Teaching and Teacher Education 59 (2016) 403–413

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Assessing teachers' metacognition in teaching: The Teacher Metacognition Inventory

Yingjie Jiang [a,*], Lin Ma [a,b], Liang Gao [a]

[a] School of Psychology, Northeast Normal University, Changchun 130024, China
[b] School of Education, Anqing Normal University, Anqing 246133, China

HIGHLIGHTS

- A six-factor structure scale, Teacher Metacognition Inventory (TMI), measuring teacher metacognition was developed.
- The TMI had satisfactory reliability and validity.
- It is recommended for use in classroom teaching practice and teacher training to benefit teacher teaching and student learning.

ARTICLE INFO

Article history:
Received 4 December 2015
Received in revised form
5 July 2016
Accepted 11 July 2016

Keywords:
Teacher metacognition
Effective teaching
Reflective teacher
Effective teacher
Teacher professional development

ABSTRACT

Metacognition plays a pivotal role in teachers' professional development. However, the absence of effective instrument for measuring teacher metacognition has hampered researches in this area. Two studies were conducted to develop a valid scale - the Teacher Metacognition Inventory (TMI). Results from Study 1 with 412 middle school teachers showed a six-factor structure, with a satisfactory reliability and convergent validity. Results from Study 2 with 204 participants supported the structure, further revealed an acceptable criterion-related validity and discriminant validity. All findings suggested that the TMI was an effective instrument and can be used to assess teacher metacognition in educational practice.

1. Introduction

In the field of teaching and learning, the ultimate goal would be to make students learn well with the assistance of teachers' effective teaching. Metacognition plays an important role in this process (Perfect & Schwartz, 2002). It is highly believed that teacher metacognition significantly affect the procedure of teacher teaching and student learning (Prytula, 2008, 2012). Hartman (2001) argued that the method to maximize instructional effectiveness was "teaching with metacognition". Moreover, knowing what teachers know about their own teaching should be a starting point for a change in the professional development of teacher (Manning & Payne, 1996). However, the lack of appropriate measurements has stymied the research on teacher metacognition. Therefore, the purpose of this research was to develop a new instrument – Teacher Metacognition Inventory (TMI) - for measuring teacher metacognition and helping teachers be aware of the level of metacognition in their teaching.

Metacognition has important significance for learning and instruction in educational research and practice (e.g. classroom teaching and learning, teacher training). In educational contexts, metacognition is continually used to explain the process by which students/teachers learn to understand their thinking, with the notion that if they can regulate their thinking effectively, they will be better learners (Perfect & Schwartz, 2002). In the last decade, there is growing consensus that metacognition is of importance in successful learning (McCormick, 2003) and efficient teaching (Ben-David & Orion, 2013; Fathima, Sasikumar, & Roja, 2014).

Teacher metacognition can facilitate student learning. Metacognition is generally defined as how individuals monitor and control their own cognitive processes, which is closely related to executive function (Fernandez-Duque, Baird, & Posner, 2000). It is

* Corresponding author. School of Psychology, Northeast Normal University, No. 5268, Renmin Road, Nanguan District, Changchun, Jilin 130024 China.
E-mail address: jiangyj993@nenu.edu.cn (Y. Jiang).

http://dx.doi.org/10.1016/j.tate.2016.07.014
0742-051X/

central to self-regulated learning (Bjork, Dunlosky, & Kornell, 2013). Individuals who can regulate self behaviors effectively will benefit their outcomes. Metacognitive instruction has been proved that it can give some positive effects on students' problem solving skills and academic achievements, and metacognitive strategies are recommended to be taught to students (Safari & Meskini, 2015). Furthermore, teachers' self-awareness is an essential precondition to enhance students' opportunities to be metacognitive (Lee, Irving, Pape, & Owens, 2015). Accordingly, if teachers intend to teach students to think in a metacognitive way, they must be metacognitive themselves, as well as be clearly aware of their metacognition levels and characteristics firstly. Therefore, a valid instrument to assess teacher metacognition is urgently needed.

Teacher metacognition can promote their professional development. Self-regulated learning ability is also essential for teachers' professional growth during their entire career (Kramarski & Michalsky, 2009). Metacognitive teaching can optimize instructional effectiveness. Therefore, teacher professional development should be started with what teachers already know about their teaching. Moreover, recent researches revealed that metacognitive intervention could enhance teachers' teaching competency (Fathima et al., 2014). Teachers themselves also expressed a common desire to continue their professional improvements toward enhancing their abilities to make the metacognition as an integrated part of the curriculum (Ben-David & Orion, 2013). In a word, if there is a feasible way to improve teacher teaching and student learning, metacognitive teaching may be more impactful. Therefore, teachers should develop a meta-perspective on their instructional activities as a prerequisite.

However, the reality is that one cannot teach what one doesn't know. Hence, a teaching metacognition instrument will be needed to assist teachers to realize their strengths and weaknesses at different aspects of teaching activities. For that reason, the present research aimed at developing a new instrument to measure teacher metacognition specially.

1.1. Metacognition and teacher metacognition

Metacognition is generally defined as how individual monitor and control their cognitive process (Young & Fry, 2008). Rooted in Brown's (1978) and Flavell's (1976, 1979) theoretical perspectives, the foundational metacognition research has emerged several decades ago. There are two main theory frameworks in the field of metacognition research. Brown's (1978) perspective, as one of the two, emphasized knowledge and regulation of cognition. However, Flavell (1979), as the other, emphasizes the person, task, and strategy variances as the sub-components of metacognitive knowledge.

Synthesized views have emerged along with the research going deeper. For instance, some researchers proposed that metacognitive knowledge comprised declarative, conditional, and procedural knowledge (Schraw & Dennison, 1994; Schraw & Moshman, 1995). Alternatively, other researchers argued that metacognitive knowledge included personal variables, task variables and strategy variables (Peverly, Brobst, & Morris, 2002; Veenman & Spaans, 2005), while skills, such as planning, monitoring, and evaluation, should be included in metacognitive regulation (Huff & Nietfeld, 2009; Moos & Azevedo, 2009). Moreover, it is worth noting that, metacognitive experiences, as a new component of metacognition, were proposed at the same time (Flavell, Miller, & Miller, 2002). According to Flavell et al. (2002), metacognitive experiences refer to cognitive or affective experiences that pertain to a cognitive enterprise. Recently, researchers tended to adopt a comprehensive view which suggested that metacognition should include metacognitive knowledge, metacognitive experiences and metacognitive skills (Nussinson & Koriat, 2008; Zohar & Barzilai, 2013; Zohar & David, 2009).

Benefit from the development of the conceptual and empirical studies on metacognition, a number of metacognition researches have been conducted under educational situation (e.g., Ku & Ho, 2010; Manasia, 2015; Prytula, 2008, 2012; Spruce & Bol, 2015; Toci, Camizzi, Goracci, Borgi, Santis, & Coscia et al., 2015; Waters & Schneider, 2010; Yerdelen-Damar, Özdemir, & Ünal, 2015). When it comes to teacher metacognition, researchers suggest that effective teachers are "more metacognitive" (Duffy, Miller, Parsons, & Meloth, 2009) or possess "adaptive metacognition" which involves both the adaptation of self and environment in response to multifarious classroom variability (Lin, Schwartz, & Hatano, 2005; Manasia, 2015). Zohar (2006) emphasized the complexity of teacher metacognition. Except monitoring and regulating their cognitive activity, teachers, compared to students, have the additional missions of improving course learning, adjusting teaching strategies, and so on (Zohar, 2006). Recently, studies on teacher metacognition have made a rapid progress. Teachers' metacognitive knowledge about person variable was further divided into three sub-variables labeled as metacognitive content knowledge, metacognitive method knowledge, and metacognitive knowledge about students' knowledge (Yerdelen-Damar, Özdemir, & Ünal, 2015).

