
This version is available at https://strathprints.strath.ac.uk/55634/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
The Assessment of Metacognition in Children Aged 4-16 Years: A Systematic Review

Louise Gascoine¹, Steve Higgins¹ and Kate Wall²

¹Durham University, School of Education
²University of Strathclyde, School of Education

louise.gascoine@durham.ac.uk; s.e.higgins@durham.ac.uk; kate.wall@strath.ac.uk

*Corresponding author Miss Louise Gascoine, School of Education, Durham University, Leazes Road, Durham, DH1 1TA, UK
louise.gascoine@durham.ac.uk

Abstract

This article presents the results of a systematic review of methods that have been used to measure or assess metacognition in children aged 4-16 years over a 20-year period (1992-2012). It includes an overview of the types of tool and methods used linked with the ages of the participants targeted and how metacognition and associated concepts are defined. 2721 records were identified through systematic searching; 525 articles or reports were full text screened, resulting in 149 included studies reporting 84 distinct tools or methods. Of these four were excluded from further analysis after appraisal for reliability, validity and replicability. The final number of methods and tools for metacognitive assessment included in the analysis is 80. The key findings of this review include:

- Self-report measures (including questionnaires, surveys and tests) comprise 61% of the included tools.
- Observational methods that do not rely on prompting to ‘think aloud’ (Think Aloud Protocols) have only been used with students aged 9 years and under;
- Information about reliability and validity is not always given or given accurately for different tools and methods;
- The definition of metacognition in a particular study relates directly to its assessment and therefore its outcomes: this can be misaligned.

Keywords:
Metacognition, systematic review, research methods
The Assessment of Metacognition in Children Aged 4-16 Years: A Systematic Review

1 Background & Aims

This article presents the results of a systematic review (Gough, Oliver, & Thomas, 2012) of methods that have been used to measure or assess metacognition in school-aged children (4-16 years). It therefore provides a synthesis of recent literature in English focussing on the measurement or assessment of metacognition, with particular relevance for education. There is a wealth of research claiming to measure or assess metacognition, but the different methods have not previously been synthesised in a systematic way. This systematic methodological review of methods therefore identifies the different tools and methods used to assess metacognition in the last 20 years and their reported reliability and validity. Additionally, this review aims to facilitate an exploration of the potential links between:

- The types of tool or method used and the ages of the participants they are used with; and
- How metacognition and associated concepts are defined and the types of tool or method used.

Before addressing these aims it is vital to consider the complexity of defining metacognition, exploring prominent debates within the wider field. What follows in the review explores how metacognition has been defined and operationalized in the included tools and methods. This review seeks to be explicit about the decision-making processes applied by the authors throughout, with an understanding that in such a vast and complicated field there may be dissension. However, by presenting the logic and rationale behind the decision making process in this review it is hoped
that the usefulness of the overview that it provides will outweigh any contention, thus providing a practical starting point for future reviews in this area.

1.1 Defining Metacognition

Prior to presenting the research questions, design and methods for this review it is important to recognise the complexity involved in defining metacognition. In order to situate this review, its questions and findings it is essential to think about how metacognition was and is defined. Specifically, how Flavell defined metacognition in 1976, how definitions have since developed and how metacognition has been operationalized in successive research.

Flavell (1976, p.232) defined metacognition as: “[referring] to one’s own knowledge concerning one’s own cognitive processes and products or anything related to them”. Metacognition has become something of a paradox, now spanning a variety of disciplines including education, psychology and linguistics. Consequently, there are many debates about what metacognition is, how it should be measured and how it develops. Wilson (1999, para 9) noted that even Flavell himself did not have a detailed proposal for defining metacognition in the late 1980s, over a decade after he first introduced the term: “Flavell (1987) admitted that: ‘none of us has yet come up with deeply insightful, detailed proposals about what metacognition is.’ (1987: 28).”

Since Flavell (1976) coined the term ‘metacognition’ there has been widening debate about what metacognition actually is and also how it can be assessed. The complexities of this have become increasingly clear over the last 30 years. Metacognition is something of a “fuzzy” concept (Wellman, 1985), when one digs below the surface of the popular practice centred definition thinking about thinking, there are many competing perspectives about metacognition and associated concepts
such as self-regulation and executive function. These competing claims about metacognition require a “multiplistic perspective” (Hofer & Sinatra, 2010: p. 117).

Executive function and executive control are terms more widely used in developmental psychology than in educational psychology (Borkowski, Chan, & Muthukrishna, 2000), but the ideas cover much of the same conceptual ground. Garner’s (2009) study for example, comparing executive function and self-regulation indicates points of overlap and difference. The executive functions of planning, impulse control, and motivational drive significantly predicted cognitive strategy use, metacognitive strategy use, and academic effort regulation. However, in Garner’s (2009) study attributional and affective components of self-regulated learning did not correlate with executive functions. It is beyond the scope of this review to include executive control and executive function instruments, although this would be a valuable next step. For an extended discussion of the relationships between metacognition, self-regulation and executive function see Moseley et al. 2005 (pp. 187-198 and pp. 243-249). The following presents a summary discussion of some of the intersections between metacognition and associated concepts (including self-regulation) and explores established subdivisions of metacognition.

1.1.1 Metacognition and Self-Regulation.

Returning to the ‘fuzziness’ of metacognition, one of the clearest aspects of metacognition is perhaps that fact that it is so multifarious. Almost 30 years after Flavell, Efklides (2008) defined metacognition by referring back to Flavell’s (1979) definition, but added that “metacognition is multifaceted”. Efklides’ (2008) definition encompassed all of metacognitive experiences, metacognitive knowledge, the monitoring of cognition, metacognitive strategies and metacognitive skills. The
inclusion of multiple concepts relating to metacognition in this definition underlines the complex and interlinked nature of metacognition. The complexity of metacognition is further increased when terms including metacognition and self-regulation are used interchangeably and without adequate or explicit consideration of their intersections and differences (Dinsmore, Alexander, & Loughlin, 2008; Hofer & Sinatra, 2010; Moseley et al., 2005; Schunk, 2008). Careless use of terminology can lead to misperception, especially if there are no clear accompanying explanations. With this in mind, the subsequent paragraphs explore intersections between metacognition and self-regulation.

Despite many questions about metacognition and its intersections with self-regulation, there is no doubt that the question of which concept (metacognition or self-regulation) is superordinate of the other is dominant in the field (Veenman, 2007; Veenman, Van Hout-Wolters, & Afflerbach, 2006). There is continued debate about where the definitions for metacognition and self-regulation, if separated, intersect or are distinct. The issue of ascendancy, or not, for metacognition and self-regulation is often the ‘elephant in the room’. Debate around what comes first and which term, if either, is dominant has spanned over two decades and it is widely recognised as remaining largely unresolved (Kistner et al., 2010; Robson, 2010; Veenman, 2007; Veenman et al., 2006). A hierarchical approach to exploring the intersections between metacognition and self-regulation is not the most appropriate approach given the recognised complexity of metacognition. The section instead adopting a “multiplistic perspective” in line with the approach taken by Hofer and Sinatra (2010).

Metacognition and self-regulation are intrinsically linked; the fuzziness of existing definitions of metacognition and self-regulation do however leave it very much open to researcher interpretation in terms of how these links are portrayed.
Debate around what comes first and which term (metacognition or self-regulation), if either, is dominant has spanned over two decades and it is widely recognised as remaining largely unresolved (Kistner et al., 2010; Robson, 2010; Veenman, 2007; Veenman et al., 2006). Veenman et al. (2006) raised pertinent questions about this relationship between metacognition and self-regulation, presenting debate about whether self-regulation is subordinate to metacognition or whether self-regulation is actually superordinate to metacognition. In 2007, Veenman noted the content of more recent definitions of self-regulation and the inclusion of metacognitive knowledge and skills within this inferring that self-regulation is the overarching concept. Boekaerts (1999) also proposed a model with self-regulation as the major construct of which the use of metacognitive knowledge and skills are a part of, but do not have the central role. In another example Veenman, Elshout, and Meijer (1997, pp. 187-188) described self-regulatory activities as “representatives of metacognitive skilfulness”, inferring that metacognition is overarching and that there is a direct link between definitions of metacognitive skilfulness and self-regulation.

Popularly regarded definitions of self-regulation, including Zimmerman’s (1995) description, state that self-regulation is more than metacognition (both knowledge and skill). This ‘more than metacognition’ stems from the notion of self-regulation as involving “students’ underlying sense of self-efficacy and personal agency” (Zimmerman, 1995, p. 220). Zimmerman asserted that these are present in addition to metacognition. Zimmerman (1995) explained the necessity of self-regulation particularly clearly, but his definition also highlighted the necessity of also having metacognitive knowledge and skill. Zimmerman (1995) emphasised however that the possession of metacognitive knowledge and skill does not infer automatic ability to self-regulate this knowledge and skill.
The perspective that self-regulation is the overarching concept may be popularly regarded in the literature, but this section has demonstrated that there are often clear links and references to the concept of metacognition, implying that this is the (under) arching or perhaps the enabling concept. For example, without metacognitive and being able to recall this, the notion of regulation in the moment would be somewhat less grounded. In 2008 Dinsmore et al., contemplated debate around defining metacognition and associated concepts in detail, their literature review focussed on the concepts of metacognition, self-regulation and self-regulated learning. Dinsmore et al., (2008) concluded that explicitly stating the differences between metacognition, self-regulation and self-regulated learning is inherently risky and that there is often a need to make inferences from literature where details were lacking or underspecified. Inferences from the literature are similarly often required in defining metacognition and facets of it (including knowledge and skilfulness). What follows explores in summary these subdivisions of metacognition.

Subdivisions of metacognition

Having noted the ‘fuzzy’ nature of metacognition (Wellman, 1985), it is important to consider the different ways that metacognition has been subdivided in the literature. Metacognition is popularly divided into two components: knowledge of cognition and regulation of cognition (Lu, 1995; Shamir, Mevarech, & Gida, 2009; Sperling, Howard, Miller, & Murphy, 2002; Wilson, 1999; Yildiz, Akpinar, Tatar, & Ergin, 2009) or meta-cognitive knowledge and skilfulness (Veenman, Kok, & Blöte, 2005; Veenman & Spaans, 2005; Veenman, Wilhelm, & Beishuizen, 2004). Linked to this division, the regulation of cognition is described by Schmitt and Sha (2009, p. 256) as “…meta-cognitive control (or regulation), and includes problem solving”. There are clear links here between popular definitions of self-regulated learning and this definition of metacognitive control (or regulation),
which may also be described as part of metacognitive skillfulness (Veenman et al., 2005). The relationship between meta-cognitive knowledge and skillfulness with meta-cognitive beliefs and experiences, particularly the affective aspects of these beliefs, are also complex (e.g. Efklides (2006)). It is not the purpose of this review to arbitrate between these affective differences, but to note them and then be as transparent as possible in data extraction about how different definitions and conceptions are related to the tools and techniques used to assess meta-cognition.

Reflecting on Flavell’s (1979) divisions, metacognition it is popularly presented as comprising three phenomena metacognitive knowledge, metacognitive experiences and metacognitive skills or skilfulness (Desautel, 2009; Efklides, 2008; Efklides & Vlachopoulos, 2012; Veenman & Elshout, 1999). Efklides (2008) and Efklides and Vlachopoulos (2012) presented distinction between these facets particularly clearly and explored interactions between them. Efklides (2008) defined the three components of metacognition as follows:

- **Metacognitive knowledge:** “declarative knowledge stored in memory and comprises models of cognitive processes, such as language, memory and so forth” (p. 278. It is also described as involving knowledge of person, task, strategy and goals. Efklides and Vlachopoulos (2012) further condensed this to knowledge of person, task and strategy.

- **Metacognitive experiences:** “what the person is aware of and what he or she feels when coming across a task and processing the information related to it (Efklides, 2001 xx, 2006)” (p. 279). Efklides and Vlachopoulos (2012) further described metacognitive experiences as including metacognitive feelings (of difficulty, satisfaction, knowing, confidence) and judgments or estimates (e.g. estimate of effort, judgement of learning).

- **Metacognitive skills:** “the deliberate use of strategies (i.e. procedural knowledge) in order to control cognition…executive control…related to metacognitive regulation; that is both monitoring and control.” (p. 280).

Efklides and Vlachopoulos (2012) referenced the definition of metacognitive
skilfulness given by Veenman and Elshout (1999) and referred to “procedural knowledge manifested in peoples behaviour” (p. 228).

Beginning with metacognitive knowledge, what follows briefly summarises debate in the field around defining metacognitive knowledge, metacognitive experiences and metacognitive skills.

The roots of conceptualising metacognitive knowledge and the division of it into person, task and strategy are popularly regarded as being initiated by Flavell (1976) (Neuenhaus, Artelt, Lingel, & Schneider, 2011). Neuenhaus et al. (2011) described the person variable in terms of self and others; task in terms of knowledge of task demands and strategy in terms of knowledge of strategies. Neuenhaus et al. (2011, p. 165) explained that Paris, Lipson, and Wixson (1983) and Brown (1978) further subdivided metacognitive about strategy into declarative, procedural and conditional knowledge as follows:

> In accordance with the three metacognitive knowledge dimensions proposed by Anne Brown (1978), they differentiated between declarative strategy knowledge, referring to knowledge on “what” measures can be taken to solve a task, procedural strategy knowledge on “how” to realize these measures, and conditional strategy knowledge regarding the circumstances of a strategies effectiveness (“when” to apply a strategy).

