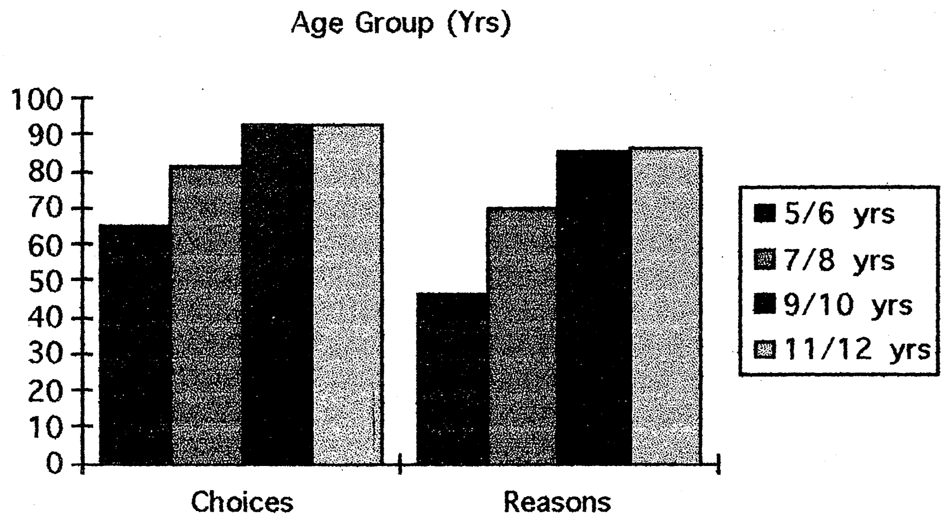


Figure 2: Combined percentage correct across Tasks 1, 2,3 and 5

When the percentages of correct choices and reasons are examined (see Figure 3) it becomes obvious that the Racing Car Task was somewhat easier for the children than the other tasks. It is also apparent that the children found it quite difficult to correctly explain their reasons for choosing a sample space in the Proportions Task. The most common error in the Proportions Task was to only compare the target colour (red) in each jar, and to ignore the amounts of yellows. In other words, the children tended not to use ratio and proportion in their reasoning, and when they did they had difficulty in explaining their thinking.

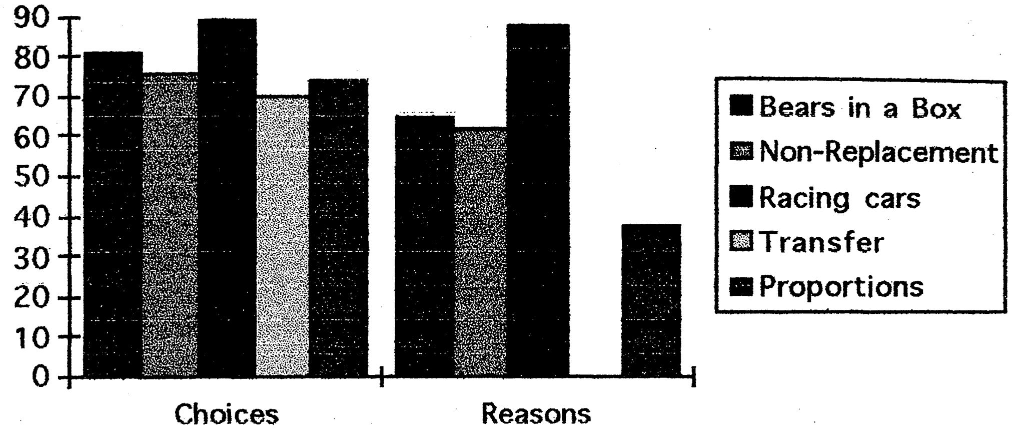


Figure 3: Summary of correct responses and reasons for each task. (Note: reasons were not a relevant feature of the 'Transfer Task')

Task 1 'Bears in a Box': In this task the children (with the exception of the 5/6 year olds) made very confident predictions about the outcome of the first draw, backed up by equally confident reasons. However, the predictions for the third and fifth draws showed a decline. For example, 9/10 years: 100%, 78.6%, 78.6%. This suggests that the children's thinking was inappropriately influenced by the frequency data provided by the outcomes of the previous draws. This is supported by the changes in the reasons given for choices.