Based on the construct suggested by Flavell and his colleagues (Flavell et al., 2002), also drawing lessons from Zohar and Barzilai's research (Zohar & Barzilai, 2013), we adopt the notion that an integrated metacognition should comprise metacognitive knowledge, metacognitive experiences and metacognitive skills.

Metacognitive knowledge consists of personal variable, task variable and strategy variable. Personal variable refers to self-knowledge including knowledge of one's strengths and weaknesses. Teachers' personal variable reflects that they are conscious of the advantages and disadvantages of their teaching ability (e.g., how well do they know about the curriculum and why this happens). Task variable includes knowledge about the range and demands of tasks, as well as knowledge about the conditions and factors that influence the tasks. As for teachers, they should have knowledge of requirements for teaching tasks, criteria of different pedagogical tasks in teaching processes, and what successful ones should be. Strategy variable refers to knowledge about specific and general cognitive strategies along with an awareness of the potential use for approaching and fulfilling certain tasks. Teachers' strategy variable contains knowledge about pedagogy and tactics that teacher use to handle various situations in class. The metacognitive aspect of such knowledge lies in knowing where it can be used and in knowing when and how to apply it. Metacognitive experiences encompass feelings, judgments, and online task-specific experience. Teachers' metacognitive experiences include both cognitive and affective feelings when they go through the entire teaching activities (e.g., a sense of achievement when complete the teaching task successfully). Metacognitive skills comprise deliberate activities and the use of strategies for effort/time allocation, planning, monitoring and regulating cognitive processing, as well as evaluating the outcomes. In educational practice, teachers with metacognitive skills ensure that they can make teaching plans, monitor teaching behaviors, regulate teaching methods, evaluate teaching performances, and reflect teaching activities automatically. The TMI was designed based on this framework.

1.2. Domain-general vs. domain-specific teacher metacognition

Another crucial issue in metacognition research is whether metacognition is domain-general or domain-specific. To figure out this question, lots of studies have been done. Unfortunately, studies so far yielded inconsistent results. With respect to domain-general metacognition, individuals should present similar levels of metacognitive abilities across tasks. Some studies have reported evidence for such a general ability (Schraw & Nietfeld, 1998; Schraw, 1998; Schraw, Dunkle, Bendixen, & Roedel, 1995; Veenman & Beishuizen, 2004; Veenman & Verheij, 2003; West & Stanovich, 1997). But other researchers found no evidence for a general metacognitive ability (Kelemen, Frost, & Weaver, 2000; Poitras & Lajoie, 2013). Recently, researches suggested that it's not an either-or issue, as metacognition can both be domain-specific and domain-general (Neuenhaus, Artelt, Lingel, & Schneider, 2011; Stel & Veenman, 2014; Veenman & Spaans, 2005; Zohar & Barzilai, 2013).

For example, Veenman and Beishuizen (2004) found a stable monotonic growth of general metacognitive ability across discovery learning tasks for geography and biology, whereas Kelemen et al. (2000) noted that there might be no general metacognitive ability for the observation that individual differences in metamemory accuracy were not stable across tasks. Moreover, researchers found that metacognitive skills of the younger students appeared to be rather domain specific, whilst those of the older ones turned out to be general by nature (Veenman & Spaans, 2005). Analogously, Stel and Veenman (2014) claimed that metacognitive skills appeared to be predominantly general by nature over the years. Although studies mentioned above were conducted with student, we can trace that some metacognitive skills could be general for teacher probably. We agree with Schraw's view that “cognitive skills tend to be encapsulated within domains or subject areas, whereas metacognitive skills span multiple domains, even when those domains have little in common” (Schraw, 1998). As for teachers, especially regarding their metacognitive skills in teaching, their metacognition might be generalized across subjects.

In this research, we emphasized the domain-general feature of teacher metacognition, and the TMI was designed to assess this domain-general teacher metacognition. Considering the whole procedure of teaching activity, teachers severed in different subjects generally have some common teaching metacognitive abilities (e.g. making teaching schedule, monitoring their own teaching behavior, reflecting on teaching quality). Teachers in different domains need some general metacognitive skills to ensure that their plans and managements in classroom would be effective (Artzt & Armour-Thomas, 1998). And this general teacher metacognition is applicable cross-discipline.

1.3. Measurement of teacher metacognition

Numerous instruments have developed to examine metacognition (Table 1). Although some instruments could be used by adults, there is rarely a special instrument designed for teacher. In fact, as far as we know, there is only one metacognition scale designed solely for teachers − Metacognitive Awareness Inventory for Teachers (MAIT), developed by Balcikanli (2011).

The MAIT was developed from Schraw and Dennison's MAI (Schraw & Dennison, 1994). The MAI is a 52-item self-report inventory and each item is rated on 5-Point Likert-type scale ranging from “1 − always false” to “5 − always true” to report respondents' level of agreement with the items. Items were classified into eight sub-components subsumed under two broader categories, namely knowledge of cognition and regulation of cognition (Schraw & Dennison, 1994). Using the sample chosen from student teachers of English Language Teaching Program, Balcikanli (2011) modified MAI. After screening 42 items from MAI, teaching aspects were added to these items. For example, the item “I ask myself periodically if I am meeting my goals” changed into “I ask myself periodically if I meet my teaching goals while I am teaching”. The modified version of MAI, entitled MAIT, contained 24 items subsumed under 6 dimensions (i.e. declarative Knowledge, procedural Knowledge, conditional Knowledge, planning, monitoring and evaluating).

However, the MAIT is incapable of measuring teacher metacognition entirely because it disregarded metacognitive experiences, which have been considered as the most significant facet of metacognition, and act “as a mediator between teaching and learning” (Ben-David & Orion, 2013). Metacognitive experiences as a significant component of metacognition have traditionally received less attention from researchers (Efklides, 2006). According to recent development of metacognition research, an integrate metacognition should consist of three aspects: metacognitive knowledge, metacognitive skills and metacognitive experiences (Zohar & Barzilai, 2013). Therefore, the present research aimed at developing a new self-report instrument, the Teacher Metacognition Inventory (TMI), which designed to assess teacher metacognition integrally.

2. Methods

2.1. Study 1

2.1.1. Purpose of Study 1

Purpose of Study 1 was twofold: (1) form a final version of the TMI by item analysis and exploratory factor analysis (EFA); (2) examine reliability and validity of the TMI.

2.1.2. Participants

Participants (N = 430) were Chinese in-service middle school teachers taking part in a teacher training program. Before the study, we obtained the approval of school scholar committees. And all participants consented to take part in the study. Data of 18 participants were excluded for further analysis because they chose the

Table 1
Scales designed for assessing metacognition.

Study	Scale
Balcikanli, 2011	Metacognitive Awareness Inventory for Teachers
Chen, Gualberto, & Tameta, 2009	Metacognitive Reading Awareness Inventory
Meijer, Veenman, & van Hout-Wolters, 2006	Taxonomy of Metacognitive Activities
Mokhtari & Reichard, 2002	Metacognitive Awareness of Reading Strategies Inventory
O'Neil & Abedi, 1996	State Metacognitive Inventory
Pereira-Laird & Deane, 1997	Reading Strategy Use
Schmitt, 1990	Metacomprehension Strategy Index
Schraw & Dennison, 1994	Metacognitive Awareness Inventory
Sperling, Howard, Miller, & Murphy, 2002	Junior Metacognitive Awareness Inventory
Taasoobshirazi & Farley, 2013	Physics Metacognition Inventory
Weinstein & Palmer, 2002	Learning and Study Strategies Inventory
Yildiz, Akpinar, Tatar, & Ergin, 2009	Metacognition Scale

same number (TMI was a five-point scale ranging from 1 "strongly disagree" to 5 "strongly agree") through the whole scale or failing to complete the inventory (more than five items were blank). Finally, 412 participants were remained.

