Beliefs about Chance in the Middle Years: Longitudinal Change

Jane M. Watson
University of Tasmania
<Jane.Watson@utas.edu.au>

Annaliese Caney
University of Tasmania
<Annaliese.Caney@utas.edu.au>

Ben A. Kelly
University of Tasmania
<Ben.Kelly@utas.edu.au>

This report considers beliefs about chance in relation to understanding of random processes and luck. Changes in chance beliefs over two- and four-year periods were measured for 265 students initially in Grades 3 and 6. Students were asked three questions on repeated occasions, based on the meaning of the word random, their beliefs about luck, and their beliefs about how luck affects winning a lottery. Change was also measured for performance on three questions related to chance measurement. The association of chance belief and chance measurement was found to be weak. Educational implications are considered.

Beliefs about chance and chance mechanisms are notoriously difficult to document consistently (e.g., Watson & Moritz, 2003), difficult to change with instruction (e.g., Fischbein & Gazit, 1984), and resistant to change over long periods of time (e.g., Fischbein & Schnarch, 1997). From young children who believe that God, fate, or mental powers determine dice outcomes (Truran, 1995) to adults who always choose tails because it is lucky for them (Watson & Moritz, 2003), non-mathematical beliefs about random processes abound in society. The meaning of "random" itself is fraught with difficulty. In social contexts it is often equated with "haphazard" and in mathematical contexts it is often used as an adjective, such as for "random sample" or "random processes." In this latter sense it is often used without further amplification. In statistics education the bench mark for discussion of the concept of random is provided by David Moore (1990) who says the following:

> Phenomena having uncertain individual outcomes but a regular pattern of outcomes in many repetitions are called *random*. 'Random' is not a synonym for 'haphazard,' but a description of a kind of order different from the deterministic one that is popularly associated with science and mathematics. Probability is the branch of mathematics that describes randomness. (p. 98)

It is unlikely that most teachers, let alone many students, can describe random in this fashion. It is essential, however, to acknowledge the importance of the concept and begin discussing it in the middle school years.

Luck is another idea that has diverse interpretations across the population and in different contexts. It is often associated with a positive aspect of fortune, for example, when expressing the belief that "tails are lucky for me" or in describing the winner of a lottery. It also, however, is sometimes used as a synonym for chance itself, as in "the luck of getting one black is a 50-50 chance, but the luck of getting a second one is even less of a chance." Responses similar to these are reported by Amir and Williams (1999), who surveyed 11-12-year-old students and found that 72% agreed that some people are luckier than others in raffles and dice games, whereas 43% believed that their own choices when tossing a coin would be better than others' choices.

A National Statement on Mathematics for Australian Schools (Australian Education Council [AEC], 1991) acknowledges the existence of the idea of "luck" in suggestions for students in Band A, when it suggests activities that "Clarify and use common expressions such as ‘being lucky’, ‘that's not fair’, ‘always’, ‘it might happen’, ‘tomorrow it will probably rain’.” (p. 166). Reference to “random” however is less specific in the document, where it is mainly used as an adjective to describe “random sample” (p. 185) or “random number generators” (p. 182). The New South Wales Board of Studies’ (2002) curriculum document for mathematics contains the outcome “students learn to recognise randomness in chance situations” (p. 75) but provides no further elucidation of the term. This is an area where intuitions are being built but there is little support from curriculum documents.

Previous Research

Research on random processes was among the earliest conducted in the field of chance and data in the 1980s (e.g., Kahneman & Tversky, 1982), using university students to explore widely held misconceptions. Later work by Green (1983) and Fischbein and Gazit (1984) began to explore school students’ conceptions. Fischbein and Gazit developed two items similar to those used in this study to explore students’ views of luck in Grades 5, 6, and 7. For a question about entering the classroom by putting the right foot first to increase the chances of getting good marks, high percents of students rejected this belief, ranging from 60% in Grade 5 to 97% in Grade 7. For a question about using winning lottery numbers again because they were lucky, rejection levels were not as high, ranging from 12% for one Grade 5 group to 54% for a Grade 9 group. Many of these students reasoned that the same numbers *could not* win again. Watson, Collis, and Moritz (1995) used the two luck questions in this study, derived from those of Fischbein and Gazit, and found increasingly appropriate levels of response in Grades 3, 6, and 9.

