

# Profiling Teacher Capacity in Statistical Thinking of National Curriculum Reform: A Comparative Study between Australia and China

Qinqiong Zhang

*Wenzhou University, CHINA*

Max Stephens

*The University of Melbourne, AUSTRALIA*

•Received 20 May 2015•Revised 10 August 2015 •Accepted 26 October 2015

In the official curriculum documents of many countries, statistical thinking have become part of the mainstream in school curriculum. We argue that teacher capacity is a key dimension in realizing essential goals for developing students' statistical literacy, reasoning and thinking in practical teaching. In this paper, a construct of Teacher Capacity was used to analyze how Australian and Chinese teachers understand and give effect to content in "Statistical thinking". The responses of the 82 teachers involved in the study to a questionnaire were analyzed qualitatively and quantitatively in terms of four criteria which form the basis of our construct of teacher capacity: Knowledge of Mathematics, Interpretation of the Intentions of the Official Mathematics Curriculum, Understanding of Students' Thinking, and Design of Teaching. These analyses gave rise to three classifications of Teacher Capacity: High, Medium and Low Capacity. Australian teachers performed slightly better on all four criteria than Chinese teachers, but there did not exist statistically significant difference. Among the four criteria, Design of Teaching appears to be the critical dimension for the implementation of curriculum reform.

*Keywords:* statistical thinking, teacher capacity, national curriculum reform, comparative study, Australia, China

## INTRODUCTION

In the official curriculum documents of many countries, statistics and statistical thinking have become part of the mainstream in school curriculum. In The Australian Curriculum: Mathematics (ACARA, 2010), "Statistics and Probability" is one of three key content areas. In its overview statement to this strand, ACARA (2010) states that: "Statistics and probability initially develop in parallel, and the curriculum then progressively builds the links between them. Students recognize

Correspondence: Qinqiong Zhang,  
College of Teacher Education, Wenzhou University, Zhejiang Province, 325035,  
People's Republic of China.  
E-mail: qqzhang922@126.com  
doi: 10.12973/eurasia.2016.1225a

and analyze data and draw inferences...They develop ... to critically evaluate chance and data concepts ...and develop intuitions about data” (p.2). A corresponding strand, Chance and Data has been present, for at least five years, in related State curriculum documents. E.g. VELs (VCAA, 2008) and the Mathematics Developmental Continuum (DEECD, 2006). China’s newly revised Mathematics Curriculum Standard for Compulsory Education (Ministry of Education of PRC , 2011) also presents a single strand entitled Statistics and Probability. In the overall objective for this content strand, it is stated that “to experience the process of collecting and dealing with data in practical problems, as well as using data to analyze questions and obtaining information” in “knowledge and skills” (p. 8); and it refers to “to experience the significance of statistical methods, to develop ideas of statistical analysis and to experience random phenomena” in “mathematical thinking” (p. 9).

These intentions are endorsed by Garfield & Ben-Zvi (2008) who point out that in contrast to traditional approaches to teaching which focus on computations of theoretical probability, new emphases are squarely focused on understanding data and development of statistical thinking and literacy (p. 7). They argue that “the goals for students at the elementary and secondary level tend to focus more on conceptual understanding and attainment of statistical literacy and thinking and less on learning a separate set of tools and procedures.” (p. 14). These goals are reflected in the National Curriculum in Australia and China, where students are expected to learn and understand that: (1) explanations supported by data are more powerful than personal opinions or anecdotes; (2) variability is natural and is also predictable and quantifiable; (3) association is not the same as causation; and (4) random sampling allows results of surveys and experiments to be extended to the population from which the sample was taken. (cf. Garfield & Ben-Zvi, 2008, p. 15).

However, the implementation of curriculum change is never simply from the top down. Teachers’ interpretations and responses at the level of practice are never simple reflections of what is contained in official curriculum documents. While curriculum documents set out broad directions for change, any successful implementation of these “big ideas” depends on teachers’ capacity to apply subtle interpretations and careful local adaptations (Datnow & Castellano, 2001). We argue that Teacher Capacity is a key dimension in realizing that goal. Meanwhile, both Australia and China published the newest professional standard for teachers that will be guidance for looking at teachers’ professional development (AITSL, 2011; Ministry of Education of PRC, 2012). As a result, research on teacher capacity will be with great significance for both teacher professionalization and curriculum implementation.

### **State of the literature**

- Shulman’s construct of teachers’ knowledge using Pedagogical Content Knowledge became the most cited literature for research in teacher knowledge. But PCK did not look at any specific subjects.
- Ball’s continuing work on PCK, using Mathematics Knowledge for Teaching focused on mathematics teachers and distinguished between Specialized Content Knowledge and Common Content Knowledge, but MKT did not value knowledge of official curriculum due to its American background. Researchers now recognize the importance of official curriculum.
- The term Teacher Capacity comes out of the literature of school improvement and curriculum reforms. It combines knowledge and disposition to act.