Subdivisions of metacognitive knowledge have been developed further than person, task and strategy and declarative, procedural and conditional. Pintrich (2002) divided strategic knowledge or metacognitive knowledge of strategy into different types of strategies, which based on the work of Weinstein & Mayer (1986) were: rehearsal (e.g. repeating words over and over to remember), elaboration (e.g. mnemonics for memory, summarising, paraphrasing…) and organisational (e.g. outlining, concept mapping, note taking).
Efklides has been at the forefront of research in metacognitive experiences since the early 2000’s. Efklides (2002) stated that metacognitive experiences are online metacognition comprise “ideas, feelings, judgments and metacognitive knowledge evoked during problem solving [...] metacognitions available in working memory” (p. 20). Whitebread et al. (2009) in their Cambridgeshire Independent Learning in the Foundation Stage Coding Framework (C.Ind.Le) coded for “Emotional and motivational regulation” (p. 80), the “[expression] of positive or negative emotional experience of a task” – this is not dissimilar to feelings of difficulty or familiarity generally described as metacognitive experiences.

Veenman and colleagues have explored metacognitive skilfulness via the reportedly online method of TAPs, examples include Prins, Veenman, and Elshout (2006), van der Stel and Veenman (2010) and (Veenman et al., 2005). In the most recent example in this group van der Stel and Veenman (2010) divide into four sub-categories: orientation, planning and systematic orderliness, evaluation and elaboration. van der Stel and Veenman (2010, p. 221) exemplified evaluation as including monitoring, whereas in other conceptualisations monitoring and evaluation are explicitly separated. Despite the majority view that MS and the associated metacognitive strategies (planning, monitoring, control, evaluation) are best assessed ‘online’ there is some evidence in the literature to support metacognitive knowledge of these metacognitive strategies. Pintrich (2002, p. 220) noted “students can have knowledge of various meta-cognitive strategies that will be useful to them in planning, monitoring, and regulating their learning and thinking”. Consequently, there is potentially an argument for the offline assessment of metacognitive knowledge of metacognitive strategies that would normally be encompassed within definitions of metacognitive skills.
Even in this brief exploration of metacognition and how it and associated concepts like self-regulation are defined, it is clear that metacognition is a multifarious concept. This multifariousness required a pragmatic and transparent approach to the research design in this review, in particular around the data extraction of how different tools and methods defined and operationalized metacognition.

2 Research Question, Design and Methods

The central research question for this review is:

- What different research or assessment tools have been used explicitly to measure or assess metacognition in school aged children (4-16 years) in the last 20 years?

The methods that have been employed in this systematic review are based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Moher, Liberati, Tetzlaff, & Altman, 2009). The rigorous nature of the PRISMA statement was adopted to maintain quality and integrity especially during the search and screening processes.

The focus of this review is on the tool or method stated by the authors as the measure or assessment of metacognition, as opposed to a more typical systematic review which focuses on the results or effects of a given metacognitive intervention or comparing the results of different interventions (Torgerson, 2003). Systematic methodological reviews to date lie mainly in the field of health and social care (e.g. Brandstätter, Baumann, Borasio, and Fegg (2012) who review ‘life assessment instruments”; or Berne et al. (2013) who look at assessment instruments for measuring cyber-bullying). We felt that the field of meta-cognition was sufficiently
broad and complex, but also readily identifiable to justify a similar methodological review.

2.1 The search process

After defining the research question and thinking about the intended parameters of the search, pilot searches using key words and strings were completed in ERIC (Education Resources Information Center) and BEI (British Education Index) in order to refine the search strategy and to limit results to a manageable numbers of records for screening. Searches were completed for eight key databases: (AEI (Australian Education Index), BEI, ERIC, First Search ECO (Electronic Collections Online), First Search Journal Articles, PsychArticles, PsychINFO and Web of Knowledge). Detailed information showing the search strings used and limits applied can be found in Appendix A.

2.2 Inclusion criteria

In order to complete the screening process in a systematic and transparent way, clear criteria for the inclusion of records from the beginning of the review process were defined in relation to the research question. The inclusion and indeed exclusion criteria were based on the categories below and Table 1 shows how they were applied. Table 1 also lists examples of records that were excluded because they did not meet the inclusion criteria listed.

- The date of record
- What is being measured in the record
- The sample population in the record
- An empirical data set being present in the record
- The language in which the record is available
2.3 The screening process

The screening process was lengthy, but rigour at this stage was important in order to maintain the integrity of the review process. Appendix B contains a table showing the exact numbers included and excluded from each database at each stage of the screening process. An inductive process was adopted so as to respond to findings within the search and screening process, but consistency was key and when decisions were made they had to be applied in the same way to all records. The first author completed the first stage screening, for this stage the title and abstract for each record were scrutinised to see if they were on topic (i.e. about metacognition or a specified closely related concept like self-regulation) and that the sample was potentially in the correct age group (i.e. school aged, age 4-16 years). To calculate inter-rater reliability 20% of the 2089 original records were double-screened in the first stage screening by the second author, an inter-rater agreement of 98% was recorded. After this initial screening, the list of records classified as unsure were reviewed by all three authors. Individual records were discussed until consensus was reached. If there was uncertainty, records were included in order that they could be looked at in more detail in the second stage screening.

Second stage screening involved detailed full text screening; this focussed primarily on the methodology sections of the records because this information would be key in the next stage of data extraction. Based on the structure used by Dignath, Buettner, and Langfeldt (2008) the records at this stage were coded for the following variables in order to include or exclude them:

- The full reference details – for ease of reference and accurate record keeping
- A definition of metacognition – was this present, and clear?
- The sample characteristics – age group and educational setting
• Methodological information – was there clear information about the method or tool that had been used? Did it appear to be replicable from the information given?

Records were included, excluded or placed into a category labelled unsure. These records (n = 39) were subsequently double screened by the second and third authors. Records were discussed until all parties reached total agreement.

2.4 Data Extraction and Quality Appraisal

Data extraction for each tool or method was performed using a template and completed from the earliest available record (with detailed methodological information) for each tool or method. In some cases this was a record that had been added to the total via citation searches. This mainly applied to records that would not have been picked up in the original searches due to falling outside of the specified dates. For example Jacobs and Paris (1987) is included as the first record detailing the Index of Reading Awareness (IRA) but was not initially identified through the systematic search process.

The template for data extraction for the 84 tools or methods in the final data extraction is illustrated in Figure 1. The data extracted in this example are for the Inventory of Metacognitive Self-Regulation (IMSR) first referred to in the data extracted records by Howard, McGee, Shia, and Hong (2000b). Tools or methods were allocated to groups according to their methodological similarities (this classification is included in Section 3). For example, which tools or methods are questionnaire based, or based on the completion of a particular task or set of tasks. These broad categories are listed below, it is important to note that tools or methods do not always exclusively fit into just one category.

1. Questionnaires, surveys, self-report, tests
2. Observational methods
3. Teacher ratings
4. Interviews and focus groups
5. Task based methods
6. Multi-method approaches

2.5 *Results of the search process*

Search results are illustrated below in Figure 2.

2.6 *Application of Inclusion Criteria*

It was evident from the initial screening of the final included records here were multiple records to data extract for particular tools or methods. For example, Think Aloud Protocol(s) (TAP(s)) were cited as a method used in 18 separate records, the Index of Reading Awareness (IRA) and the Motivated Strategies for Learning Questionnaire (MSLQ) were individually cited in 12 and 9 included records each respectively. Therefore, rather than data extracting from each of the 152 (149 post reliability and validity checking) included records they were summarised in terms of the tool or method that they used. Similar tools were data extracted concurrently, the method or tool that had been used was identified and data were extracted under the heading of the tool or method. Some records uniquely cited a tool or method, these records were data extracted individually. In total 36 studies were excluded during the data extraction phase because it was realised that they did not contain sufficient data for analysis (including not focussing on the assessment of metacognition), they duplicated information available in other records, or because due to human error detail had been missed that would have excluded them earlier.
2.7 Results of the Quality Appraisal

An appraisal of the reliability, validity and replicability appraisal of the tools or methods as part of the final data extraction was important, given the methodological focus of this review. Tools were excluded at this stage because they were not replicable (i.e., there was not sufficient published information to make replication possible), or if there was no information given or available regarding both reliability and validity.

What follows in Table 2 is based on Coffield, Moseley, Hall, and Ecclestone (2004) analysis of learning styles instruments. It presents each of the 84 tools and methods included after the final screening; it indicates whether or not they are replicable and highlights the different types of reliability and validity reported. These have been divided into the eight most frequent main types in the included records:

- Reliability: Internal consistency, test-retest and inter-rater
- Validity: Construct, face, content, criterion and ecological

Some of the included records list ways of reporting reliability and validity data that are not reported in the above list. One example is that of parallel forms reliability. Sperling et al. (2002) focuses on testing two forms of the same tool in one experiment; the Junior Metacognitive Awareness Inventory (JrMAI), versions A and B. None of the records with less commonly reported forms of reliability were excluded, all of these examples contained other types, too many to report within the scope of Table 2 and this review.

Records were deemed replicable if they referenced other records that replicated the tool in part or full, or in the case of computer programmes if the method was based in a computer programme or a software package it was assumed that it could therefore be replicated through use of the software. Five tools or methods that
did not meet the replicability criterion and/or had no information about reliability and/or had no information about validity were excluded at this stage and are shaded in the table.

The final number of included tools is 80. Although four methods or tools were excluded at this final stage (excluded tools shaded grey in Table 2), this only led to three records being excluded from the final total. Fortunato, Hecht, Tittle, and Alvarez (1991) (HISQ, item number 15 in Table 2) had been added in as a citation search so its exclusion was reflected in the numbers given in Figure 2. The three excluded records (with reasons shown in Table 2) were:

- Carr, Alexander, and Folds-Bennett (1994) – A strategy card sort and individual interviews (item number 76 in Table 2).
- Erbas and Okur (2012) – Clinical interview (item number 7 in Table 2)
- Rahman, Yasin, Ariffin, Hayati, and Yusoff (2010) – Metacognitive skills and metacognitive development questionnaire (item number 47 in Table 2).

3 Summary of findings relating to the methods used

The purpose of this review is threefold: to present an overview of the field of assessing metacognition, highlight the main trends and themes with examples from the included records and provide context for the methodological questions that this review raises. Summarising and describing the results of the review with 149 included records (including 13 records added via citation searching, see Figure 2) was undertaken using synthesis tables to identify patterns in data and then a narrative synthesis to describe the key themes and findings. These relate to the issues identified in the literature about the assessment of metacognition and in particular the types of
methods used, the use of tools across multiple age groups and information about the reported reliability and validity of the methods and tools.

Table 3 comprehensively lists the 80 tools categorised into the six groups identified in Section 2.4, alongside the references for the included tools and a short rationale for the categorisation of the tool. This rationale is particularly important for tools that potentially cross the six categories outlined in Section 2.4. Table 3 also briefly describes the tool alongside the definition of metacognition given for each tool or method. For included tools with particularly high numbers of diverse records (e.g. TAPs) a summary is provided with reference to multiple included citations. The authors acknowledge that this is not ideal, but that in a review aiming to summarise the field in an accessible manner it serves a purpose.

The categories adopted in this review are not exclusive; some tools could be described in more than one of the categories. For example, the Multi-Method Interview (MMI) (Wilson, 1999, 2001) is clearly described as both being multi-method and an interview. The MMI was allocated to the multi-method category because although a clinical interview is part of this multi-method approach, the problem-based interview is firmly situated alongside other methods including observation, video and audio recordings. Another example of a tool that could cross the categorisation in this review are the Swanson Metacognitive Questionnaire (SMQ), although it has questionnaire in the title it was administered in the earliest record included (Table 3) as an interview by Swanson (1990). Other examples are described within Table 3. Despite the complexity of the included tools and inevitable overlap, the groupings described did provide a clear means to explore, compare and critically evaluate the findings of this review.
The categorisation of the included tools as listed above facilitated the identification of trends and themes across the six categories. The foremost findings of this review relate to age: the age ranges different tools have been used with, the methodological differences between tools used with different ages and potential limitations of this. Age related findings and associated debates about metacognition are the focus of this summary, but it is important to note other equally significant findings of this review.

Firstly, relating to the methodological limitations of included tools is the fact that comprehensive information regarding reliability and validity is not always provided, or reported accurately (Table 2). That is not to discourage the development of new tools and methods, but rather to encourage a broad understanding of what exists in the field and the importance of being able to reliably validate tools and their findings.

Secondly, the majority of the included assessments of metacognition in education are based in the subjects of Mathematics, Literacy (first language) and Science (see Table 5). This focus on ‘core’ subjects is not surprising, schools are often judged by their students’ attainment in these subjects and research supports a positive link between metacognitive awareness, positive student outcomes and attainment (Akyol, Sungur, & Tekkaya, 2010; Dignath et al., 2008; Higgins, Hall, Baumfield, & Moseley, 2005; Prins et al., 2006).

The prevalence of self-report measures (including questionnaires and surveys) is one of the key findings in this review, a finding from which the age related trends identified in this review were derived. Self-report measures comprise 61% of the 80 included tools. Tools categorised as multi-method were the smallest group, only 4%. Of the other categories applied, observation based methods accounted for 8%, teacher ratings 6%, interviews 14% and task-based methods 8% of the total. Table 3
exemplifies this pattern; 149 records were included 186 times, within these 186 references 120 were references to distinct ages using a self-report measure, questionnaire, survey or test. The predominance of self-report in the field of assessing metacognition is clear, leading us to question:

- Why are self-report measures dominant in this field?
- What other types of tools have been used less often?
- What the methodological limitations of different types of tools and their use with differing age ranges?

Self-report measures are perceived as easy to use and as placing little in the way of time demands in terms of their application. Sperling et al. (2002) asserted that self-report inventories are perhaps the least problematic in terms of measuring metacognitive processing, that they are useful on a large scale and for identifying learners that require intervention, as well as being useful for theoretical research. Sperling et al. (2002) clearly identified that there is a gap in research using self-report inventories of metacognition, in terms of their lack of use with younger learners.

