Educational Studies

ISSN: 0305-5698 (Print) 1465-3400 (Online) Journal homepage: http://www.tandfonline.com/loi/ceds20

Performance of sixth graders in Hong Kong on a number sense three-tier test

Ka Luen Cheung & Der-Ching Yang

EDUCATIONAL STUDIES
https://doi.org/10.1080/03055698.2018.1516631

Performance of sixth graders in Hong Kong on a number sense three-tier test

Ka Luen Cheung[a] and Der-Ching Yang[b]

[a]Department of Mathematics and Information Technology, The Education University of Hong Kong, Tai Po, New Territories, Hong Kong; [b]Graduate Institute of Mathematics and Science Education, National Chiayi University, Chiayi, Taiwan

ABSTRACT

This study examined the performance of sixth graders in Hong Kong on a number sense three-tier test (NSTTT) and identified their possible misconceptions about number sense. The NSTTT comprises a content tier, which assesses the content knowledge of number sense; a reason tier, which assesses the reasons for the first-tier responses; and a confidence tier, which assesses students' confidence in their responses to questions in the first two tiers. The study participants were 942 sixth graders in Hong Kong. The results indicated that the percentage of correct answers (and confidence indices) of the first two tiers (i.e. the content tier and reason tier) of the sixth graders was 53.96% (4.27). In addition, the sixth graders showed the highest performance on "ability to recognise number size" and the lowest on "ability to judge the reasonableness of computational results". Moreover, the results revealed that some students showed a low performance on number sense but demonstrated extremely high confidence levels, indicating that they may have significant misconceptions regarding number sense and may lack of number sense. Furthermore, this paper discusses the educational implications of the findings.

ARTICLE HISTORY
Received 29 March 2018
Accepted 23 August 2018

KEYWORDS
Hong Kong; number sense; sixth grade; three-tier test

Introduction

Number sense is considered to be the key ingredient and a major topic in the mathematics curriculum (National Council of Teachers of Mathematics [NCTM] 2000; Nickerson and Whitacre 2010; Sengul and Gulbagci 2012, 2014; Lin, Yang, and Li 2016; Yang and Li 2017; Yang, Yu and Lin, 2016; Verschaffel, Greer, and De Corte 2007). Due to the importance of number sense in mathematics, this topic has attracted a growing amount of attention and research worldwide (McIntosh et al. 1997; NCTM 2000; Sengul and Gulbagci 2012; Lin et al. 2016; Yang et al 2016; Verschaffel et al. 2007). Correspondingly, the importance of number sense can be seen from various ministries of education and bodies all around the globe that are striving hard to create a reformation in learning and instruction of mathematics with the emphasis of number sense component (Johnny and Mohamed 2010; NCTM 2000). For instance, the reforms of the school mathematics curriculum in Hong Kong have emphasised

CONTACT Der-Ching Yang dcyang@mail.ncyu.edu.tw Graduate Institute of Mathematics and Science Education, National Chiayi University, 85, Wen Lung, Ming-Hsiung, Chiayi 621, Taiwan

the importance of developing teaching and learning of number sense, in which the reforms highlighted the needs for teachers to provide instructions and activities that lead to conceptual understanding for students to develop and foster number sense (The Hong Kong Curriculum Development Council 2000, 2014, 2017).

Analysing students' performance on number sense is crucial and critical to the teacher to discover students' understanding of number sense component (Akkaya 2016; Li and Yang 2010; Şengül and Gülbağcı 2014; Yang and Li 2013). The results of assessing performance on number sense could be used and considered to provide the early predictors and supports for interventions in teaching and learning number sense before children fall seriously behind in school in mathematics (Gersten, Jordan, and Flojo 2005; Jordan, Glutting, and Ramineni 2010). However, there is no study in the literature assessing the performance of elementary students in Hong Kong (particularly, the sixth graders). On the other hand, it is important to have knowledge measures that can diagnose misconceptions in order to accurately assess students' performance and learning (Durkin and Rittle-Johnson 2015; McNeil and Alibali 2005; Resnick et al. 1989). A misconception can persist over a long period of time and become entrenched and adverse to change (Eryilmaz 2002; McNeil and Alibali 2005). Thus, the misconceptions should be diagnosed earlier in order to develop and provide teaching strategies and instructions for developing better conceptual knowledge for students, particularly on number sense. Therefore, the authors believe that it would be beneficial for the teacher and researcher to assess the students' performance and misconceptions on number sense in order to address a preliminary strategy and teaching development involving number sense.

Students in Hong Kong consistently showed a high performance on international mathematics assessment tests such as the 2015 Trends in International Mathematics and Science Study (TIMSS), in which the eighth and fourth graders in Hong Kong ranked fourth and second among 39 and 49 countries respectively (Mullis et al. 2016) and the 2015 Programme for International Student Assessment (PISA) reported that Hong Kong ranked ninth among 72 participating countries and economies (Organization for Economic Co-operation & Development (OECD) 2016). The previous studies have shown that number sense is an accurate predictor of future mathematics achievement (Jordan et al. 2010) and that a moderate relationship exists between number sense and mathematics achievement (Yang, Li, and Lin 2008). However, to the best of our knowledge and review of the literature, only one study has examined the number sense performance in Hong Kong which assessed the performance of preschool children (approximately 6 years old) (Aunio et al. 2004) and reported that children in Hong Kong outperform the children in Finland on number sense (particularly on relational and counting skill tasks), and the students in Singapore perform better than in Hong Kong. These issues emerged the motivation to conduct the current study. Accordingly, this study would shed light on the students' performance and misconceptions of sixth graders on number sense in Hong Kong.

This study was aimed at investigating the students' performance in Hong Kong on number sense and scrutinising the students' misconceptions on number sense development. We examined the number sense performance and confidence index of students in Hong Kong who were aged approximately 10–12 years. The research questions were as follows:

(1) How do sixth graders in Hong Kong perform on the number sense three-tier test (NSTTT) and how confident they are in giving answers and reasons for their answers in the test?
(2) What is the distribution of sixth graders in Hong Kong across number sense performance results and confidence levels?
(3) What are the prevalence of and the difference in students' misconceptions about number sense?

Background

Number sense and related studies

Number sense is such an important topic that many researchers have discussed it globally (Berch 2005; Markovits and Sowder 1994; McIntosh, Reys, and Reys 1992; McIntosh et al. 1997; Menon 2004; Reys and Yang 1998; Yang et al. 2008; Verschaffel et al. 2007). Lin et al. (2016) defined number sense as "a person's general understanding of numbers and operations and the ability to handle daily life situations involving numbers, including the use of flexible and efficient strategies." Students with nifty number sense tend to exhibit sense-making approach, planning and control, and flexibility and appropriateness sense of reasonableness, when performing mental computations (Johnny and Mohamed 2010). In the present study, number sense components were defined based on the definitions in previous studies (Lin et al. 2016) and were as follows: (C1) ability to understand the basic meaning of numbers and operations, (C2) ability to recognise number size, (C3) ability to use multiple representations of numbers and operations, (C4) ability to recognise the relative effects of operations on numbers and (C5) ability to judge the reasonableness of computational results.