For the total 412 participants, the majority of them were females (72.1%) and from urban school districts (65.3%). They taught different grade in junior school (32.5% grade one, 43.0% grade two, and 24.5% grade three). The age of the participants ranged from 25 to 59 ($M = 37.41$, $SD = 7.66$). Participants had either a bachelor's degree (89.6%) or a master's degree (8.0%), with only 2.4% at the high school level. Years of teaching experience ranged from 1 to 35 ($M = 13.47$, $SD = 7.64$). The distribution of severed subject was as follows: Chinese (35.0%), maths (28.9%), English (27.7%), other (8.5%).

2.1.3. Instrument and procedure

A two-step procedure was used to compile items of the TMI. In the first phase, we reviewed literature related to definition and structure of metacognition, and adopt the three-component model for teacher metacognition (Ben-David & Orion, 2013).

The first component is teacher metacognitive knowledge consisting of knowledge of persons, tasks and strategies. All these knowledge were considered in the context of classroom teaching. Knowledge of persons refers to teachers' self-knowledge about the variables that influence his/her cognitive activities, as well as knowledge about the cognition of students. Knowledge of tasks refers to knowledge about how the nature of teaching task conditions, demands and goals influences cognitive activities. Knowledge of strategies refers to teachers' knowledge about thinking, instructing and problem-solving strategies that they might use in order to improving students learning. The second component is teacher metacognitive experiences. It refers to the emotional or affective experiences that pertain to teachers' cognitive activity in teaching. The third component is teacher metacognitive skill comprised of four sub-components (i.e. planning, monitoring, evaluating and debugging in teaching activities). Planning refers to setting teaching goals, selecting appropriate instruction strategies, and making predictions for the teaching activity. Monitoring indicates online awareness of cognitive process and performance of teaching activity. Evaluation is an assessment of the teaching products or efficiency. Debugging entails teachers' self-checking, reflection and re-correction of their teaching practice, which enable teacher to check the errors and correct them.

By taking all the components of metacognition into account, and drawing on the experiences of MAIT (Balcikanli, 2011), we compiled 53 items and developed an initial version of the TMI.

In the second phase, all 53 items were sent to three experts (one professor, two lecturers, with research interests in metacognition) to get advice on the content validity of the inventory. Based on their feedback, several wording issues of items were addressed.

After doing all of this, the initial 53-item TMI was administered to participants. Responses were given on a 5-point Likert-type scale, ranging from 1 "Strongly Disagree" to 5 "Strongly Agree". Data was collected from the teachers using a paper-and-pencil format, taking no more than 30 min.

Data analysis was threefold: (1) analyze items discrimination index, and eliminate the items with poor discrimination; (2) explore structure of the TMI by conducting EFA and parallel analysis; (3) examine convergent validity and reliability of the TMI.

2.1.4. Results

2.1.4.1. Missing data. According to Rubin (1976), three types of missing exist in academic research: missing completely at random, missing not at random, and missing at random. If the percent of missing data is minimal or the data is missing completely at random, statistical analyses could be largely robust to missing data (Muthén, Kaplan, & Hollis, 1987). Although the percent of missing data in the present study was only 1.73%, all missing data were imputed with "3", which indicated the middle extent of the agreement (i.e., "Neutral"), in order to ameliorate estimation bias and increase statistical accuracy.

2.1.4.2. Item analysis. The single best measure of effectiveness of an item is its ability to separate participants who vary in their degree of knowledge of the material tested and their ability to use it. Within several indexes that successfully compute item discrimination, discrimination index is a popular and valid measure for item quality (Fred, 1973). Item discrimination indicates the extent to which participants perform well on an item corresponds to perform well on the whole test. Since all items in a test are intended to cooperate to generate an overall test score, any item with bad discrimination undermines the test.

The Discrimination Index (D), a strict indicator of item discrimination, can be calculated by ranking the participants according to total score and then selecting the top 27 percent and the lowest 27 percent in terms of total score. For each item, the percentage of participants' response in the upper and lower groups is calculated. Specifically, D is computed from 27 percent high and low scoring groups in the test. Subtract the number of successes by the low group on the item from those by the high group, and divide this difference by the size of a group. The range of this index is from +1 to −1. Using Truman Kelley's "27 percent of sample" group size, values of 0.4 and above are regarded as high and less than 0.2 as low by Ebel (1954, 1972). More specifically, Ebel (1972) recommended guidelines of classical test theory (CTT) item analysis about D (i.e., 0.0−0.19 = to be revised; 0.2−0.29 = acceptable; 0.3−0.39 = good; >0.4 = excellent), and suggested that items with low or negative D should be discarded.

The TMI was a 5-point Likert-type instrument, different from dichotomous variable scale. Item in the TMI was not highly right or wrong for each participant, the way which item discrimination normally treated did not work. To address this problem, we followed Brown's (Brown, 2000) trick to calculate this kind of item discrimination. The formula of the item discrimination index we used was as follows:

D = average score of 27% upper group/5 - average score of 27% lower group/5.

After the discrimination index, D, was calculated, we compared D with 0.2. If D is greater than 0.2, the item will be reserved. Instead, if D is smaller than 0.2, the item will be dropped. Based on this criterion of psychometry, 18 items (i.e., item 1, 2, 3, 5, 10, 13, 16, 19, 20, 25, 26, 32, 35, 39, 41, 42, 50, 53) were eliminated, and 35 items ($D > 0.2$) were retained. In order to explore the structure of these items, EFA was conducted with the left 35 items.

2.1.4.3. Exploratory factor analysis. Exploratory factor analysis (EFA) is a popular technique to model latent factors, and particularly appropriate for scale development where little theoretical basis exists for specifying the number and patterns of common factors. Considering the nature of metacognition is multifaceted and complex (see Duffy et al., 2009, p. 242; Valli & Buese, 2007), EFA is adopted to identify the underlying relationships between the measured variables of metacognition by using IBM SPSS Statistics 22.0. According to suggestions about absolute sample size for EFA: 100 = poor, 200 = fair, 300 = good, 500 = very good, 1000 or more = excellent (Comrey & Lee, 2013), the sample size ($n = 412$) in this study is appropriate.

Prior to EFA, Kaiser-Meyer-Olkin measure of sampling adequacy (KMO-test) and Barlett's Sphericity test were conducted to identify the appropriateness of the data for factor analysis. According to

Kaiser (1974), KMO value of the scale: smaller than 0.50 = unacceptable, greater than 0.50 = miserable, greater than 0.60 = mediocre, greater than 0.70 = middling, greater than 0.80 = meritorious, greater than 0.90 = marvelous.

When considering the types of factoring, principal component analysis (PCA) and principal axis factoring (PAF) are two generally recommended extraction methods used to estimate the factor loadings and unique variances of the model. PFA assumes that components are uncorrelated and the communality of each item sums to 1 over all components, implying that each item has 0 unique variance, while PAF allows the variance of each item to be composed to be a function of both item communality and nonzero unique item variance. Briefly, PAF takes into account the random error that is inherent in measurement and seeks the least number of factors which can account for the common variance of a set of variables, whereas PCA fails to do so. Brown (2009) recommended using PAF when theoretical ideas about relationships between variables exist, whereas PCA should be used if the goal of the research is to explore patterns in the data. Taking all this into consideration, PAF was applied to identify underlying constructs of the items in this study.

In addition, factor rotation is used to improve interpretability and utility of the factor matrix, and the goal of it is to rotate factors in multidimensional space to arrive at a solution with best simple structure. Orthogonal rotation increases the meaning of the factors. Given simplicity and conceptual clarity as the advantage of orthogonal rotation, orthogonal rotation and varimax were used in the present study.

To confirm the factors of metacognition, procedure of EFA ran several times. Each time we ran the procedure to decide which/if any items should be eliminated next based on: (1) item did not have a strong loading on one factor (>0.4); (2) number of items in the factor was less than three (Field, 2013). Finally, 5 items were eliminated (i.e., item 8, 12, 21, 34, 43) and 30 items were remained. A four-factor structure was modeled. However, this structure showed mixed components loaded on one factor. For example, items originally designed for planning and reflection were loaded on one factor. This made the interpretation of each factor complicated. Taking all this into consideration, we limited the extract number of factors at seven using PCA. Then, only one item (i.e., item 27) loaded on factor 7, and one other item's content (i.e., item 24) mismatched the others. So we deleted this two items and limited the extract number of factors at six using PCA again with the left 28 items. Finally, a six-factor structure was modeled (Table 2).