Leutwyler (2009) identified “one-sided criticism” (p. 115) about the credibility of self-report measures and the differences between pro and retrospective self-report and online measures. However, he also affirms the importance of recognizing the differences between which facets of metacognition measures actually explore. Is it therefore possible that this multi-faceted approach to assessing metacognition applies not only to different methods and whether they are online or offline, but also to the different stages of the development of metacognition for children of different ages? Desoete (2008, p. 204) states “how you test is what you get”. This review serves to highlight the importance of distinguishing, but yet appreciating the different aspects of metacognition explored by different methods.
Acknowledging the different contributions of different tools facilitates something of a puzzle like approach to exploring the development of metacognition in school-aged children. With such variety of age ranges within the included tools, questions about the development of metacognition and claims made using different measures require careful consideration.

Examining categories and the age ranges of individual records reveals some interesting patterns. Self-reports, questionnaires, surveys and tests have only been used with students over the age of 7 years in the included records (Table 4). In contrast, observational methods have been used with participants aged 4-8 years (including TAPs the range extends up to 15 years, but TAPs have only been used with students as young as 6 years). Teacher completed ratings have a range of 4 – 16 years, and interviews and focus groups 4 – 14 years. Task based methods have a range of 7 – 16 years. Within these broader categories the age ranges (within 4 – 16 years) of the majority of the individual tools are much smaller. For example, the RAC (8-10 years), MMI (11-12 years), MSTRAT (12-14 years) and C.Ind.Le (4-5 years).

Younger students lacking in the samples of the included records (particularly for self-report based tools) and the age ranges of other types of tools identified raises debate about: the age(s) at which metacognition is observable or recordable, the demands and understanding associated with completing a self-report measure and the development of metacognition. For example, with regards to self-report measures: is it that metacognition has not developed and therefore is not recordable in this age group, or more likely is it that the practicalities of using a self-report measure with this age group present challenges (e.g. literacy demands or the level of understanding required to complete)?
The age at which metacognition develops and is observable or recordable is a continuing debate within the field of metacognition, conflicting evidence is presented to support the development of metacognition at different ages. Hofer and Sinatra (2010) propose that unlike many linear perspectives of children’s development, metacognitive development is far from linear or one-dimensional. The complexities of the development of metacognition are clear and require a “multiplistic perspective…in which competing knowledge claims can be adjudicated and supported with evidence” (Hofer & Sinatra, 2010: p. 117). Similarly, Kuhn (2000) stated that it is helpful to have a developmental framework within which to explore metacognition, but that it is also essential to consider that there can be a wide variety of influences (e.g. the social context of learning).

Within the records included in this review there is a range of opinions regarding evidence of metacognition and its development. Leutwyler (2009, p. 112) asserts that children aged as young as 3 years old show “the first roots of metacognition”. Similarly Whitebread et al. (2010) concluded that using their observation based methods “enabled the clear identification of early metacognitive skills in young children” (p. 237) and Wall (2008) presented evidence of both metacognitive knowledge and metacognitive skilfulness in children as young as 4 and 5 years old. Nonetheless it is made clear that the findings of these studies relating to age and metacognition (with younger children) are contrary to established belief in the literature. Established belief has asserted that metacognitive skills in particular do not emerge until much later than this at aged eight years or even beyond this (Bartsch, Horvath, & Estes, 2003; Kuhn, 1999b; Veenman et al., 2004).

Looking at individual tools and methods in Appendix C (the data from which Table 4 was compiled) we can see that few tools of the 80 included have been used
across a wide age range. Each of the included tools and methods span no more than
nine of the recorded ages (3 years to 16 years) apart from TAPs and PVTs. TAPs have
been used with students aged 6 to 15 years and PVTs have been used with students
aged 4 – 13 years in the included studies. The wider age range of TAPs and PVTs
requires further examination of the differences between these two methods and other
tools in the same categories.

TAPs are described in the included literature as an online method where evidence
of metacognition is derived from an instruction to ‘think aloud’ whilst engaging in an
activity, e.g. problem solving. In the example of Veenman et al. (2005) this occurs
whilst participants are solving maths problems individually, a uniform prompt to
think aloud was added if participants fell silent. Veenman et al. (2005) assert that
thinking aloud does not hinder cognitive and metacognitive processes but merely slow
them down. Wall (2008) explains that PVTs are a visual tool, comprising a template
that forms part of a mediated interview, which is often completed as part of a focus
group and sometimes in a whole class situation. The templates comprise a picture of a
learning situation (including a person or group of people) that has speech and thought
bubble(s) in which the students write during and after discussion in the focus group.
The learning situations range from working in a group or pair to using ICT (Wall,
Higgins, & Packard, 2007). PVTs are inherently retrospective; the situations depicted
facilitate student reflection on past experience.

The distinction between the perceived online nature of TAPs and the assumed
reflective nature of PVTs is an interesting point to debate. This is explored further in
3.1 but it is interesting to note that PVTs are not explicitly described as either online
or offline (or indeed prospective or retrospective). If TAPs do indeed slow down
cognitive and metacognitive processes does this disadvantage and therefore exclude
the use of TAPs with younger students? The complexity of the demands on working memory (for the task being observed and completing the TAPs) may well prove challenging for younger students. This potentially complex need to ‘think aloud’ whilst learning may be why TAPs have not been used with students under 6 years or age.

PVTs appear to have advantages here in that they are completed in focus groups, perhaps mediating the pressure on individual students and recognising the social context of learning in school-aged children. PVTs are a visual tool, the picture representation of learning scenarios in PVTs may well appeal to younger students. Observation based methods observing regular classroom activity (without TAPs) have similar advantages in terms of their use with younger students (e.g. Classroom Coding System, CASE@KS1 and C.Ind.Le). The absence of additional demand(s) that may be added to a learning experience by requesting that students externalise internal metacognitive and cognitive processes verbally. It is important to consider if the slowing down associated with ‘think aloud’ could alter the trajectory that the learning episode being observed would have taken without this forced externalisation.

Unlike TAPs other observation methods included in this review do not seem to place explicit demands (i.e. to ‘think aloud’) on participants; rather they appeared to focus on observing behaviour/listening to dialogue. For example, unlike the C.Ind.Le (Whitebread et al., 2005; Whitebread et al., 2009), TAPs have direct researcher input in the form of request to think aloud. Whereas for the C.Ind.Le (Whitebread et al., 2009), video was used to record children participating in “interesting and productive” (p. 70) activities, but during this time there was no researcher input in terms of requests to ‘think aloud’ as in TAPs. Observation was completed of regular classroom
activity with the classroom teacher and video was retrospectively analysed for evidence of metacognitive or self-regulatory events.

Although originally grouped with teacher ratings, observation based methods were explored separately due to significant methodological differences. The five teacher rating tools: CHILD 3–5 checklist; Teacher Rating; The Teacher Rating; RSSRL and MKQ were different in one key way. Both observation and teacher rating rely on third party (i.e. researcher or teacher and not the learner to report evidence of metacognitive or self-regulatory activity). However, the included teacher ratings were checklists completed retrospectively and based on teacher experience, rather than reflection on a single learning episode or the observation of a particular ‘live’ task.

The CHILD 3-5 checklist (Whitebread et al., 2005) and Teacher Rating (Sperling et al., 2002; Sperling, Richmond, Ramsay, & Klapp, 2012) involved teachers rating their students retrospectively on a scale of 1 – 6 (Always – Never for the CHILD 3-5) for metacognition; the rating in both examples was assisted by examples given for each point on the scale of student behaviours. The Teacher Rating (Desoete, 2008) is a 20 item rating scale, described as a teacher questionnaire and again is not explicitly linked to a task. The RSSRL comprises a 12-item behaviour frequency 5-point scale is similarly not associated with observing behaviour in particular task. The ratings in the RSSRL a more general reflection based on day-to-day classroom activity for the ‘observed’ students. The MKQ focuses on the “declarative, procedural, and conditional knowledge of the application of strategies” (Metallidou & Vlachou, 2010, p. 780), again a teacher rating that is based on retrospective and generalized reflection as opposed to a specific task. Aside from Child 3-5 (Whitebread et al., 2009) the other teacher ratings are all used with children aged 7 or older, this may imply that assessing metacognition in children younger than this is more specialized or rather
that there is a link between the methodology by which metacognition is assessed and the outcomes of this.

The retrospective nature of the teacher ratings mentioned above and their associated reliance on the reflections of classroom teachers is distinct from the included observation based methods including TAPs, The Classroom Coding System, CASE@KS1, C.Ind.Le, Private Speech Coding and Self Directed Learning Instrument. These are all observations focused on specific tasks and observation is recorded while the task takes place and/or is video taped for later analysis. These observations are typically not completed by the regular class teacher, but rather by researchers who in some instances are specially trained. To give a contrasting example from another category (interviews and focus groups), Wall (2008) cited the use of Pupil Views Templates (PVTs) in a national Learning to Learn project (Higgins et al., 2007) where they have been used by school staff (teachers) to elicit pupil views as well analyzed as by researchers for evidence of metacognition.

Continuing on the theme of why different tools have been used with different age ranges, it is important to consider demands additional to those on working memory already discussed in relation to TAPs. Returning to the predominance of self-report measures and their use with students aged 7 years and over, one wonders if the literacy and reading demands of completing a self-report play a role. It is important to consider the potentially high literacy demands of questionnaires, surveys and self-report measures on respondents. The very nature of self-report implies a level of ability for the respondent in terms of literacy. If intervention is applied, for example the researcher or another non-participating individual reading out the questions and/or answer options, then at what point does a self-report questionnaire or survey become an interview or mediated interview? Additionally, if varying literacy levels across
respondents imply varying levels of understanding of what a self-report measure is asking, does lack of understanding mean a lower ‘score’ and therefore less evidence of metacognition? This debate resonates to questions around the common internal states assumed by observation, if a student does not have the literacy level to understand and/or complete a self-report fully this does not mean that they are not metacognitive in learning situations.

Once again we return to Desoete’s (2008) mantra that ‘how you test is what you get’. The definition of metacognition relates not only to the outcomes of a study but is also intrinsically linked to the tool or method and how it measures or assesses metacognition. How you test is what you get (Desoete, 2008), but how you define metacognition is also what you get and, in the planning and execution of empirical research influences how you test. For example, if one method or tool has a limited age range or the literacy demands are too high for younger students to participate, findings need to be moderated by this. Assertions about developmental trends in metacognition need to be considered alongside the tools or methods that have been used to ascertain them, the age range of the participants in a given study and any potential methodological limitations of this given study. With this in mind it is important to revisit defining metacognition, the implications of this on assessment of metacognition and one of the most commonly made distinctions between tools and methods (whether they are online or offline)?

3.1 Defining metacognition: in relation to the method and is the measure online or offline?

Defining metacognition and its associated concepts is not an easy task. It is important to recognise that different groups of tools and particular techniques and methods can define metacognition in very different ways. Table 3 lists for each of the
80 included tools a brief summary of the definition of metacognition explored in the included records for each tool. Table 3 is a valuable resource to explore the links (or lack of) between the tool described and what it seeks to measure in comparison to the definition of metacognition (or the associated concept) that is presented. The MSTRAT (Roeschl-Heils, Schneider, & van Kraayenoord, 2003), categorised as a test in this review, is an example of a tool where the definition of metacognition is hard to precisely determine, making it more difficult for the reader to draw definite links between this and the tool being applied.

Some of the definitions listed in Table 3 show similarity between different tools. For example, two self-report measures the Metacognitive Awareness of Reading Strategies Inventory (MARS-I) and the Metacognitive Awareness Inventory (MAI) both have similar definitions of metacognition based on the reflection on and monitoring of learning, including understanding of learning and an individuals’ control of their own learning. In contrast records concerning TAPs (described as online) often define metacognition in relation to its relevance as a predictor of learning, they make the same distinction as research using PVTs (not explicitly described as online or offline) between metacognitive knowledge and metacognitive skillfulness. Related to this is whether or not a tool is “administered either prospectively, concurrently, or retrospectively to performance on a learning or problem-solving task” (Desoete, 2009, p. 436). Examples of prospective tools in this review are the Inventory of Metacognitive Self-Regulation (IMSR), Metacognitive Ability Self-report Questionnaire (MASQ), Prospective Assessment of Children (PAC) and Metacognitive Awareness Inventory (MAI). This debate is often presented under the umbrella of debate around the distinction between online and offline methods: what tools measure and how, as well as the different tools or methods in
each category (on-line or off-line) and why they fit into it (Saraç & KaraKelle, 2012; Tillema, van den Bergh, Rijlaarsdam, & Sanders, 2011).

Concurrent methods include TAPs, which is also commonly described as an online technique (Desoete, 2007; Mateos, Martín, Villalón, & Luna, 2008). However as Mateos et al. (2008, p. 695) rightly point out, “while think-aloud protocols are considered one of the most effective tools we have for gaining access to the online cognitive processing of readers and writers, they have certain well-known limitations (e.g., Ericsson & Simon, 1993).” There is room for further debate here, as it could be argued that as soon as a researcher asks a participant to stop, think about and articulate out loud the processes behind their learning that they are actually being forced to be retrospective so the previously presumed [on-line] “reflection-in-action” (Schön, 1983) becomes [offline] reflection-on-action when a student is asked to stop and think aloud. This reflection and its subsequent influence on learning via self-regulatory processes could mean that TAPs are indeed and can remain concurrent throughout the process but this would depend on the tightness of the feedback loop when a learner reflects on their own learning. The degree to which forced reflection on their learning made ‘aloud’ then makes it retrospective and then how the reflection then does or does not influence their behavior in the remainder of the task requires significant consideration. Other examples of retrospective tools or methods include the Retrospective Assessment of Children (RAC) (Desoete, 2007, 2008) and the majority of the included interviews and task-based methods.