Number sense plays a key role in our daily lives (Dehaene 1997; Li and Yang 2010) and in mathematics education (Berch 2005; Dunphy 2007; Griffin 2004; McIntosh, Reys, and Reys 1992; National Council of Teachers of Mathematics [NCTM] 2000; Sowder 1992; Yang and Li 2013). Many international studies have confirmed that helping children develop number sense is crucial (Berch 2005; Dunphy 2007; Li and Yang 2010; Sood and Jitendra 2007; McIntosh et al. 1997; Yang and Li 2013; Verschaffel et al. 2007). The lack of number sense often leads to mathematics learning disabilities (Jordan et al. 2007; Dyson, Jordan, and Glutting 2013). Students' ability to exercise number sense flexibly enables the long-term retention of acquired knowledge (Jordan et al. 2010; Yang et al. 2008).

In the past two decades, numerous studies on number sense have been conducted globally (Dunphy 2007; Jordan et al. 2010; Li and Yang 2010; Lin et al 2016; Markovits and Sowder 1994; Menon 2004; McIntosh et al. 1997; Reys and Yang 1998; Sowder 1992). These studies have used instruments including paper-and-pencil tests, interviews and web-based two-tier tests. For example, Reys and Yang (1998) showed that the performance of sixth and eighth graders in Taiwan was significantly higher on written computation than on similar questions related to number sense. The findings revealed that the sixth and eighth graders in Taiwan use few number sense-based methods (e.g. the use of benchmarks) when responding to number sense-related questions. In addition, Menon (2004) indicated that students' number sense tends to decrease as they progressed through the grades. Menon (2004) also indicated an almost complete absence of estimation strategy and an increasing reliance on written computation as students progressed from grade four to grade seven.

Two-tier and three-tier-related studies

Li and Yang (2010) and Lin et al. (2016) designed a number sense web-based two-tier test to assess the number sense performance of elementary school children. The first tier of this test assesses the children's responses to number sense-related questions, and the second-tier assesses their reasons for their choices in the first tier. The given reasons for each answer tier were collected from the students in the two past decades through paper-and-pencil data collection and interview of earlier studies (see Li and Yang 2010; Lin et al 2016; Reys and Yang 1998; Yang 2005, Yang et al 2008; Yang and Li 2017). The reasons tier were developed and created by analysing the students' responses in all data collection of the earlier studies. Therefore, the authors only provided the relevant and compatible reasons for students' choice in the reason-tier for each answer-tier. In the test design, students could only see the set of reasons that corresponded to their choice in the answer tier (usually two or three different reasons) and these corresponding options were based on students' most frequent misconceptions (Lin et al. 2016; Yang and Li 2017). Students were forced to focus on these compatible options in the reason tier instead of posting their own reason. The authors believed that allowing students to focus on fewer but more relevant and compatible would provide more meaningful results.

The two-tier test provides the quantitative advantage of being able to collect a large amount of data on students' number sense with the limited consumption of time, paper and effort. It provides the qualitative advantage of examining students' explanations for their answer choices and assessing the possible causes of their misconceptions at the same time (Lin et al. 2016; Yang and Li 2017). Lin et al. (2016) revealed that sample students do not perform as well on the number sense web-based two-tier test as they do on their school tests. Approximately 45% of these students provided correct answers to questions in the answer tier, and approximately 23% applied number sense-based methods to answer questions in the reason tier. A high percentage of students had misconceptions about number sense, probably due to the limited opportunities offered to learn about number sense at school (Yang et al. 2008). In addition, the two-tier test integrates the advantages of quantitative and qualitative methods (Lin et al. 2016; Yang and Li 2017), implying that it can be used to examine students' number sense performance and to identify related misconceptions.

The number sense two-tier (i.e. answer and reason tiers) test did not measure students' confidence in their answers. Answering with a low level of confidence indicates a lack of knowledge, regardless of whether the answer is correct or incorrect (Caleon and Subramaniam 2010; Clement, Brown, and Zietsman 1989; Stankov and Crawford 1997). Answering correctly with a high level of confidence indicates a thorough understanding of the relevant concepts (Caleon and Subramaniam 2010). By contrast, answering incorrectly with a high level of confidence indicates the existence of misconceptions about the subject (Caleon and Subramaniam 2010). The number sense two-tier test can be further strengthened by including self-ratings of students' confidence in their answers and the reasons for their answers. The confidence tier provides a measure of students' confidence in their responses to items. Additionally, the web-based NSTTT designed by Yang and Li (2017) demonstrated high reliability and validity. The results of the present study revealed that some sample students showed a low performance on number sense but exhibited extremely high confidence, indicating that these students may have significant misconceptions and may lack number sense. In addition, this study showed that a third tier (i.e. the confidence rating) could be added to the two-tier test to mitigate its weaknesses.

Methods

Samples

We selected 942 sixth graders (approximately 11–12 years old) in Hong Kong to assess their performance on the NSTTT. Students were selected with roughly equal portions from three categories (high, medium and low) of primary schools. In fact, there are three types of secondary schools in Hong Kong: bands 1–3. High: primary school has 50% or higher portion of students who can promote to band 1 secondary schools in the recent 3 years (up to 2015). Medium: primary school students have 30–50% students who can promote to band 2 secondary schools in the recent 3 years. Low: primary school students have less than 30% students who can promote to band 3 secondary schools in the recent 3 years. In addition, students were also selected from a broad range of family backgrounds, covering various parent occupations, incomes and educational levels.

Instrument

An example of the NSTTT is illustrated in Figure 1 (instrument). The NSTTT comprises three tiers: the first tier (i.e. the answer tier) assesses the content knowledge of number sense, the second tier (i.e. the reason tier) assesses the reasons for the first-tier responses and the third tier (i.e. the confidence tier) assesses sample students' confidence in their responses to questions in the first and second tiers (Lin et al 2016; Yang and Li 2017). For more specifically, the first tier measures students' answers to number sense-related questions (i.e. an assessment of students' understanding of relevant content knowledge). The second tier assesses students' reasons for their choices in the first tier (i.e. an examination of students' explanatory knowledge). The third tier assesses students' certainty about their responses in the first and second tiers (i.e. an assessment of students' degree of certainty in their answers and reasoning). The NSTTT contains five number sense components, each of which contains eight items. Thus, the NSTTT comprises a total of 40 items. In addition, the third tier of the NSTTT is rated using a 5-point Likert scale, ranging from very confident to very unconfident; this tier examines students' confidence in their responses to questions in the first and second tiers. This produces ratings of students' confidence in their answers to the questions in the content and rationale tiers of the test (Yang and Li 2017).