Result of EFA formed a final six-factor, 28-item version of the TMI. All six factors combined explained 62.051% of the variance. Specifically, factor 1 to 6 explained 12.718%, 11.814%, 11.307%, 10.315%, 8.032%, and 7.866% respectively after the rotation. Furthermore, it should be noted that each time conducting EFA, value of KMO was greater than 0.93 and was significant at level of 0.05.

After considering the theoretical structure of metacognition and analyzing items in each factor, we named the six subscales as teacher metacognitive experiences (TME, Factor 1), metacognitive knowledge about pedagogy (MKP, Factor 2), teacher metacognitive reflection (TMR, Factor 3), metacognitive knowledge about self (MKS, Factor 4), teacher metacognitive planning (TMP, Factor 5), and teacher metacognitive monitoring (TMM, Factor 6).

2.1.4.4. Parallel analysis. In order to re-verify the number of factors of the TMI after EFA, a parallel analysis (PA) was conducted with the final 28 items. PA, a method based on the generation of random variables, compares the observed eigenvalues extracted from the correlation matrix with those obtained from uncorrelated normal variables (Ledesma & Valero-Mora, 2007), is one of the most important methods for determining the number of factors to retain in EFA (Hayton, Allen, & Scarpello, 2004; Thompson & Daniel, 1996). The rationale underlying PA is that significant components from observed data with a valid potential factor structure would have bigger eigenvalues than the parallel components derived from random uncorrelated data which having the same number of variables and sample size (Lautenschlager, 1989). Currently, Glorfeld (1995) recommended using a given percentile of eigenvalues (e.g., the 95th of the distribution of eigenvalues), not the mean of those, derived from the random data.

The ViSta program was used to complete PA. ViSta is a free and open statistical system focused on statistical visualization methods (Young, Valero-Mora, & Friendly, 2006). It offers PA two models and two computation methods. Principal axis factor analysis and normal data simulation were chosen respectively in this study. The cut-off percentile was 95th and the number of samples to be simulated was 100 as default. Then, observed and random eigenvalues were calculated (Table 3).

Results in Table 3 indicated that the first 6 observed eigenvalues are larger than those random generated by PA (for both the mean and 95th percentile criteria), and suggested that six components should be retained for interpretation. This re-verified the number of factors of TMI and attested that the six-factor structure of the TMI was reasonable.

2.1.4.5. Convergent validity analysis. In order to confirm the validity of the 28-item TMI, convergent validity, as one of the important indicators of construct validity, was analyzed. Correlations between subscales and the TMI, as well as the inter-correlation of the subscales were calculated (Table 4).

The correlation coefficient, estimating relationships between all components of the TMI, were computed in the correlation matrix. On one hand, correlation coefficients between the subscales were all significant, ranging from 0.321 to 0.706. Results showed that relationships between all sub-components of the TMI were positive. This indicated there was covariation between the dimensions, and these six dimensions were indeed measuring the same content (i.e., teacher metacognition). On the other hand, correlation coefficients between subscales and the TMI ranged from 0.740 to 0.841 and were all significant. These results showed relatively strong relations between subscales and the TMI, indicating that all six dimensions were well contributive to the TMI and matched the integrated construct consistently. In sum, results of correlation analysis demonstrated a good convergent validity of the TMI.

2.1.4.6. Reliability analysis. To evaluate the internal consistency, Cronbach's alpha coefficient was calculated. The analysis showed good internal consistency of the TMI ($\alpha = 0.936$). The Cronbach's alpha coefficient of the six sub-scales (i.e., TME, MKP, TMR, MKS, TMP, and TMM) were 0.784, 0.812, 0.839, 0.769, 0.771 and 0.820, respectively. The average inter-item correlation of the TMI was 0.343 ($SD = 0.011$), which suggested the items have good discrimination once again. The average corrected item-total correlation (CITC) of the TMI was 0.534 ($SD = 0.059$), indicated the items had good consistency. The value of Cronbach's alpha if item deleted ranging from 0.927 to 0.932, attested that the internal consistency would not increase when removed any item in the TMI. Collectively, results suggested that the TMI had satisfactory reliability.

2.1.5. Discussion

The primary aim of Study 1 was to develop an original instrument to measure teacher metacognition using item discrimination, EFA, PA, and reliability analysis. Study 1 examined the psychometric properties of the TMI and showed promising results. Firstly, all 28

Table 2
Final factor loading matrix for EFA model.

Item no.	Inventory item	Factor					
		1	2	3	4	5	6
47	I always worry about students feel tedious in my classroom.	0.757					
37	I am worried that I can't control the pace of classroom teaching well.	0.742					
28	When my classroom teaching fails, I always feel anxious.	0.666					
51	When I successfully complete the classroom teaching task, I feel very relaxing.	0.642					
33	When I make a satisfactory teaching program, I feel unquenchable excited.	0.561					
46	I clearly know demonstration can make the abstract knowledge concrete.		0.727				
49	I know exactly catechetic method can inspire students to think.		0.677				
45	I know that group discussion do not apply to the case when time is short for teaching.		0.654				
48	I know very well that interact with students can make them concentrate.		0.630				
23	I re-evaluate the appropriateness of my teaching goals after each lesson.			0.746			
7	I reflect on whether my teaching design is appropriate after each lesson.			0.632			
38	I ask myself how well I have accomplished my teaching goals after each lesson.			0.629			
15	I ask myself if I have considered other possible teaching methods after each lesson.			0.516			
22	I reflect on the teaching effect after each lesson.			0.510			
52	I reflect on whether my teaching performance is proper after each lesson.			0.490			
40	I re-evaluate to what extent the teaching goals have been met after each lesson.			0.429			
17	I know well about the concepts, principles and methods of the subject I teach.				0.774		
18	I can quickly adjust my condition before I begin the lesson.				0.643		
9	I know very well about why I have some certain advantages in teaching.				0.596		
44	I am well aware of my weaknesses in teaching.				0.485		
31	I prepare for the unexpected situations that may arise in the classroom.					0.720	
30	I always set a specific teaching goal for each lesson.					0.637	
29	I design the specific teaching program in advance for each lesson.					0.582	
11	I pay attention to the changes of my emotion in class.						0.646
6	I check teaching progress periodically to figure out whether it meets my expectation.						0.610
4	I ask myself about how well I am doing while I am teaching.						0.584
14	I ask myself periodically if my teaching method is applicable while I am teaching.						0.534
36	I check regularly to what extent students comprehend the content while I am teaching.						0.402

Note: all item number in this table was original in Study 1. And they were all renumbered from 1 to 28 in Study 2.

Table 3
Observed and random eigenvalues.

Factor	Observed eigenvalue	Random mean eigenvalue	Random 95th percentile eigenvalue
1	9.870	0.588	0.674
2	2.172	0.514	0.565
3	0.950	0.458	0.509
4	0.537	0.408	0.452
5	0.436	0.363	0.397
6	0.362	0.323	0.361
7	0.267	0.282	0.315
8	0.238	0.288	0.279
…			

Table 4
Pearson correlation matrix of the TMI.

	TME	MKP	TMR	MKS	TMP	TMM	TM
TME	–						
MKP	0.490**	–					
TMR	0.380**	0.629**	–				
MKS	0.321**	0.614**	0.677**	–			
TMP	0.401**	0.553**	0.706**	0.596**	–		
TMM	0.690**	0.623**	0.536**	0.500**	0.509**	–	
TM	0.740**	0.807**	0.827**	0.745**	0.756**	0.841**	–

Note: $N = 412$. TME = teacher metacognitive experiences. MKP = metacognitive knowledge about pedagogy. TMR = teacher metacognitive reflection. MKS = metacognitive knowledge about self. TMP = teacher metacognitive planning. TMM = teacher metacognitive monitoring. TM = teacher metacognition.
$^{**}p < 0.01$ (two-tailed).

items retained had an acceptable discrimination (i.e., D > 0.2). Secondly, items' estimated factor pattern loadings on six factors were all large (i.e., >0.4). Thirdly, the analysis of Cronbach's alpha coefficient indicated good reliability of the TMI and all six subscales. Although encouraging results had achieved, some issues still needed to be interpreted.

