Chinese Students' Understanding of Likelihood

Li Jun and Lionel Pereira-Mendoza
Division of Mathematics, School of Science, NIE

Abstract

Recent research in the area of probability has explored the development of probabilistic thinking of students from primary to post-secondary. However, most of the research has been undertaken in the West, and there has been no research on students in China, where probability is not introduced until Grade 12 (a very different situation from that in the West). This paper presents an analysis of twenty-four questionnaire items designed to focus on students' understanding of certain, impossible, and possible events. Students' misconceptions are identified and discussed. The analysis reported here is part of a larger study of Chinese students' understanding of probability.

Background

There has been considerable research into students' thinking and misconceptions in probability (Fischbein, Nello, & Marino, 1991; Green, 1983; Jones, Langrall, & Thornton, 1997; Lecoutre, 1992; Moritz, Watson & Pereira-Mendoza, 1996; Shaughnessy, 1983; Williams & Amir, 1995), but none of this research has been undertaken on students in China. *A more complete list of references is available from the authors.*

In this paper, the authors are interested in students' understanding of certain, impossible, and possible events. Specifically, what misconceptions do they have and what are the cognitive levels of students' reasoning in terms of the SOLO taxonomy (Biggs & Collis, 1991) are the research questions discussed in this paper.

Methodology

Sample

The 567 students in this study were enrolled in 12 classes in 7 schools in Shanghai. This consisted of 174 grade 6 students, 209 grade 8 students, and 184 grade 12 students (the only group who had received prior instruction in probability). Four classes were selected at each grade; half from ordinary schools and half from advanced schools (the official curriculum in the two kinds of schools is same). It should be noted that for grade 6 there should be no difference between advanced and ordinary schools. However, one of the grade 6 classes drew its students from all over the city, and these students might be above average in ability.

Questionnaires

Nine different questionnaires (a total of 83 items) were used in each class. This paper reports the data on the first 24 items (see Table 1). Each student was asked to answer one questionnaire that contained two or three of the items analysed in this paper. They were required to explain the reasons for their answer. These questions include some one-die tasks that had been used by Fischbein et al. (1991) and were extended to one pack of cards tasks and three dice/three packs of cards tasks. The 24 items were developed along two parallel dimensions, structure and context. For example, the one die tasks (1A1~1A6) had almost the same mathematical structure as the one pack of cards tasks (2A1~2A6), the main difference being the context; one die tasks (1A1~1A6) parallels three dice tasks (1B1~1B6) on the surface, but their mathematical structures are very different. The individual tests were designed to contain different versions of the "same" item (context or structure was changed). The final component of each item asked students to indicate their level of confidence in their response. Finally, since younger students have no prior instruction in probability, the word "outcome" was used in the questions instead of the formal term "event".

Procedure

A short set of instructions was given to each class prior to the administration of the questionnaire. The major points in these instructions were: (1) showing them a normal six-sided die; (2) showing them a pack of cards without picture cards; (3) explaining that in the multiple-choice items if they did not agree with any of the answers, they could give an alternative; and (4) telling them that they had to give explanations of their answers for each question.

Forty-five to fifty minutes were allowed for the test, but most of the students finished comfortably within this time limit.

Interviews

After a class completed the questionnaires, a small group of students from that class were selected to be interviewed the next day. The interviewees were given their own test paper and sometimes they were required to answer other parallel problems. The interviewer read out the problems. The main purpose of the interviews was to check what the students actually meant in their written answers and get additional evidence on their misconceptions. The interviews were audiotaped.