Treatment of data

Following previous studies on the number sense two-tier test (Li and Yang 2010; Lin et al. 2016) and three-tier test (Caleon and Subramaniam 2010; Cetin-Dindar and Geban

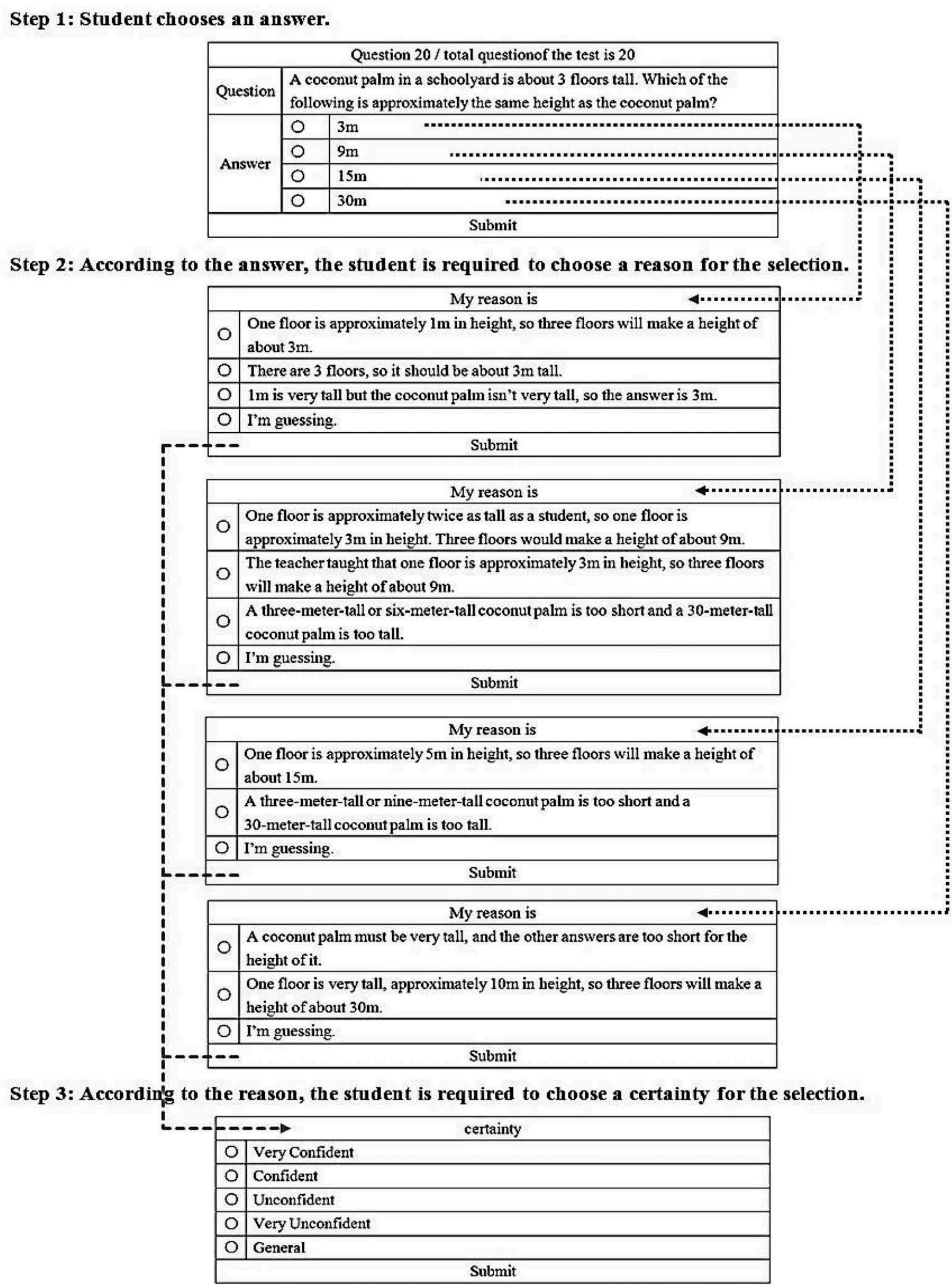


Figure 1. Example of the NSTTT; adapted from the author's previous study (Yang and Li 2017).

2011; Pesman and Eryilmaz 2010; Yang and Li 2017), the scoring rules were defined, as illustrated in Table 1. Yang and Li (2017) categorised students' number sense performance (defined as their performance on the first and second tiers) into four levels: (1) high number sense, defined as a mean score of 6 and higher (the highest score was 8); (2) high–medium number sense, defined as a mean score of less than 6 but equal to or

Table 1. Three-tier test scoring.

The first and second tiers					
Answer options		Correct answer 4 points			Wrong answer 0
Reason options	NS-based 4	Rule-based 2	Misconception 1	Guessing 0	0
Score given	8	6	5	4	0
The third tier					
CRI	Very confident	Confident	Neutral	Unconfident	Very unconfident
Score given	5	4	3	2	1

Note: the three-tier test scoring was adapted from the author's prior work (Yang and Li 2017).

greater than 4.8; (3) medium–low number sense, defined as a mean score of less than 4.8 but equal or higher than 3 and (4) low number sense, defined as a mean score of less than 3.

In addition, students' third-tier results (i.e. the confidence index) were categorised into three levels: (1) high confidence, defined as a mean score of 3.3 and higher; (2) medium confidence, defined as a mean score of 2.9–3.3 and (3) low confidence, defined as a mean score of 2.9 and lower. Previous studies have suggested that high scores for number sense with high confidence ratings indicate a thorough understanding of the relevant concepts, whereas low scores for number sense with low confidence ratings imply a lack of knowledge of the relevant concepts (Hasan, Bagayoko, and Kelley 1999; Pesman and Eryilmaz 2010). In addition, these studies have reported that low scores on the first two tiers coupled with high confidence indicate the existence of misconceptions about the relevant concepts. Moreover, high scores for number sense coupled with low confidence have been interpreted from two perspectives: (1) the results reflected students' personalities and (2) the test answers consisted of a preponderance of guesses (Hasan, Bagayoko, and Kelley 1999; Pesman and Eryilmaz 2010).

In this study, the existence of significant misconceptions about number sense among students was evaluated by examining their related responses. Based on previous studies (Caleon and Subramaniam 2010; Tan et al. 2002), incorrect response rates on the first and second tiers that exceeded 18.3% (which comprised 12% of total first tier results and 10% of total second tier results) were defined as significant misconceptions. According to Caleon and Subramaniam (2010), if the mean confidence scores associated with significant misconceptions were less than 2.9 (out of 5), then the significant misconceptions were defined as spurious misconceptions. If the mean confidence ratings associated with the significant misconceptions were more than 2.9 (out of 5), then the significant misconceptions were defined as genuine misconceptions. The genuine misconceptions included moderate and strong misconceptions. Mean confidence ratings of 2.9–3.3 were defined as moderate misconceptions, whereas mean confidence ratings of 3.3 or higher were defined as strong misconceptions. This classification helped us determine whether sample students had real misconceptions or a lack of number sense.