As for strategies existed for dealing with weighted items in item analysis, there are also alternative methods (Brown, 1996). The method we adopted in the present study was the easiest to understand and carry out, with no confusing interpretation (Brown, 2000). Items with loading lower than 0.4 were eliminated from the inventory, because this loading was considered as too low to be accepted (Hair, Black, Babin, & Anderson, 2010). These steps ensured strong factorial validity of the TMI. In addition, a noticeable result is the good preliminary construct validity, given the six dimensions were well contributive to the integrated construct of the

TMI.

When referring to the content of each factor as the result of EFA, there was an item (i.e., item 24, "I know teachers should help students master the strategies and methods suitable for learning the subject") loaded onto factor which did not match the other items. This situation is extremely common because metacognition is a relational rather than a definite concept as noted by Zohar and David (2009). Additionally, results of PA re-verified the number of factors of TMI and suggested the current structure was reasonable. Nevertheless, given that more evidence regarding the psychometric properties of the TMI is necessary, Study 2 was carried out to test criterion-related validity, discriminant validity, and the factorial validity of the TMI with a separate sample.

2.2. Study 2

2.2.1. Purpose of Study 2

Study 2 was conducted to provide more evidence for the validity of the TMI. More specifically, criterion-related validity and discriminant validity were tested, while confirmatory factor analysis (CFA) was also conducted to check the stability of the six-factor structure TMI.

2.2.2. Participants

Data was collected from a sample of 226 Chinese in-service middle school teachers taking part in a teacher training program same as that of Study 1. But participants in Study 2 were different from that of Study 1. All participants consented to take part in the study. Data of participants who: (1) answered the scale with the same number, (2) failed to write the demographic information about the year of their teaching experience, or (3) did not complete the scales (more than three items were blank in each scale), were excluded for analysis. Finally, 22 participants' data were discarded, remaining 204 participants.

For these 204 participants, 73.5% were females and 67.6% from urban school districts. They severed in different grades of junior school (38.2% grade one, 38.7% grade two, and 23.0% grade three). Participants ranged in age from 26 to 59 ($M = 35.46$, $SD = 8.55$). They earned either Bachelors (66.7%) or Masters (26.5%) degrees, with 6.9% no higher education degree. Participants ranged in teaching experience from 1 to 35 ($M = 9.61$, $SD = 8.92$). The distribution of severed subject was as follows: Chinese (32.4%), maths (30.4%), English (22.5%), other (14.7%).

2.2.3. Instruments

2.2.3.1. Teacher Metacognition Inventory (TMI). TMI was a self-report scale, including 28 items. These item were classified into six subscales: (1) Teacher metacognitive experience (TME, 5 items), (2) Metacognitive knowledge about pedagogy (MKP, 4 items), (3) Teacher metacognitive reflection (TMR, 7 items), (4) Metacognitive knowledge about self (MKS, 4 items), (5) Teacher metacognitive planning (TMP, 3 items), and (6) Teacher metacognitive monitoring (TMM, 5 items). Participants responded to each item using a rating scale, ranging from "1 (strongly disagree)" to "5 (strongly agree)". All items were listed in Table 2 and were renumbered from 1 to 28.

2.2.3.2. Teachers' teaching motivation scale (TTMS). The TTMS, developed by Wang (2004), estimating teacher motivation in their teaching activities, was used to test the criterion-related validity of the TMI. The TTMS, a 5-point Likert-type self-report scale, consisted of 30 items, grouped into six subscales: major interests and hobbies, career interests and hobbies, preferences of students, social expectation and approval, competition with colleague, and approval and respect from students. Each of the subscales included 5 items. Participants responded to each item ranging from "1 (strongly disagree)" to "5 (strongly agree)".

2.2.4. Procedure

Participants were required to respond to each item "considering their current situation". The two scales (i,e,. the TMI and TTMS) were both paper-and-pencil format, completed in classroom, taking no more than 30 min altogether.

2.2.5. Results

2.2.5.1. Missing data. The missing data in Study 2 was 1.35%, handled in the same manner as Study 1. All missing data were filled with "3 (Neutral)".

2.2.5.2. Criterion-related validity. To investigate criterion-related validity of the TMI, the entire sample of participants completed both the TMI and TTMS. Taking the TTMS as the criterion of teacher metacognition, Pearson product-moment correlation coefficient between scores of the TMI and TTMS was calculated. Relation between the TMI and TTMS appeared a relatively high correlation and was statistically significant ($r = 0.620$, $p < 0.01$). This positive correlation indicated a predictive relationship, which can be exploited in practice, between the TMI and TTMS.

2.2.5.3. Discriminant validity. Previous researches have revealed that teachers with more teaching experiences would be more metacognitive than those who had less teaching experiences (Stewart, Cooper, & Moulding, 2007). Based on these findings, we compared the differences between teachers with four years teaching experience or less (group 1, $n = 106$) and teachers with five years teaching experience or more (group 2, $n = 98$). Results confirmed the hypothesis, for the total score of the TMI for group 1 ($M = 122.245$, $SD = 10.376$) was significantly lower than group 2 ($M = 125.459$, $SD = 5.268$), $t\ (158.395) = 2.82$, $p < 0.05$. Furthermore, differences of scores of six subscales between the two groups were also compared. Consistent with the result of total score, results showed group 1 was significantly lower than group 2 on TME, MKS and TMR ($ps < 0.05$). Moreover, on TMP, difference between group 1 ($M = 12.943$, $SD = 1.644$) and group 2 ($M = 13.337$, $SD = 1.573$) was marginally significant ($p = 0.083$). But group 1 and group 2 were not significantly different ($ps > 0.05$) on TMM and MKP. In sum, these results indicated that teachers with more teaching experiences would be more metacognitive than those who had less teaching experiences, and the TMI had good discriminant validity.

2.2.5.4. Confirmatory factor analysis. We conducted a CFA with maximum likelihood estimation on the sample using AMOS 21 software to test the structure of the TMI. A variety of statistics was used and CFA analysis revealed an adequate to good model fit, $\chi^2\ (335) = 645.399$, $p < 0.001$, $\chi^2/\text{df} = 1.927$, CFI = 0.886, IFI = 0.888, TLI = 0.871, PNFI = 0.702, PCFI = 0.785, RMSEA = 0.068, C.I. = 0.060–0.075. Although χ^2 value was statistically significant for the model, χ^2 statistic is sensitive to sample size and, therefore, may reject well-fitting models. For this reason, less emphasis was placed on the χ^2 value compared to the other model fit statistics, while χ^2/df and RMSEA indices were primarily relied (Bollen, 1989; Browne & Cudeck, 1993) to assess model fit. Good model fit is indicated by a χ^2/df values lower than 2 or RMSEA value lower than 0.05, while RMSEA values between 0.05 and 0.08 are considered reasonable fit (Hu & Bentler, 1999; Schweizer, 2010). These two indices in the present study suggested an acceptable to good model fit. In addition, estimated factor loadings were all greater than 0.4 (between 0.498 and 0.842) and statistically significant at 0.001. Collectively, these results supported the six-dimensional structure of the TMI.

2.2.6. Discussion

The purpose of Study 2 was to provide more evidence on validity of the TMI. Correlation analysis manifested a significant correlation between the TMI and TTMS. Previous researches have also proved the positive correlation between teachers' metacognition and motivation (Santisia, Magnanob, Hichya, & Ramacib, 2014; Vrieling, Bastiaens, & Stijnen, 2012). Moreover, participants with more teaching experiences appeared more metacognitive than those who had less teaching experiences. This is consistent with the finding "metacognition improves significantly with years of teaching experience" (Stewart et al., 2007). These results suggested that the TMI had good discriminant validity.