Table 1. Twenty-four Questionnaire Items

1A A six-sided normal die is rolled once. Please indicate whether the following outcome is impossible, possible or certain: (1) the number rolled is an even number; (2) the number rolled is smaller than 7; (3) the number rolled is bigger than 6; (4) the number rolled is 2; (5) the number rolled is 6; (6) the number rolled is not 6. [6 items]
1B Three six-sided normal dice are rolled once. Please indicate whether the following outcome is impossible, possible or certain: (1) all the numbers rolled are three even numbers; (2) all the numbers rolled are smaller than 7; (3) all the numbers rolled are bigger than 6; (4) all the numbers rolled are 2; (5) all the numbers rolled are 6; (6) none of the numbers rolled is 6. [6 items]
2A There is a pack of cards without the picture cards. One card is drawn from the pack of shuffled cards. The Ace is worth one. Please indicate whether the following outcome is impossible, possible or certain: (1) the card drawn is heart or diamond; (2) the card drawn is smaller than a Jack; (3) the card drawn is bigger than a 10; (4) the card drawn is the 2 of diamonds; (5) the card drawn is a 10; (6) the card drawn is not a10. [6 items]
2B There are three packs of cards without the picture cards. One card is drawn from each pack of shuffled cards. The Ace is worth one. Please indicate whether the following outcome is impossible, possible or certain: (1) all the three cards drawn are hearts or diamonds; (2) all the three cards drawn are smaller than a Jack; (3) all the three cards drawn are bigger than a 10; (4) the three cards drawn are the 2 of diamonds, 5 of diamonds and the 8 of spades; (5) all the three cards drawn are the 2 of diamonds; (6) all the three cards drawn are different in both number and suit. [6 items]

Table 2. Summary Description of SOLO Level Related to Different Contexts

SOLO Level		ELE whether the target outcome is on the list of all possibilities	CLA measure probability	SUB Judge by experiences or beliefs
IK		-	-	Believes outcome is controllable, egotistic response
C	U	lists an incomplete set of outcomes for a one-stage experiment; limited knowledge of uncertainty leads to confuse "certain" and "impossible" with "possible"	uncertainty expressed with very simple, general reasons such as "according to the theory of probability"	Interprets uncertainty as one's luck; qualifies uncertainty; limited knowledge of uncertainty leads to confuse "certain" and "impossible" with "possible"
	M	lists a complete set of outcomes for a one-stage experiment but not able to do it so well for a three-stage case; gives two examples to support neither "impossible" nor "certain"	attempts to quantify uncertainty such as indicates equiprobability of all the possible outcomes or although quantifies correctly, takes "possible" as "certain"	Assigns an evaluated chance value to an experiment
	R	lists a complete set of outcomes (unequally-likely) for a three-stage experiment	assigns a correct chance value to a one-stage experiment (simple event); attempts to quantify uncertainty for a one-stage experiment (compound event) or although quantifies correctly, takes "possible" as "certain"	-
F	U	-	assigns a correct chance value to a one-stage experiment (compound event); attempts to quantify uncertainty for a three-stage case, but eventually gives up	-
	M	-	quantifies probability for a three-stage case by a wrong formula, or although quantifies correctly, takes "possible" as "certain"	-
	R	-	Quantifies probability in a three-stage case by using classical probability	-

Analysis

All the students' responses to the 24 items were assigned to 3 contexts: elements of a sample space (ELE), classical definition of probability (CLA), and subjective estimation of chance (SUB). Then the SOLO model was used to describe students' understanding of probability. Ikonic, concrete-symbolic, and formal levels' responses were observed in this study. The characteristics of the different SOLO levels are shown in Table 2. The responses under each context were labelled "ikonic" level (IK), "concrete-unistructural" level (C-U), "concrete-multistructual" level (C-M), "concrete-relational" level (C-R), "formal-unistructural" level (F-U), "formal-multistructual" level (F-M), or "formal-relational" level (F-R). If a student gave more than one reason to explain his/her choice, the response was assigned to the highest level. If several reasons led to different choices, the response was labelled with the highest level of the reasons that led to his/her final choice.