4 Some implications

This synthesis of tools and methods used to measure metacognition in school-aged children is important for wider research on metacognition, as there is not a
current review in this area looking systematically at the assessment of metacognition. This review has raised important questions, such as about the age groups with which different methods of assessing metacognition are used.

There are wider debates about the age at which metacognition is present. This is clearly contestable, as we found 20 tools or methods purporting to assess metacognition in participants aged 4 – 7 years, indeed 11 tools or methods assessing metacognition or closely associated concepts in the youngest age group of 4-5 years. Evidence gathered by Wall (2008) indicates that evidence of metacognitive skilfulness, as gathered using PVTs, appears at an earlier age than previously thought, in children as young as 4 and 5 years old. In contrast, Bartsch et al. (2003) discuss the difficulties that children of this age have in recognizing how and when knowledge is acquired and Kuhn (1999a) argued that metacognitive knowledge could be present at a much younger age than metacognitive skilfulness, which she states does not develop until aged 10-12. Similar to Wall (2008), Leutwyler (2009, p. 112) makes reference to children aged three showing “the first roots of metacognition” and Whitebread et al. (2009) have observed young children showing emergent metacognitive behaviours.

The relationship of method to finding may be crucial. In terms of implications for the assessment of metacognition in young children, tools that combine or triangulate observed behaviours, and link these with tools that help to elicit declarative knowledge (rather than skilfulness).

As noted in the introduction, this review has focussed on instruments which operationalized meta-cognition explicitly and it was beyond the scope of the review to include measures of executive control and executive function, though this would clearly be a valuable next step, as would a systematic mapping of comparative use of such research tools (see Garner, 2009, for example).
From this review we can also see how tools or methods have changed and been adapted over time, sometimes to form completely new tools. For example, Wolters (1996) describes a conditional knowledge questionnaire that is adapted from two other tools: the IRA and the MSLQ. The IRA is again mentioned by Schmitt and Sha (2009) when discussing the IMA which is also in part based on the IRA. In addition there are crucial connections between how metacognition is defined in relation to a tool or method and how this definition is then linked to what is being measured. It is important in evaluating the findings of metacognitive assessments to understand what a particular tool or method purports to measure, how this related to the type of tool and the data collected to ensure it is well aligned with the definition of metacognition adopted. This alignment or congruence of definition, of tool, findings resulting from its use with wider claims made about metacognition are essential for the further development of the field.

Acknowledgements:

This work was supported by a Durham Doctoral Studentship in the School of Education, Durham University.
References

(* = Records included in the systematic review)


Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Should we be using learning styles? What research has to say to practice.* London: Learning and Skills Research Centre (Great Britain) (LSRC).


Dermitzaki, I. (2005). Preliminary investigation of relations between young...
students' self-regulatory strategies and their metacognitive experiences. 
* Dermitzaki, I., & Efklides, A. (2001). Age and gender effects on students' 
evaluations regarding the self and task-related experiences in mathematics. In 
S. Volet & S. Järvelä (Eds.), *Motivation in learning contexts: Theoretical 
advances and methodological implications*. (pp. 271-293). Elmsford, NY US: 
Pergamon Press.

* Dermitzaki, I., & Efklides, A. (2003). Goal orientations and their effect on self- 
concept and metacognition in adolescence. *Psychology: The Journal of the 

* Desautel, D. (2009). Becoming a Thinking Thinker: Metacognition, Self-Reflection, 

Psychology, 5*(3), 705-730.


* Desoete, A. (2009). Metacognitive prediction and evaluation skills and 
mathematical learning in third-grade students. *Educational Research and 
Evaluation, 15*(5), 435-446. doi: 10.1080/13803610903444485


mathematical problem solving. *Learning and Instruction, 16*(1), 12-25. doi: 
10.1016/j.learninstruc.2005.12.003


Enhance mathematical problem solving? *Journal of Educational Psychology, 95*(1), 188-200. doi: 10.1037/0022-0663.95.1.188

learn self-regulated learning strategies most effectively? A meta-analysis on 

Conceptual Lens on Metacognition, Self-Regulation, and Self-regulated 

* Dowson, M., & McInerney, D. M. (2004). The Development and Validation of the 
Goal Orientation and Learning Strategies Survey (GOALS-S). *Educational and Psychological Measurement, 64*(2), 290-310. doi: 
10.1177/0013164403251335

Efklides, A. (2001 xx). Metacognitive experiences in problem solving: Metacognition, 
(Eds.), *Trends and prospects in motivation research* (pp. 297-322). Dordrecht, 
The Netherlands: Kluwer.

Chambres, M. Izaute & P.-J. Marescaux (Eds.), *Metacognition: Process, 
Function and Use* (pp. 19-34). Massachusetts: Springer.


Gough, D, Oliver, S., & Thomas, J. (2012). *An Introduction to Systematic Reviews.*


* Howard, B. C., McGee, S., Shia, R., & Hong, N. S. (2001). The Influence of Metacognitive Self-Regulation and Ability Levels on Problem Solving (pp. 8-8).


Available from EBSCOhost psyh database.


Montague, M., & Applegate, B. (1993). *MIDDLE SCHOOL STUDENTS*


Short, J. A. (2002). *Examining the cognitive and metacognitive strategies of first-grade journal writers in a literature-based classroom*. (62), ProQuest Information & Learning, US. Retrieved from


* van der Stel, M., & Veenman, M. V. J. (2008). Relation between intellectual ability and metacognitive skillfulness as predictors of learning performance of young students performing tasks in different domains. Learning and Individual Differences, 18(1), 128-134. doi: 10.1016/j.lindif.2007.08.003


Table 1: Inclusion and Exclusion Criteria with Examples of Excluded Records

<table>
<thead>
<tr>
<th>Category</th>
<th>Rationale</th>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
<th>Examples of excluded records</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• Record specifies it is measuring metacognition or a closely related concept and there is a clear definition of what is being measured</td>
<td>• Metacognition or closely associated concept not being measured or the definition of metacognition is not clear or clearly linked to the measurement outcomes</td>
<td>Morgan and Brooks (2012) – The focus is on scaffolding and not metacognition.</td>
</tr>
<tr>
<td>What is being measured?</td>
<td>The focus of the review is metacognition and closely related and defined concepts</td>
<td>• Measured in the first language of the participants</td>
<td>• Not measured in the first language of the participants</td>
<td></td>
</tr>
<tr>
<td>Sample population</td>
<td>The sample population must fall within the defined age group (4-16 years) and be normally or average achieving in mainstream education in order that there is a degree of homogeneity in the samples for the different included tools or methods</td>
<td>• Participants aged 4-16 years (at least 50%)</td>
<td>• Participants not 4-16 years</td>
<td>Hanson and Williams (2008) – This contains a higher education sample, not in the range of 4-16 years</td>
</tr>
<tr>
<td>(age, setting, normally achieving)</td>
<td></td>
<td>• Mainstream school</td>
<td>• Not mainstream school setting</td>
<td>Montague and Applegate (1993) – The sample is entirely comprised of students with additional (special) needs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Cross section of students (average achieving or cross section of abilities)</td>
<td>• More than 50% of students identified as having additional needs or being gifted</td>
<td></td>
</tr>
<tr>
<td>Data set and methodology</td>
<td>The record needs to have an empirical data set to be included (unless the first example of a particular tool with detailed explanation of that tool or method)</td>
<td>Empirical data needs to be collected and there must be a clear and replicable tool or method</td>
<td>No empirical data or the methodology is not clear or replicable</td>
<td>Feldhusen and Goh (1995) and Vermunt and Vermetten (2004) – both excluded as they do not contain an empirical data set</td>
</tr>
<tr>
<td>Language of the record</td>
<td>Time and financial constraints did not allow for records to be translated if they were not readily available in English. Every effort made to obtain, including contacting authors.</td>
<td>Record readily available in English</td>
<td>Record not readily available in English</td>
<td>Mañá, Vidal-Abarca, Domínguez, Gil, and Cerdán (2009) and Yalçın and Karakaş (2008) – the authors of this review were not able to obtain copies in English or translate within their given timescale and budget.</td>
</tr>
</tbody>
</table>
Table 2: The reliability, validity & replicability for each of the data extracted tools or methods (n = 84)

<table>
<thead>
<tr>
<th>Tools or methods</th>
<th>Reliability</th>
<th>Validity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internal consistency</td>
<td>Test-retest</td>
</tr>
<tr>
<td>1. Bandura’s Self Efficacy for Self-Regulated Learning Scale</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>2. Cambridgeshire Independent Learning in the Foundation Stage Coding Framework (C.Ind.Le)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>3. Checklist of Independent Learning Development 3-5 (Child 3-5)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>4. CA (Child Assessment)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>5. CDR (Cognitive Developmental aRithmetics test)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>6. Classroom Coding System</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>7. Clinical Interview (Erbas and Okur, 2012)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>8. Clinical Interview (Pappas, Ginsberg and Jiang, 2003)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>9. Computer based measure of metacognitive skilfulness</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>10. Concept maps</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>11. Conditional knowledge</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>12. EPA2000 (Evaluation and Prediction Assessment)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>13. Epistemic metacognition measure</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>14. Goal Orientation and Learning Strategies Survey (GOALS-S)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>15. How I Study Questionnaire (HISQ)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16. Index of Metacognitive Awareness about</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Tools or methods</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Internal consistency</td>
<td>Test-retest</td>
</tr>
<tr>
<td>Writing (IMAW)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>17. Index of self-efficacy for writing (ISEW)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>18. Index of Self-Regulated Writing (ISRW)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>19. Index of Reading Awareness (IRA)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>20. Index of Science Reading Awareness (ISRA)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>21. Individual interview – strategy use and</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>metacognition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Integrated Learning Assessment</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>23. Interview about Metacognitive Awareness (IMA)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>24. Interview from the Munich Longitudinal Study ...</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>25. Inventory of Metacognitive Self-Regulation (IMSR)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>26. Junior Metacognitive Awareness Inventory (JrMAI)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>27. Knowledge and skills questionnaire</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>28. Learning strategies assessed by journal writing</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>29. Learning Through Reading Questionnaire (LTRQ)</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>30. Metacognition Applied to Physical Activities Scale (MAPAS)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>31. Metacognition of Nature of Science Scale (MONOS)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>32. Metacognition Scale</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Tools or methods</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Internal</td>
<td>Test-retest</td>
</tr>
<tr>
<td>33. Metacognitive Processes in Physical Education Questionnaire (MPIPEQ)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>34. Metacognitive Ability Self-report Questionnaire (MASQ)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>35. Metacognitive Attribution Assessment (MAA)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>36. Metacognitive Awareness Inventory (MAI)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>37. Metacognitive Awareness of Reading Strategies Inventory (MARI)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>38. Metacognitive experiences</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>39. Metacognitive Interview (Lu, 1995)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>40. Metacognitive Interview (MCI) (Lefevre, 1995)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>41. Metacognitive Knowledge in Mathematics Questionnaire (MKMQ)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>42. Metacognitive Knowledge Monitoring Assessment (KMA)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>43. Metacognitive Knowledge Questionnaire</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>44. Metacognitive Knowledge Test (needs to move)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>45. Metacognitive Questionnaire</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>46. Metacognitive Skills and Knowledge Assessment (MSA)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>47. Metacognitive skills and metacognitive development questionnaire</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>48. Metacognitive Strategies (MSTRAT)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>49. Metacomprehension Strategy Index</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>Tools or methods</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Internal consistency</td>
<td>Test-retest</td>
</tr>
<tr>
<td>(MSI)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>50. Motivated Strategies for Learning Questionnaire (MSLQ)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>51. Multi method assessment of metacognitive behaviours</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>52. Multi-Method Interview (MMI)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>53. Observation (CASE@KS1)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>54. Original standardized test for metacognition</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>55. Private speech coding</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>56. Problem solving interview</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>57. Prospective Assessment of Children (PAC)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>58. Pupil Views Templates (PVTs)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>59. Questionnaire about Learning in Mathematics (QLM)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>60. Questionnaire about Learning Slovene Language (QLSL)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>61. Questionnaire about metacognitive beliefs</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>62. Questionnaire based on Think Aloud</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>63. Rating Student Self-Regulated Learning Outcomes: A Teacher Scale</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>64. Reading Strategy use scale (RSU scale)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>65. Retrospective Assessment of Children (RAC)</td>
<td>✓</td>
<td>-</td>
</tr>
<tr>
<td>66. Retrospective Questionnaire Interview</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Tools or methods</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>------------------------------------------------------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td></td>
<td>Internal consistency</td>
<td>Test-retest</td>
</tr>
<tr>
<td>(RQI)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>67. Self Regulated Learning Scale (SRL)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>68. Self report metacognitive learning strategies</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>69. Self-Assessment in Metacognitive Comprehension Strategies Reading Survey</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>70. Self-Directed Learning Instrument</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>71. Self-Efficacy and Metacognition Learning Inventory – Science (SEMLI-S)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>72. Self-efficacy for Learning Form (SELF)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>73. Self-Regulated Learning Strategies Measurement Questionnaire</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>74. Self-report for cognitive and metacognitive learning strategies</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>75. State Metacognitive Inventory</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>76. Strategy card sort, individual interviews</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>77. Strategy knowledge in the domain of Chemistry</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>78. Swanson Metacognitive Questionnaire (SMQ)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>79. Teacher Rating (Sperling et al. 2002)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>80. The Teacher Rating (Desoete, 2008)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>81. Think About Reading Index (TARI)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>82. Think Aloud Protocol(s) (TAP/TAPs)</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>83. Worksamples Interview</td>
<td>-</td>
<td>✓</td>
</tr>
<tr>
<td>Tools or methods</td>
<td>Reliability</td>
<td>Validity</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>-------------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Internal consistency</td>
<td>Test-retest</td>
</tr>
<tr>
<td>84. Würzburg Metamemory Test</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Table 3: Summary Data Extraction for the 80 included tools