Criterion-related validity was tested using teachers' teaching motivation as the criterion of teacher metacognition. In fact, a large number of studies have demonstrated that there is a significant positive correlation between metacognition and motivation (Landine & Stewart, 1998; Santisia et al., 2014; Vrieling et al., 2012; Zepeda, Richey, Ronevich, & Nokes-Malach, 2015). "Motivation and metacognition are strongly intertwined" (Hull & du Boulay, 2015). For this reason, it is reasonable to choose teaching motivation as the criterion for testing the criterion-related validity of teacher metacognition. As for discriminant validity, experienced teachers were supposed to think and perform more metacognitively than inexperienced teachers (Stewart et al., 2007). Moreover, metacognitive knowledge and skills are conducive to promote teacher professional development (Adams & Mabusela, 2014; Graham & Phelps, 2003; James, 1987). Therefore, the significant difference between teachers with less teaching experience and teachers with more teaching experience provided a convincing evidence for discriminant validity of the TMI.

Additionally, the CFA results manifested an acceptable to good model fit of the TMI. CFA is a powerful statistical tool can often be used in the process of scale development to examine the latent structure of an instrument (Brown, 2015). Nevertheless, some clarifications were required for the CFI and TLI indices. Values of CFI and TLI were below 0.90 in this study, while Hu and Bentler (1999) suggest those higher than 0.90 are considered acceptable. This may be partly caused by the small sample size. When sample size is small ($N < 250$), some combinational rules are less preferable (Hu & Bentler, 1999). Although it is common to modify model fit that using the modification indices, as Browne (2001) suggests, this virtually results in CFA becoming partly exploratory. This might be inappropriate as the model modifications may produce a significantly better fitting model simply as a result of a capitalization on chance (MacCallum, Roznowski, & Necowitz, 1992). Moreover, the use of CFI as structural equation modeling fit indexes is, to some extent, problematic (Rigdon, 1996). Rigdon (1996) argues that CFI seems to be appropriate in more exploratory contexts, whereas RMSEA appears to be a better choice "when researchers wish to determine whether a given model fits well enough to yield interpretable parameters and to provide a basis for further theory development" in more confirmatory contexts. Collectively, examining the multiple indices of CFA, results indicated that the six-dimension structure of the TMI was reasonable and model fit was adequate.

In sum, Study 2 found further evidence for validity of the TMI. More specifically, results of Study 2 indicated the six-factor structure of TMI was acceptable, and had good criterion-related validity as well as discriminant validity.

3. General discussion and conclusion

The aim of present research was to construct an original instrument - Teacher Metacognition Inventory (TMI) - measuring teacher metacognition. Two studies were conducted to test and validate the TMI with two separate samples. The TMI aimed to estimate the whole facets of teacher metacognition in teaching activities. Study 1 utilized EFA to explore the structure of the TMI and formed a final 28-item version. Meanwhile, convergent validity and reliability of the final version were examined. In order to examine the validity in-depth, Study 2 further tested the criterion-related validity and discriminant validity of the TMI, as well as the stability of the six-factor structure. Collectively, the two studies provided converging evidence of adequate psychometric properties for the 28-item TMI.

Study 1 brought out a six-dimension structure of the TMI from the final solution of EFA. Although the MAIT also included six factors: Declarative Knowledge, Procedural Knowledge, Conditional Knowledge, Planning, Monitoring, and Evaluating (Balcikanli, 2011), it is different from the TMI. Considering the nature of metacognition, metacognitive experiences, which were not measured in the MAIT, were a very important component of metacognition (Efklides, 2006; Flavell et al., 2002). Efklides (2006) argued that measures including metacognitive experiences can increase the reliability and validity of the measurement. Consequently, compared to the MAIT, the TMI, consisted of six subscales, could measure teacher metacognition more integrally and effectively.

Although Study 1 has yielded out a set of satisfactory outcomes, it was accompanied with limitations. A potential limitation is relationships between the subscales. Magnitude of correlation between subscales might suggest some dimensions were relatively dependent. And this might lead to a poor interpretation to the structure of the TMI to some extent. Future research is needed to replicate these results to confirm the structure in practice.

Study 2 provided further evidences for the validity of the TMI. Results of CFA revealed that the six-factor structure TMI was stable, and showed an adequate to good model fit. Convergent and discriminant validity are both considered as subtypes of construct validity. But, neither one alone is sufficient for establishing construct validity. Given that convergent validity has been examined in Study 1, Study 2 tested the discriminant validity. Differences between teachers with less teaching experience and those with more teaching experience were analyzed. Findings were consistent with previous study (Stewart et al., 2007). Teachers with more teaching experience appeared to be more metacognitive than those who have less teaching experience. In addition, results of this study revealed a significant correlation between teacher metacognition and motivation. Along with the previous research, these findings confirmed the positive relationship between metacognition and motivation (Santisia et al., 2014; Vrieling et al., 2012). As noted by Papleontiou-Louca (2003), "metacognition refers to both people's awareness and control, not only of their cognitive processes, but of their emotions and motivations as well." At that point, future research could explore the ways to improve teacher metacognition by enhancing their teaching motivation and further examine the relation between these two variances.

This study is of great significance for education research and practice related to metacognitive teaching and learning.

First, the present study is beneficial for more efficient interaction between teaching and learning. Classroom teaching requires frequent interaction between teachers and students. Expert teachers are capable of monitoring their class more automatically. The TMI was designed for teachers to self-evaluate their teaching abilities including monitoring the efficiency of their teaching performance. Some items in the TMI, for example "I check regularly to what extent students comprehend the content while I am teaching", can assess how sensitive teachers are to students' learning performance in classroom. According to Lee et al. (2015), teachers' awareness is often a means of helping students to become self-regulated learners. Teachers who get high rank at such items might be capable of regulating their teaching activities in time. Briefly, the TMI findings can help teachers regulate their teaching methods timely and dynamically to optimize their teaching quality and facilitate students' learning in classroom practice.

Second, the TMI plays an important role in helping teachers realize their strengths and weaknesses in teaching activities, and that would benefit reflective teaching. Based on the results of this study, the TMI is capable of assessing different aspects of teacher metacognition, including planning, monitoring, reflecting, experiences during teaching, knowledge about pedagogy. Hence, the TMI can be used as a check list for reflective teaching and make teachers' self-inspection in their teaching performance more accessible. For example, “I ask myself if I have considered other possible teaching methods after each lesson”. Besides, teachers also can compare the different performance patterns between themselves and proficient teachers (Han, Cetin, & Matteson, 2016), or conduct an analysis of their classroom-capturing videos (Toci et al., 2015), with the aid of the TMI, to reflect their teaching. Consequently, all of these will benefit teachers to make themselves more metacognitive and professional, as that reflective teaching practices could improve teachers' metacognitive teaching abilities (Robinson, Anderson-Harper, & Kochan, 2001; see also; Adams & Mabusela, 2014).

Third, the availability of this valid multifaceted teacher metacognition scale may also have important practical implications on making teacher training more specific. Ordinarily, most present teaching improvement programs tend to focus on more general teaching techniques or teaching theories. Some metacognition training is confronted with the lack of appropriate materials (Ben-David & Orion, 2013). The TMI can help teacher trainers design more specific teacher-training contents to ameliorate the training effects. Seeing that items in the TMI are all elaborate contents (e.g., I know exactly catechetic method can inspire students to think), and covered diverse aspects of teaching activities, it can provide references for teacher trainers to make the training contents more concrete and functional.