### **Contribution of this paper to the literature**

- This paper researches mathematics teachers considering the background of national curriculum reform in Australia and China. It identified teacher capacity as a key dimension in practical teaching of statistical thinking.
- Our construct of teacher capacity is based on four criteria, including knowledge of the intention of the official mathematics curriculum. And out of the four, Design of Teaching is the most significant dimension. This finding was consistent with an earlier parallel research on algebraic thinking.
- The results of this paper support that teacher capacity needs to be explored further focusing on specific contents in different subjects.

## RELEVANT RESEARCH OF TEACHER CAPACITY

While the term “teacher capacity” is not widely used in mathematics education research, it has clear connections with the research of “Pedagogical Content Knowledge” by Shulman (1986; 1987) and “Mathematical Knowledge for Teaching” by Ball, Thames & Phelps (2008).

### Shulman’s model

Shulman (1987) identified *pedagogical content knowledge* as the category most likely to distinguish the understanding of the content specialist from that of the expert teacher. The importance given to PCK suggests that what is needed in mathematics teaching is not just knowledge of the subject, or general knowledge of pedagogy, but rather a combination of both. However, after twenty five years of exposure to Shulman’s thinking, Petrou & Goulding (2011) conclude that: “Although Shulman’s work was ground-breaking and his ideas continue to influence the majority of research in the field, later researchers in the same tradition argue that it is not sufficiently developed to be operationalised in research on teacher knowledge and teacher education” (p12). We note that Shulman did not write specifically for mathematics teaching, but for all teaching subjects; and that his categories tend to reflect the educational context of the USA where there was no national curriculum.

### Michigan model

Ball et al. (2008), while sympathetic to Shulman, prefer to use the term Mathematical knowledge for teaching (MKT). Within this idea, they identify four constituent domains or categories: (1) Common content knowledge (CCK) defined as the mathematical knowledge and skill used in settings other than teaching; (2) Specialized content knowledge (SCK) as the mathematical knowledge and skill unique to teaching specific topics; (3) Knowledge of content and students (KCS) defined as knowledge that combines knowing about students and knowing about mathematics; and (4) Knowledge of content and teaching (KCT), which combines knowing about teaching and knowing about mathematics.

Among these four domains discussed by Ball et al. (2008), CCK is a primary component of mathematical knowledge, and needs to be combined with a teacher’s SCK, the subject matter knowledge needed for teaching specific mathematics content or topics. KCS and KCT are both intended to describe distinct knowledge for teaching. However, “content” used in the four categories may refer to: today’s worksheet, or this year’s textbook, or what is contained in official curriculum documents. In this sense, KCT may not be too far removed from Shulman’s category of Curriculum knowledge under which he includes teachers’ having a grasp of relevant materials and programs. While these knowledge domains are intended to anticipate classroom use, their instructional consequences are only implied. What is more, what appears to be a common feature of both Ball et al. (2008) and Shulman (1986; 1987) is a an interpretation of “curriculum” and “curriculum knowledge” which may be based too closely on their USA experience, where curriculum knowledge can be interpreted simply as “the particular grasp of the materials and programs that serve as ‘tools of the trade’ for teachers” (Shulman,1987, p. 8); and Ball et al. (2008) do not seem to have moved beyond this.

### Limitations of research on PCK and MKT

Ruthven (2011) has presented four distinct conceptualizations of Mathematical knowledge for teaching – *Subject knowledge differentiated*; *Subject knowledge contextualized*; *Subject knowledge interactivated* and *Subject knowledge mathematised* – each of which is intended to move forward debate about and

research – but in different directions. These four lines of thinking show that Mathematical knowledge for teaching is no longer a single unified idea. Researchers also need to be aware of the limitations of some or all of these four approaches: (1) all four have a strong focus on how to improve pre-service teachers' mathematical knowledge needed for their teaching in the future; (2) apart from the first framework adopted by Petrou & Goulding (2011), the other three do not appear to place a strong emphasis on the way in which official mathematics curriculum documents are intended to guide teaching in many countries; (3) apart from the first framework, the other three tend to view knowledge for teaching mathematics in general terms, rather than considering the specific areas of mathematical content important for curriculum reform; (4) all four theoretical frameworks are not easy to conceptualize into empirically conducted in research. Our own position on Mathematical knowledge for teaching is closest to that of Petrou & Goulding (2011). We use this framework to inform our construct of teacher capacity, and to show where it differs from that of Ball et al. (2008), especially in its stronger links to research on curriculum reform and school change.

### **Teacher capacity model**

The term “Teacher capacity” comes out of the literature of school improvement, school leadership and system reform (McDiarmid, 2006; Fullan, 2010). When used in this context, teacher capacity usually relates to teachers' ability to understand and act on the reforms that policy makers are seeking to implement (Spillane, 1999). It is close to our definition of Teacher Capacity as professionally informed judgment and disposition to act. Researchers such as Floden, Goetz & O'Day (1995) emphasize that teacher capacity is multidimensional and evolving. Firstly, they argue that teachers' ability to assist students in learning is dependent on teachers' own knowledge, which includes knowledge of the subject matter, knowledge of curriculum, knowledge about students and knowledge about general and subject-specific pedagogy; secondly, they argue that, while knowledge interacts with skills, there is a considerable gap between what teachers believe they should be doing in the classroom and their ability to teach in the desired ways; and thirdly, they point to the importance of dispositions, since enacting reform requires having the dispositions to meet new standards for student learning and to make the necessary changes in practice.