<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records (The record in bold is the primary record cited)</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child Assessment CA</td>
<td>CA</td>
<td>Desoete (2009) – need to add other references too</td>
<td>The CA is 12-item rating scale about metacognitive knowledge for children with a 7 point Likert Scale (used prospectively and retrospectively).</td>
<td>“…the knowledge, awareness, and deeper understanding of one’s own cognitive processes and products” (p. 436)</td>
</tr>
<tr>
<td>Cognitive Developmental aRithmetics test</td>
<td>CDR</td>
<td>Desoete (2008); Desoete (2009); Desoete &amp; Roeyers (2006a)</td>
<td>90-item test for assessment of arithmetic &amp; metacognitive experiences.</td>
<td>Metacognitive experiences: “…metacognitive feelings, metacognitive judgments/estimates, and online task-specific knowledge. Metacognitive experiences make the person aware of his or her cognition and trigger control processes that serve the pursued goal of the self-regulation process (Efklides, 2008)” (p. 436).</td>
</tr>
<tr>
<td>Conditional knowledge measure (part of a larger questionnaire)</td>
<td>-</td>
<td>Wolters (1996)</td>
<td>To assess conditional knowledge of strategy use. Included strategies were adapted from both the IRA and MSLQ.</td>
<td>Definition centres on the relationship between metacognitive knowledge and self-regulated learning. Metacognition described as a prominent component of models of self-regulation.</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
  • Metacognitive knowledge of person, task and strategy. Also subdivided into declarative and procedural.  
  • Metacognitive skills  
  • Metacognition as an important variable in arbitrating learning. |
| Goal Orientation and Learning Strategies Survey GOALS-S | Participants are required to respond to items using a Likert scale              | Dowson and McInerney (2004)                                                     | A survey “designed to measure three academic goals, five social goals, three cognitive strategies and three metacognitive strategies” (p. 293). A series of statements responded to using a 5-point Likert scale. | Metacognitive strategies described as monitoring, planning and regulating.  
Full descriptions are given in Table 1 in the record, 18 items for metacognitive strategies in total. |
| Index of Metacognitive Awareness about Writing IMAW | Participants are required to respond to items using a Likert type scale.       | De Kruif (2000)                                                                 | Designed to assess metacognitive knowledge of the writing process.  
It has items for planning, translating and reviewing. | A model of self-regulated writing where the model of self-regulation assumes the integration of three components that determine the degree of self-regulated learning:  
  • Metacognitive knowledge  
  • Motivation (described as self-efficacy)  
  • Strategy use |
<p>| Index of Self-Efficacy for Writing ISEW         | Participants are required to respond to items using a Likert type scale.       | De Kruif (2000)                                                                 | Linking social cognitive theory and self-efficacy for self-regulated writing performance. It has items for planning, translating and revising. | |</p>
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records (The record in bold is the primary record cited)</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index of Reading Awareness</td>
<td>IRA</td>
<td>Multiple choice self-report</td>
<td>observation, self-judgement and self-reaction.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bouffard (1998); Bouffard &amp; Vezeau (1998); Jacobs &amp; Paris (1987); McBride-Chang &amp; Chang (1995); Meloth &amp; Deering (1992); Osbourne (1998); Pereira-Laird &amp; Deane (1997); Schmitt &amp; Sha (2009); Sperling, Howard, Miller &amp; Murphy (2002); Swanson &amp; Trahan (1996); van Kraayenoord &amp; Paris (1996); van Kraayenoord &amp; Schneider (1999)</td>
<td>Multiple-choice index to explore children’s knowledge of reading strategies. A focus on conditional knowledge – understanding of when/why strategies are applied.</td>
<td>Metacognition as self-regulated thinking. What people know and the application of knowledge to tasks. Wellman (1985) and the fuzziness of defining metacognition was noted.</td>
</tr>
<tr>
<td>Index of Science Reading Awareness</td>
<td>ISRA</td>
<td>Multiple-choice self-report</td>
<td>The first 63 items of the ISRA – a measure of metacognitive awareness of declarative, procedural and conditional domains in relation to being a successful reader in science.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Craig &amp; Yore (1998); Holden (1997); Yore, Craig &amp; Maguire (1998).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integrated Learning Assessment</td>
<td>ILA</td>
<td>Written responses required in order to assess metacognition (test format)</td>
<td>Developed to ‘measure’ the degree to which students use “cognitive and meta-cognitive skills” (p. 2). The tool was used in relation to reading in biology and history. Students asked to describe their reading process.</td>
<td>Metacognition and use of strategies (reading strategies in particular). The degree to which students have awareness of their thought processes.</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Inventory of Metacognitive Self-Regulation IMSR | The IMSR is a self-report completed using a Likert scale. | Howard, McGee, Hong, & Shia (2000), Howard, McGee, Shia & Hong (2000, 2001), Parcel (2005) | Developed to examine metacognitive monitoring and regulatory skills. | Metacognitive self-regulation broken down into five facets:  
  - Knowledge of cognition  
  - Objectivity  
  - Problem representation  
  - Subtask monitoring  
  - Evaluation  
Metacognition and problem solving (predictors of problem solving). |
| Junior Metacognitive Awareness Inventory (JrMAI), JrMAI | Described as a self-report inventory | Ciascai & Lavinia (2011); Huber (2012); Kim & Pederson (2010); Lemberger & Clemens (2012); Schwartz, Anderson, Hong, Howard & McGee (2004); Sperling et al. (2002); Sperling, Richmond, Ramsay & Klapp (2012) | JrMAI version A and B developed from the MAI (Schraw & Dennison, 1994). Both versions were self-report inventories with slightly differing response scales. | Metacognitive knowledge and regulation described as part of self-regulatory abilities. Distinction made between metacognitive and self-regulatory skills. |
| Measurement of Metacognition (Skills and knowledge) and Student Intelligence (Knowledge and skills questionnaire) - | Authors state explicitly that the measure is a questionnaire. | de Jager, Jansen, & Reezigt (2005) | Two part questionnaire looking at five stages (before reading, during reading, repairing misunderstanding, after reading)  
Part 1: Metacognitive skills (22 questions, sometimes and no).  
Part 2: Metacognitive knowledge (12 questions, two possible answers to choose from for each question). | Separate parts of the questionnaire for metacognitive knowledge and metacognitive skills (reference made back to Flavell (1976). The role of teachers in developing metacognition. |
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Questionnaire</td>
<td></td>
<td>(2011)</td>
<td>questions on learning and self-regulating strategies.</td>
<td>(including feedback). Self-Regulating Strategies: planning, monitoring (learning and work progress/methods), adjusting (working with text &amp; re-reading, linking information, work management), emotion/motivation control, self-evaluating.</td>
</tr>
<tr>
<td>Metacognition Applied to Physical Activities Scale</td>
<td>MAPAS Questionnaire</td>
<td>Settanni, Magistro, &amp; Rabaglietti (2012)</td>
<td>10-item questionnaire with 4 possible responses from 1 (completely disagree) to 4 (agree completely).</td>
<td>Refers back to Flavell (1979) – knowledge of own cognition and the relative control of this. Two components of metacognition – knowledge about cognition (declarative, procedural and conditional) and regulation of cognition.</td>
</tr>
<tr>
<td>Metacognition of Nature of Science Scale</td>
<td>MONOS Described as a survey</td>
<td>Peters (2008); Peters &amp; Kisantas (2010)</td>
<td>A 16 item survey to test student perceptions: attitude about the science, use of metacognition in observation, use of metacognition in data collection, use of metacognition in measurement, ability to explain reasoning in making conclusions. 5-point scale to answer: 5 (agreed with the statement) – 1 (disagreed with the statement).</td>
<td>Metacognition as executive functions to control actions or recognise patterns of thinking and evaluate them (Weinert, 1987). Metacognition as the monitoring or control of cognition.</td>
</tr>
<tr>
<td>Metacognition Scale</td>
<td>-</td>
<td>Participants required to respond using a Likert scale.</td>
<td>Yildiz, Akpınar, Tatar, &amp; Ergin (2009)</td>
<td>A 40-item scale, delivered using a 4-point Likert scale from ‘every time’ (4 points) to ‘never’ (1 point).</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Metacognitive Processes in Physical Education Questionnaire</td>
<td>MPIPEQ</td>
<td>Theodosiou, Mantis, &amp; Papaioannou (2008)</td>
<td>A measure of metacognitive knowledge and regulation. Eight scales of the MPIPEQ used in this study: declarative knowledge, procedural knowledge, conditional knowledge, information management, planning, self-monitoring, problem solving strategies and evaluation.</td>
<td>Metacognition as an important element of self-regulation. Distinction made between metacognitive knowledge and experiences. Referred back to Flavell (1979). Metacognitive skills also described (as partially independent of intellectual ability). The difference between online and offline measures of metacognition is noted.</td>
</tr>
<tr>
<td>Metacognitive ability self-report questionnaire</td>
<td>-</td>
<td>Panaoura &amp; Panaoura (2006); Panaoura &amp; Philippou (2003, 2007)</td>
<td>An inventory based on the idea of the MAI and the JrMAI to measure metacognitive ability in mathematics (for young pupils). Two-part questionnaire: part 1 measured metacognitive abilities in mathematics (30 items on a five point Likert-scale). Part 2 was about cognitive ability in problem solving (questions to answer before and after attempts at solving problems presented that they read).</td>
<td>Metacognition as a multidimensional construct – two (basic) dimensions are metacognitive knowledge and self-regulation of cognition. Metacognitive knowledge as including knowledge of person, task and strategy. Metacognitive regulation as the processes that coordinate cognition. The two constructs are seen as interdependent.</td>
</tr>
<tr>
<td>Metacognitive Attribution Assessment</td>
<td>MAA</td>
<td>Despete, Roeyers, &amp; Buysse (2001)</td>
<td>A 13-item attribution rating scale based on the work of Carr &amp; Jessup (1995). With the MAA children evaluate internal stable (e.g. ability), internal nonstable (e.g. effort), external stable (e.g. planning, monitoring, evaluation).</td>
<td>Referred back to Flavell (1976). Metacognitive knowledge – declarative, procedural and conditional or strategic. Executive control or metacognitive skills (planning, monitoring, evaluation).</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metacognitive Awareness Inventory (MAI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metacognitive Awareness of Reading Strategies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metacognitive experiences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Included records (The record in bold is the primary record cited)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantwell &amp; Andrew (1998, 2002); Kesici, Erdogan, &amp; Ozteke (2011); Schwart &amp; Dennison (1994); Sungur &amp; Senler (2009); Symons &amp; Reynolds (1999)</td>
</tr>
<tr>
<td>Boudreauz (2008); Huber (2012); Law (2009); Mokhtari &amp; Reichard (2002); Morley (2010)</td>
</tr>
<tr>
<td>Dermitzaki (2005); Dermitzaki &amp; Efklides (2001, 2003); Efklides &amp; Tsiota (2002)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Description of the tool</th>
</tr>
</thead>
<tbody>
<tr>
<td>task characteristics) and external nonstable (e.g. luck). Evaluation completed by ranking using a 4-point scale.</td>
</tr>
<tr>
<td>52 item self-report inventory. Eight scales: declarative knowledge, procedural knowledge, conditional knowledge, planning, information management strategies, monitoring, debugging strategies and evaluation of learning.</td>
</tr>
<tr>
<td>Can be administered individually or in groups but students rate on their own. The inventory requires students to think about strategies that they have awareness of having used when reading.</td>
</tr>
<tr>
<td>Three kinds of metacognitive experiences recorded when solving a mathematics problem. Participants asked to rate twice (before and after) on 4-point scales, exploring: difficulty before, difficulty after, effort before, effort after, correctness before, correctness after.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refers back to the distinction made between knowledge of cognition and regulation of cognition.</td>
</tr>
<tr>
<td>Links made between metacognitive awareness and strategy use.</td>
</tr>
<tr>
<td>Focuses on metacognitive awareness (of reading strategies). Metacognitive awareness of cognitive and motivational processes while reading. Knowledge of cognition and self-control mechanisms to monitor and regulate text comprehension.</td>
</tr>
<tr>
<td>Metacognitive experiences – online judgements/estimates of feelings, ideas and thoughts about a current task. This includes feeling of Knowing (FOK), feeling of familiarity (FOF), feeling of difficulty (FOD), feeling of confidence (FOC), judgement of learning (JOL), estimate of effort expenditure (EOE) and estimate of solution correctness (EOC).</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Metacognitive Knowledge in Mathematics Questionnaire</td>
</tr>
<tr>
<td>Metacognitive Questionnaire</td>
</tr>
<tr>
<td>Metacognitive Skills and Knowledge Assessment</td>
</tr>
<tr>
<td>Metacognitive Strategies</td>
</tr>
<tr>
<td>Metacomprehension Strategy</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td><strong>Index</strong></td>
</tr>
<tr>
<td><strong>Motivated Strategies for Learning Questionnaire</strong></td>
</tr>
<tr>
<td><strong>Metacognitive Knowledge Test</strong></td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------------------------------</td>
</tr>
<tr>
<td>Prospective Assessment of Children</td>
</tr>
<tr>
<td>Questionnaire about Learning Slovene Language</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Questionnaire about metacognitive beliefs</td>
</tr>
<tr>
<td>Questionnaire based on Think Aloud</td>
</tr>
<tr>
<td>Reading Strategy use scale</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Self Regulated</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>Learning Scale</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Self report metacognitive learning strategies</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------</td>
</tr>
<tr>
<td>Strategies Reading Survey</td>
</tr>
<tr>
<td>Self-Efficacy and Metacognition Learning Inventory – Science</td>
</tr>
<tr>
<td>Self-efficacy for Learning Form, SELF</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Self-Regulated Learning Strategies Measurement Questionnaire (SRLSMQ)</td>
</tr>
<tr>
<td>Name of tool</td>
</tr>
<tr>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Think About Reading Index</td>
</tr>
<tr>
<td>Würzburg Metamemory Test</td>
</tr>
</tbody>
</table>