Results

1. Identification of the three kinds of events (impossible, possible and certain)

The results shown in Table 3 support the conclusions of Fischbein et al (1991) that "the majority of the subjects are able to identify the three kinds of events" and that "*the lowest rate of adequate answers...is that referring to certain events*" (p.527). Compared to "certain" and "possible", the term "impossible" seems easier than "certain", but more difficult than "possible". The correct percentages for the three packs of cards items were generally lower than the other items, due to the sophistication of the sample spaces.

Table 3. Correct Percentages of Choice to Each Item *

Items	G6	G8	G12	Items	G6	G8	G12
1A1 (p)	0.70	0.92	0.94	1A2 (c)	0.55	0.83	0.94
2A1	0.80	0.91	1.00	2A2	0.57	0.54	0.66
1B1	0.75	0.83	0.86	1B2	0.47	0.74	0.80
2B1	0.65	0.85	0.83	2B2	0.61	0.70	0.66
1A3 (i)	0.63	0.81	0.66	1A4 (p)	0.68	0.87	0.90
2A3	0.78	0.70	0.90	2A4	0.82	0.92	0.95
1B3	0.71	0.79	0.77	1B4	0.79	0.83	0.90
2B3	0.63	0.75	0.80	2B4	0.73	0.79	0.86
1A5 (p)	0.90	0.92	0.85	1A6 (p)	0.79	0.87	0.90
2A5	0.75	0.83	0.77	2A6	0.89	0.96	0.95
1B5	0.90	0.90	0.83	1B6	0.79	0.87	0.86
2B5	0.55	0.71	1.00	2B6	0.83	0.88	0.94

* approximately 20 students at each level answered each item

2. Selection of the three contexts

When answering "certain events" and "impossible events" items more than 75% of the students applied an ELE approach. In answering "possible events" items all three approaches, ELE, CLA and SUB were used. The three approaches were used almost equally in one-die items (1A), but CLA was used progressively less from one pack of cards and three dice (2A and 1B) items, to three packs of cards items (2B). The likelihood in the 24 items can be measured by classical probability, but the subjects were not asked to do so. However, some students, especially those from Grade 12, who had received some instruction in classical probability, applied this approach in their explanations. Some of the younger students were able to measure chance value correctly for one die or one pack of cards items, but as their intuitive knowledge of calculating a probability of a compound event was very limited they usually were unable to do so on three dice or three packs of cards items.

3. Cognitive levels in the explanations

Only about 3% students in this study gave ikonic responses. Some of them think the outcome of a trial is certain and can be controlled or predicted, while other students make judgements based on their own convictions or preferences. Most of the responses were classified at the U, M, or R of concrete symbolic mode.

4. Misconceptions

Seven types of misconceptions were observed in this study (see Table 4). All the misconceptions recorded by Fischbein et al. (1991) were observed, together with two others (Type 6 and 7) not reported in their research.

Type 1: Students take "rare" to mean "impossible", "highly frequent" to mean "certain"

This is the most common student misconception. Students with this misconception believe an outcome with a small chance is impossible and an outcome with a large chance is certain.

As Table 4 shows, this misconception declines with age, but 6 12^{th} grade students (who had had instruction in probability) still retain this misconception. Thirty-seven students used this approach once and 2 students used it twice. The authors speculate that this misconception could have arisen by students using "logical" thinking. Konold (1991) described students trying to predict an outcome for a single trial. They evaluate probability values in terms of their proximity to the anchor values of 100%, 0%, and 50%, which have the respective meanings of "yes", "no", and "I don't know". So if the likelihood of an outcome is very small, almost zero, they believe the outcome is impossible, especially for a single experiment.

Type 2: Students take "uncertain outcomes" to mean "possible events"

Students believe that since a trial has many possible outcomes, namely it is uncertain, that the target outcome is "possible", resulting in an incorrect answer when the target outcome is impossible or certain. There appears to be three causes for this confusion. Students

a) neglect the given conditions such as all the numbers on a die should from 1 to 6.
b) only consider one or several possible outcomes when making a judgement.
c) consider all outcomes but do not realise it is impossible that the target outcome will not happen.