There is little reported in the literature about students’ conceptions of the idea of random itself. Moritz, Watson, and Pereira-Mendoza (1996) asked the question used in this study, “What things happen in a random way?”, of students in Australia and Singapore and found increasing performance with grade level. In a different study, Watson and Kelly (2003) also included the question “What does random mean?” with the above question for students in Grades 7 and 9. In looking at change for students involved in specific instruction in chance and data, Watson and Kelly found that students in Grades 7 and 9 improved in their ability to define and illustrate the term after instruction. This level was retained after two years but the longitudinal change was not significantly different than that for students not involved in the instruction. Watson and Caney (in press) also considered understanding of the random concept for 99 students involved in face-to-face interviews, some of whom were part of the large study from which this report is taken. Only 15 of these students however, are included in the current data set. To our knowledge, no study has considered the longitudinal change in chance beliefs for middle school students over two consecutive two-year periods.

Research Questions

In a curriculum environment where traditionally there has been more emphasis on the measurement of chance (e.g., finding numerical probabilities), of interest in this study are the levels of understanding associated with chance beliefs that are expressed in words not numbers. What levels of understanding of the concepts of random and luck are displayed across the middle years of schooling and how does understanding change over time? How does this compare with change in the understanding associated with chance measurement skills that are more fundamentally a part of the mathematics curriculum and is there an association between the two aspects of probabilistic thinking?

Method

Participants. As part of a larger five-year study of students' understanding of chance and data in 13 state government primary, secondary, and district schools in Tasmania, 148 students in Grade 3 and 117 students in Grade 6 were surveyed on three separate occasions in years 1, 3, and 5 of the project. Students completed the same survey each time and hence data reflect one cohort of students in Grades 3, 5, and 7, and another in Grades 6, 8, and 10. Although many more students were involved in the larger study, in considering longitudinal change, only those students who completed the survey on all three occasions are included in this report.

Tasks and Procedure. As part of an instrument developed to assess student understanding of statistical concepts (Watson, 1994), three survey questions that explored students' descriptive understanding of Chance Beliefs based on random processes and luck, and three questions investigating understanding of Chance Measurement were administered to students (see Figure 1). The Chance Measurement questions are more typically a part of the chance and data curriculum and had been analysed longitudinally for these students earlier (Watson & Moritz, 1998). In all schools, 45 minutes of school class time was allocated to administer the larger survey of which these survey questions were part. Extra questions were asked of students in Grades 6, 8, and 10.

Chance Belief Questions

Q1. What things happen in a 'random' way?

Q2. Every morning, James gets out on the left side of bed. He says that this increases his chances of getting good marks. What do you think?

Q3. One day Claire won Tattslotto with the numbers 1; 7; 13; 21; 22; 36. So she said she would always play the same group of numbers, because they were lucky. What do you think about this?

Chance Measurement Questions

Q4. Consider rolling one six-sided die. Is it easier to throw
(1) a one, or
(6) a six, or
(=) are both a one and six equally easy to throw? Please explain your answer.

Q5. A mathematics class has 13 boys and 16 girls in it. Each pupil's name is written on a piece of paper. All the names are put in a hat. The teacher picks out one name without looking. Is it more likely that
(b) the name is a boy, or
(g) the name is a girl, or
(=) are both a girl and a boy equally likely? Please explain your answer.

Q6. Box A contains 6 red and 4 blue marbles and Box B contains 60 red and 40 blue marbles. Each box is shaken. You want to get a blue marble, but you are only allowed to pick out one marble without looking.
Which box should you choose?
(A) Box A (with 6 red and 4 blue).
(B) Box B (with 60 red and 40 blue).
(=) It doesn't matter. Please explain your answer.

Figure 1. Chance Belief and Chance Measurement Questions used in the survey (adapted from Fischbein & Gazit, 1984).

Data Coding and Analysis. Responses to the Chance Belief items were coded by the second author in line with the criteria set in Watson and Caney (in press) and Watson, Collis, and Moritz (1995). Codings were confirmed and disagreements resolved with either the first or third authors. The levels of response associated with the coding were influenced by the Biggs and Collis (1982) Structure of Observed Learning Outcomes (SOLO) and the statistical appropriateness of the responses. Non-responses were coded 0; ikonic responses reflected intuitive beliefs susceptible to balance and order (Code 1); unistructural responses contained single ideas about random behaviour or luck or an "anything can happen" belief (Code 2); multistructural responses recognised regularity in chance but were inconsistent or included sequential ideas in describing random (Code 3); relational responses showed integration of ideas about chance outcomes not influenced by balance or order (from Table 2, Watson & Caney, in press). Examples of responses associated with each code are given in the Results. As well, the distributions of responses across codes for each grade are presented. Although the same students were involved in years 3 and 5 of the study, the data give an indication of change across grades as well as time. The maximum score possible on the three Chance Belief items combined was 12. The mean scores out of 12 for each grade are presented to give an indication of overall change for the subscale over time in relation to the two cohorts. Paired *t*-tests were used to judge progress for each cohort.