There are clear parallels here with Ball et al. (2008) who make the equally strong point that any definition of Mathematical knowledge for teaching (MKT) should begin with teaching, not teachers. Any such definition must be “concerned with the tasks involved in teaching and the mathematical demands of these tasks (our emphasis). Because teaching involves showing students how to solve problems, answering students' questions, and checking students' work, it demands an understanding of the content of the school curriculum” (p. 395).

## **METHODOLOGICAL POSITION**

### **Theoretical framework**

Our construct of Teacher Capacity, as professionally informed judgments and dispositions to act, is intended to capture a common ground between movements for school system and curriculum reform and the construct of Mathematical knowledge for teaching elaborated by Ball et al. (2008). Four criteria inform our theoretical model.

Criterion A – **Knowledge of Mathematics** – is intended to be applied to the tasks that the students have completed or are being asked to complete. Knowledge of

Mathematics is intended to capture the key mathematical ideas for teaching specific content.

Criterion B – **Interpretation of the Intentions of the Official Mathematics Curriculum** – is concerned with how teachers relate what is mandated or recommended in official curriculum documents of China and Australia to what their students are being asked to learn. This Criterion differs from MKT (Ball et al., 2008) in giving a greater emphasis to official curriculum documents and teachers' willingness to use them in planning instruction.

Criterion C – **Understanding of Students' Mathematical Thinking** – is directly concerned with teachers' capacity to interpret and differentiate between what students actually do (or did) and to anticipate what they are likely to do. It implies that teachers are able to recognize the typical errors that students make and what mathematical thinking led to these errors.

Consequently, Criterion D – **Design of Teaching** – places a clear emphasis on teachers' capacity to design teaching in order to move students' thinking forward and to respond to specific examples of students' thinking in the light of official curriculum documents. Criterion D is intended to give greater emphasis to how teachers use their professionally informed judgment to design practical teaching on specific topics.

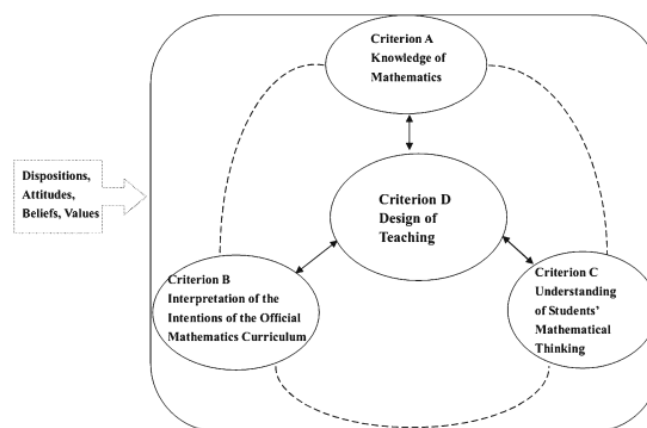
Among the four criteria above, criterion D is intended to give greater emphasis to how teachers use their professionally informed judgment to design practical teaching on specific topics. Our model of teacher capacity is shown in Figure 1 (Zhang & Stephens, 2013).

As shown in Figure 1, each of the four criteria is focused on teachers' professionally informed judgments and dispositions to act, distinguishing them from the four knowledge domains of Ball et al. (2008). Criterion D (design of teaching) is put forward as the central component in our model. As the instructional embodiment of teacher capacity, design of teaching rests on strong connections with the other three criteria. However, we also anticipate that, when separated from criterion D, the inter-relationships between the other three criteria will not be as strong as their relationship to criterion D (design of teaching) (Zhang & Stephens, 2013).

What is more, each criterion of above was elaborated in terms of four specific indicators (see Table 1).

### The research instrument

Teachers were invited to complete a written questionnaire consisting of two parts. Part A has four questions which were based on tasks developed in previous research, containing some situations relating to statistical thinking that students are expected to meet.