Procedure

For data collection, students completed an online three-tier test that was divided into 2 parts, each of which included 20 questions. Each question contained an answer selection, reasons selection and confidence selection, and the time limits for answering were 40, 60 and 20 s, respectively. Thus, each question required 120 s. Therefore, completing each part of the online test required approximately 45 min. In addition, regarding how did the author assign the time to these three categories of students, we explained it from three different perspectives:

(1) The research project aimed at the studying of number sense of primary 6 students in Hong Kong, the research team would like to cover the whole spectrum of the schools in Hong Kong, based on their academic performance, in the limited samples that we can take. Therefore, we selected sample schools from the high, medium and low categories according to the common classification systems in Hong Kong which should be an appropriate selection approach that can cater the needs of the research.
(2) The assigning time was the same for these three categories of students. The details about the time allocation for the test questions have been explained in the paper. Based on previous studies of the online test, the given duration for the test is reasonable and appropriate for primary 6 students. Actually, no complaints about the assigned time for the online test were received during or after the test conducted in Hong Kong. Majority of the students can finished the test within the given time limit. Also, we think "uniform duration of time" for these three categories of students should be the right regulation to keep the fairness to them.
(3) Finally, we would like to point out that our test questions are not "computational in nature" but putting emphasis on the "conceptual understanding and application" of number sense of students as well as their "confidence levels" in giving responses to the questions. The given time will not affect much on students' responses as not too many computational and procedural steps needed to be done before making the responses to the questions. The given time should be enough for a student to think about on what they should choose to respond for the questions based on the previous studies.

Reliability and validity

The authors' previous study (Yang and Li 2017) reported that the Cronbach's α coefficients of the three-tier test for each number sense component and the entire test are 0.856, 0.840, 0.873, 0.857, 0.823 and 0.902. The results of the SEM-based construct reliability of the three-tier test for each component and the whole entire are 0.827, 0.836, 0.840, 0.828, 0.839 and 0.905. This shows that the NSTTT has good reliability.

The NSTTT was developed by the authors based on the earlier studies of the two-tier test (Li and Yang 2010; Yang et al. 2008). The construction of the questions in the test was based on the five-component number sense framework of earlier studies (Li and Yang 2010; Yang et al. 2008). The choices of reason-tier for each question were collected from earlier studies by paper-and-pencil tests, open-ended questions and interviews (e.g. Reys and Yang 1998; Yang 2005; Yang et al. 2008). To determine whether students could fully understand each question and whether the questions used in the test were appropriate, 20 students were initially interviewed for content clarity. Furthermore, we revised several questions to rectify their ambiguous contents and make these questions sufficiently clear for sixth-grade students. To ensure the questions and answers were representative and not beyond the curriculum scope for sixth-grade students, two mathematics researchers and three elementary school teachers were invited to review the questions. They unanimously agreed that the NSTTT was effectively designed, reflected number sense and was not beyond the curriculum scope for sixth-grade students (Yang and Li 2017). This shows that the test has good content validity.

Results

Hong Kong sixth graders' performance on NSTTT

Table 2 illustrates the descriptive statistics related to Hong Kong sixth graders' performance on the NSTTT. The results revealed that the students of both grades performed the highest on "ability to recognise number size [C2]" (62.19% of sixth graders), with the highest mean confidence score 4.41 out of 5 for sixth graders. In addition, the students performed the lowest on "ability to judge the reasonableness of computational results [C5]" (43.06% of sixth graders), with the lowest mean confidence score (4.17 out of 5 for sixth graders). The results also illustrated that the percentage of total correct answers (mean confidence score) for sixth graders was 55.34% (4.27 out of 5).

Among the sixth graders, the percentages of total correct answers to number sense-related questions were lower than their mathematics achievement in school and international mathematics assessments (e.g. PISA and TIMSS). In addition, the mean confidence scores for each component and the entire test were more than 4, implying that these students were highly confident in their responses to number sense-related questions.

Distribution of sixth graders in Hong Kong across number sense performance results and confidence levels

Table 3 illustrates the results of sixth graders' number sense performance and confidence across different levels. The results revealed that 129 (13.7% of the total) sixth graders exhibited both high number sense and high confidence. Thus, these students have a thorough understanding of number sense because they exhibited both high performance on number sense-related questions and high confidence in their answers to the questions. The results also

Table 2. Descriptive statistics of sixth graders' performance on NSTTT (N = 942).

Component	Number of item	Mean of both tiers (Correct %)	SD	Mean of confidence (SD)
		6th	6th	6th
C1	8	35.89 (56.08%)	14.27	4.30 (1.14)
C2	8	39.80 (62.19%)	12.05	4.41 (1.07)
C3	8	36.92 (57.69%)	15.5	4.22 (1.90)
C4	8	32.51 (50.79%)	12.4	4.27 (1.17)
C5	8	27.56 (43.06%)	12.33	4.17 (1.22)
Total score		172.68 (53.96%)	55.34	4.27 (1.16)

Note. The highest score is 8 for each item of both tiers. C1: Ability to understand the basic meaning of numbers and operations; C2: ability to recognise number size; C3: ability to use multiple representations of numbers and operations; C4: ability to recognise the relative effects of operations on numbers; C5: ability to judge the reasonableness of computational results.

Table 3. Distribution of students across number sense performance results and confidence levels.

Distribution of students	High confidence	Medium confidence	Low confidence	Total
High number sense	129	3	1	133
High–medium number sense	236	7	5	248
Medium–low number sense	347	20	17	384
Low number sense	148	13	16	177
Total	860	43	39	942

Note. High number sense: 6 and higher; high–medium number sense less than 6 and higher than or equal to 4.8; medium–low number sense: less than 4.8 and higher than or equal to 3; low number sense: <3. High confidence: 3.3 and higher; medium confidence: 2.9–3.3; low confidence: 2.9 and lower.

revealed that 16 (1.7% of the total) sixth graders exhibited low number sense and low confidence, implying that these students lack number sense. Relevant learning opportunities, such as number sense-related courses, should be provided to these students.

In addition, 148 (15.7% of the total) sixth graders exhibited a low number sense and high confidence, implying that these students have strong misconceptions about number sense-related concepts. Moreover, 347 (36.8% of the total) sixth graders exhibited medium–low number sense and high confidence, implying that they also have strong misconceptions about number sense-related concepts.