According to the data, this misconception decreases with age, possibly because the thinking of older students is not so superficial and they are able to distinguish "uncertain outcomes" with "possible events".

Type 3: Students confuse "possible outcomes" with "certain events" and "impossible events"

This is another misconception confusing possible, certain and impossible. Although all these students know that several possible outcomes might occur (some of them even calculated the probability value of the target outcome first to explain that the outcome is possible) they do not chose "possible" as the answer. There are two kinds of explanations observed in the study.

a) students take "it will not always happen" to mean an "impossible event"

This misconception mainly existed in younger normal students and it disappeared for the better grade 8 students and grade 12 students.

b) students take "it will happen after enough trials" to mean a "certain event"

Misconception b) is very unusual as it mainly existed in better or older students (see Table 4). Five grade 12 students used it once and eight used it twice. Some students think that if you do an experiment once the outcome might not occur, but if you repeat the experiment many many times, the outcome is certain to happen. This is supported by the results of the interviews, where all three students with this misconception changed their choices from "certain" to "possible" after be reminded the target outcome was for one trial.

Type 4: Students believe all outcomes have an equal chance, even in situations where this is not true

Very few students misused "equiprobability" (see Table 4). This doesn't mean few Chinese students have this bias ("equiprobability" was observed in other parts of the study), rather it is an artefact of the items selected for this paper.

Type 5: Students decide on their subjective beliefs or personal likes

Students use their experience with games or their personnel beliefs to make judgements. There are two kinds of explanations used by the students. They

a) think the situation can be controlled to lead to a specific outcome or that the outcome can be predicted.
b) based their decisions on their own convictions or likes.

This kind of response is at the lowest cognitive level observed in this study. There were only 13 out of 567 students who made decisions in this way.

Type 6: Students interpret the likelihood of a compound event as a combination of the likelihood of its elements.

This type of misconception is referred to as a "compound approach". In this part of the study, the following explanation was used (other kinds of explanations occurred for other problems on the questionnaire)

a) suppose the likelihood of each of outcome 1, outcome 2 and outcome 3 is p%, then the likelihood of outcome 1, outcome 2 and outcome 3 occurring in combination is p%, too.

One girl from an advanced school explained her thinking in this way.

The authors believe that the "compound approach" is the result of logical thinking. Depending on the task, this approach may or may not lead to a wrong answer. Although there was only one example of a student using the compound approach inappropriately, there were other examples of the compound approach being used appropriately. For example, if all three outcomes are individually possible, then the combination of all three is possible. As with the situation of "equiprobability", there were only limited examples of this due to the nature of the items.

Type 7: Students misuse the "drawer principle" in predicting outcomes of trials

Two students used this approach, namely if there are n possible outcomes totally, in no more than $n+1$ trials, the target outcome will be certain. As with the "compound approach", this bias is caused by logical thinking.

5. Language Problem

There were eight 6^{th} grade students, four 8^{th} grade students and one 12^{th} grade student who felt it was difficult selected one of the given options when the situation was uncertain (see Table 4). No such difficulty was observed for impossible events and certain events.

Some of them choose both "impossible" and "possible" to express their uncertainty, while others left all three choices blank and wrote their own answer. They wrote "sometimes it occurs, sometimes it doesn't" or "uncertain".

Other students had a different problem with the language, usually leading to incorrect answers. For example, when students select the "impossible" option they interpret it as "impossible to predict".

Table 4. Misconceptions observed in this study*

Types of misconceptions	Grade 6		grade 8		grade 12	
	Normal (n=124)	advanced (n=50)	normal (n=108)	advanced (n=101)	normal (n=87)	advanced (n=97)
Type1	17	3	6	7	5	1
Type2 a)	11	3	2	3	1	1
b)	0	1	2	1	0	0
c)	2	2	4	1	0	0
Type3 a)	10	1	2	0	0	0
b)	5	1	1	6	5	8
Type4	2	1	0	0	1	0
Type5 a)	1	1	1	2	2	0
b)	1	0	1	2	2	0
Type6 a)	0	0	0	1	0	0
Type7	0	0	0	2	0	0
Feel difficult in expressing uncertainty	7	1	4	2	0	1

* These numbers are a lower bound. Other students may have these misconceptions, but they are not included since they made a choice without stating their reasons.