**Observation based methods**
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records (The record in bold is the primary record cited)</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom coding system for Children’s self-regulatory behaviours</td>
<td>-</td>
<td>Explicitly described as observation</td>
<td>Neitzel (2004); Neitzel &amp; Stright (2003); Stright, Neitzel, Sears &amp; Hoke-Sinex (2001)</td>
<td>Children’s awareness of their thinking assessed via observation (of children talking about their thinking). Coding system focussed on five areas: attention to instructions, seeking help, monitoring progress, involvement in class, metacognitive talk. Metacognitive task and strategy information. Self-regulatory behaviours in the classroom Metacognition as underlying self-regulation.</td>
</tr>
<tr>
<td>Observation (CASE@KS1)</td>
<td>-</td>
<td>Explicitly described as observation</td>
<td>Larkin (2006)</td>
<td>Observations of collaborative group work where children were working on CASE@KS1 activities. Qualitative approach to coding based on Flavell (1979) (cognitive monitoring). Metacognitive knowledge (person, task and strategy), metacognitive experiences, goals (or tasks) and actions (or strategies). Metacognition as thinking the ability to reflect on one’s own thinking (also to monitor and control consciously thinking). Metacognition as important for problem solving. Notes the problems that lack of clarity in defining metacognition has created in the field.</td>
</tr>
<tr>
<td>Cambridgeshire Independent Learning in the Foundation Stage Coding Framework (C.Ind.Le)</td>
<td>Video recorded observation of learning activities</td>
<td>Whitebread et al. (2005); Whitebread et al. (2009)</td>
<td>Video recorded observation of innovative learning activities. The coding framework (verbal and nonverbal indicators of metacognition and self-regulation) focussed on: metacognitive knowledge, psychological approaches to independent learning, inclusive of metacognitive experience, metacognitive knowledge and self-regulation. Metacognitive knowledge - individual’s knowledge about person, task and</td>
<td></td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------------------------</td>
<td>------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Private speech coding</td>
<td>-</td>
<td>Daugherty &amp; Logan (1996)</td>
<td>The private speech of the participating children coded according to its semantic characteristics.</td>
<td>Study exploring relationships between metacognitive processing and creative ability.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Coding scheme was extended and modified from Manning (1991): (a) task irrelevant speech; non-facilitative, (b) task relevant speech, (c) task relevant speech, (d) coping/reinforcing speech, (e) solving speech and (f) metacognitive speech.</td>
<td>Flavell (1987) – individuals’ awareness of how they are thinking. Metacognitive processing important for perception of problem demands and constraints. References to Vygotsky.</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records</td>
<td>Description of the tool</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Think Aloud Protocol(s)</td>
<td>TAP/TAPs</td>
<td>There is most often a task to promote metacognition (or associated concept) but the data analysed relies on observation (videoed or live) of what participants say.</td>
<td>Desoete (2007): TAP applied during word problem solving tasks (Reference made to Veenman &amp; Spaans, 2005). Desoete (2007): Metacognitive knowledge (person, task and strategy) and declarative knowledge, procedural knowledge. van Kraayenoord &amp; Schneider (1999): Metacognition as important in academic achievement, comprising knowledge and control of cognition. Veenman, Kok, &amp; Blöte (2005) Metacognition as a predictor of learning. Metacognitive skillfulness and metacognitive knowledge distinction made (knowledge as declarative knowledge about relationship between person, task and strategy characteristics, skillfulness as procedural knowledge of regulation and control of one’s learning).</td>
<td></td>
</tr>
<tr>
<td>Teacher ratings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metacognitive Knowledge Questionnaire</td>
<td>-</td>
<td>Explicitly described as a teacher rating</td>
<td>Metallidou &amp; Vlachou (2010) 9 item questionnaire (based on Carr &amp; Kurtz, 1990) concerning the degree to which children have declarative, procedural and conditional knowledge of the Focus on self-regulated learning (SRL) – the way students initiate, monitor and control their own learning. Relationships between motivational, cognitive and metacognitive</td>
<td></td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------</td>
<td>---------------------------------------------------------------</td>
<td>-------------------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td><strong>Checklist of Independent Learning Development 3-5</strong></td>
<td>CHILD 3–5</td>
<td>The CHILD 3-5 is completed by teachers</td>
<td>A 22-item checklist highlighting key elements of independent learning in children aged 3-5. The following headings: emotional, ProSocial, cognitive, motivational.</td>
<td>Three related elements of metacognition: metacognitive experience, metacognitive knowledge and self-regulation (Brown, 1987). See also the definition from Whitebread et al. (2009) as detailed for the C.Ind.Le.</td>
</tr>
<tr>
<td><strong>Rating Student Self-Regulated Learning Outcomes: A Teacher Scale</strong></td>
<td>RSSRL</td>
<td>Explicitly states completion of the RSSRL by teachers</td>
<td>RSSRL is a 12-item teacher scale; teachers use the RSSRL to evaluate the frequency of behaviours indicative of self-regulated learning.</td>
<td>Self-regulated learning. Self-regulated learners as: self-efficacious, autonomous and (intrinsically) motivated. Self-regulated learning strategies including self-evaluation, organising and transforming, goal-setting and planning, seeking information, keeping records and self-monitoring, environmental structuring, self-consequences, rehearsing and memorising, seeking assistance and reviewing.</td>
</tr>
<tr>
<td><strong>Teacher Rating</strong></td>
<td>-</td>
<td>Explicitly described as a teacher rating</td>
<td>Teachers rated students on a scale of 1-6 (high metacognition or low metacognition – five behavioural descriptors were provided for each).</td>
<td>Self-regulated learning &amp; metacognition. Flavell (1979) – metacognition as metacognitive knowledge (person, task and strategy) and metacognitive experiences (including feelings of understanding).</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>The Teacher Rating</td>
<td>-</td>
<td>Explicitly stated as teacher completed</td>
<td>Teacher rating created for this research. A 20 item rating scale teacher-questionnaire about metacognitive prediction, planning, monitoring and evaluation skills.</td>
<td>Metacognitive knowledge and skills.</td>
</tr>
<tr>
<td>Interviews and focus groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Clinical Interview                  | -                                                      | Explicitly described as an interview                             | Individual interview (conducted by a clinical interviewer). Interview lasting approximately 30 minutes. The interview questions centred on mathematical tasks.                                                                              | Metacognition as comprising three main components: recognition of mistakes, adaptability and awareness and expression of thought. |}

<p>| Epistemic Metacognition Measure (retrospective interview) | -                                                      | Mason, Boldrin, &amp; Ariasi (2010)                                   | A retrospective interview comprising 4 questions (literature base identified as Hofer, 2000). Aim of interview to seek reflection about four epistemic beliefs (in one common learning situation). | Epistemic thinking as a metacognitive process. Study focuses on topic-specific epistemic beliefs |</p>
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual Interview, Strategy use and metacognition</td>
<td>-</td>
<td>Explicitly described as structured interview. There are tasks involved so this could also be described as task based.</td>
<td>Thronsden (2011) Metacognitive questions followed every second arithmetic item. Structured interview. Responses scored 0, 1, or 2 dependant upon the quality of the response. Categories – procedural knowledge, declarative knowledge and situational knowledge.</td>
<td>Academic self-regulation – skilled self-regulation occurs when cognitive, metacognitive and motivational components are fully integrated. Strategy use in problem solving. Strategy selection as metacognitive, metacognition distinguishing between knowledge and regulation.</td>
</tr>
<tr>
<td>Interview about Metacognitive Awareness</td>
<td>IMA</td>
<td>Explicitly described as an interview</td>
<td>Schmitt &amp; Sha (2009) 8 questions modified from IRA and the Metacognitive Interview (Schmitt, 1998) A qualitative rubric was used for analysis.</td>
<td>Metacognition as both awareness and regulation of strategic skills. Metacognitive knowledge (person, task and strategy) – declarative, procedural and conditional. Metacognitive regulation – problem solving, self-monitoring, self-correcting.</td>
</tr>
<tr>
<td>Interview from Munich Longitudinal Study on the Genesis of Individual Competencies</td>
<td>-</td>
<td>Explicitly described as an interview.</td>
<td>Lockl &amp; Schneider (2006) An interview to assess declarative metamemory knowledge comprising the following items: preparation, retrieval, study time, number of items, colour of hair (irrelevant), random vs. categorised order.</td>
<td>Metacognition as knowledge and regulation. The role of metacognitive vocabulary. Relationship between metacognition and theory of mind.</td>
</tr>
<tr>
<td>Metacognitive</td>
<td>-</td>
<td>Explicitly stated</td>
<td>Lu (1995) Questions about understanding</td>
<td>Metacognitive knowledge (person, task</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interview</td>
<td>as being an interview.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Included records** (The record in bold is the primary record cited)


**Description of the tool**

A combination of original answers and 5-point rating scale type questions, and strategy variables) and metacognitive experiences (conscious experiences). Also describes metacognition as knowledge about cognition and regulation of cognition. Regulation of cognition referring to planning, monitoring and checking.

**How the concept being measured is described (e.g. metacognition, self-regulation...)**

- Original Standardized Test for Metacognition

**Description**

Although described as a test this tool is administered more like an interview where children are for example shown objects and asked questions about recall.

- Pupil Views Templates

**Description**

PVTs are a mediated interview (the visual template is a meditational tool). A three-way interaction between pupil, research and stimulus (PVT template). Speech and thought bubbles prompt children to talk about what they are thinking.

**How the concept being measured is described (e.g. metacognition, self-regulation...)**

The links between cognitive skills and metacognition. Cognitive skills explored using the Moseley et al. (2005) model (frameworks for thinking) and metacognition explored as metacognitive knowledge and metacognitive skillfulness (Veenman et al., 2005).
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrospective Questionnaire Interview</td>
<td>RQI</td>
<td>Explicitly stated as being an interview</td>
<td>Deductive coding scheme based on Moseley et al. (2005) model of frameworks for thinking and Veenman et al. (2005) conceptualisation of metacognition (knowledge and skilfulness)</td>
<td>Declarative knowledge (if students addressed ‘what’ they did). Procedural knowledge (if students addressed ‘how’ they accomplished with a particular strategy). Conditional knowledge (if students addressed ‘why’ and ‘when’ particular strategies were selected). Cognitive strategies are also defined.</td>
</tr>
<tr>
<td>Swanson Metacognitive Questionnaire</td>
<td>SMQ</td>
<td>Administered as a structured interview by Swanson (1990) completed individually by participants in Sperling et al. (2012).</td>
<td>Questionnaire in two sections (7 questions in total) – first the formation of ideas, secondly strategies used by the writers (children) to transpose ideas to text. Alongside students completed a daily journal – one student per day administered the questionnaire interview.</td>
<td>Knowledge and control of one’s thinking and learning activities. The distinction (or not) of metacognition from the general aptitude of learners is unclear.</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Task based methods</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Computer based measure of metacognitive skilfulness</td>
<td>-</td>
<td>The data comes from Logfiles recording what participants do on a computer based task.</td>
<td>Computerized inductive learning tasks. Logfiles were scored automatically (by the computer) on two measures of metacognition: the mean number of variables changed per experiment and the frequency of scrolling back (to earlier experiments). Both mean number of variables and frequency of scrolling back were taken as indicators of metacognitive skilfulness.</td>
<td>Metacognitive knowledge (declarative knowledge about the relationships between person, task and strategy) and skilfulness distinction (procedural knowledge for regulation and control over learning activities).</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Concept maps</td>
<td>-</td>
<td>Ritchard, Turner, &amp; Hadar (2009)</td>
<td>Students creating concept maps as part of their routine classroom activity.</td>
<td>Being metacognitive as being aware of one’s own cognitive resources. The importance of task demands, planning, monitoring and control are highlighted. Metastrategic knowledge.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Prompt: “What is thinking? When you tell someone you are thinking, what kind of things might actually be going on in your head?”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>An inductive approach to coding:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Associative responses – described actual people, places and things.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Emotional responses – an affective connection to thinking.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Strategic responses – mentioned specific or general action when engaging in thinking processes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Meta-responses – focussed on epistemology, understanding and conceptualisations of building knowledge.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Measures of prior knowledge, motivational goal orientation, learning outcomes, quantity, and quality of learning strategies</td>
<td></td>
</tr>
</tbody>
</table>
| Name of tool | Rationale for category | Included records (The record in bold is the primary record cited) | Description of the tool | How the concept being measured is described (e.g. metacognition, self-regulation...)
<p>|
|-------------|------------------------|----------------------------------------------------------------|-------------------------|--------------------------------------------------------------------------------|
| Metacognitive Knowledge Monitoring Assessment | KMA | Task based because the post-test KMA is explicitly linked to performance in a task. | Osborne (1998), Tobias &amp; Everson (1996) | KMA to ask students to estimate knowledge of words or ability to solve maths problems. Estimates are then compared with actual performance to generate a score. Metacognition as monitoring, evaluating and making plans for own learning. Metacognitive process – knowledge, monitoring and control. |
| Problem solving interview | - | Although described as an interview the interview itself focuses on tasks that were completed (and video recorded), the interview would not happen without these tasks. | Carr &amp; Jessup (1995); Carr &amp; Jessup (1997) | Children videoed solving problems (20 addition and 20 subtraction). Strategy use was observed and children asked about their strategy use. Discrepancies resolved with children via discussion. Metacognitive knowledge about mathematics strategy assessed immediately after initial strategy use (e.g. why did you use that way for this problem?). Strategy specific metacognitive knowledge. Metacognitive knowledge about strategies as a predictor of use of strategies and performance. |
| Strategy knowledge in the domain of Chemistry | = | A computer-based task underlies this approach. | Scherer &amp; Tiemann (2012) | Tasks designed by referring to PISA problem solving framework (OECD., 2010 September). Computer based. 20 items in five final tasks to measure strategy knowledge. Students required evaluating the appropriateness to solve a given problem or the Metacognitive factors including strategy knowledge. Metacognition as a key competency in education. Metacognitive knowledge about strategies as a predictor of problem solving competency. |</p>
<table>
<thead>
<tr>
<th>Name of tool</th>
<th>Rationale for category</th>
<th>Included records (The record in bold is the primary record cited)</th>
<th>Description of the tool</th>
<th>How the concept being measured is described (e.g. metacognition, self-regulation...)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metacognitive Interview</td>
<td>MCI</td>
<td>Although it has interview in the title this method is explicitly described as self-report, it is completed as a task based (problem solving) self-report.</td>
<td>Lefevre (1995) Seven page self-report measure to investigate metacognitive knowledge, metacognitive skill, and monitoring and self-awareness.</td>
<td>Awareness of knowledge and use of strategies (discussion around both cognitive and metacognitive). Flavell (1979) – awareness of own cognitive processes and the ability to regulate them. Three (interactive) variables of metacognitive knowledge: individual (self-knowledge), task (or information processing demands) and strategy variables (monitoring strategies).</td>
</tr>
<tr>
<td>Multi method assessment of metacognitive behaviours</td>
<td>-</td>
<td>Explicitly described as a multi-method tool</td>
<td>Shamir, Mevarech, &amp; Gida (2009) Metacognitive behaviours assessed via a combination of methods – interviews (self-reports post task), online observations during the task. Coding based on grounded analysis – behaviours reflecting metacognition. Declarative metacognitive behaviours assessed immediately after children performed the task – e.g. “Please tell me what you</td>
<td>Metacognition as cognition about cognition refers to Nelson and Narens (1990) distinction between object and meta-cognitive level of cognition (relationship between monitoring and control). Knowledge about cognition (declarative, procedural and conditional knowledge Regulation of cognition (planning, information management, monitoring, debugging and evaluation – during</td>
</tr>
<tr>
<td>Name of tool</td>
<td>Rationale for category</td>
<td>Included records (The record in bold is the primary record cited)</td>
<td>Description of the tool</td>
<td>How the concept being measured is described (e.g. metacognition, self-regulation...)</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Multi-Method Interview</td>
<td>MMI</td>
<td>Explicitly described as multi-method, an interview but also has elements of being task-based.</td>
<td>The MMI included a problem based clinical interview, the interview included card sorting and self-reporting. The problem was a mathematics problem – action cards were sequenced according to how participants solved the problem. The problem solving was videoed and then replayed to participants and they were asked to check the sequence of their cards as they watched the video. The sequence of metacognitive actions was used to hypothesise about individual metacognitive behaviour.</td>
<td>Lack of clarity in defining metacognition in the field is recognised. The importance of metacognition for learning is acknowledged. Metacognition as multidimensional, generally including interrelated components: knowledge of cognition, regulation of cognition. Metacognitive functions: metacognitive awareness, metacognitive evaluation and metacognitive regulation.</td>
</tr>
</tbody>
</table>
Table 4: Percentage of instances of different groups of tools or methods being used for each age

