Eurasia Journal of Mathematics, Science & Technology Education, 2016, **12**(1), 41-55

Diagnosing Students' Misconceptions in Number Sense via a Web-Based Two-Tier Test

Yung-Chi Lin
National Changhua University of Education, TAIWAN
Der-Ching Yang & Mao-Neng Li
National Chiayi University, TAIWAN

•*Received 10 March 2015* • *Accepted 14 September 2015* •*Published online 12 Oct 2015*

A web-based two-tier test (WTTT-NS) which combined the advantages of traditional written tests and interviews in assessing number sense was developed and applied to assess students' answers and reasons for the questions. In addition, students' major misconceptions can be detected. A total of 1,248 sixth graders in Taiwan were selected to participate in this study. Results showed that the average percentage of correct answers was about 45%. Among the students who chose correct answers, about 22.9% of them used a number sense method to solve problems. In addition, students' misconceptions are classified by content domains. The major contributions of the WTTT-NS are to (1) avoid students using written computations answering number sense questions; (2) present a whole picture of students' misconceptions and the weight of these misconceptions; (3) include the strengths of quantitative and qualitative methods; (4) identify students' "true understanding" (a correct answer based on their correct understanding instead of guessing) by exploring reasons for their choices. In sum, the WTTT-NS is a new worthwhile method to assess students' number sense competence.

Keywords: sixth graders, misconceptions, number sense, two-tier test

INTRODUCTION

Assessing number sense is an important and challenging job for mathematics educators, researchers, and curriculum designers (Chrysostomou, Pitta-Pantazi, Tsingi, Cleanthous & Christou, 2013; Durkin & Rittle-Johnson, 2014). Past studies have been stressed on paper-and-pencil tests and/or interviews (Dunphy, 2007; Markovits & Sowder, 1994; McIntosh et al., 1997; Menon, 2004; Reys & Yang, 1998; Şengül, & Gülbağcı, 2012; Yang, 2005; Yang, Reys, & Reys, 2009). However, these methods wasted time and paper, required strenuous effort to analyze data, and were restricted to classrooms and individual students. Most importantly, using these methods cannot provide a whole picture of students' understanding of number

Correspondence: Der-Ching Yang,
Graduate Institute of Mathematics and Science Education, National Chiayi University, #85 Wen Lung, Ming-Shiung, Chiayi, Taiwan, R. O. C.
E-mail: dcyang@mail.ncyu.edu.tw
doi: 10.12973/eurasia.2016.1420a

Copyright © 2016 by iSER, International Society of Educational Research
ISSN: 1305-8223

sense. That is, using a paper-and-pencil test can collect lots of data but lack of students' explanations for their responses; conducting an individual interview can obtain students' thinking process but it is a time-consuming job and it may not be useful for a large-scale study to detect a large number of students' number sense and misconceptions.

A web-based two-tier test provided in this study may overcome the existed methods' limitations. It included the first-tier (answer-tier), which measures content knowledge; the second-tier (reason-tier), which assesses a reason for the first-tier response (Chou, Tsai & Chan, 2007; Yang & Li, 2013). It provided an opportunity not only to analyze students' number sense understanding, but also to explore their thinking process as it involved number sense. In particular, number sense is generally thought as a kind of higher order thinking (Sood & Mackey, 2014). If we do not assess students' explanations for their answers (reason-tier) but only the accuracy of their answers (answer-tier), we probably cannot fully understand one's number sense. Moreover, even though the two-tier test might have its potential for the mathematics education, there still few studies about two-titer test in the mathematics education. A reason for the less emphasis in may be due to its difficulty in designing reason-tier, since there are many approaches to a correct answer in solving mathematical problems.

State of the literature

- Assessing number sense is an important and challenging job for researchers.
- Two-tier test included the first-tier (answer-tier), which measures content knowledge; the second-tier (reason-tier), which assesses a reason for the first-tier response
- Number sense is generally thought as a kind of higher order thinking. If we do not assess students' explanations for their answers (reason-tier) but only the accuracy of their answers (answer-tier), we probably cannot fully understand one's number sense.

Contribution of this paper to the literature

- Revealing how to design and implement a two-tier test in the mathematics education filed, particularly in the topic of number sense;
- By using the two-tier test, the results of this study were able to provide a whole picture of students' understanding of number sense, particularly their misconceptions;
- Students most frequent misconceptions (termed as significant misconceptions) were identified. More contributions of this study were discussed in the final section.

Misconceptions (or termed as alternative conceptions) related to number sense were identified in the earlier studies (e.g., Durkin & Rittle-Johnson, 2014). However, these misconceptions were usually scattered in the studies of fractions, decimals, estimations, operations and so on (e.g., Widjaja, Stacey & Steinle, 2011). It seems to lack a study of comprehensively detecting these misconceptions in a number sense test. By our online two-tier test and large-scale study, the whole picture of these misconceptions can be identified and the weight of them can be reported as well (how many percent the students chose these misconceptions). Obtaining theses information will be valuable for teachers to teach number sense.

Therefore, the purpose of this study is to develop a web-based two-tier number sense test to detect students' number sense and identify their most frequent misconceptions (significant misconceptions). The research questions for this study are as follows:

1. How do sixth graders perform on the web-based two-tiered test for number sense?
2. What are the major methods used by sixth graders to solve number sense problems?
3. What significant misconceptions can be identified by the number sense questions on the web-based two-tier test?