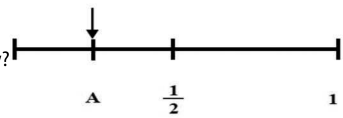
Prevalence of and difference in students' misconceptions on number sense

According to previous studies (Caleon and Subramaniam 2010; Tan et al. 2002; Yang and Li 2017), the incorrect response rates for the first and second tiers that exceeded 18.3% (which comprised 12% of total first tier results and 10% of total second tier results) were defined as significant misconceptions. In addition, the mean confidence scores higher than 3.3 (associated with significant misconceptions) were defined as strong misconceptions. The results revealed that sixth graders have strong misconceptions about 17 out of 32 items; the examples of which are illustrated in Table 4.

The data in C1-2 indicated that 26.2% (>18.3% and mean confidence score of 4.33 > 3.3) of sample students misunderstood the meaning of "ate $\frac{1}{4}$ of half of the pizza". Therefore, they selected the answer "Because he ate $\frac{1}{4}$ of half of the pizza, which can be expressed as $\frac{1}{2} - \frac{1}{4} = \frac{1}{8}$ of the pizza". The relatively high mean confidence score indicated that students were overconfident in selecting the answer, and many students did not understand the meaning of "ate $\frac{1}{4}$ of half of the pizza". In C2-7, students were asked to arrange 0.6, $\frac{11}{12}$ and $\frac{12}{24}$ from the largest to smallest. Although 36.9% (>18.3% and mean confidence score of 4.61) of students selected the correct answer, their reasons for doing so indicated strong misconceptions. They believed that "Because $\frac{11}{12}$ means 1 part left, 0.6 means 4 parts left, $\frac{12}{24}$ means 12 parts left, $\frac{11}{12}$ has the fewest parts left, and $\frac{12}{24}$ has the most parts left, so the answer is $\frac{11}{12} > 0.6 > \frac{12}{24}$". Obviously, these students lacked an understanding of fractions.

In addition, the data in C3-8 revealed that 36.9% (>18.3% and mean confidence score of 4.27 > 3.3) of sample students selected the incorrect reason: "There are four lines, so the point A is $\frac{1}{4}$". This strong misconception further indicated that these students lacked an understanding of fractions. In C4-3, 20.1% (>18.3% and mean confidence score of 4.45 > 3.3) of sample students selected the incorrect reason "Multiplication/division

Table 4. Examples of strong misconception about each number sense component among sixth graders (N = 942).

NS	Questions	Reason	N	Misc (%)	Con
C1-2	Bobby bought a pizza and ate $\frac{1}{4}$ of half of the pizza. How much pizza did he eat?	Because he ate $\frac{1}{4}$ of half of the pizza, which can be expressed as $\frac{1}{2} - \frac{1}{4} = \frac{1}{8}$ pizza	247	26.2	4.62
C2-7	Please align 0.6, $\frac{11}{12}$, $\frac{12}{24}$ from the large to the small	Because $\frac{11}{12}$ means 1 part left, 0.6 means four parts left, $\frac{12}{24}$ means 12 part left, $\frac{11}{12}$ means the fewest to be left, $\frac{12}{24}$ means left at most, I get $\frac{11}{12} > 0.6 > \frac{12}{24}$	348	36.9	4.64
C3-8	Which picture can be the best to represent the point A in the picture of the number line below? 	There are four lines, so the point A is $\frac{1}{4}$	346	36.7	4.71
C4-3	Four students tried to solve this problem: 1234 ÷ 5 × 6. Whose solution is correct? Student A: 1234 ÷ 5 × 6 = 1234 ÷ (5 × 6); Student B: 1234 ÷ 5 × 6 = 1234 × 6 ÷ 5; Student C: 1234 ÷ 5 × 6 = 1234 ÷ 6 × 5; Student D: 1234 ÷ 5 × 6 = 5 × 6 ÷ 1234	Multiplication/Division takes precedence over addition/ subtraction, so multiplication (5 × 6) should be done prior to division	189	20.1	4.29
C5-2	A pack of candies has 12 candies, Sandy and her 5 classmates bisect 2 packs of candies. How many packs does everyone have?	Because 12 × 2 ÷ 6 = 4, so there are 4 packs	231	24.5	4.62

Note. NS: Number sense; N: the number of students; Misc: misconceptions; Con: confidence rate.

takes precedence over addition/subtraction, so multiplication (5 × 6) should be done prior to division". Although students were aware that multiplication/division should be applied before addition/subtraction, they misunderstood that multiplication should be applied prior to division, implying that students did not totally understand the roles of multiplication, division, addition and subtraction in the relevant computation. The data in C5-2 revealed that 24.5% (>18.3% and mean confidence score of 4.40 > 3.3) of sample students selected the incorrect reason "Because 12 × 2 ÷ 6 = 4, so there are 4 packs". These examples implied that students correctly selected the correct answers, but they did not understand the meaning of computational results.

Discussion and conclusion

This study examined the performance of the sixth graders in Hong Kong on the NSTTT and their possible misconceptions about number sense. The results indicated that the sixth graders in Hong Kong performed the highest on "ability to recognise number size", with the highest mean confidence score. By contrast, all sixth graders performed the lowest on "ability to judge the reasonableness of computational results", with the lowest mean confidence score. These findings are consistent with those of previous studies (Lin et al. 2016; Reys and Yang 1998; Yang et al. 2008). Yang et al. (2008) argued that Taiwanese students' performance was the highest on "ability to recognise number size" and the lowest on "ability to judge the reasonableness of computational results" because their mathematics textbooks presented more questions related to number size recognition than to judging the reasonableness of computational results. Correspondingly, a review of mathematics textbooks used by students in Hong Kong also revealed that the textbooks contained a greater number of questions related to number size recognition than to judging the reasonableness of computational results (Yang and Li 2017).

Previous studies have highlighted that mathematics textbooks are important resources for providing learning opportunities to students (Cai et al. 2011; Fan, Zhu, and Miao 2013; Hiebert and Grouws 2007; National Research Council 2004; Schmidt et al. 2001; Stein, Remillard, and Smith 2007; Tornroos 2005; Wijaya, Van den Heuvel-Panhuizen, and Doorman 2015). Stein, Remillard and Smith (2007, 360) stated that "curricula differ in significant ways and differences impact student learning". Textbooks determine what mathematical content is taught, thereby affecting the learning opportunities provided to students. Thus, it is reasonable to infer that the performance of sample students in Hong Kong was the highest on number size recognition and the lowest on judging the reasonableness of computational results because the mathematics textbooks used by students in Hong Kong provide more opportunities to learn about number size recognition than to learn about judging the reasonableness of computational results. The findings of this study also support those of previous studies, which have demonstrated that the learning opportunities provided affect students' learning outcomes (Cai et al. 2011; Fan, Zhu, and Miao 2013; Stein, Remillard, and Smith 2007; Tarr et al. 2013; Tornroos 2005; Wijaya, Van den Heuvel-Panhuizen, and Doorman 2015). In addition, the results revealed that the confidence indices of sample students were the highest on number size recognition and the lowest on judging the reasonableness of computational results. This indicated that sample students have more opportunities to learn about number size recognition and therefore more confident in answering number size recognition-related questions.