<table>
<thead>
<tr>
<th>Category</th>
<th>Number of tools</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-report, tests, surveys, questionnaires</td>
<td>49 (61%)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>31%</td>
<td>61%</td>
<td>70%</td>
<td>67%</td>
<td>71%</td>
<td>72%</td>
<td>74%</td>
<td>77%</td>
<td>77%</td>
<td>85%</td>
</tr>
<tr>
<td>Observation</td>
<td>6 (8%)</td>
<td>33%</td>
<td>46%</td>
<td>50%</td>
<td>13%</td>
<td>15%</td>
<td>12%</td>
<td>10%</td>
<td>6%</td>
<td>11%</td>
<td>9%</td>
<td>10%</td>
<td>11%</td>
<td>0%</td>
</tr>
<tr>
<td>Teacher ratings</td>
<td>5 (6)</td>
<td>11%</td>
<td>8%</td>
<td>0%</td>
<td>13%</td>
<td>4%</td>
<td>0%</td>
<td>4%</td>
<td>4%</td>
<td>5%</td>
<td>5%</td>
<td>2%</td>
<td>3%</td>
<td>4%</td>
</tr>
<tr>
<td>Interviews &amp; focus groups</td>
<td>11 (14%)</td>
<td>44%</td>
<td>38%</td>
<td>50%</td>
<td>38%</td>
<td>11%</td>
<td>12%</td>
<td>13%</td>
<td>12%</td>
<td>5%</td>
<td>8%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Task based</td>
<td>6 (8%)</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>47%</td>
<td>4%</td>
<td>4%</td>
<td>4%</td>
<td>3%</td>
<td>3%</td>
<td>7%</td>
<td>9%</td>
<td>12%</td>
</tr>
<tr>
<td>Multi-method</td>
<td>3 (4%)</td>
<td>11%</td>
<td>8%</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
<td>1%</td>
<td>4%</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>TOTAL % for each age*</td>
<td></td>
<td>9 (2%)</td>
<td>13 (2%)</td>
<td>13 (2%)</td>
<td>17 (3%)</td>
<td>46 (8%)</td>
<td>57 (10%)</td>
<td>70 (12%)</td>
<td>84 (15%)</td>
<td>74 (13%)</td>
<td>64 (11%)</td>
<td>61 (11%)</td>
<td>36 (6%)</td>
<td>26 (5%)</td>
</tr>
</tbody>
</table>

NOTES:
- 80 tools, 149 included records and 567 references to age (Records may have referred to multiple tools or methods; the age/age range in each record that used a particular tool or method were counted individually for each tool in a record).
- The figure in brackets is the number of instances this type of tool or method was used with this age
- * = Total number of records referencing this age group for all of the 567 references to different ages, extracted from the 149 included records.
### Table 5: Additional subject focus (where specified)

<table>
<thead>
<tr>
<th>Method type</th>
<th>Questionnaires, surveys, self-report, tests</th>
<th>Observational methods</th>
<th>Teacher ratings</th>
<th>Interviews &amp; focus groups</th>
<th>Task-based methods and tests</th>
<th>Multi-method tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mathematics</td>
<td>24%</td>
<td>0%</td>
<td>20%</td>
<td>9%</td>
<td>17%</td>
<td>33%</td>
</tr>
<tr>
<td>Literacy (first lang.)</td>
<td>24%</td>
<td>0%</td>
<td>0%</td>
<td>18%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Science</td>
<td>4%</td>
<td>17%</td>
<td>0%</td>
<td>0%</td>
<td>17%</td>
<td>0%</td>
</tr>
<tr>
<td>Computer/ internet</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Physical education</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Religious education</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Language learning</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>History</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Multiple subjects</td>
<td>12%</td>
<td>33%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>No additional focus</td>
<td>22%</td>
<td>50%</td>
<td>80%</td>
<td>73%</td>
<td>67%</td>
<td>67%</td>
</tr>
<tr>
<td>Totals</td>
<td>49 tools</td>
<td>6 tools</td>
<td>5 tools</td>
<td>11 tools</td>
<td>6 tools</td>
<td>3 tools</td>
</tr>
</tbody>
</table>
## Appendix A: Search strategy for all databases for searches conducted on 15.11.2012

<table>
<thead>
<tr>
<th>Database &amp; provider</th>
<th>Search string</th>
<th>Limits applied</th>
<th>n</th>
<th>n - duplicates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Australian Education Index (AEI)</strong></td>
<td>Pro Quest ab(metacognit* OR meta-cognit*) AND ab(measure OR assess* OR evaluate OR evaluat*) AND ab(student OR pupil OR school OR child OR children)</td>
<td>Date: After 1 January 1992</td>
<td>225</td>
<td>207</td>
</tr>
<tr>
<td><strong>British Education Index (BEI)</strong></td>
<td>Pro Quest ab((metacognit* OR meta-cognit*)) AND ab(measure) OR ab(assess*) OR ab(evaluate OR evaluat*) AND ab(student OR pupil OR school OR child OR children)</td>
<td>Date: After January 01 1992; Language: English; Age group: Adolescents (13-17), All children, Children (0-12 years), Infants (0-2), Pre-school children (2-4/5), Young children (0-8)</td>
<td>234</td>
<td>233</td>
</tr>
<tr>
<td><strong>ERIC</strong></td>
<td>Pro Quest ab(metacognit* OR meta-cognit*) AND ab(measure OR assess* OR evaluate OR evaluat*) AND ab(student OR pupil OR school OR child OR children)</td>
<td>Date: After January 01 1992; Language: English; Education level: Early childhood education, Elementary education, Elementary secondary education, Grade 1, Grade 10, Grade 11, Grade 12, Grade 2, Grade 3, Grade 4, Grade 5, Grade 6, Grade 7, Grade 8, Grade 9, High schools, Intermediate grades, Junior high schools, Kindergarten, Middle schools, Preschool education, Primary education, Secondary education</td>
<td>397</td>
<td>266</td>
</tr>
<tr>
<td><strong>First Search Article First</strong></td>
<td>ECO (kw: metacognit* OR kw: meta-cognit*) and (kw: measure OR kw: assess* OR kw: evaluate OR kw: evaluat*) and (kw: student OR kw: pupil OR kw: school OR kw: child OR kw: children)</td>
<td>Date: Yr 1992-2012</td>
<td>17</td>
<td>6</td>
</tr>
<tr>
<td><strong>Psych Articles</strong></td>
<td>Ebsco-host AB ( metacognit* OR meta-cognit* ) AND AB ( measure OR assess* OR evaluate OR evaluat* ) AND AB ( student OR pupil OR school OR child OR children )</td>
<td>Year of publication: from 1992 – 2012; Age: Childhood (Birth – 12 years); School age (6-12 Years); Adolescence (13-17)</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td><strong>PsycINFO</strong></td>
<td>Ebsco-host AB ( metacognit* OR meta-cognit* ) AND AB ( measure OR assess* OR evaluate OR evaluat* ) AND AB ( student OR pupil OR school OR child OR children )</td>
<td>Year of publication: from 1992 – 2012; Age: Childhood (Birth – 12 years); School age (6-12 Years); Adolescence (13-17 years); Preschool age (2-5 years)</td>
<td>624</td>
<td>615</td>
</tr>
</tbody>
</table>

**Total:** 2721 2089
**Appendix B: Numbers of records per database searched for each stage in search and screening.**

<table>
<thead>
<tr>
<th>Database searched</th>
<th>Total records</th>
<th>Post de-duplication</th>
<th>Excluded (First screening)</th>
<th>Records remaining (after first screening)</th>
<th>Not available</th>
<th>Excluded (Second screening)</th>
<th>Records forward to data extraction</th>
<th>Records excluded during data extraction</th>
<th>Records excluded (reliability, validity...)</th>
<th>Total number of included records</th>
</tr>
</thead>
<tbody>
<tr>
<td>AEI</td>
<td>225</td>
<td>207</td>
<td>173</td>
<td>34</td>
<td>12</td>
<td>19</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>BEI</td>
<td>234</td>
<td>233</td>
<td>231</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>ERIC</td>
<td>397</td>
<td>266</td>
<td>198</td>
<td>68</td>
<td>18</td>
<td>32</td>
<td>18</td>
<td>5</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>First Search Article First</td>
<td>17</td>
<td>6</td>
<td>6</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>First Search ECO</td>
<td>282</td>
<td>147</td>
<td>109</td>
<td>38</td>
<td>0</td>
<td>14</td>
<td>24</td>
<td>8</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Psych Articles</td>
<td>17</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>PsycINFO</td>
<td>624</td>
<td>615</td>
<td>335</td>
<td>280</td>
<td>6</td>
<td>159</td>
<td>115</td>
<td>21</td>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td>Web of Knowledge</td>
<td>925</td>
<td>615</td>
<td>512</td>
<td>103</td>
<td>4</td>
<td>84</td>
<td>15</td>
<td>2</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Citations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2721</strong></td>
<td><strong>2089</strong></td>
<td><strong>1564</strong></td>
<td><strong>525</strong></td>
<td><strong>40</strong></td>
<td><strong>310</strong></td>
<td><strong>175</strong></td>
<td><strong>36</strong></td>
<td><strong>3</strong></td>
<td><strong>149</strong></td>
</tr>
</tbody>
</table>
### Appendix C: Summary table – the ages each included tool or method has been used with