BACKGROUND

Number sense components

Even though number sense is a relatively complex term, the emphasis on developing children's number sense has been broadly highlighted in mathematics education (Durkin & Rittle-Johnson, 2014; Verschaffel et al., 2007; Yang & Li, 2013; Yang, Reys, & Reys, 2009). Several researchers defined number sense and its components in different ways (Gersten, Jordan, & Flojo, 2005; Jordan et al., 2010; Verschaffel et al., 2007), but most of the number sense-related literature (Berch, 2005; Markovits & Sowder, 1994; McIntosh et al., 1997; Reys & Yang, 1998; Yang, 2005) showed that understanding the meaning of numbers, recognizing relative number magnitude, recognizing the relative effect of operations on numbers, and judging the reasonableness of a computational result were the key elements of number sense. Accordingly, this study defined number sense components as the following:

F1. Understanding the meaning of numbers

It means an understanding of the base-ten number system (whole numbers, fractions, and decimals), including place value, number patterns, and the use of multiple ways to represent numbers (McIntosh, Reys & Reys, 1992; Yang & Li, 2013).

F2. Recognizing relative number size

It means the recognition of the relative size of numbers. For example, when comparing fractions, students are able to use meaningful ways to solve the problem, such as the same numerator, same denominator, transitive, and residual (Cramer, Post, & delMas, 2002), without depending on standard written methods (such as finding the least common denominator suggested in the mathematics curriculum).

F3. Being able to use different representations

It means the ability to switch among different representations and use the most appropriate representation (Faulkner, 2009). For example, children should know $\frac{1}{4}$ can be represented in different forms, $\frac{2}{8}$, 0.25, 25%.

F4. Recognizing the relative effects of an operation on numbers

When children are asked to determine the result of $\frac{14}{25}\times\frac{7}{17}$, they know $\frac{14}{25}$ is less than 1 and $\frac{7}{17}$ is less than $\frac{1}{2}$, so they can conclude that the result will be less than $\frac{1}{2}$. Students should make sense of the operations and understand that multiplication can increase or decrease a fraction (McIntosh et al., 1992).

F5. Being able to judge the reasonableness of a computational result

It means children can mentally apply estimation strategies to problems without using written computation (McIntosh et al., 1992; Sowder, 1992; Yang & Li, 2013).

Misconception-related studies

Misconceptions related to numbers and operations have been widely discussed over the past few decades. In this section, we organize the students' common misconceptions about decimals, fractions, operations, and estimations.

Regarding decimals, there are two key misconceptions, "longer is larger" (e.g., 4.03 > 4.3) and "shorter is larger" (e.g., 0.2 > 0.25) (Durkin & Rittle-Johnson, 2014; Resnick et al., 1989). Studies (Sarwadi & Shahrill, 2014; Shahrill, 2013; Stacey & Steinle, 1998; Widjaja, Stacey & Steinle, 2011) have identified several possible reasons for these two misconceptions. The most common misconceptions include misusing the "whole number rule", "zero rule", "reciprocal rule", and "incorrectly generalizing the term of place value." For example, students believe 4.03 > 4.3 because 403 > 43, and 0.2 > 0.25 because $\frac{1}{2} > \frac{1}{25}$. The zero rule leads students to believe that 2.03 and 2.3 are the same number because zeros after the decimal point can be disregarded. In addition, students may also have the misconception of "number density." For example, R. Reys et al. (1999) found that students often did not realize that there were infinite decimals between two decimals on the number line (they thought there was no decimals between 1.52 and 1.53). Students misuse the rule of "no whole number between two consecutive whole numbers" by also applying it to decimals.

Students' misconceptions of fractions are discussed widely. For example, in comparing fractions, students may think $\frac{5}{6}$ and $\frac{8}{7}$ are equivalent because they only consider the difference between numerators and denominators (Clarke & Roche, 2009). The thinking of whole numbers without considering the size of the denominator and the ratio of the numerator to denominator is known as "gap thinking" (Pearn & Stephens, 2004). Another misconception caused by whole numbers occurs in making separate comparisons of the numerators and the denominators (Behr, Wachsmuth, Post, & Lesh, 1984).

Some misconceptions are related to confusion between decimals and fractions. For example, Stacey and Steinle (1998) indicated that students may incorrectly interpret decimals as reciprocals of whole numbers or as other fractions (e.g., 2.64 as $2\frac{1}{64}$). When adding and subtracting fractions, the most common misconception occurs when students think they can simply add or subtract the numerators and denominators (e.g., $\frac{1}{4}+\frac{2}{3}=\frac{3}{7}$). Moreover, many students cannot explain "$2 \div \frac{1}{2}$" as "how many $\frac{1}{2}$s are there in 2?" (Newstead & Murray, 1998). This will result in inverting the dividend instead of the divisor, or inverting both the dividend and the divisor, when dividing fractions (Tirosh, 2000).

Number sense in Taiwan's mathematics textbooks

Inappropriate application of whole-number schemas, the way and sequence of the content presented to students, is also identified as a factor that causes students' number sense misconceptions (Newstead & Murray, 1998). Therefore, knowing the way of number sense presented in the curriculum is helpful for analyzing students' misconceptions. Only few problems related to number sense are found in one of the three published textbooks (Yang, Li, & Lin, 2008). However, the activities in Taiwanese mathematics textbooks highly focus on written computation. Previous studies in Taiwan showed that textbooks of fifth and sixth grade put more emphasis on procedural knowledge of the topic of fractions and correctness of answers (Reys & Yang, 1998; Yang, 2005).

METHOD

Participants

The participants were 1,248 sixth-grade students from 25 elementary schools in Taiwan. The sample population was classified into four demographic areas: Northern, Central, Southern, and Eastern. The selected student number ratio corresponds to the actual student number ratio across these four areas. We selected at most three classes at each school. The participants came from a variety of elementary schools in both metropolitan and rural cities and counties in Taiwan, and across different demographic areas, so they may, to a certain extent, be said to represent a good cross-section of Taiwan's sixth-grade population.