Forth, the TMI will be conducive to teacher instructors formulating more individualized guidance to improve teachers' self-regulated teaching. Metacognitive intervention strategies, such as planning, monitoring and evaluating, can enhance teachers' teaching competency (Fathima et al., 2014). The TMI can be used to explore teachers' different metacognition characteristics. Teacher instructors can observe the teaching processes of lecturers, like peer observation (Tenenberg, 2016), using the TMI as an analytical framework to discriminate their proper and improper teaching behaviors, further to guide them in appropriate ways. In addition, given that the TMI is in a position to compare the composition between lower and higher metacognition level teachers, teacher instructors may confirm what dimensions that can make a key difference in effective teaching. Subsequently, they can formulate more targeted programs for various level teachers and make the training more purposeful and efficient.

There are some limitations to our study. The first limitation relates to the generalizability. Considering that two sample of participants were all in-service middle school teachers, the generalizability of the TMI was restricted. More specifically, none of the two studies addressed the issue of invariance across groups (e.g., area, grade levels, etc.). For instance, it remains unknown that to what extent the TMI is an effective instrument across elementary school teachers and college teachers, or pre-service teachers. This question was presented for future researches. Furthermore, providing supplemental validity evidences in various ways would strengthen the appropriateness and effectiveness for using the TMI. Another limitation of our study concerns the nature of metacognition. Given that metacognition is very complicated, any certain scale is impossible to measure it radically, and so was the TMI. The TMI was designed for measuring teachers' general metacognition to some extent and helping teachers to be aware of their metacognition level in their teaching. Future research may replenish and further improve the TMI when using it in practice.

To conclude, teacher metacognition, symbolizing levels of self-awareness and self-regulation in teaching activities, plays an important role in teacher professional development and has been recognized as a powerful factor that impacts the quality of instruction in classrooms. Teachers who have metacognitive awareness and metacognitive abilities will benefit both their teaching and student learning. The present research provided an effective instrument for measuring teacher metacognition. Results from the two studies suggested that the TMI has competent psychometric properties, and can be used reliably and validly in educational practice.

References

Adams, J. D., & Mabusela, M. S. (2014). A metacognitive approach to teacher Development: Supporting national professional diploma in education (Npde) students. *Mediterranean Journal of Social Sciences, 5*(15), 289–296.

Artzt, A. F., & Armour-Thomas, E. (1998). Mathematics teaching as problem solving: A framework for studying teacher metacognition underlying instructional practice in mathematics. *Instructional Science, 26*, 5–25.

Balcikanli, C. (2011). Metacognitive awareness inventory for teachers (MAIT). *Electronic Journal of Research in Educational Psychology, 9*(3), 1309–1332.

Ben-David, A., & Orion, N. (2013). Teachers' voices on integrating metacognition into science education. *International Journal of Science Education, 35*(18), 3161–3193.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *vol. 1. Advances in instructional psychology* (pp. 77–165). Hillsdale, NJ: Erlbaum Associates.

Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River. NJ: Prentice Hall Regents.

Brown, J. D. (2000). Statistics corner. Questions and answers about language testing statistics (how can we calculate item statistics for weighted items?). Shiken: JALT testing & evaluation SIG. *Newsletter, 3*(2), 19–21. Retrieved from the World Wide Web at http://jalt.org/test/bro_6.htm. on 5 July 2015.

Brown, J. D. (2009). Principal components analysis and exploratory factor analysis - Definitions, differences and choices. *Shiken: JALT Testing & Evaluation SIG Newsletter, 13*(1), 26–30.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilford Publications.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*, 111–150.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Chen, M. H., Gualberto, P. J., & Tameta, C. L. (2009). The development of metacognitive reading awareness inventory. *TESOL Journal, 1*, 43–57.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.

Duffy, G. G., Miller, S., Parsons, S., & Meloth, M. (2009). Teachers as metacognitive professionals. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 240–256). New York: Taylor & Francis.

Ebel, R. L. (1954). Procedures for the analysis of classroom tests. *Educational and Psychological Measurement, 14*, 352–364.

Ebel, R. L. (1972). *Essentials of educational measurement* (1st ed.). New Jersey: Prentice Hall.

Efklides, A. (2006). Metacognitive experiences: The missing link in the self-regulated learning process. *Educational Psychology Review, 18*(3), 287–291.

Fathima, M. P., Sasikumar, N., & Roja, M. P. (2014). Enhancing teaching competency of graduate teacher trainees through metacognitive intervention strategies. *American Journal of Applied Psychology, 2*(1), 27–32.

Fernandez-Duque, D., Baird, J. A., & Posner, M. I. (2000). Executive attention and metacognitive regulation. *Consciousness and Cognition, 9*(2), 288–307.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics 4th*. London: Sage publications.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231–236). Hillsdale, NJ: Erlbaum.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. *American Psychologist, 34*, 906–911.

Flavell, J. H., Miller, P. H., & Miller, S. A. (2002). *Cognitive development* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Fred, P. (1973). Validity of the discrimination index as a measure of item quality.