The association of outcomes for the Chance Belief items with the Chance Measurement items is considered using the total score for each subscale. The Chance Measurement items were recoded from the codes of Watson & Moritz (1998), following the same ordering of levels of response for Questions 4 and 5 but using codes from 0 to 4, and combining the highest three levels of Watson and Moritz for Question 6 to a Code 4 response. This recoding equated all appropriate numerical explanations to the Chance Measurement questions (equivalent to "1/6" for Question 4, "16/29" for Question 5, and "equal because the ratio of colours is the same" for Question 6). The total possible score on Chance Measurement was hence also 12. Due to this recoding, a similar comparison over time with paired *t*-tests was completed for the Chance Measurement subscale. It is hence possible to compare change for the two subscales as well as calculate a Pearson product-moment correlation for each grade.

Results

Table 1 contains the distribution of codes for each grade level for each code for Q1 to Q3. As can be seen, although there was a tendency for improvement with grade level on the three items, this was not strong. Significant, however, was the monotonic decline in non-responses to the random question (Q1) with grade. Table 2 contains illustrative responses for each coding level for each item, annotated by grade level. As can be seen in Table 1, Code 2 responses were modal for nearly every grade on each survey. Code 2 for Q2 and Q3 amounts to rejecting luck but without a strong justification for doing so. The percent of students in the younger cohort with a Code 2 or higher level response for getting out of bed on the left side to be lucky (Q2) rose from 67% in Grade 3 to 89% in Grade 7. For the other cohort it rose from 79% in Grade 6 to 95% in Grade 10. For the lottery question (Q3), 67% of Grade 3 rejected luck with at least a Code 2 response, rising to 82% in Grade 5 but dropping back to 80% in Grade 7. For Grade 6 students, 81% rejected luck, which dropped to 79% in Grade 8 and rose to 94% in Grade 10. Combining data from all groups, half of students gave Code 2 responses to both questions when answering them on the same survey.

Table 1
Distribution of Response Codes for Chance Beliefs by Grade for Each Item (n=148 for Grades 3, 5, and 7; n=117 for Grades 6, 8, and 10)

Grade	Q1 (Random)					Q2 (Bed & Luck)					Q3 (Lottery)				
	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
3	108	31	7	2		12	37	93	6		12	37	99		
5	57	31	59	1		3	21	108	16		6	21	116	5	
7	31	27	82	8		3	14	108	23			29	103	15	1
6	44	29	44			7	17	75	16	2	3	19	87	8	
8	24	18	67	8		4	6	84	20	3	3	21	79	14	
10	10	12	77	15	3	3	3	79	27	5	1	6	76	30	4

Table 2
Responses Illustrating Each Response Level by Code for Each Chance Belief Item Annotated by Grade

Code	Q1 (Random)	Q2 (Bed & Luck)	Q3 (Lottery)
0	"I have not heard of it before so I don't know what it means." [Grade 3]	"I don't know." [Grade 3]	"I am not sure." [Grade 3]
1	"Quickly." [Grade 6] "Things in order." [Grade 5]	"I think it would be good." [Grade 3] "Sometimes it could happen." [Grade 6]	"It's lucky." [Grade 6] "If she was lucky before she might get lucky again." [Grade 9]
2	"Tattslotto." [Grade 7] "In any order." [Grade 5]	"That it's not true." [Grade 7] "It wouldn't matter which side he got up on." [Grade 7]	"Superstitious." [Grade 10] "It's unlikely that the same group of numbers will ever come up again." [Grade 9]
3	"The lotto is random because you don't know what number will be picked." [Grade 7] "When there is an even chance of things occurring such as a competition draw." [Grade 10]	"No, it really depends on how well you study or prepare for tests or other things." [Grade 7] "I think it's stupid. Getting out of bed on a certain side can't increase your chance of getting good marks. If it makes him feel confident, fine, but it has nothing to do with intelligence." [Grade 6]	"I think that the numbers aren't lucky because they are picked randomly." [Grade 8]
4	"Things that happen randomly happen with no preference if something is picked randomly it is picked without prejudice." [Grade 10]	"This cannot be true physically, but as he thinks this, he psychologically expects to do well if he gets out on the left side of bed, therefore this makes him try harder." [Grade 10]	"Tattslotto numbers are chosen randomly so she would have an equal chance with other numbers." [Grade 9]

Table 3 shows the change in mean scores on the total score for the Chance Beliefs subscale over the two two-year periods. Each cohort improved over each period, with all paired *t*-tests being statistically significant. These changes represent an improvement of between about half a standard deviation for most pairs and one standard deviation for the Grade 3/5 comparison. It is interesting to note, however, that there is no difference between Grades 5 and 6, or Grades 7 and 8, across the two cohorts.