The instrument

The items for the web-based two-tier test (WTTT-NS) were written based on the large amount of interview data in our previous study on number sense as well as literature on number sense misconceptions (e.g., Durkin & Rittle-Johnson, 2014). Based on these earlier studies, the authors have developed sufficient number sense test questions and collected interview data from more than 300 participants over the past 10 years (Reys & Yang, 1998; Yang, 2005; Yang & Li, 2013). The items on the web-based two-tier test were created and their related reasons were designed according to children's frequent answers in those interviews. The final web-based two-tier test for number sense includes 50 items covering the five number sense components, with 10 items for each component. The content of the items consisted of questions about fractions, decimals, whole numbers, and their operations. The first-tier of each item was a multiple-choice question with four choices (answer-tier). The second-tier of each item contained three or four reasons for each choice (reason-tier). The reasons for a correct choice (in the answer-tier) were usually based on "a number sense-based method", "a rule-based method", "misconceptions", or "guessing." But the reasons for an incorrect choice in the A-tier (answer-tier) were only "misconceptions" and "guessing." A sample question is presented in Figure 1.

Several steps were taken to ensure the validity and reliability of the instrument. Before the actual online test was given, two experts in the field of mathematics education reviewed the WTTT-NS items for face and content validity, and five experienced elementary school teachers and 20 fifth graders were selected to clarify the wording of each statement. A few items were revised based on the results from this initial evaluation process. Furthermore, 70 fifth graders were chosen to participate in an online pilot study. Except for some minor modifications found in the keying process, no considerable changes were made.

The Cronbach α of the web-based two-tier test for the sixth graders was 0.877, and the construct reliability indexes derived from SEM analysis was 0.897. To ensure the designed items were representative and not beyond the curriculum scope usually taught to sixth graders in Taiwan, three elementary school teachers and mathematics educators were invited to review the test items. All of them agreed that these test items were representative and appropriate for sixth graders. Furthermore, data analysis showed that the difficulty indexes of the test items were between 0.26-0.79 and the discrimination power was between 0.22-0.77.

The above analysis implied that the web-based two-tier test for sixth graders was a reliable and valid measure to students' performance on number sense.

Procedure of implementing the WTTT-NS

Due to the time constraint of 40 minutes per class and children's limited attention span, the test questions were divided into two sections (40 mins/persection). Part I and Part II each contained 25 items. The procedure to take the WTTT-NS test was as follows: (1) Log on to the online testing system; (2) Key in personal information; (3) Review the test rules for the test; (4) Display one practice item for the testee; (5) Conduct the formal WTTT-NS test.