The results also revealed that among sixth graders, the percentages of correct answers to questions in the first two tiers were lower than their mathematics scores in school and international mathematics assessments such as PISA (OECD 2016) and TIMSS (Mullis et al. 2016). These results are similar to those of previous studies, which have indicated low student performance on number sense-related tests (Lin et al. 2016; Markovits and Sowder 1994; McIntosh et al. 1997; Purnomo, Kowiyah, and Assiti 2014; Yang et al. 2008). Simultaneously, the mean confidence scores for each component and the entire test were higher than 4 among the sixth graders, implying that students were overconfident in their responses to the number sense-related questions. In addition, the results indicated that sample students might have significant misconceptions or lack of knowledge of number sense, which also supports the findings of previous studies (Hasan, Bagayoko, and Kelley 1999; Pesman and Eryilmaz 2010).

There were 36.8% of sixth graders who exhibited medium–low number sense and high confidence, indicating that they may also have misconceptions about number sense-related concepts, as implied by their unsatisfactory performance and high confidence. This finding is consistent with that of previous studies (Hasan et al. 1999; Pesman and Eryilmaz 2010), which reported that students may have strong misconceptions about number sense but do not know that they may have such misconceptions. Organization for Economic Co-operation and Development [OECD] (2013) reported that students in countries with higher mean performance in mathematics were more likely to report confidence in their ability to solve a range of pure and applied mathematical problems. Alves-Martins et al. (2002), who investigated 838 secondary students in the United States, observed a significant relationship between confidence (i.e. self-esteem) and academic performance (i.e. achievement). Gok (2014), who examined students' achievement, skill and confidence in using stepwise problem-solving strategies, observed a relationship between students' performance and confidence.

Following previous studies (Caleon and Subramaniam 2010; Tan et al. 2002; Yang et al. 2016), this study defined students with significant misconceptions as those whose percentage of incorrect responses to questions in the first two tiers of the NSTTT was more than 18.3%. The results indicated that sixth graders had strong misconceptions about 17 out of 32 items. In addition, the sixth graders had strong misconceptions about 13 out of 32 items, implying that misconceptions may persist over a long period (Durkin and Rittle-Johnson 2015; Eryilmaz 2002; Sarwadi and Shahrill 2014) and correcting misconceptions is difficult without appropriate instruction. For example, Table 4 illustrates that the sixth graders had common and persistent misconceptions involving concepts related to the application of fractions to solve word problems, judging the magnitude of decimals and fractions, and use of fractions. These misconceptions may be attributed to students' lack of understanding of the basic definitions of decimals and fractions and to the overgeneralisation of their knowledge of whole numbers to solve decimal- and fraction-related questions (DeWolf and Vosniadou 2014; Durkin and Rittle-Johnson 2015; Irwin 2001; Resnick et al. 1989). This study demonstrated that some sixth graders in Hong Kong exhibited strong misconceptions about half of the number sense-related questions. Previous studies have indicated that students' misconceptions are probably persistent and may hinder their future learning (Durkin and Rittle-Johnson 2015; Stafylidou and Vosniadou 2004; Vamvakoussi and Vosniadou 2004) unless teachers intervene appropriately by using suitable examples or activities to correct their misconceptions (Sarwadi and Shahrill 2014). Therefore, the authors suggest that effective number sense-related activities should be designed and integrated into classrooms. In particular, when designing such activities, measures should be implemented to alleviate strong or weak misconceptions. The authors believe that teaching is more effective when misconceptions are identified, challenged and alleviated.

In conclusion, based on our findings, we believe that the NSTTT has several advantages: (1) it provides the advantages of a two-tier test, including the simultaneous collection of quantitative and qualitative information; (2) it measures students' confidence in their responses and (3) it distinguishes misconceptions resulting from the lack of knowledge. The test can be used to distinguish strong misconceptions, weak misconceptions and non-significant misconceptions among students. We hope that these findings on students' number sense performance and misconceptions can be used to inform Hong Kong's educational policies and future research.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Ministry of Science and Technology, Taiwan: [Grant Number MOST 105-2511-S-415-003-MY3].

Notes on contributors

Dr. Ka Luen Cheung is an Assistant Professor at the Education University of Hong Kong, Hong Kong. His primary research interest is on Nonlinear Partial Differential Equations. He also conducts research in Mathematics Problem Solving, Number Sense, Geometric Thinking, E-Learning in Mathematics and STEM education.

Dr. Der-Ching Yang is a Distinguished Professor at National Chiayi University, Taiwan (R.O.C). His research interests focus on Number Sense, International Comparative Study, Mathematics Textbook, and Curriculum Development.

References

Akkaya, R. 2016. "An Investigation into The Number Sense Performance of Secondary School Students in Turkey." *Journal of Education and Training Studies* 4 (2): 113-123. doi: 10.11114/jets.v4i2.1145.

Alves-Martins, M., F. Peixoto, M. Gouveia-Pereira, V. Amaral, and I. Pedro. 2002. "Self-Esteem and Academic Achievement among Adolescents." *Educational Psychology: An International Journal of Experimental Educational Psychology* 22 (1): 51–62. doi:10.1080/01443410120101242.

Aunio, P., J. Ee, S. Eng, A. Lim, J. Hautamaki, and J. E. H. Van Luit. 2004. "Young Children's Number Sense in Finland, Hong Kong and Singapore." *International Journal of Early Years Education* 12 (3): 195–216. doi:10.1080/0966976042000268681.

Berch, D. B. 2005. "Making Sense of Number Sense: Implications for Children with Mathematical Disabilities." *Journal of Learning Disabilities* 38 (4): 333–339. doi:10.1177/00222194050380040901.

Cai, J., N. Wang, J. C. Moyer, and B. Nie. 2011. "Longitudinal Investigation of the Curriculum Effect: An Analysis of Student Learning Outcomes from the LieCal Project." *International Journal of Educational Research* 50 (2): 117–136. doi:10.1016/j.ijer.2011.06.006.

Caleon, I., and R. Subramaniam. 2010. "Development and Application of a Three-Tier Diagnostic Test to Assess Secondary Students' Understanding of Waves." *International Journal of Science Education* 32 (7): 939–961. doi:10.1080/09500690902890130.

Cetin-Dindar, A., and O. Geban. 2011. "Development of A Three-Tier Test to Assess High School Students' Understanding of Acids and Bases." *Procedia-Social and Behavioral Sciences* 15: 600–604. doi:10.1016/j.sbspro.2011.03.147.