**Figure 1.** Model of Teacher Capacity

**Table 1. Four criteria and associated indicators**

Criterion A – Knowledge of Mathematics	Criterion C – Understanding of Students’ Mathematical Thinking
<p>(1) Is the teacher able to solve the theoretical mathematical probability problem (Q1a) and be able to understand relationship between chance of real events and sample size (Q3)?</p> <p>(2) Does the teacher consistently understand the variability of theoretical probability always happens in natural events in real life (Q1b, 1d), and the variability has a certain range close to the theoretical probability (Q1c, 1e, 1f)?</p> <p>(3) Does the teacher understand the meaning of “variability” by giving specific certain information (Q2)?</p> <p>(4) Does the teacher recognize that the difference between association and causation (Q4)?</p>	<p>(1) Is the teacher able to anticipate students’ common difficulties and misconceptions on Question 1 (e) in questionnaire?</p> <p>(2) Does the teacher give clear and reasonable explanations to students’ incorrect answers?</p> <p>(3) Is the teacher able to discriminate between students’ different levels of understanding statistics and probability according to their answers, especially discriminating between incorrect answers?</p> <p>(4) Does the teacher recognize the importance of using familiar contexts, such as coin tossing or rolling dice, to help students understand the statistical features of (less familiar) situations that contain similar statistical characteristics.</p>
Criterion B – Interpretation of the Intentions of Official Mathematics Curriculum	Criterion D – Design of Teaching
<p>(1) Does the teacher realize that “statistical thinking” should be valued in teaching and learning beyond the solutions of probability problems or does the teacher refer to relevant statements on statistical thinking in the official curriculum documents?</p> <p>(2) Does the teacher understand and support the intention of the curriculum of helping students understand key ideas of statistical thinking such as theoretical probability, sampling, sample size and inevitable variability in actual data, rather than calculating theoretical probabilities?</p> <p>(3) Does the teacher think it important to consider statistics and probability by linking natural events and real life?</p> <p>(4) Does the teacher show in his/her descriptions of developing students’ ability to read and understand data and information which is important for their further learning and future life?</p>	<p>(1) In design of teaching, does the teacher focus on the important key conceptions of statistical thinking (theoretical probability, sampling, sample size and inevitable variability in actual data, as well as using familiar contexts to simulate real world events), not focusing too much on general teaching strategies or overall descriptions on statistics and probability?</p> <p>(2) Does the teacher have the subsequent plan in next one or several lessons to respond students’ incorrect answers in Question 1 (e)?</p> <p>(3) Does the teacher have a longer-term plan to consistently develop students’ deep understanding of statistical thinking (see 1 above), not just aiming to have students correctly calculate theoretical probability problems?</p> <p>(4) Does the teacher, in his/her teaching, give concrete examples that are familiar and easy for students to understand to help them understand statistical thinking and its relationships with real life?</p>

Question 1 was adapted from Shaughnessy et al. (2004):

A gumball machine has 100 gumballs in it. 20 are yellow, 30 are blue, and 50 are red. The gumballs are all mixed up inside the machine.

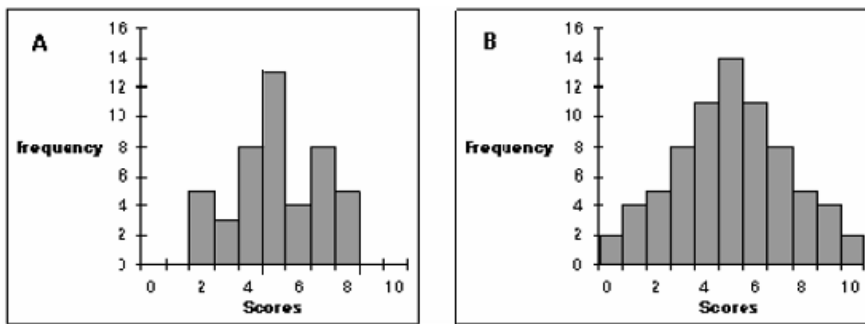
(a) Suppose you do the following experiment: you pick out a handful of 10 gumballs, count the reds and write down the number of red gumballs in one handful. How many reds do you expect to get?

(b) You replace the handful of 10 gumballs back in the machine and mix them up again. Now you draw another handful of 10 gumballs. Would you expect to get the same number of reds in every handful if you did this several times? Briefly describe why.

(c) How many reds would surprise you in a handful of ten? Why would that surprise you?

(d) If each time a handful of 10 gumballs is taken, these are replaced and remixed before taking another handful again, what do you think is likely to occur for the numbers of red gumballs that come out for a sequence of five handfuls? Please write the number of reds in each handful here.

(e) Look at these possibilities that some students have written down for the numbers they thought likely when they answered question d. Which one of these lists do you think best describes what is most likely to happen? Circle it. (A. 8,9,7,9,10; B. 3,7,5,8,5; C. 5,5,5,5,5; D. 2,4,3,4,3; E. 3,0,9,2,8; F. 7,7,7,7,7). Why do you think the list you chose best describes what is most likely to happen?



**Figure 2.** Two distributions

(f) In the above six repetitions of the experiment, what do you think will be the highest and lowest number of reds in one handful? Please discuss briefly why you think this.

Question 2 was adapted from Meletiou, & Lee (2002):

Look at the histogram of the two distributions shown in Figure 2.

Which of the two distributions you think has more variability? (a)

Distribution A (b) Distribution B

Briefly describe why you think this.

Question 3 was adapted from Garfield, & Gal (1999):

Half of all newborns are girls and half are boys. Hospital A records an average of 50 births a day. Hospital B records an average of 10 births a day. On a particular day, which hospital is more likely to record 80% or more female births?

(a) Hospital A (with 50 births a day); (b) Hospital B (with 10 births a day); (c) The two hospitals are equally likely to record such an event; (d) There is no basis for predicting which hospital would have that percentage of female births. Give a brief explanation of why you think like this.