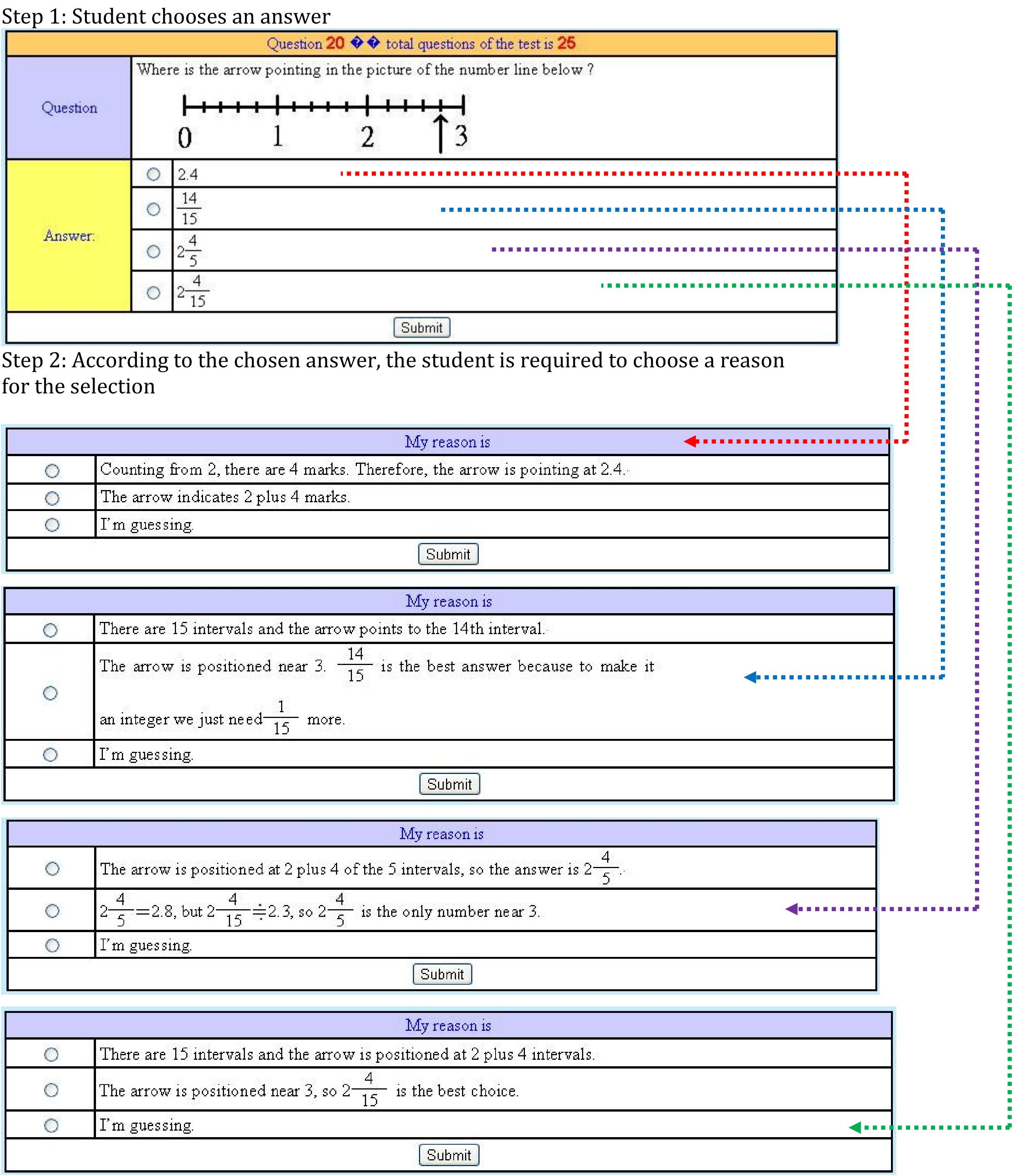

Figure 1. An example from the web-based two-tier test for number sense

Data analysis

The WTTT-NS included two sections of answers and reasons. After students have chosen their choices, their choices were coded automatically by the WTTT-NS. Students' selected answers were coded as either correct or incorrect (1 or 0); their selected reasons were coded as a number sense-based method, a rule-based method, misconception, or guessing (we have designed different choices corresponding to these four categories in the WTTT-NS). These automatic saved codings were then analyzed by SPSS 17.0 for descriptive statistics.

For identifying students' most frequent misconceptions, we defined "significant misconceptions" (Caleon & Subramaniam, 2010). It implies students may have more serious difficulties in these significant misconceptions. We used 7% as a criterion for significant misconceptions because it is above the percentage of students who may select the A-R options by chance. WTTT-NS contained 4 options in the A-tier and 3 or 4 reasons for each option in the A-tier. In average, we had total 14 options in the A-R tier and $\frac{1}{14}$ are quite closed to 7%.

RESULTS

Students' performance on the WTTT-NS test

Table 1 summarized the descriptive statistics of the sixth graders' performance on the WTTT-NS test. Results showed that the average percentage of correct answers (A-tier) was about 45% and the correct answers with the use of a number sense-based method (R-tier) was about 22.9%, which indicated that less than one fourth of these sixth graders correctly utilized number sense on the test questions. That less students used number sense on the test suggested the sixth graders' number sense was not fully developed.

The diagnostic results in Table 1 showed that around 50% of the students had misconceptions in solving these items (F2 and F3 are 48.3% and 48.2% respectively, and they are close to 50%; F5 was the highest percentage of misconceptions). Additionally, the results showed that only 23% of the sixth graders were able to adopt a number sense-based method to correctly solve these problems, which indicated that these students' performance on number sense was relatively poor. This may be attributed to the overemphasis on written computation and the lack of conceptual understanding of number sense (Sood & Mackey, 2014).

In addition to the above results, over 60% of the students had misconceptions on F5 (Being able to judge the reasonableness of the computational results). This component had the highest percentage of misconceptions among the five number sense components. In reviewing the textbooks used in Taiwan (Yang et al., 2008),

Table 1. The descriptive statistics for each number sense component (N=1248)

	A-tier		R-tier		
	Correct%	**NS-based[b]**	**Rule-based[b]**	**Misconceptions**	**Guessing**
F1[a]	46%	23.1%	9.6 %	52.0 %	15.3 %
F2	55%	29.1%	6.5 %	50.1 %	14.3 %
F3	46%	23.2%	15.5 %	48.3 %	13.0 %
F4	44%	21.6%	14.5 %	48.2 %	15.7 %
F5	36%	17.5%	9.6 %	64.3 %	15.3 %
Total	45%	22.9%	10.3 %	52.6 %	14.2 %

Note. F1=Understanding the meaning of numbers; F2=Recognizing the magnitude of numbers; F3=Being able to use different representations; F4=Recognizing the relative effects of an operation on numbers; F5=Being able to judge the reasonableness of the computational results; [a]Each number sense component includes 10 items; [b]NS-based=number sense-based method; Rule-based=written computation method; NS-based and Rule-based choices are only for those who can choose the correct answer in the A-tier.

few questions can be found in textbooks and few practices can be seen in mathematics classes that were related to F5. It is probably one of the key reasons why students perform poorly on this number sense component.

Summary of the most-committed misconceptions on number sense

Based on the analyses of WTTT-NS, the most frequently misconceptions are summarized in Figure 2.

Students' most frequent misconceptions were categorized into five groups (Figure 2). Results showed that most of the misconceptions came from inappropriately applying the whole number schema. Therefore, two groups, "Fractions were treated as whole numbers" (FW) and "Decimals were treated as whole numbers" (DW), were created. Under FW and DW, both had three sub-groups (e.g., denoted as FW1, FW2, and FW3). "Confusion with decimals and fractions" (CDF) meant students misunderstood the relationship between fractions and decimals, which had one sub-group. "Incorrect operation" (IO) referred to misconceptions that were related to applying the incorrect computational methods or misunderstanding the meaning of operations, which contained six sub-groups. "Incorrect intuition" (IN) meant students' misconceptions may come from incorrect intuition, which included four sub-groups.

Table 2 reported more detail about student's misconceptions. Similar to the literature (Durkin & Rittle-Johnson, 2014; Newstead & Murray, 1998; Yang, 2005), inappropriate application of whole number schemas and misunderstanding of the meaning of the operation were the most common factors that cause children's misconceptions. For example, there were high response percentages in problems