<table>
<thead>
<tr>
<th>Tool or method</th>
<th>Primary Citation</th>
<th>Total records</th>
<th>Number of ages</th>
<th>Age in Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bandura’s Self Efficacy for Self-Regulated Learning Scale</td>
<td>Zimmerman et al. (1992)</td>
<td>3</td>
<td>7</td>
<td>2 2 2 2 2 1 1 2</td>
</tr>
<tr>
<td>2. Cambridgeshire Independent Learning in the Foundation Stage Coding Framework (C.Ind.Le)</td>
<td>Whitebread et al. (2009)</td>
<td>2</td>
<td>3</td>
<td>1 2 1</td>
</tr>
<tr>
<td>3. Checklist of Independent Learning Development 3-5 (Child 3-5)</td>
<td>Whitebread et al. (2009)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>4. Child Assessment (CA)</td>
<td>Desoete (2009)</td>
<td>2</td>
<td>3</td>
<td>1 1</td>
</tr>
<tr>
<td>5. Cognitive Developmental aRithmetics test (CDR)</td>
<td>Desoete and Roeyers (2006a)</td>
<td>2</td>
<td>3</td>
<td>1 1</td>
</tr>
<tr>
<td>6. Classroom Coding System</td>
<td>Stright et al. (2001)</td>
<td>4</td>
<td>5</td>
<td>2 2 2 1 1</td>
</tr>
<tr>
<td>7. Clinical Interview</td>
<td>Pappas et al. (2003)</td>
<td>2</td>
<td>3</td>
<td>2 2 1</td>
</tr>
<tr>
<td>8. Computer based measure of metacognitive skilfulness</td>
<td>Veenman et al. (2004)</td>
<td>1</td>
<td>7</td>
<td>1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>9. Concept maps</td>
<td>Ritchhart et al. (2009)</td>
<td>1</td>
<td>9</td>
<td>1 1 1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>10. Conditional knowledge (part of a questionnaire)</td>
<td>Wolters (1996)</td>
<td>1</td>
<td>5</td>
<td>1 1 1 1 1 1</td>
</tr>
<tr>
<td>12. Epistemic metacognition measure</td>
<td>Mason et al. (2010)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>14. Index of Metacognitive Awareness about Writing (IMAW)</td>
<td>De Kruif (2000)</td>
<td>1</td>
<td>3</td>
<td>1 1 1</td>
</tr>
<tr>
<td>15. Index of self-efficacy for writing (ISEW)</td>
<td>De Kruif (2000)</td>
<td>1</td>
<td>3</td>
<td>1 1 1</td>
</tr>
<tr>
<td>16. Index of Self-Regulated Writing</td>
<td>De Kruif (2000)</td>
<td>1</td>
<td>3</td>
<td>1 1 1</td>
</tr>
<tr>
<td>Tool or method</td>
<td>Primary Citation</td>
<td>Total records</td>
<td>Number of ages</td>
<td>Age in Years</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>(ISRW)</td>
<td></td>
<td></td>
<td></td>
<td>4 5 6 7 8 9 10 11 12 13 14 15 16</td>
</tr>
<tr>
<td>17. Index of Reading Awareness (IRA)</td>
<td>Jacobs and Paris (1987)</td>
<td>12</td>
<td>7</td>
<td>4 6 8 8 6 3 1</td>
</tr>
<tr>
<td>18. Index of Science Reading Awareness (ISRA)</td>
<td>Yore et al. (1998)</td>
<td>3</td>
<td>6</td>
<td>2 2 3 3 3 2</td>
</tr>
<tr>
<td>19. Individual interview – strategy use and metacognition</td>
<td>Thronsd (2011)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>20. Integrated Learning Assessment</td>
<td>Silver et al. (2011)</td>
<td>1</td>
<td>3</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>21. Interview about Metacognitive Awareness (IMA)</td>
<td>Schmitt and Sha (2009)</td>
<td>1</td>
<td>6</td>
<td>1 1 1 1 1 1</td>
</tr>
<tr>
<td>22. Interview from the Munich Longitudinal Study ...</td>
<td>Lockl and Schneider (2006)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>23. Inventory of Metacognitive Self-Regulation (IMSR)</td>
<td>Howard et al. (2000b)</td>
<td>4</td>
<td>7</td>
<td>3 4 4 3 3 3 2</td>
</tr>
<tr>
<td>24. Junior Metacognitive Awareness Inventory (JrMAI)</td>
<td>Sperling et al. (2002)</td>
<td>7</td>
<td>9</td>
<td>1 4 4 4 3 3 3 2</td>
</tr>
<tr>
<td>25. Knowledge and skills questionnaire</td>
<td>de Jager et al. (2005)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>26. Learning strategies assessed by journal writing</td>
<td>Glogger et al. (2012)</td>
<td>1</td>
<td>3</td>
<td>1 1 1</td>
</tr>
<tr>
<td>27. Learning Through Reading Questionnaire (LTRQ)</td>
<td>Butler et al. (2011)</td>
<td>1</td>
<td>5</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>28. Metacognition Applied to Physical Activities Scale (MAPAS)</td>
<td>Settanni et al. (2012)</td>
<td>1</td>
<td>4</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>29. Metacognition of Nature of Science Scale (MONOS)</td>
<td>Peters (2008)</td>
<td>2</td>
<td>4</td>
<td>1 1 2 2</td>
</tr>
<tr>
<td>30. Metacognition Scale</td>
<td>Yildiz et al. (2009)</td>
<td>1</td>
<td>4</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>31. Metacognitive Processes in Physical Education Questionnaire (MPIPEQ)</td>
<td>Theodosiou et al. (2008)</td>
<td>1</td>
<td>7</td>
<td>1 1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>Tool or method</td>
<td>Primary Citation</td>
<td>Total records</td>
<td>Number of ages</td>
<td>Age in Years</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>33. Metacognitive Attribution Assessment (MAA)</td>
<td>Desoete et al. (2001)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>34. Metacognitive Awareness Inventory (MAI)</td>
<td>Schraw and Dennison (1994)</td>
<td>5</td>
<td>6</td>
<td>1 2 2 3 3 3 3</td>
</tr>
<tr>
<td>35. Metacognitive Awareness of Reading Strategies Inventory (MARI)</td>
<td>Mokhtari and Reichard (2002)</td>
<td>4</td>
<td>8</td>
<td>1 2 3 3 2 2 1 1</td>
</tr>
<tr>
<td>36. Metacognitive Awareness of Reading Strategies Inventory (MARI)</td>
<td>Dermitzaki and Efklides (2001)</td>
<td>4</td>
<td>9</td>
<td>1 1 1 1 3 3 2 2 2</td>
</tr>
<tr>
<td>37. Metacognitive Interview</td>
<td>Lu (1995)</td>
<td>1</td>
<td>4</td>
<td>1 1</td>
</tr>
<tr>
<td>38. Metacognitive Interview (MCI)</td>
<td>Lefevre (1995)</td>
<td>1</td>
<td>7</td>
<td>1 1 1 1 1 1 1</td>
</tr>
<tr>
<td>39. Metacognitive Knowledge in Mathematics Questionnaire (MKMQ)</td>
<td>Efklides and Vlachopoulos (2012)</td>
<td>1</td>
<td>5</td>
<td>1 1 1 1 1 1</td>
</tr>
<tr>
<td>40. Metacognitive Knowledge Monitoring Assessment (KMA)</td>
<td>Tobias and Everson (1996)</td>
<td>2</td>
<td>3</td>
<td>1 1</td>
</tr>
<tr>
<td>41. Metacognitive Knowledge Questionnaire</td>
<td>Metallidou and Vlachou (2010)</td>
<td>1</td>
<td>3</td>
<td>1 1 1</td>
</tr>
<tr>
<td>42. Metacognitive Knowledge Test</td>
<td>Neuenhaus et al. (2011)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>43. Metacognitive Questionnaire</td>
<td>Okamoto and Kitao (1992)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>44. Metacognitive Skills and Knowledge Assessment (MSA)</td>
<td>Desoete et al. (2001)</td>
<td>3</td>
<td>4</td>
<td>1 1 2 2</td>
</tr>
<tr>
<td>45. Metacognitive Strategies (MSTRAT)</td>
<td>Roeschl-Heils et al. (2003)</td>
<td>1</td>
<td>3</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>46. Metacomprehension Strategy Index (MSI)</td>
<td>Schmitt (1990)</td>
<td>9</td>
<td>8</td>
<td>1 4 5 4 1 1 4 4</td>
</tr>
<tr>
<td>Tool or method</td>
<td>Primary Citation</td>
<td>Total records</td>
<td>Number of ages</td>
<td>Age in Years</td>
</tr>
<tr>
<td>---------------</td>
<td>------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>47. Motivated Strategies for Learning Questionnaire (MSLQ)</td>
<td>Pintrich and De Groot (1990)</td>
<td>9</td>
<td>7</td>
<td>2 4 6 3 5 3</td>
</tr>
<tr>
<td>48. Multi method assessment of metacognitive behaviours</td>
<td>Shamir et al. (2009)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>50. Observation (CASE@KS1)</td>
<td>Larkin (2006)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>51. Original standardized test for metacognition</td>
<td>Kreutzer et al. (1975)</td>
<td>3</td>
<td>7</td>
<td>1 2 3 3 2 2 2</td>
</tr>
<tr>
<td>52. Private speech coding</td>
<td>Daugherty and Logan (1996)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>53. Problem solving interview</td>
<td>Carr and Jessup (1995)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>54. Prospective Assessment of Children (PAC)</td>
<td>Desoete (2007)</td>
<td>2</td>
<td>3</td>
<td>2 1 1</td>
</tr>
<tr>
<td>55. Pupil Views Templates (PVTs)</td>
<td>Wall (2008)</td>
<td>3</td>
<td>10</td>
<td>1 1 1 1 1 2 3 1 1</td>
</tr>
<tr>
<td>56. Questionnaire about Learning in Mathematics (QLM)</td>
<td>Peklaj and Vodopivec (1998)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>57. Questionnaire about Learning Slovene Language (QLSL)</td>
<td>Peklaj (2001)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>58. Questionnaire about metacognitive beliefs</td>
<td>van der Zee et al. (2006)</td>
<td>2</td>
<td>3</td>
<td>2 2 2</td>
</tr>
<tr>
<td>59. Questionnaire based on Think Aloud</td>
<td>Schellings (2011)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>60. Rating Student Self-Regulated Learning Outcomes: A Teacher Scale</td>
<td>Zimmerman and Martinez-Pons (1988)</td>
<td>2</td>
<td>6</td>
<td>1 1 1 1 1 1</td>
</tr>
<tr>
<td>61. Reading Strategy use scale (RSU scale)</td>
<td>Pereira-Laird and Deane (1997)</td>
<td>1</td>
<td>5</td>
<td>1 1 1 1 1</td>
</tr>
<tr>
<td>Tool or method</td>
<td>Primary Citation</td>
<td>Total records</td>
<td>Number of ages</td>
<td>Age in Years</td>
</tr>
<tr>
<td>---------------------------------------------------</td>
<td>--------------------------------</td>
<td>---------------</td>
<td>----------------</td>
<td>--------------</td>
</tr>
<tr>
<td>62. Retrospective Assessment of Children (RAC)</td>
<td>Desoete (2007)</td>
<td>2</td>
<td>3</td>
<td>2 1 1</td>
</tr>
<tr>
<td>63. Retrospective Questionnaire Interview (RQI)</td>
<td>Short (2001)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>64. Self Regulated Learning Scale (SRL)</td>
<td>Prupas (1995)</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>65. Self report metacognitive learning strategies</td>
<td>Leutwyler (2009)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>66. Self-Assessment in Metacognitive Comprehension Strategies Reading Survey (SAMS)</td>
<td>Pinto (2009)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>67. Self-Directed Learning Instrument</td>
<td>Hwang (1999)</td>
<td>2</td>
<td>4</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>68. Self-Efficacy and Metacognition Learning Inventory – Science (SEMLI-S)</td>
<td>Thomas et al. (2008)</td>
<td>1</td>
<td>4</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>69. Self-efficacy for Learning Form (SELF)</td>
<td>Zimmerman and Kitsantas (2005)</td>
<td>3</td>
<td>6</td>
<td>1 1 2 3 1 1 1</td>
</tr>
<tr>
<td>70. Self-Regulated Learning Strategies Measurement Questionnaire</td>
<td>Eom (1999)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>71. Self-report for cognitive and metacognitive learning strategies</td>
<td>Wolters (1999)</td>
<td>2</td>
<td>5</td>
<td>1 1 2 1 1</td>
</tr>
<tr>
<td>72. State Metacognitive Inventory</td>
<td>O’Neil and Abedi (1996)</td>
<td>5</td>
<td>8</td>
<td>2 2 3 1 1 2 1 1</td>
</tr>
<tr>
<td>73. Strategy knowledge in the domain of Chemistry</td>
<td>Scherer and Tiemann (2012)</td>
<td>1</td>
<td>2</td>
<td>1 1</td>
</tr>
<tr>
<td>74. Swanson Metacognitive Questionnaire (SMQ)</td>
<td>Swanson (1990)</td>
<td>3</td>
<td>5</td>
<td>2 2 2 1 1</td>
</tr>
<tr>
<td>75. Teacher Rating</td>
<td>Sperling et al. (2002)</td>
<td>2</td>
<td>5</td>
<td>1 1 2 2 1</td>
</tr>
<tr>
<td>Tool or method</td>
<td>Primary Citation</td>
<td>Total records</td>
<td>Number of ages</td>
<td>Age in Years</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------------------------</td>
<td>---------------</td>
<td>---------------</td>
<td>--------------</td>
</tr>
<tr>
<td>76. The Teacher Rating</td>
<td>Desoete (2008)</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>77. Think About Reading Index (TARI)</td>
<td>Schreiber (2003)</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>78. Think Aloud Protocol(s) (TAP/TAPs)</td>
<td>Veenman et al. (2005)</td>
<td>19</td>
<td>10</td>
<td>1 1 5 6 7 5 8 6 6 4</td>
</tr>
<tr>
<td>79. Worksamples Interview</td>
<td>van Kraayenoord and Paris (1997)</td>
<td>1</td>
<td>4</td>
<td>1 1 1 1</td>
</tr>
<tr>
<td>80. Würzburg Metamemory Test</td>
<td>van Kraayenoord and Schneider (1999)</td>
<td>2</td>
<td>6</td>
<td>1 1 1 1 1 1</td>
</tr>
</tbody>
</table>
Figure 1: Example of data extraction for one of the included tools (IMSR).
**Figure 2:** Flow diagram showing numbers of records throughout searching, screening, and data extraction, based on the PRISMA flow diagram (Moher et al., 2009)