Question 4 was adapted from Garfield, & Gal (1999):

For one month, 500 elementary students kept a daily record of the hours they spent watching television. The average number of hours per week spent watching television was 28. The researchers conducting the study also obtained report cards for each of the students. They found that the students who did well in school spent less time watching television than those students who did poorly.

Which of the following statements is (are) correct? (a) The sample of 500 is too small to permit drawing conclusions; (b) If a student decreased the amount of time spent watching television, his or her performance in school would improve; (c) Even though students who did well watched less television, this doesn't necessarily mean that watching television hurts school performance; (d) One month is not a long enough period of time to estimate how many hours the students really spend watching television; (e) The research demonstrates that watching television causes poorer performance in school; (f) I don't agree with any of these statements. For one statement that you agree with, explain why you think that way. For one statement that you disagree with, explain why you think that way.

Part B of the questionnaire had three questions which asked teachers to consider teaching implications arising from the questions in Part A. Specifically they were asked to consider common misunderstandings and difficulties for students in the Part A questions; how the key mathematical ideas or critical points presented in these questions are

addressed in their respective country's official curriculum documents; and how to design some lessons to help students to understand these key ideas.

### **The participants**

There were 17 Australian secondary and primary schools randomly selected in both urban and rural regions in Melbourne. Up to four teachers of Year 6 or Year 7 in each participating school were invited to complete the questionnaire. The Australian sample consisted of 41 Australian teachers, 28 secondary teachers and 13 from primary schools. The China sample comprised 41 teachers randomly selected from training programs in Chongqing, Hangzhou and Wenzhou. Twenty eight were secondary and 13 were primary teachers.

### **DATA ANALYSIS: QUALITATIVE**

The following examples provide evidence of Chinese and Australian teachers' (coded either as Teacher *n* CH or Teacher *n* AU) responses with respect to each of the four criteria. Teachers in both countries showed their different levels of understanding on all four criteria.

#### **Criterion A (Knowledge of Mathematics)**

Teacher 57 AU responded to Question 3 of Part A: "Hospital B is more likely to record the 80% as it has a much smaller population... Larger samples or more trials give results that are closer to theoretical probability." This teacher clearly demonstrated understanding of the relationship between sample size and variation from theoretical probability.

However, Teacher 31 AU answered: "Both hospitals are equally likely to record 80% female births because the probability is the same for each birth to be a boy or a girl." This teacher considered this problem as a completely theoretical probability question and did not realize variation exists and the sample size will influence the variation. And, Teacher 28 CH didn't identify the key point of this question by saying "it is random and no absolute result".

#### **Criterion B (Interpretation of the Intentions of Official Mathematics Curriculum)**

When referred to official mathematics curriculum, Teacher 35 CH said "This is the typical question representing thinking of probability and statistics. In the stage of primary school, statistics is more important. The main content of statistics is data processing, not to infer or guess with (theoretical) probability..."

Meanwhile, some teachers like Teacher 53 AU just listed several headings that are used in curriculum documents such as "measurement, chance and data" and some related ideas such as "calculating theoretical probabilities". And Teacher 24 CH referred to "including mathematical thinking of abstraction, transformation, modeling and etc.", but could not identify any specific mathematical thinking implied in the questionnaire.

#### **Criterion C (Understanding of Students' Mathematical Thinking)**

When teachers were required to comment on students' answers in Question 1(e) of Part A, Teacher 21 AU said "Students who choose C do not consider the variation but understand the basic principles of chance. Students who choose A, D, E and F, ... do not understand the basic principles of chance." This teacher gave reasonable explanations for each response of students and was able to discriminate students' different thinking level on statistics.



But these “typical” answers of students were confusing for some teachers. For example, Teacher 4 CH initially thought that “Students’ understanding is there are more red balls”, but then pointed that “Students will think all outcomes are possible, it’s difficult to judge”. This teacher did not understand the various misconceptions embedded in the “typical” answers.

### **Criterion D (Design of Teaching)**

When discussed on how to help students understand the critical mathematical thinking of Question 3 of Part A, Teacher 54 AU articulated teaching plans: “There are many activities that can be carried out using counters, coins and dice to simulate certain events. In the case of babies being born male or female, tossing a coin 10 times and recording Heads as female and Tails as male could be done. If every student performs the 10 tries, I would have enough data to compare and expect a good range including possibly 80% female. I could then compare individual trials of 10 to collective trials by putting together 5 groups of 10 results and comparing the male and female numbers and hopefully show that the results tend more to 50:50 female: male”, concluding “(one) would need to get across the idea that when an experiment is conducted many times over, certain patterns are likely to appear.” This teacher correctly focused on the critical points and designed very elaborate simulation – coin tossing – which is more familiar to students, not just talking about general teaching strategies.