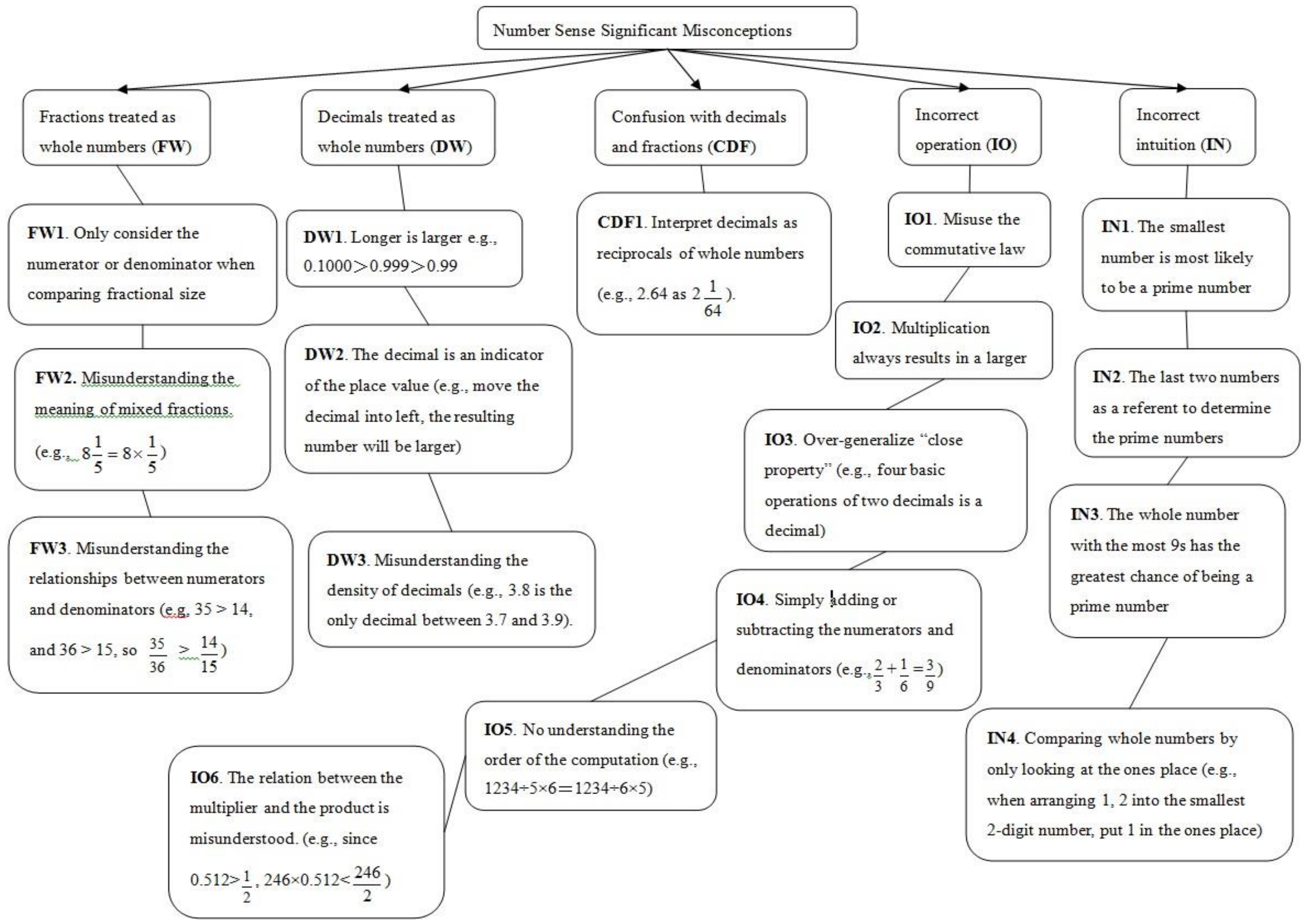

Figure 2. The map of students' misconceptions

Table 2. Zoom in the students' most-committed misconceptions (significant misconceptions) on number sense

Category	Misconceptions (Reasons chosen in the R-tier)	% *
FW1	15<36, so $\frac{14}{15} > \frac{35}{36}$	20.3
	If you cut up a pizza, the ratio of each piece in 15 pieces is larger than the ratio of each piece in 36 pieces.	13.8
	$\frac{18}{37}$ means that you have 18 "units," $\frac{17}{33}$ means that you have only 17 "units." 18＞17, so $\frac{18}{37}$ is larger.	9
FW2	$8\frac{1}{5}=\frac{8}{5}=8\times\frac{1}{5}$	9
FW3	Both the difference between 14 and 15 and the difference between 35 and 36 are equal to 1, so $\frac{35}{36} = \frac{14}{15}$	8.6
	35 > 14, and 36 > 15, so $\frac{35}{36} > \frac{14}{15}$	8.1
CDF1	18.37＞17.33, so $\frac{18}{37} > \frac{17}{33}$	8.7
DW1	0.1000＞0.999＞0.99 due to 1000＞999＞99.	16.7
DW2	If the decimal is moved to the left, it will be closer to the hundreds and thousands place. Therefore the resulting number will be greater than the original number.	14.7
	If the decimal is moved to the left, there will be more numbers after the decimal, so the new number will be greater than the original number.	15.8
DW3	There should be more decimals between 2.1 and 2.4 because the range of the two numbers is broader than the range between 3.7 and 3.9	11.5
	3.8 is the only decimal between 3.7 and 3.9, but 2.2 and 2.3 are both decimals between 2.1 and 2.4.	16
IO1	1234÷5×6＝1234÷(5×6)	35
IO2	The product of 246×0.512 will be larger than 246.	18
IO3	Four basic operations of two decimals is a decimal	10
IO4	$\frac{2}{3}+\frac{1}{6}=\frac{3}{9}$	11.8
IO5	1234÷5×6＝1234÷6×5 since 5x6 is equal to 6x5.	16
IO6	0.512 was greater than a half, so the product of 246×0.512 will be smaller than half of 246.	20
IN1	14717 is the smallest number among 82752, 14717, 69939, and 37315, so it is a prime number.	10
IN2	Which of the following is most likely to be a prime number? (1)82752 (2)14717 (3) 69939 (4)37315? Students reasoned that 52, 39, and 15 are not prime numbers; therefore, 14717 is a prime number.	27
IN3	69939 has more 9s than 82752, 14717 and 37315, so it is a prime number.	7.4
IN4	If you have to arrange 8, 4, 1, 6, 9 into the smallest 5-digit number, which number will be put in the "ones" place? 1 was the answer because 1 has the least value.	29

*% indicates response percentage of the whole samples; *, We use 7% as a criterion for significant misconceptions because it is above the percentage of students who may select the A-R options by chance; FW= fractions treated as whole numbers; DW= decimals treated as whole numbers; CDF= confusion with decimals and fractions; IO=incorrect operation; IN=incorrect intuition*

that revealed the following misconceptions: "15 < 36, so $\frac{14}{15} > \frac{35}{36}$" (20.3%), "0.1000＞0.999＞0.99 due to 1000＞999＞99" (16.7%), "1234 ÷ 5 × 6＝1234 ÷ (5 × 6)" (35%), and "The product of 246 × 0.512 will be larger than 246" (18%). Additionally, students' incorrect intuition was also one of the factors to the relatively high percentage in students' misconceptions. For example, the percentage of misconceptions in the categories of IN2 and IN4 were 27% and 29%, separately.