Clement, J., D. Brown, and A. Zietsman. 1989. "Not All Preconceptions are Misconceptions: Finding "Anchoring Conceptions" for Grounding Instruction on Students' Intuitions." *International Journal of Science Education* 11 (5): 554–565. doi:10.1080/0950069890110507.

Dehaene, S. 1997. *The Number Sense: How the Mind Creates Mathematics*. London: Oxford University Press.

DeWolf, M., and S. Vosniadou. 2014. "The Representation of Fraction Magnitudes and the Whole Number Bias Reconsidered." *Learning and Instruction* 37: 39–49. doi:10.1016/j.learninstruc.2014.07.002.

Dunphy, E. 2007. ""The Primary Mathematics Curriculum: Enhancing Its Potential for Developing Young Children's Number Sense in the Early Years at School." *Irish Educational Studies* 26 (1): 5–25. doi:10.1080/03323310601125088.

Durkin, K., and B. Rittle-Johnson. 2015. "Diagnosing Misconceptions Revealing Changing Decimal Fraction Knowledge." *Learning and Instruction* 37: 21–29. doi:10.1016/j.learninstruc.2014.08.003.

Dyson, N. I., N. C. Jordan, and J. Glutting. 2013. "Number Sense Intervention for Low-Income Kindergartners at Risk for Mathematics Difficulties." *Journal of Learning Disabilities* 46 (2): 166–181. doi:10.1177/0022219411410233.

Eryilmaz, A. 2002. "Effects of Conceptual Assignments and Conceptual Change Discussions on Students' Misconceptions and Achievement regarding Force and Motion." *Journal of Research in Science Teaching* 39 (10): 1001–1015. doi:10.1002/tea.10054.

Fan, L., Y. Zhu, and Z. Miao. 2013. "Textbook Research in Mathematics Education: Development Status and Directions." *ZDM – the International Journal on Mathematics Education* 45 (5): 633–646. doi:10.1007/s11858-013-0539-x.

Gersten, R., N. C. Jordan, and J. R. Flojo. 2005. "Early Identification and Intervention for Students with Mathematics Difficulties." *Journal of Learning Disabilities* 38 (4): 293-304. doi:10.1177/00222194050380040301.

Gok, T. 2014. "Students' Achievement, Skill and Confidence in Using Stepwise Problem-Solving Strategies." *Eurasia Journal of Mathematics, Science & Technology Education* 10 (6): 617–662. doi:10.12973/eurasia.2014.1223a.

Griffin, S. 2004. "Building Number Sense with Number Words: A Mathematics Programs for Young Children." *Early Childhood Research Quarterly* 19 (1): 173–180. doi:10.1016/j.ecresq.2004.01.012.

Hasan, S., D. Bagayoko, and E. L. Kelley. 1999. "Misconceptions and the Certainty of Response Index (CRI)." *Physic Education* 34 (5): 294–299. doi:10.1088/0031-9120/34/5/304.

Hiebert, J. S., and D. Grouws. 2007. "The Effects of Classroom Mathematics Teaching on Students' Learning." In *Second Handbook of Research on Mathematics Teaching and Learning*, edited by F. J. Lester, 371–404. Charlotte, NC: Information Age.

Irwin, K. C. 2001. "Using Everyday Knowledge of Decimals to Enhance Understanding." *Journal for Research in Mathematics Education* 32 (4): 399–420. doi:10.2307/749701.

Johnny, J., and M. Mohamed. 2010. "Investigating Number Sense Among Students." *Procedia Social and Behavioral Sciences* 8: 317–324. doi:10.1016/j.sbspro.2010.12.044.

Jordan, N. C., D. Kaplan, M. N. Locuniak, and C. Ramineni. 2007. "Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories." *Learning Disabilities Research & Practice* 22 (1): 36–46. doi:10.1111/j.1540-5826.2007.00229.x.

Jordan, N. C., J. Glutting, and C. Ramineni. 2010. "The Importance of Number Sense to Mathematics Achievement in First and Third Grades." *Learning and Individual Differences* 20 (2): 82–88. doi:10.1016/j.lindif.2009.07.004.

Li, M. N., and D. C. Yang. 2010. "Development and Validation of a Computer-Administered Number Sense Scale for 5th-Grade Children in Taiwan." *School Science and Mathematics* 110 (4): 220-230.

Lin, Y. C., D. C. Yang, and M. N. Li. 2016. "Diagnosing Students' Misconceptions in Number Sense via a Web-based Two-Tier Test." *Eurasia Journal of Mathematics, Science & Technology Education* 12 (1): 41-55.

Markovits, Z., and J. T. Sowder. 1994. "Developing Number Sense: An Intervention Study in Grade 7." *Journal for Research in Mathematics Education* 25 (1): 4–29. doi:10.2307/749290.

McIntosh, A., B. J. Reys, and R. E. Reys. 1992. "A Proposed Framework for Examining Basic Number Sense." *For the Learning of Mathematics* 12 (3): 2–8.

McIntosh, A., B. J. Reys, R. E. Reys, J. Bana, and B. Farrel. 1997. *Number Sense in School Mathematics: Student Performance in Four Countries*. Perth, Australia: Mathematics, Science, & Technology Education Centre, Edith Cowan University.

McNeil, N. M., and M. W. Alibali. 2005. "Why Won't You Change Your Mind? Knowledge of Operational Patterns Hinders Learning and Performance on Equations." *Child Development* 76 (4): 883–899. doi:10.1111/j.1467-8624.2005.00884.x.

Menon, R. 2004. "Elementary School Children's Number Sense." *EBSCO Host database*. Accessed 10 January 2008. http://www.cimt.plymouth.ac.uk/journal/default.htm

Mullis, I. V. S., M. O. Martin, P. Foy, and M. Hooper. 2016. "TIMSS 2015 International Results in Mathematics." Boston College, TIMSS & PIRLS International Study Centre. accessed 15 January 2008. http://timssandpirls.bc.edu/timss2015/international-results/

National Council of Teachers of Mathematics [NCTM]. 2000. *The Principles and Standards for School Mathematics*. Reston, VA: NCTM.

National Research Council. 2004. *On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematic Evaluation*. J. Confrey & V. Stohl (Eds.). Washington, DC: National Academy Press.

Nickerson, S. D., and I. Whitacre. 2010. "A Local Instruction Theory for the Development of Number Sense." *Mathematical Thinking and Learning* 12 (3): 227-252. doi: 10.1080/10986061003689618.

Organization for Economic Co-operation & Development (OECD). 2016. *PISA 2015 Results in Focus*. Paris: OECD.

Organization for Economic Co-operation & Development [OECD]. 2013. "Ready To Learn: Students' Engagement, Drive And Self-Beliefs – Volume III." *OECD database*. accessed 26 October 2017. http://www.oecd.org/pisa/keyfindings/PISA2012-Vol3-Chap4.pdf

Pesman, H., and A. Eryilmaz. 2010. "Development of a Three-Tier Test to Assess Misconceptions about Simple Electric Circuits." *The Journal of Educational Research* 103 (3): 208–222. doi:10.1080/00220670903383002.