Teacher 19 CH indicated that the teaching focus was about “statistical knowledge”, but offered no discussion of any specific statistical concepts, giving only very general teaching strategies like having “students conduct various kinds of experiments... they need practical manipulations to explore possibility”. Likewise, Teacher 2 focused only on “understanding of fractions, percentages and decimals. I would introduce whole numbers and equivalence and converting decimals to percentages.” This teacher referred only to some general strategies like “open ended questions including ratio of boys and girls.”

### **DATA ANALYSIS: QUANTITATIVE**

By assigning a score of 1 if one of the four indicators was evident in a teacher’s response, and 0 if it was omitted from their response or answered inappropriately, it was possible to construct a score of 0 to 4 for each criterion, and hence a maximum score of 16 across the four criteria. We allowed for the possibility that teachers might provide convincing alternative indicators to the four indicators listed in each criterion.

The two researchers operated independently to score teachers’ responses; then a careful confirmative check took place in order to resolve any difference. A high degree of consistency was present in the initial grading by the two graders, where, in less than 30 cases of 0/1 grading, only minor differences occurred. Any resulting differences in grading the 82 responses across the four criteria were easily resolved by consensus.

### **A summary for Chinese and Australian samples**

For the 41 Chinese teachers, the highest score was 14 and the lowest score was 3, with a median score of 8. For Australian teachers, the highest score was 15 and the lowest was 4, with a median score of 9. The respective mean scores were 8.34 (Chinese) and 9.27 (Australian) with standard deviations 2.70 and 2.63 respectively. Table 2 shows means and deviations that were calculated for each of the Criteria and total score. Of the four criteria, Criterion D (Design of Teaching) had the lowest mean (1.77) followed by Criterion B (Interpretation of the Official Mathematics

**Table 2.** Means for each criterion and global means and deviations

Sample	Criterion A	Criterion B	Criterion C	Criterion D	Total
Chinese(41)	2.42(0.67)	2.02(0.96)	2.27(0.71)	1.66(0.97)	8.37(2.70)
Australian(41)	2.72(0.86)	2.21(0.83)	2.49(0.76)	1.85(0.93)	9.26(2.63)
CH & AU(82)	2.56(0.82)	2.11(0.90)	2.38(0.76)	1.77(0.97)	8.82(2.79)

**Table 3.** Classifications of teacher capacity

Capacity	Chinese	Australian
High	7(17.1%)	8(19.5%)
Medium	26(63.4%)	26(63.4%)
Low	8(19.5%)	7(17.1%)

Curriculum) with the mean of 2.11, followed by Criterion C (Understanding of Students' Mathematical Thinking) with the mean of 2.38. Criterion A (Knowledge of Mathematics) had the highest mean at 2.56.

Australian teachers scored slightly higher on all four criteria than their Australian counterparts, but there was no statistically significant difference. On Criterion A, Australian teachers were clearer on the understanding critical concepts in statistics, especially in distinguishing variability from theoretical probability; on Criterion B, Chinese teachers paid more attention to methods to calculate possibility or chance that students need to learn, but Australian teachers were more focused on the development of how to deal with data in practical situations; on Criterion C, Australian teachers performed better on anticipating difficulties and misunderstandings that students might encounter; on Criterion D, Australian teachers were more likely to locate the key statistical idea in the hospital question in Part A, and could show in more practical ways how to develop related statistical thinking.

### Three classifications of Teacher Capacity

Three sub-categories of our construct of teacher capacity were created, with the boundaries set on the basis of the qualitative analysis of teachers' responses as discussed earlier. These were High capacity (score 11-16), Medium capacity (score 6-10) and Low capacity (score 0-5). These classifications using the two samples are shown in Table 3.

There were one more High Capacity teachers in Australian sample than in Chinese sample (respectively 8 and 7); one less Australian teachers were classified as Low Capacity than Chinese teachers (respectively 7 and 8). In both Chinese and Australian samples, Medium Capacity group was the biggest group which included 26 teachers out of 41, that was more than 60%.

High Capacity teaching of statistical thinking was evident in nearly 20% of Chinese and Australian teachers' responses to the questionnaire. It was shown by a clear understanding of the critical thinking and concepts in statistics of the four questions of Part A; relating the tasks to relevant curriculum documents; by high interpretative skills when applied to each of the six possible answers of students' work in Question 1(e); and by an extensive range of ideas for designing and implementing a teaching program to support the development of students' statistical thinking. Medium Capacity was shown by approximately 60% of teachers who, while possessing knowledge and skills supportive of these directions, clearly need to increase their current levels of professional knowledge and skills. In both samples, Low Capacity was evident in a minority of teachers – nearly 20% – who appeared unable to express a clear articulation of the mathematical nature of the tasks, or what differentiated the six students' answers in Question 1(e). These

**Table 4.** Bivariate correlations between each criterion and the total and any two of four criteria

		A	B	C	D	total
<b>A</b>	Pearson Correlation	1	.384**	.506**	.553**	.749**
	Sig. (2-tailed)		.000	.000	.000	.000
<b>B</b>	Pearson Correlation	.384**	1	.476**	.648**	.792**
	Sig. (2-tailed)	.000		.000	.000	.000
<b>C</b>	Pearson Correlation	.506**	.476**	1	.601**	.786**
	Sig. (2-tailed)	.000	.000		.000	.000
<b>D</b>	Pearson Correlation	.553**	.648**	.601**	1	.885**
	Sig. (2-tailed)	.000	.000	.000		.000
<b>total</b>	Pearson Correlation	.749**	.792**	.786**	.885**	1
	Sig. (2-tailed)	.000	.000	.000	.000	

\*\**. Correlation is significant at the 0.01 level (2-tailed).*

teachers were unable to point with any confidence to how the tasks related to what is contained in official curriculum documents, and found it difficult to describe how they would plan a program of teaching to foster students' statistical ideas.