Clearly, from the Figure 2 and Table 2, we can know many sixth grade Taiwanese students had various significant misconceptions such as fractions, decimals, whole number-related questions, and less than one-fourth of students could use a number sense method to solve the number sense-related questions. For example, when the children were asked to compare $\frac{17}{33}$ and $\frac{18}{37}$, about one-fourth of them believed that "the lower the denominator is, the larger the fraction will be, i.e., 33 < 37, so $\frac{17}{33} > \frac{18}{37}$." Moreover, about 20% of students thought that "$\frac{17}{33}$ is divided into 33 parts, while $\frac{18}{37}$ is divided into 37 parts, so $\frac{17}{33} > \frac{18}{37}$," or "18 > 17 and 37 > 33, so $\frac{18}{37} > \frac{17}{33}$." This was most likely due to the fact that they did not understand the definition of a fraction clearly. Children did not see a fraction as a quotient relation between two numbers. Instead, they saw the fraction as two separate whole numbers.

In addition, less than 10% of the students could flexibly apply $\frac{1}{2}$ as a benchmark to determine fractional size. This result was similar to the findings in the study by Yang (2005). Furthermore, incorrect applying operation or misunderstanding the meaning of the operation led to another misconception. For example, about two-thirds of the sixth graders believed that multiplication or addition makes the result larger. The misconceptions found in this study were likely due to the highly focused drill and practice seen in mathematics teaching in Taiwan (Reys & Yang, 1998; Yang et al., 2008).

Students who were skilled in written computation do not necessarily had a better development in number sense (Rey & Yang, 1998; Menon, 2004). For instance, when students were asked to compare $\frac{14}{15}$ and $\frac{35}{36}$ without using paper and pencil, they generally tried to use the traditional written algorithm to arrive at their answer. If students were accustomed to using the written method, their mind functions will be fixed and thus it will be difficult for them to use an alternate strategy, such as the residual strategy ($\frac{14}{15}+\frac{1}{15} = \frac{35}{36}+\frac{1}{36}$, since $\frac{1}{15} > \frac{1}{36}$. hence $\frac{1}{15} < \frac{35}{36}$) to solve problems. The residual strategy is, in fact, a reversal of thinking. If we are going to help children effectively develop number sense, school teachers and textbooks should support children with meaningful learning opportunities to learn mathematical concepts and should gradually reduce the chance to use written computation.

DISCUSSION AND CONCLUSION

This study adopted a web-based two-tier test to diagnose sixth graders' number sense performance and identify significant misconceptions. The web-based two-tier test provided an opportunity not only to analyze students' number sense understanding, but also to explore their thinking process as it involved number sense. The items in the test were designed to investigate different elements of number sense used by these sixth graders in answering the items. The average percentage of the students' misconceptions detected by the instrument was about 50%, and only about 20% of students used the number sense-based method to answer the items. The results were quite similar to previous studies (Yang, 2005;

Reys et al., 1999). This confirmed our online two-tier test is quite reliable and could provide appropriate information for researchers and teachers.

This study reported a web-based two-tier test taken by Taiwanese sixth graders; strikingly, these results were similar to the findings of investigating many elementary and middle grade students in other countries such as Australia, England, Kuwait, Israel, Sweden, and the United States. Reys & Yang(1998) claimed that students who were skilled in written computation and their exact answers do not necessarily have well-developed number sense. In addition, Markovits and Sowder (1994) reported that "Few students exhibit number sense when solving arithmetic problems in schools" (p. 4) (in the United States), which was similar to the report of Office for Standards in Education (2008) that students relied on written computation and did not incline to use informal and mental strategies. The poor performance on the use of number sense with the web-based two-tier test also confirmed an earlier study (Yang, 2005) that mathematics teaching in Taiwan places a great emphasis on the written methods; this seems to limit students' thinking and most likely hinders their development of number sense.

Taiwanese students usually relied highly on written computation (Yang, 2005). Providing calculation as a reason offered a good distractor to detect students' misconception. The needed calculations in our items were usually quite complex. Students were less likely to really solve the problems by using calculations. However, from the research results, we still found that a number of (around 10%) students justified their answers by formal written methods. It was evident again that students in Taiwan highly preferred to use standard written algorithms (Yang, 2005).

Results showed that students did not perform well in number sense competence on the WTTT-NS. Particularly, a high percentage of students revealed their misconceptions on the meaning of number operations as they gave reasons for their answers. This was most likely because of the limited opportunities offered at school or at home to learn number sense (Yang et al., 2008). Studies on Taiwanese textbooks pointed out that the textbooks put more emphasis on teaching written methods and thus teachers paid no attention on the development of number sense (Reys & Yang, 1998; Yang & Li, 2013).

Additionally, the mathematics textbooks used by students will influence what mathematics content was taught as well as the ways in which it was presented. This supported the documents reported in related articles (Stein, Remillard, & Smith, 2007; Tarr et al., 2008). Textbooks have a direct influence on teachers' teaching and students' learning. Moreover, teachers in Taiwan do not know how to teach number sense since the teacher education program does not highlight the importance of number sense, and number sense programs are not integrated into the teacher education training. Therefore, it is not surprising that these students, who are taught by the teachers and under the educational system in Taiwan, perform poorly on number sense test.