Purnomo, Y. W., F. A. Kowiyah, and S. S. Assiti. 2014. "Assessing Number Sense Performance of Indonesian Elementary School Students." *International Education Studies* 7 (8): 74–84. doi:10.5539/ies.v7n8p74.

Resnick, L. B., P. Nesher, F. Leonard, M. Magone, S. Omanson, and I. Peled. 1989. "Conceptual Bases of Arithmetic Errors: The Case of Decimal Fractions." *Journal for Research in Mathematics Education* 20 (1): 8–27. doi:10.2307/749095.

Reys, R. E., and D. C. Yang. 1998. "Relationship between Computational Performance and Number Sense among Sixth- and Eighth-grade Students in Taiwan." *Journal for Research in Mathematics Education* 29 (2): 225-237.

Sarwadi, H. R. H., and M. Shahrill. 2014. "Understanding Students' Mathematical Errors and Misconceptions: The Case of Year 11 Repeating Students." *Mathematics Education Trends and Research* 2014 (2014): 1–10. doi:10.5899/2014/metr-00051.

Schmidt, W. H., C. C. McKnight, R. T. Houang, H. Wang, D. E. Wiley, L. S. Cogan, and R. G. Wolfe. 2001. *Why Schools Matter: A Cross-National Investigation of Curriculum and Learning*. San Fransisco, CA: Jossie-Bass.

Şengül, S., and D. H. Gülbağcı. 2014. "The Strategies of Mathematics Teachers When Solving Number Sense Problems." *Turkish Journal of Computer and Mathematics Education (TURCOMAT)* 5 (1): 73–88. doi:10.16949/turcomat.67936.

Sengul, S., and D. H. Gulbagci. 2012. "An Investigation of 5th Grade Turkish Students' Performance in Number Sense on the Topic of Decimal Numbers." *Procedia-Social and Behavioral Sciences* 46: 2289–2293. doi:10.1016/j.sbspro.2012.05.472.

Sood, S., and A. K. Jitendra. 2007. "A Comparative Analysis of Number Sense Instruction in Reform-Based and Traditional Mathematics Textbooks." *The Journal of Special Education* 41 (3): 145–157. doi:10.1177/00224669070410030101.

Sowder, J. T. 1992. "Estimation and Number Sense." In *Handbook of Research on Mathematics Teaching and Learning*, edited by D. A. Grouws, 371–389. New York: Macmillan.

Stafylidou, S., and S. Vosniadou. 2004. "The Development of Students' Understanding of the Numerical Value of Fractions." *Learning and Instruction* 14 (5): 503–518. doi:10.1016/j.learninstruc.2004.06.015.

Stankov, L., and J. D. Crawford. 1997. "Self-Confidence and Performance on Test of Cognitive Abilities." *Intelligence* 25 (2): 93–109. doi:10.1016/S0160-2896(97)90047-7.

Stein, M. K., J. Remillard, and M. S. Smith. 2007. "How Curriculum Influences Student Learning." In *Second Handbook of Research on Mathematics Teaching and Learning*, edited by F. Lester Jr., 319–369. Charlotte, NC: Information Age.

Tan, K. C. D., N. K. Goh, L. S. Chia, and D. F. Treagust. 2002. "Development and Application of A Two-Tier Multiple-Choice Diagnostic Instrument to Assess High School Students' Understanding of Inorganic Chemistry Qualitative Analysis." *Journal of Research in Science Teaching* 39 (4): 283–301. doi:10.1002/tea.10023.

Tarr, J. E., D. A. Grouws, O. Chavez, and V. M. Soria. 2013. "The Effects of Content Organization and Curriculum Implementation on Students' Mathematics Learning in Second-Year High School Courses."." *Journal for Research in Mathematics Education* 44 (4): 683–729. doi:10.5951/jresematheduc.44.4.0683.

The Hong Kong Curriculum Development Council. 2000. *Learning to Learn: The Way Forward in Curriculum Development*. Hong Kong: Author.

The Hong Kong Curriculum Development Council. 2014. "*Mathematics Education Key Learning Area*." *Hong Kong: Author*. Accessed 08 January 2018 http://334.edb.hkedcity.net/doc/eng/curriculum/Math%20C&A%20Guide_updated_e.pdf

The Hong Kong Curriculum Development Council. 2017. "Mathematics Education Key Learning Area Curriculum Guide (Primary 1 – Secondary 6)." Hong Kong: Author. Accessed 08 January 2018) http://www.edb.gov.hk/attachment/en/curriculum-development/renewal/ME/ME_KLACG_eng_draft_2017_04.pdf.

Tornroos, J. 2005. "Mathematics Textbooks, Opportunity to Learn and Student Achievement." *Studies in Educational Evaluation* 31 (4): 315–327. doi:10.1016/j.stueduc.2005.11.005.

Vamvakoussi, X., and S. Vosniadou. 2004. "Understanding the Structure of the Set of Rational Numbers: A Conceptual Change Approach." *Learning and Instruction* 14 (5): 453–467. doi:10.1016/j.learninstruc.2004.06.013.

Verschaffel, L., B. Greer, and E. De Corte. 2007. "Whole Number Concepts and Operations." In *Second Handbook of Research on Mathematics Teaching and Learning*, edited by F. K. Lester Jr., 557–628. Charlotte, NC: Information Age.

Wijaya, A., M. Van den Heuvel-Panhuizen, and M. Doorman. 2015. "Opportunity-To-Learn Context-Based Tasks Provided by Mathematics Textbooks." *Educational Studies in Mathematics* 89 (1): 41–65. doi:10.1007/s10649-015-9595-1.

Yang, D. C. 2005. "Number Sense Strategies used by Sixth Grade Students in Taiwan." *Educational Studies* 31 (3): 317-334. doi: 10.1080/03055690500236845.

Yang, D. C., M. C. Yu, and M. N. Li. 2016. "The Development and Application of Number Sense Three-Tier Test for Fourth Graders." Paper presented at the Second Asian Conference on Education & International Development, Japan, April 3- 6.

Yang, D. C., and M. N. Li. 2013. "Assessment of Animated Self-Directed Learning Activities Modules for Children's Number Sense Development." *Journalof Educational Technology and Society* 16 (3): 44-58.

Yang, D. C., and M. N. Li. 2017. "A Study of Fifth Graders' Performance on the Three-Tier Number Sense Test." Paper presented at the 41st Annual Meeting of the International Group for the Psychology of Mathematics Education, Singapore, July 17- 22.

Yang, D. C., M. N. Li, and C. I. Lin. 2008. "A Study of the Performance of 5th Graders in Number Sense and Its Relationship to Achievement in Mathematics." *international Journal of Science and Mathematics Education* 6 (4): 789-807