### Interrelations of teacher capacity

Table 4 (calculated using SPSS 19.0 English version) shows the bivariate correlation between any two of the four criteria as well as that between each criterion and the total for the 82 Chinese and Australian samples.

First of all, all the four criteria have clear and strong contributions between .749 and .885 to the total score at the 0.01 level (2-tailed). And furthermore, all pairings of the four criteria have a significant correlation, supporting our theory that all four components of teacher capacity are interrelated. However, we cannot know whether the correlation between any two criteria was direct or was influenced by a third variable (Pallant 2001, p. 130). As a result, we statistically controlled one of the four criteria, then carried out four partial correlation analyses between any two of the other three criteria. When Criterion A (Knowledge of Mathematics) was controlled, the partial correlations between the other three criteria–B/C (.354), B/D (.567), and C/D (.447)–were all significant at 0.01 level. The result was the same when Criterion B (Interpretation of the Intentions of the Official Mathematics Curriculum) was controlled – A/C (.398), A/D (.433) and C/D (.437). When Criterion C (Understanding of Students' Mathematical Thinking) was controlled, the partial correlation between criteria A and D (.362) and that between criteria B and D (.515) were significant at the 0.01 level, only partial correlation between criteria A and B (.188) was the exception.

However, it showed very different results when Criterion D (Design of Teaching) was considered as the statistically controlled variable. We can see from Table 5, correlation between Criterion A and Criterion C was significant (.260) at 0.05 level, both of the rest pairs had no statistically significant correlations: A/B (.039), B/C (.142).

The results were slightly different from a previous research (Zhang & Stephens, 2013), but it was still obvious that after statistically controlling any one of the other three criteria, Criterion D still had significant and strongest correlation with the other two at 0.01 level. However, when Criterion D was statistically controlled, the relationship between any two of the other criteria became much weaker as shown in Table 5. A multi-step linear regression analysis confirmed the importance of Criterion D: On its own, Criterion D accounted for 88.5 % of the variance of our model. This also endorses our placing of Design of Teaching at the center of the

**Table 5.** Partial correlations (criterion D as controlled variable)

Control Variables			A	B	C
D	A	Correlation	1.000	.039	.260*
	B	Correlation	.039	1.000	.142
	C	Correlation	.260*	.142	1.000

\*. Correlation is significant at 0.05 level

model shown in Figure 1, where it is informed by the other three criteria, which are only weakly related when dissociated from Design of Teaching.

## CONCLUSIONS

Our construct of Teacher Capacity, presented here as teachers' professionally informed judgments and dispositions to act, connects to but differs from earlier research into Pedagogical Content Knowledge by Shulman (1986; 1987) and Mathematical knowledge for teaching by Ball et al. (2008). Here Teacher Capacity was investigated in terms of Knowledge of Mathematics, Interpretation of the Intentions of Official Curriculum documents, Understanding of Students' Thinking and Design of Teaching to foster the underlying mathematical ideas. Performance on each criterion was ascertained using a precise set of indicators that were related to the specific mathematical tasks, students expected thinking in relation to those tasks, the relationship between the tasks and official curriculum documents, and teachers' ability to design explicit teaching sequences to support students' learning.

Design of Teaching, informed by the other three criteria, appears to be the critical dimension for the implementation of curriculum reform; and the criterion that most clearly distinguishes between different levels of teachers' capacity to enact reform. Our construct of Teacher capacity strongly reflects the view that effective implementation of any curriculum reform depends on teachers' subtle interpretations of official curriculum documents and their professional dispositions to act on those ideas, which go well beyond general descriptions or statements of intent that are usually embodied in official curriculum advice.

Our construct of teacher capacity was built on earlier research into mathematical knowledge for teaching by MKT (Ball et al., 2008; Hill et al., 2008). However, to make the conclusions of this study compatible to their key idea of mathematical knowledge for teaching would appear to require a major reframing of their category, knowledge of content and teaching. In light of this study, that category appears to be too static and less suited to direct attention to design of teaching, which we have interpreted as understanding key aspects of national curriculum reform and knowing how to enact these aspects in practice. Our construct seems more suited to capture this key feature of curriculum reform.