Several earlier studies (Griffin, 2004) demonstrated that children's number sense can be promoted through well-designed programs that have integrated number sense into activities. Teachers, too, play a key role in helping children develop number sense. This study suggested that mathematics textbooks should put much more emphasis on the acquisition of this skill and integrate number sense into related activities to help children in their development. Teacher education in Taiwan should stress the importance of number sense and assist teachers to realize the instruction of number sense. If we want to improve children's mathematical skills and development, then actions should be taken to raise the quality of mathematics textbooks by the inclusion of number sense questions; moreover, teachers' knowledge and teaching skills of the topic must be encouraged.

Besides, it was obvious that the web-based two-tier test used in this study, which was easily administered and time-efficient, not only can assist researchers in collecting a great deal of students' responses and data on number sense, but also can screen their misconceptions.

In the past 15 years, all of the studies on students' number sense relied mainly on paper-and-pencil tests and interviews. These methods were time-consuming as well as required much manual labor, and the number of students that could be assessed was limited. Furthermore, the assessment of number sense usually asked the children not to use algorithms in those tests but it was not easily enforced on the paper-and-pencil format. However, the web-based test can prevent children from using paper and pencil and can encourage them to answer the questions on their own.

THE MAJOR CONTRIBUTIONS OF THIS STUDY

The major contributions of this study were further discussed in the following directions:

1. The WTTT-NS was an innovative instrument in the field of mathematics education, particularly in assessing number sense.
2. It integrated the advantages of both quantitative (paper-and-pencil tests) and qualitative (interviews) methods.
3. It can summarize the students' number sense performance and identify the frequent misconceptions.
4. It can detect students' "true understanding" by creating "I'm guessing" as an option in R-tier.
5. It encouraged students to use different approaches except written computation through the WTTT-NS.
6. It provided immediate feedback after the test.

Developing and applying the two-tier test in number sense are not only new but also significant to the mathematics education. Unlike science education, few two-tier test related studies are used in mathematics education (e.g., Ang & Shahrill, 2014), even though the two-tier test might have its potential for the mathematics education. A reason for the less emphasis in mathematics education may be due to its difficulty in designing R-tier, since there are many approaches to a correct answer in solving mathematical problems.

To create a possible reason in the R-tier for all options in the A-tier is difficult. In most cases, once an option in the A-tier is chosen, the other incorrect reasons in the R-tier are often easily eliminated. Therefore, in this study, we did not create reasons for all options in the A-tier which actually was the typical format in the science education. Instead, we created reasons for each option in A-tier. The results of this study have shown that this design is useful. More importantly, assessing number sense, unlike assessing other common mathematical content, cannot only check their correct answers (a quantitative approach). We probably need to know the students' reasons for their answers (a qualitative approach). As shown in this study, even though one can choose a correct answer in A-tier, it is still possible to lack number sense because of choosing a non-number sense method. From this concern, developing and applying a two-tier test is more powerful to ensure the validity of the instrument.

Regarding the R-tier, we also specifically developed "I'm guessing" as an option for each option in the A-tier. This makes us easily to distinguish students' answers based on guessing from answers, misconceptions, or true understandings. There is little research discussing students' guessing or confidence levels in the number sense test. The design of our two-tier test allows us to do so. When students choosing "I'm guessing" as their justification, it may have two implications: students are not familiar with the mathematical concepts containing in the problems (particularly, those who chose incorrect answers and choose guessing in the R-tier); students have no confidence in what they are choosing (particularly, for those who chose the correct answers but choose guessing in the R-tier). The former one is even worse than the students who chose misconceptions because it implies that students are lack of mathematics knowledge on what are testing. The latter one deals with the issue about confidence in students' learning.

Studies have shown that East Asian students have low confidence in their academic self-report (Martin et al., 2008) but how this result will affect their learning is not clear. Our study seems to show a direction for the further research. That is, students may not explicitly justify their correct answers due to their low confidence. The WTTT-NS successfully identified about 15% of the students justifying their answers by choosing "guessing." From this result, we should notice that some students did not develop number sense due to the lack of some basic mathematics knowledge.

Although studies have shown Taiwanese students have low performance on the number sense and much relied on the written computation in solving problems, these results were usually based on the small sample size and those samples usually restricted in a local district (Yang et al., 2008). In contrast, this study involved large samples and the samples were selected from different areas (e.g., the North, Center…) in Taiwan. Therefore, the results of this study are more representative of the Taiwanese elementary students and they are more able to be generalized.

Moreover, through the use of the WTTT-NS, students can get an instant feedback after the test. That is, the screen will show how many problems are successfully and unsuccessfully solved. Students are allowed to click on the problems to review their answers and reasons. The immediate feedback allows students to review their incorrect answers and reasons. It may help students to revise the incorrect concepts.

ACKNOWLEDGEMENTS

This paper is a part of a research project supported by the Ministry of Science and Technology, Taiwan with grant no. MOST 102-2511-S-415-002-MY3. Any opinions expressed here are those of the author and do not necessarily reflect the views of the Ministry of Science and Technology, Taiwan.

REFERENCES

Ang, L. H. & Shahrill, M (2014). Identifying students' specific misconceptions in learning probability, *International Journal of Probability and Statistics*, 3(2), 23-29.

Behr, M. J., Wachsmuth, I., Post, T. R., & Lesh, R. (1984). Order and equivalence of rational numbers: A clinical teaching experiment. *Journal for Research in Mathematics Education*, *15*, 323-341.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*, 333-339.

Caleon, I. S., & Subramaniam, R. (2010). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. *Research in Science Education, 40*(3), 313-337.