83% of the 5/6 year olds said they expected the next six draws to produce a different sequence, but surprisingly only 63.6% of the 11/12 year olds made this prediction. Most children, (including 75% of the 5/6 yrs) understood the box was shaken to maintain the randomness of draws.

Non-replacement task: Table 2 shows the typical increase of appropriate probabilistic thinking with age. Of note in this set of results is the fact that the youngest age group failed to recognise the equally likely situation that arose when there was only one of each colour left in the box. Most of these children were able to state that one of each colour remained in the box, but insisted on naming one as being more likely to be drawn.

Table 2. Non-replacement: Percentages correct of total possible correct

Age Group (Years)	Correct Choices	Correct Reasons	Recognised Equally Likely
5/6	60%	30%	0%
7/8	73.1%	57.7%	50%
9/10	83.3%	75%	60%
11/12	86.4%	81.8%	75%

Racing Cars Task : All age groups were very successful in identifying the best sample space (spinner) to use to produce a specified winner of a game (eg. 5/6 years 91.7% and 11/12 years 100%). With the exception of the 5/6 years group (66.7% correct choice, and 58.3% correct reasons) there was 100% correct responses for identifying the most likely result of the game by looking at a biased spinner. The same number of correct reasons were supplied as choices for most questions. Again the 5/6 year olds had difficulty in recognising the 'equally likely' (58.3% correct) sample space. They also had difficulty in identifying the 'impossible for green to win' (41.7% correct) sample space.

When asked how certain they were about the specified colour winning the game with the specified spinner, 88.9% of 11/12 year olds were able to explain that because of the element of chance any colour represented on the spinner could win the game. However, when discussing the spinner they had chosen to try to make a colour lose the race only 50% applied the same type of reasoning. None of the 5/6 year olds acknowledged the element of chance when responding to these questions.

Transfer task: All age groups had little difficulty in constructing 'equally likely' sample spaces, with many children stating that the number of bears placed into the box was irrelevant as long as each colour was represented in equal amounts. However, the children found constructing an 'imposable for green' sample space more difficult (5/6 yrs 50%, 7/8 yrs 75%, 9/10 yrs 92.3%. 11/12 yrs 80% of responses correct). Although almost half the children managed to construct a sample space with the coloured bears that resembled the biased Spinner B (4:2:1:1), only 21.3% of the responses were mathematically correct. Analysis of the strategies used by the children to determine the numbers of bears revealed four distinct categories: non-comparison, measurement, ordering and fractional thinking. All of the correct responses, except one, arose from fractional thinking.

Proportions: In general the children (sample of 23) found this the hardest task, with 73.3% of choices being correct and 37.6% of the reasons correct. The children performed best in jar choice where comparing the target colours was an appropriate strategy (1 red, 4 yellows vs 3 red, 4 yellows), but most were unable to give a complete explanation (ie. note that the yellows were equal in both jars). The children had particular difficulty in recognising the proportionally equal sample spaces (1 red, 4 yellow vs 4 red, 8 yellow).

Symmetrical Sample Spaces: During the interview the children were given five opportunities to identify sample spaces where each outcome was equally represented. In general, this presented little difficulty in the Racing Car and Transfer tasks, but considerable difficulty in the Non-replacement and proportions tasks. For example, the 7/8 years olds responses were 100% correct in the Transfer Task, but only 33.3% in the Proportions task.

Discussion

One way of discussing these preliminary results is to view the interview tasks in terms of whether the random generator presented the sample spaces in a spatial model (sectors on spinners) or in a numerical model (discrete items, bears, in containers). This can be combined with whether the questions required consideration of a single sample space or the comparison of two or more sample spaces.

Figure 4. Relationship between type of random generator and type of probability judgement.