Table 3
Means, Standard Errors and Paired t-values for Chance Beliefs across Two Years for Each Cohort (all p-values less than 0.0001)

	Grade 3	Grade 5	Grade 7	Grade 6	Grade 8	Grade 10
Mean	3.56	4.76	5.39	4.76	5.50	6.40
Std Err.	0.103	0.105	0.112	0.149	0.141	0.155
t		9.27	5.24		4.74	5.55

In order to compare change in Chance Beliefs with change in Chance Measurement understanding, the means and standard errors for the Chance Measurement subscale are given for each grade in Table 4. The mean scores (out of 12) are higher for each grade than for Chance Beliefs and the differences over two-year periods, although significant, are not as large as those for the Chance Belief items. Although the Grade 5 and 6 mean difference was larger for Chance Measurement ($p<.02$) than for Chance Belief, again the Grade 7 and 8 difference was negligible.

Table 4.
Means, Standard Errors and Paired t-*values for Chance Measurement across Two Years for Each Cohort*

	Grade 3	Grade 5	Grade 7	Grade 6	Grade 8	Grade 10
Mean	5.05	6.30	7.67	7.03	7.75	8.84
Std Err.	0.207	0.206	0.171	0.231	0.236	0.209
t		4.28	5.12		2.17	3.44
p		$p<.0001$	$p<.0001$		$p<.02$	$p<.0004$

The relationship of students' scores on the Chance Belief items with their scores on the Chance Measurement items can be seen in Table 5 containing the correlations for the six grades. For the younger cohort this represents only 4.5% of shared variation in Grade 3 up to 11.3% in Grade 7. For the other cohort, the variance explained ranges from 26.9% to 29.2%.

Table 5.
Correlation of Total Scores for Chance Belief Questions with Chance Measurement Questions for Each Grade

Grade 3	Grade 5	Grade 7	Grade 6	Grade 8	Grade 10
.326	.211	.336	.519	.528	.540

Discussion

The discussion will consider the research questions and the educational implications of the outcomes. Considering first the levels of Chance Belief displayed across the grades observed in this study, the modal response was nearly always (17 out of 18 cases) the unistructural level (Code 2), and for 15 cases a majority of students could give a single example or idea associated with random or disagree with the idea of luck as an influence on events. That more students could not go further in describing random or explaining why luck is not operative is disappointing. Out of a total possible score of 12 on the Chance Belief subscale, the mean by Grade 10 only reached 6.40. Although there was no direct intervention by researchers over the years of the project, the chance and data curriculum was being implemented in Tasmania.

The change in relation to Chance Beliefs within the two cohorts of students across the four years of the study is encouraging, if not as great in absolute terms as desired. It is interesting to note, however, in the light of other studies involving middle school students (e.g., Callingham & McIntosh, 2002) that there is little or no difference between Grade 5 in the younger cohort and Grade 6 in the older group, or between Grades 7 and 8 in the two groups. Although students were in the same schools across the state little else links the students. Two possible explanations could be suggested. One would be associated with a middle school plateau performance effect at the end of primary school or the beginning of high school. The other would be that the implementation of the curriculum brought the performance of the Grade 5 students in year 3 closer to the performance of the Grade 6 students in year 1, and the performance of the Grade 7 students in year 5, closer to that of the Grade 8 students in year 3. This would be an encouraging outcome but it is not possible to continue the comparison into year 5 of the project.

In considering the difference in improvement for the two subscales of chance understanding, two features are of interest. One is that the levels of performance in relation to Chance Measurement are higher, and the other is that the change over the two-year periods is less pronounced for the Chance Measurement items. These observations may be related in that more classroom experiences are likely to be employed earlier in the primary years in relation to Chance Measurement and a ceiling effect may occur in the middle school. The fact that the highest mean is 8.84 out of 12 in Grade 10 may be a realistic expectation for these items but one might desire more.