Chou, C., Tsai, C. C., & Chan, P. S. (2007). Developing a web-based two-tier test for internet literacy. *British Journal of Educational Technology*, *38*(2), 369-372.

Chrysostomou, M., Pitta-Pantazi, D., Tsingi, C., Cleanthous, E., & Christou, C. (2013). Examining number sense and algebraic reasoning through cognitive styles. *Educational Studies in Mathematics, 83*(2), 205-223.

Clarke, D. M., & Roche, A. (2009). Students' fraction comparison strategies as a window into robust understanding and possible pointers for instruction. *Educational Studies in Mathematics, 72*, 127-138.

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.

Durkin, K., & Rittle-Johnson, B. (2014). Diagnosing misconceptions: Revealing changing decimal fraction knowledge. *Learning and Instruction*. doi: 10.1016/j.learninstruc.2014.08.003

Faulkner, V. (2009). The components of number sense: An instructional model for teachers. *Teaching Exceptional Children, 41*(5), 24-30.

Gersten, R., Jordan, N. C., Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 293–304.

Griffin, S. (2004). Building number sense with number words: A mathematics program for young children. *Early Childhood Research Quarterly, 19*, 173-180.

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences, 20*, 82-88.

Markovits, Z. & Sowder J. T. (1994). Developing number sense: An Intervention study in grade 7. *Journal for Research in Mathematics Education, 25*(1), 4-29.

Martin, M. O., Mullis, I. V. S., Foy, P., Olson, J. F., Erberber, E., Preuschoff, C., & Galia, J. (2008). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill: MA: TIMSS & PIRLS International Study Center, Boston College.

McIntosh, A., Reys, B. J., & Reys, R. E. (1992). A proposed framework for examining basic number sense. *For the Learning of Mathematics, 12*(3), 2-8.

McIntosh, A., Reys, B. J., Reys, R. E., Bana, J., Farrel, B. (1997). *Number Sense in School Mathematics: Student Performance in Four Countries*. Perth: Edith Cowan University.

Menon, R. (2004). *Elementary school children's number sense. International Journal for Mathematics Teaching and Learning*. Retrieved 10 January 2008 from: http://www.cimt.plymouth.ac.uk/journal/ramamenon.pdf

Office for standards in Education (2008). *Mathematics: Understanding the score*. London, UK: Author.

Newstead, K. & Murray, H. (1998). Young students' constructions of fractions. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22nd Conference of the International Group for the Psychology of Mathematics Education, 3*, 295-303. Stellenbosch, South Africa.

Pearn, C. & Stephens, M. (2004). Why do you have to probe to discover what Year 8 students really think about fractions. In I. Putt, R. Faragher & M. McLean. (Eds.). *Mathematics education for the third millennium: Towards 2010. Proceedings of the 27th Annual Conference of the Mathematics Education Research Group of Australasia.Townsville*, 27-30 June. (pp. 430-437).

Resnick, L. B., Nesher, P., Leonard, F., Magone, M., Omanson, S. and Peled, I. (1989). Conceptual bases of arithmetic errors: the case of decimal fractions. *Journal for Research in Mathematics Education, 20*(1), 8-27.

Reys, R., Reys, B., McIntosh, A., Emanuelsson, G., Johansson, B., & Yang, D. C. (1999). Assessing number sense of students in Australia, Sweden, Taiwan, and the United States. *School Science and Mathematics, 99*(2), 61-70.

Reys, R. E. & Yang, D. C. (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. & Shahrill, M. (2014). Understanding students' mathematical errors and misconceptions: The case of year 11 repeating students. *Mathematics Education Trends and Research, 2014*, 1-10.

Shahrill, M. (2013). Clustering of decimal misconceptions in primary and secondary classes. *International Journal of Humanities and Social Science, 3*(11)*, 58-65.*

Sood, S., & Mackey, M. (2014). Number sense instruction: A comprehensive literature review. *World Journal of Education, 4*(5), 58-67.

Sowder, J. T. (1992). Estimation and number sense. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp.371-389). New York: Macmillan.

Stacey, K. & Steinle, V. (1998). Refining the classification of students' interpretations of decimal notation. *Hiroshima Journal of Mathematics Education, 6*, 1-21.

Stein, M. K., Remillard, J., & Smith M. S. (2007). How curriculum influences student learning. In Frank K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp.319-369). NCTM: Information Age.

Şengül, S. & Gülbağcı, H. (2012). Evaluation of number sense on the subject of decimal numbers of the secondary stage students in Turkey. *International Online Journal of Educational Sciences, 4*(2), 296-310.

Tarr, J. E., Reys, R. E., Reys, B. J., Chavez, O., Shih, J., & Osterlind, S. J. (2008). The impact of middle grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39*(3), 247-280.

Tirosh, D. (2000). Enhancing prospective teachers' knowledge of children's conceptions: The case of division of fractions. J*ournal for Research in Mathematics Education, 31*(1), 5-25

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 557-628). Charlotte, NC: Information Age Publishing.

Widjaja, W., Stacey, K., & Steinle, V. (2011). Locating negative decimals on the number line: Insights into the thinking of pre-service primary teachers. *The Journal of Mathematical Behavior, 30*, 80–91.

Yang, D. C. (2005). Number sense strategies used by sixth grade students in Taiwan, *Educational Studies, 31*(3), 317-334.

Yang, D. C. & Li, M. N (2013). Assessment of animated self-directed learning activities modules for children's number sense development. *Journal of Educational Technology and Society, 16*(3), 44-58.

Yang, D. C., Li, M. N., & Lin, C. I. (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.

Yang, D. C., Reys, R. E., & Reys, B. J. (2009). Number sense strategies used by pre-service teachers in Taiwan, *International Journal of Science and Mathematics Education, 7*(2), 383-403.

◈◇◈