	Single sample space	Comparing sample spaces
Spatial model random generator	Racing cars	Racing cars
	Transfer	
Numerical model random generator	Bears in a box Non-replacement	Proportions

The matrix shows that each cell contains at least one of the interview tasks, with the Transfer Task occupying a unique bridging position between the spatial model and the numerical model. The results showed that the children found the Racing Car task easiest, suggesting that spatial model random generators are easier for them to interpret regardless of whether they are required to work with one sample space or to make comparisons between sample spaces. The children managed the Bears in a Box and Non-replacement tasks reasonably well suggesting that numerical models are not difficult to deal with when only a single sample space needs to be considered. The children found the Proportions Task a little more difficult, particularly considering they were usually unable to explain their choices appropriately. This task demanded that attention be paid to number in the comparison of two sample spaces. Similarly, the Transfer Task demanded attention to number and, similarly the children found this more difficult than the other tasks. This is consistent with the conceptual framework's position that children may be able to make sound probability estimates, but have difficulty with quantification of probabilities.

In terms of Piagetian theory, the spatial model random generator tasks allow probability to be dealt with spatio-temporally, while the numerical models would be dealt with logico-arithmetically. When comparing colours within spinners and comparing different spinners, children are able to make simple area comparisons, or as Hoemann and Ross (1982) say, make 'magnitude estimations'. A similar strategy can be applied when looking at numerical models by noticing the size of colour groups without dealing in precise numbers. It should also be noted that some children used numbers, particular fractions, to help explain their choices of spinners. A closer analysis of explanations and language may reveal some patterns in the voluntary use of numbers. It is interesting that the children found it very difficult to recognise equal probability in numerical models, but quite easy in spatial models.

There is obviously a strong link between age and the development of probabilistic thinking, though the details of this link to various aspects of tasks and reasoning are not yet clear. The inconsistency of responses in the 5 to 8 years age range highlights the tentative nature of emerging understandings. The SOLO Taxonomy (Biggs & Collis, 1982) may be useful for exploring this area. If the youngest children lacked sufficient working memory (Brainerd, 1981) to make successful probability estimates, then a decline in performance between the Bears in a Box task and the Non-replacement task would be expected, because the children had to remember what colours were left the box. However this was not the case.

There is obviously much yet to analyse, particularly in the reasons and explanations given by the children and in the patterns of their thinking. Currently, another, more diverse proportions task is being used in interviews at a western Sydney school with the purpose of further exploring numerical reasoning when comparing sample spaces.

References

Acredolo, C., O'Connor, J., Banks, L., & Horobin, K. (1989). Children's ability to make probability estimates: Skills revealed through application of Anderson's measurement methodology. *Child Development, 60*, 933-945.

Biggs, J., & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York: Academic Press.

Brainerd, C. (1981). Working memory and the developmental analysis of probability judgements. *Psychological Review 88*, 463-502.

Carlson, J. (1969). Children's judgements as related to age, intellegence, socio-economic level, and sex. *Human Development, 12*, (192-203).

Falk, R., Falk, R., & Levin, I. (1980). A potential for learning probability in young children. *Educational Studies in Mathematics, 11*, 181-204.

Fischbein, E. (1975). *The intuative sources of probabilistic thinking in children*. Dordrecht: Reidel.

Green,D.R. (1983). School pupil's probability concepts.*Teaching Statistics, 5*(2), 34-42.

Hoemann, H., & Ross, M. (1982). Childrens concepts of chance and probability. in C. Bruind (Ed.), *Children's Logical and Mathematical Cognition: Progress in Cognitive Development Research* (pp.95-121). New York: Springer-Verlag.

Kapadia, R., & Borovcnik, M. (1991). The educational perspective. In R. Kapadia & M. Borovcnik (Eds.) *Chance Encounters: Probability in Education*. (pp.73-105). Dordrecht: Kluwer Academic.

Perner, J. (1979). Discrepant results in experimental studies of young children's understanding of probability. *Child Development, 50*, 1121-1127.

Piaget, J. & Inhelder, B. (1975) *The origin of the idea of chance in children*. London: Routledge & Kegan Paul. (Original work published 1951).

Pumphrey, P. (1968). The growth of the schema of proportionality. *British Journal of Educational Psychology, 38*, 202-204.

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.