## DISCUSSION

In an earlier parallel study of teacher capacity on algebraic thinking (Zhang & Stephens, 2013), Chinese teachers scored slightly higher than Australian on two criteria of Teacher Capacity, but in this study, they scored slightly lower on all four criteria, supporting our position that Teacher Capacity should not be considered generally, but specifically to different content areas in the curriculum. And what is more, in both studies, all four components of teacher capacity were effective in

distinguishing between different levels of capacity, but Design of Teaching was the most powerful in distinguishing between the three levels.

The above findings had clear implications for our further research into teacher capacity: First, we argued that teacher capacity should not be discussed generally, but specific content focused. So, what will be the connections of one teacher's capacity on different mathematical content? From several teachers who were involved in both studies, we knew that teacher capacity were not necessarily completely the same. But there still existed certain relations. For example, one teacher with high capacity (scored 15) in algebraic thinking having high-medium capacity in statistical thinking (scored 11), and another teacher who scored high in statistical thinking (scored 14) was just getting a medium score (9) in algebraic thinking. As a result, we are confident to have the assumption that one teacher with high capacity on one mathematical content area might possibly have medium capacity in another, but it might not be very likely for a teacher with high capacity in one content area but having low capacity in another area. Second, as we argued in the beginning of this paper, teacher capacity is a key dimension in realizing intentions of the curriculum reform and implementing these goals into practical teaching in mathematics classrooms. So, how do we identify the different ways in which High Capacity teachers design teaching and how do we use these findings to improve the capacity of Medium and Low Capacity teachers? And third, this is a comparative study, so what are the affects that "culture" possibly brings?

## ACKNOWLEDGEMENT

This study was supported by 2013 Youth Fund Projects of Humanities and Social Science Research of Ministry of Education, People's Republic of China (Grant No. 13YJC880114), 2013 General Financial Grant from the 54<sup>th</sup> Session China Postdoctoral Science Foundation (Grant No. 2013M541689), and 2013 General Financial Grant from the Jiangsu Province Postdoctoral Science Foundation (Grant No. 1301129C).

## REFERENCES

- Australian Curriculum and Reporting Authority (2010). *The Australian Curriculum: Mathematics*. Sydney: Author.
- Australian Institute for Teaching and School Leadership (AITSL). (2011). *National professional standards for teachers*. Melbourne: Author.
- Ball, D., Thames, M., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389-407.
- Datnow, A. & Castellano, M. (2001). Teachers' responses to success for all: how beliefs, experiences and adaptations shape curriculum. *American Educational Research Journal*, 37, 775-800.
- Department of Education and Early Childhood Development (2006). *Mathematics developmental continuum*. Melbourne: author.
- Floden, R., Goertz, M. & O'Day, J. (1995). *Capacity building in systemic reform*. *Phi Delta Kappan*, 77(1), 19-21.
- Fullan, M. (2010). *All Systems Go: The change imperative for whole system reform*. SAGA Publications Inc.
- Garfield, J. B., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning*. Milton Keynes, England. Springer.
- Garfield, J. & Gal, I. (1999). Assessment and statistics education: Current challenges and directions. *International Statistical Review*, 67, 1-12.
- McDiarmid, B. (2006). *Rethinking teacher capacity*. Available at: <http://scimath.unl.edu/MIM/rew2006/Powerpoints/REW06McDiarmid.ppt>.
- Meletiou, M. & Lee, C. (2002). Teaching students the stochastic nature of statistical concepts in an introductory statistics course. *Statistical Education Research Journal*, 1(2), 22-37.

- Ministry of Education of PRC (2011). *Mathematics curriculum standards for compulsory education (2011 version)* (in Chinese). Beijing: Beijing Normal University Press.
- Ministry of Education of PRC (2012). *Notification of Publication of "Kindergarten Teacher Professional Standard", "Primary Teacher Professional Standard" and "Secondary Teacher Professional Standard"* (in Chinese). Available at: [http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s6991/201212/xxgk\\_145603.html](http://www.moe.gov.cn/publicfiles/business/htmlfiles/moe/s6991/201212/xxgk_145603.html).
- Pallant, J. (2001). *SPSS survival manual*. Crows Nest NSW: Allen & Unwin.
- Petrou, M. & Goulding, M. (2011). Conceptualising teachers' mathematical knowledge in teaching. In T. Rowland. & K. Ruthven (Eds.) *Mathematical knowledge in teaching* (pp. 9-25). Dordrecht: Springer.
- Ruthven, K. (2011). Conceptualising mathematical knowledge in teaching. In T. Rowland. & K. Ruthven (Eds.) *Mathematical knowledge in teaching* (83-98). Dordrecht: Springer.
- Shaughnessy, Canada, & Ciancetta. (2004). Types of student reasoning on sampling tasks. In B. Philips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics: Developing a Statistically Literate Society*. Voorburg: International Statistical Institute.
- Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-22.
- Spillane, J. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, 31, 143-175.
- Victorian Curriculum and Assessment Authority (2008). *Victorian essential learning standards (Mathematics)*. Melbourne: author.
- Zhang, Q., & Stephens, M. (2013). Utilising a construct of teacher capacity to examine national curriculum reform in mathematics. *Mathematics Education Research Journal*. 25(3), 481 – 502.