Discussion

The 6th and 8th grade students had no formal school instruction in probability. Their explanations show their intuitive understanding of the meaning of impossible, possible, and certain events. Their approaches reflect their own methods of dealing with uncertainty. As Table 3 shows, most of students are able to distinguish the three kinds of events in one die/one pack of cards tasks. They have most difficulty with "certain events", "impossible events", and multiple-setting tasks. The answers given by students in grade 12 are not as good as might have been expected. There are two implications for teaching. First, teaching students how to identify certain situations from uncertain situations needs to be an integral part of the programme, even at grade 12. Second, students need to be presented with single and multiple-stage tasks involving possible, impossible and certain events.

Misconception type 1, "rare" means "impossible", is the most common for each grade. Since it is related to the "outcome approach" - a stable, common misconception (Konold, 1989) - action needs to be taken to replace it with a correct interpretation.

Language problem should also be addressed within the programme. Words such as certain, possible, and impossible have specific meanings within probability that do not exactly parallel their real world use. Instruction needs to be cognisant of the fact that students may have predetermined meanings for these words, and it is necessary to incorporate this information into the classroom activities.

A more general comment, which is not specific to probability, is that some students had difficulty in writing explanations. Since being able to communicate ideas in both written and oral form is an important goal of mathematics education, students need experience with writing the reasons for an answer.

Some Chinese students, especially those from advanced schools, are used to solving difficult problems. Consequently, they did not believe the test items would be so easy, so they added some unusual conditions to the original problems. When faced with the task "the number rolled is greater than 6", rather than say "impossible" they invented conditions such as die split into two pieces and "7" was rolled, leading to a conclusion that it was "possible". This could be the reason for some of the incorrect responses. This possibility is reinforced by the fact that when interviewed some students changed to a correct answer. A final note is that other students appear to be influenced by outside factors such as TV, which lead to the conclusions such as the ability to control events.

References

Biggs, J.B., and Collis, K.F. (1991). Multimodal learning and the quality of intelligent behavior. In Rowe, H.A.H. (Ed.), Intelligence: Reconceptualization and Measurement (pp. 57-75).

Fischbein, E., Nello, M.S. and Marino, M.S. (1991). Factors affecting probabilistic judgements in children and adolescents. *Educational Studies in Mathematics*, 22, 523-549.

Green, D.R. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. In Grey, D.R., Holmes, P., Barnett, V. and Constable, G.M. (Eds.), *Proceedings of the First International Conference On Teaching Statistics, Volume 2.* (pp. 766-783).

Jones, G.A., Langrall, C.W. and Thornton, C.A. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics*, 32, 101-125.

Konold, C. (1989). Informal conceptions of probability. *Cognition and Instruction*, 6(1), 59-98.

Lecoutre, M. (1992). Cognitive models and problem spaces in "purely random" situations. *Educational Studies in Mathematics*, 23, 557-568.

Moritz, J.B., Watson, J.M. and Pereira-Mendoza, L. (1996). The language of statistical understanding: an investigation in two countries.

Shaughnessy, J.M. (1983). Misconceptions of probability; systematic and otherwise: Teaching probability and statistics so as to overcome some misconceptions. In Grey, D.R., Holmes, P., Barnett, V. and Constable, G.M. (Eds.), *Proceedings of the First International Conference On Teaching Statistics, Volume 2.* (pp. 784-801).

Williams, J.S. and Amir, G.S. (1995). 11-12 year old children's informal knowledge and its influence on their formal probabilistic reasoning. ERIC ED 387256.