The low association of the scores for the Chance Beliefs subscale with the Chance Measurement subscale again allows several possible explanations. It is likely that the focus in the classroom is more specifically on the measurement of chance leading to numerical answers rather than on discussion of ideas like random and luck generally. Students hence may not have thought very much about these ideas and in answering the survey questions were relying on their out-of-school experiences. Some students may also be less likely to expand on verbal descriptions in written survey contexts. Watson and Caney (in press), for example, found that students responded on average one level higher in an interview context than on surveys, when describing the meaning of the term random. The difference in the percent of variation explained for the two cohorts is difficult to reconcile. Perhaps one might expect less consistency of response in Grade 3 with the non responses to Q1 but for this group of children it continues on to Grade 7. Similar values across time for each cohort indicate at least consistently varied performance!

The outcomes of this longitudinal study suggest that it is important at the middle school level to discuss chance beliefs as well as chance measurement. Since items of both types were coded on the basis of a structural model and statistical appropriateness, the better performance in relation to chance measurement suggests that with assistance to structure appropriate understanding, performance in relation to chance beliefs could be expected to improve. In particular, discussion should focus on beliefs about the nature of random processes and the reasons why luck is not an adequate explanation for events. These are closely connected as the highest level responses to the lottery question (Q3) rely on the random nature of the lottery process with all combinations of outcomes being equally likely and no bias involved. It is important to go beyond calculations of probabilities of chance events to develop an appreciation of the underlying phenomena, in order for descriptive understanding to reach the level of numerical performance.

Acknowledgements. The data used in this study came from projects A79231392 and W0009108 funded by the ARC with the current analysis funded by Grant No. DP020867.

References

Amir, G. S., & Williams, J. S. (1999). Cultural influences on children's probabilistic thinking. *Journal of Mathematical Behavior, 18*(1), 85-107.

Australian Education Council. (1991). *A national statement on mathematics for Australian schools.* Carlton, VIC: Author.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.

Callingham, R., & McIntosh, A. (2002). Mental computation competence across years 3 to 10. In B. Barton, K. C. Irwin, M. Pfannkuch & M. O. J. Thomas, (Eds.), *Mathematics education in the South Pacific* (Proceedings of the 25th annual conference of the Mathematics Education Research Group of Australasia, Auckland, pp. 155-163). Sydney: MERGA.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? An exploratory research study. *Educational Studies in Mathematics, 15*, 1-24.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematical Education, 28*, 96-105.

Green, D. R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In D. R. Grey, P. Holmes, V. Barnett, & G. M. Constable (Eds.), *Proceedings of the 1st International Conference on Teaching Statistics* (Vol. 2, pp. 766-783). Sheffield, England: Teaching Statistics Trust.

Kahneman, D., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases.* Cambridge: Cambridge University Press.

Moore, D. S. (1990). Uncertainty. In L. S. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

Moritz, J. B., Watson, J. M., & Pereira-Mendoza, L. (1996, November). *The language of statistical understanding: An investigation in two countries*. A paper presented at the Joint ERA/AARE Conference, Singapore. Retrieved 25 March, 2004, from http://www.swin.edu.au/aare/96pap/morij96.280

New South Wales Board of Studies. (2002). *Mathematics: Years 7-10 syllabus*. Sydney: Author.

Truran, K. (1995). Animism: A view of probability behaviour. In B. Atweh & S. Flavel (Eds.), *Galtha* (Proceedings of the 18th annual conference of the Mathematics Education Research Group of Australasia, Darwin, pp. 537-541). Sydney: MERGA.

Watson, J. M. (1994). Instruments to assess statistical concepts in the school curriculum. In National Organizing Committee (Ed.), *Proceedings of the Fourth International Conference on Teaching Statistics. Volume 1* (pp. 73-80). Rabat, Morocco: National Institute of Statistics and Applied Economics.

Watson, J. M., & Caney, A. (in press). Development of reasoning about random events. *Focus on Learning Problems in Mathematics*.

Watson, J. M., Collis, K. F., & Moritz, J. B. (1995). Children's understanding of luck. In B. Atweh & S. Flavel (Eds.), *Galtha* (Proceedings of the 18th annual conference of the Mathematics Education Research Group of Australasia, Darwin, pp. 550-556). Sydney: MERGA.

Watson, J. M., & Kelly, B. A. (2003, December). *The vocabulary of statistical literacy*. Refereed paper presented at the Annual Conference of the Australian Association for Research in Education, Auckland.

Watson, J. M., & Moritz, J. B. (1998). Longitudinal development of chance measurement. *Mathematics Education Research Journal. 10*(2), 103-127.

Watson, J. M., & Moritz, J. B. (2003). Fairness of dice: A longitudinal study of students' beliefs and strategies for making judgments. *Journal for Research in Mathematics Education, 34*, 270-304.