Journal of Educational Measurement, 10(3), 227–231.
Glorfeld, L. W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement, 55*, 377–393.
Graham, A., & Phelps, R. (2003). Being a teacher: Developing teacher identity and enhancing practice through metacognitive and reflective learning processes. *Australian Journal of Teacher Education, 27*(2), 11–24.
Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). New Jersey: Pearson Prentice Hall.
Han, S., Cetin, S. C., & Matteson, S. M. (2016). Examining the pattern of middle grade mathematics teachers' performance: A concurrent embedded mixed methods study. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(3), 387–409.
Hartman, H. J. (2001). Teaching metacognitively. In H. J. Hartman (Ed.), *Metacognition in learning and instruction: Theory, research and practice* (pp. 149–172). New York: Kluwer Academic Publishers.
Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods, 7*, 191–205.
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.
Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*(2), 161–176.
Hull, A., & du Boulay, B. (2015). Document motivational and metacognitive feedback in SQL-Tutor. *Computer Science Education, 25*(2), 238–256.
James, C. (1987). In *Cognition and metacognition in teachers' professional development. Paper presented at the annual meeting of the american educational research association*. Washington, DC. (ERIC Document Reproduction Service No. ED282844).
Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36.
Kelemen, W. L., Frost, P. J., & Weaver, C. A., III (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory and Cognition, 28*, 92–107.
Kramarski, B., & Michalsky, T. (2009). Investigating preservice teachers' professional growth in self-regulated learning environments. *Journal of Educational Psychology, 101*(1), 161–175.
Ku, K. Y. L., & Ho, I. T. (2010). Meta-cognitive strategies that enhance critical thinking. *Metacognition and Learning, 5*(3), 251–267.
Landine, J., & Stewart, J. (1998). Relationship between metacognition, motivation, locus of control, self-efficacy, and academic achievement. *Canadian Journal of Counseling, 32*(3), 200–212.
Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24*, 365–395.
Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation, 12*(2), 1–11.
Lee, S. C., Irving, K., Pape, S., & Owens, D. (2015). Teachers' use of interactive technology to enhance students' metacognition: Awareness of student learning and feedback. *Journal of Computers in Mathematics & Science Teaching, 34*(2), 175–198.
Lin, X., Schwartz, D. L., & Hatano, G. (2005). Toward teachers' adaptive metacognition. *Educational Psychologist, 40*(4), 245–255.
MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modification in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111*, 490–504.
Manasia, L. (2015). Creating A-HA moments in teaching practice. Routine versus adaptive metacognition behaviors in teachers. In *Proceedings of the scientific conference AFASES* (pp. 1255–1262).
Manning, B. H., & Payne, B. D. (1996). *Self-talk for teachers and students: Metacognitive strategies for personal and classroom use*. Needham Heights, MA: Allyn & Bacon.
McCormick, C. B. (2003). Metacognition and learning. In I. B. Weiner, Series Ed., W. M. Reynolds, & G. E. Miller (Eds.)*, vol. Eds.. Handbook of Psychology: Vol. 7. Educational psychology* (pp. 79–102). Hoboken, NJ: Wiley.
Meijer, J., Veenman, M. V. J., & van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving: Development of a taxonomy. *Educational Research and Evaluation, 12*, 209–237.
Mokhtari, K., & Reichard, C. A. (2002). Assessing students' metacognitive awareness of reading strategies. *Journal of Educational Psychology, 94*(2), 249–259.
Moos, D. C., & Azevedo, R. (2009). Self-efficacy and prior domain knowledge: To what extent does monitoring mediate their relationship with hypermedia learning? *Metacognition and Learning, 4*, 197–216.
Muthén, B. O., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika, 55*, 107–122.
Neuenhaus, N., Artelt, C., Lingel, K., & Schneider, W. (2011). Fifth graders metacognitive knowledge: General or domain-specific? *European Journal of Psychology of Education, 26*, 163–178.
Nussinson, R., & Koriat, A. (2008). Correcting experience-based judgments: The perseverance of subjective experience in the face of the correction of judgment. *Metacognition and Learning, 3*, 159–174.
O'Neil, H. F., & Abedi, J. (1996). Reliability and validity of a state metacognitive Inventory: Potential for alternative assessment. *Journal of Educational Research, 89*(4), 234–245.
Papleontiou-Louca, E. (2003). The concept and instruction of metacognition. *Teacher Development: An international Journal of Teachers' Professional Development, 7*(1), 9–30.
Pereira-Laird, J. A., & Deane, F. P. (1997). Development and validation of a self-report measure of reading strategy use. *Reading Psychology: An International Journal, 18*, 185–235.
Perfect, T. J., & Schwartz, B. L. (2002). *Applied metacognition*. New York: Cambridge University Press.
Peverly, S. T., Brobst, K., & Morris, K. S. (2002). The contribution of reading comprehension ability and metacognitive control to the development of studying in adolescence. *Journal of Research in Reading, 25*, 203–216.
Poitras, E. G., & Lajoie, S. P. (2013). A domain-specific account of self-regulated learning: The cognitive and metacognitive activities involved in learning through historical inquiry. *Metacognition and Learning, 8*, 213–234.
Prytula, M. P. (2008). *Scholarship epistemology: An exploratory study of teacher metacognition within the context of successful learning communities*. University of Saskatchewan.
Prytula, M. P. (2012). Teacher metacognition within the professional learning community. *International Education Studies, 5*(4), 112–121.
Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling, 3*(4), 369–379.
Robinson, E. T., Anderson-Harper, H. M., & Kochan, F. K. (2001). Strategies to improve reflective teaching. *Journal of Pharmacy Teaching, 8*(4), 49–58.
Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*, 581–592.
Safari, Y., & Meskini, H. (2015). The effect of metacognitive instruction on problem solving skills in iranian students of health sciences. *Global Journal of Health Science, 8*(1), 150–156.
Santisia, G., Magnanob, P., Hichya, Z., & Ramacib, T. (2014). Metacognitive strategies and work motivation in Teachers: An empirical study. *Procedia-Social and Behavioral Sciences, 116*(21), 1227–1231.
Schmitt, M. C. (1990). A questionnaire to measure children's awareness of strategic reading processes. *The Reading Teacher, 43*, 454–461.
Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*, 113–125.
Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460–475.
Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology, 87*, 433–444.
Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review, 7*(4), 351–371.
Schraw, G., & Nietfeld, J. (1998). A further test of the general monitoring skill hypothesis. *Journal of Educational Psychology, 90*, 236–248.
Schweizer, K. (2010). Some guidelines concerning the modeling of traits and abilities in test construction. *European Journal of Psychological Assessment, 26*(1), 1–2.
Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*(1), 51–79.
Spruce, R., & Bol, L. (2015). Teacher beliefs, knowledge, and practice of self-regulated learning. *Metacognition and Learning, 10*(2), 245–277.
Stel, M. V. D., & Veenman, M. V. J. (2014). Metacognitive skills and intellectual ability of young adolescents: A longitudinal study from a developmental perspective. *Journal of Psychology of Education, 29*(1), 117–137.
Stewart, P. W., Cooper, S. S., & Moulding, L. R. (2007). Metacognitive development in professional educators. *The Researcher, 21*(1), 32–40.
Taasoobshirazi, G., & Farley, J. (2013). Construct validation of the physics metacognition inventory. *International Journal of Science Education, 35*(3), 447–459.
Tenenberg, J. (2016). Learning through observing peers in practice. *Studies in Higher Education, 41*(4), 756–773.
Thompson, B., & Daniel, L. G. (1996). Factor analytic evidence for the construct validity of scores: A historical overview and some guidelines. *Educational and Psychological Measurement, 56*(2), 197–208.
Toci, V., Camizzi, L., Goracci, S., Borgi, R., Santis, F. D., Coscia, L., et al. (2015). Designing, producing and exemplifying videos to support reflection and metacognition for in-service teacher training. *Journal of E-Learning and Knowledge Society, 11*(2), 73–89.
Valli, L., & Buese, D. (2007). The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal, 44*(3), 519–558.
Veenman, M. V. J., & Beishuizen, J. J. (2004). Intellectual and metacognitive skills of novices while studying texts under conditions of text difficulty and time constraint. *Learning and Instruction, 14*, 621–640.
Veenman, M. V., & Spaans, M. A. (2005). Relation between intellectual and metacognitive skills: Age and task differences. *Learning and Individual Differences, 15*(2), 159–176.
Veenman, M. V. J., & Verheij, J. (2003). Technical students' metacognitive skills: Relating general vs. Specific metacognitive skills to study success. *Learning and Individual Differences, 13*, 259–272.
Vrieling, E., Bastiaens, T., & Stijnen, S. (2012). Effects of increased self-regulated learning opportunities on student teachers' motivation and use of metacognitive skills. *Australian Journal of Teacher Education, 37*(8), 102–117.
Wang, M. (2004). *The study on influence of new curricular reform to teachers' sense of teaching efficacy, teaching motivation and occupational stress*. Southwest Normal University. A Chinese Master's dissertation, published in China National Knowledge Infrastructure (CNKI). Retrieved 27 July 2016 from: http://www.

cnki.net/KCMS/detail/detail.aspx?QueryID=6&CurRec=1&recid=&filename=2004085533.nh&dbname=CMFD9904&dbcode=CMFD&pr=&urlid=&yx=&v=MjI3NThSOGVYMUx1eFlTN0RoMVQzcVRyV00xRnJDVVJMeWZZT1JvRnlubFU3M0pWMTI3R3JPd0c5VFBySkViUEk.

Waters, H. S., & Schneider, W. (2010). *Metacognition, strategy use, and instruction*. New York, NY: Guilford Press.

Weinstein, C. E., & Palmer, D. R. (2002). *LASSI user's manual: For those administering the learning and study strategies inventory*. Clearwater: H & H Pub.

West, R. F., & Stanovich, K. E. (1997). The domain specificity and generality of overconfidence: Individual differences in performance estimation bias. *Psychonomic Bulletin and Review, 4*, 387–392.

Yerdelen-Damar, S., Özdemir, Ö. F., & Ünal, C. (2015). Pre-service physics teachers' metacognitive knowledge about their instructional practices. *Eurasia Journal of Mathematics, Science & Technology Education, 11*(5), 1009–1026.

Yildiz, E., Akpinar, E., Tatar, N., & Ergin, O. (2009). Exploratory and confirmatory factor analysis of the metacognition scale for primary school students. *Kuram ve Uygulamada Eğitim Bilimleri, 9*(3), 1591–1604.

Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning, 8*(2), 1–10.

Young, F. W., Valero-Mora, P., & Friendly, M. (2006). *Visual statistics seeing data with dynamic interactive graphics*. NJ: Wiley and Sons.

Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. J. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology, 107*(4), 954.

Zohar, A. (2006). The nature and development of teachers' metastrategic knowledge in the context of teaching higher order thinking. *Journal of the Learning Sciences, 15*, 331–377.

Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. *Studies in Science Education, 49*(2), 121–169.

Zohar, A., & David, A. B. (2009). Paving a clear path in a thick forest: A conceptual analysis of a metacognitive component. *Metacognition and Learning, 4*(3), 177–195.