Developmental Review 34 (2014) 344–377

Contents lists available at ScienceDirect

Developmental Review

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

Defining and measuring conceptual knowledge in mathematics

Noelle M. Crooks [a,*], Martha W. Alibali [b]

[a] Broward College, Davie, FL, USA
[b] University of Wisconsin, Madison, WI, USA

ARTICLE INFO

Article history:
Received 30 September 2013
Revised 1 October 2014
Available online 6 November 2014

Keywords:
Mathematical thinking
Conceptual knowledge
Equivalence
Cardinality
Inversion

ABSTRACT

A long tradition of research on mathematical thinking has focused on procedural knowledge, or knowledge of how to solve problems and enact procedures. In recent years, however, there has been a shift toward focusing, not only on solving problems, but also on conceptual knowledge. In the current work, we reviewed (1) how conceptual knowledge is defined in the mathematical thinking literature, and (2) how conceptual knowledge is defined, operationalized, and measured in three mathematical domains: equivalence, cardinality, and inversion. We uncovered three general issues. First, few investigators provide explicit definitions of conceptual knowledge. Second, the definitions that are provided are often vague or poorly operationalized. Finally, the tasks used to measure conceptual knowledge do not always align with theoretical claims about mathematical understanding. Together, these three issues make it challenging to understand the development of conceptual knowledge, its relationship to procedural knowledge, and how it can best be taught to students. In light of these issues, we propose a general framework that divides conceptual knowledge into two facets: knowledge of general principles and knowledge of the principles underlying procedures.

* Corresponding author. Broward College, Behavioral Sciences, 3501 SW Davie Rd, Building 1/151, Davie, FL 33314, USA. Fax: 954-201-6173.

E-mail address: nmcrooks@wisc.edu (N.M. Crooks).

http://dx.doi.org/10.1016/j.dr.2014.10.001
0273-2297/

Research on mathematical thinking has typically divided mathematics knowledge into two types: procedural knowledge and conceptual knowledge (e.g., Hiebert, 1986).[1] In many mathematical domains, research has focused on procedural knowledge, typically defined as knowledge of sequences of steps or actions that can be used to solve problems (e.g., Rittle-Johnson & Siegler, 1998). In line with this theoretical definition, the way in which procedural knowledge is measured has become relatively standardized: participants solve a set of problems, and a score is calculated based on how many correct answers they obtain or based on the specific procedures they use to arrive at those answers. In recent years, however, the number of studies focused on procedural knowledge has been eclipsed by a growing literature on conceptual knowledge (see Star, 2005). There has been a shift toward studying, not only how people solve problems, but also their understanding of mathematical concepts, more broadly.

This shift in research, from a focus on procedures to a focus on conceptual knowledge, mirrors a similar trend in the mathematics education community. Mathematics curricula in the US have traditionally emphasized teaching children problem-solving procedures, with less emphasis on teaching the conceptual basis of the skills being learned (Stigler & Hiebert, 1999). However, recent reform efforts – as reflected, for example, in the standards from the National Council of Teachers of Mathematics and in the Common Core State Standards – have placed comparable emphasis on students having integrated conceptual and procedural knowledge (e.g., National Council of Teachers of Mathematics, 2000; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). The general consensus, in research on mathematical thinking and in mathematics education, is that having conceptual knowledge confers benefits above and beyond having procedural skill.

The literature suggests a number of specific ways in which conceptual knowledge might prove useful. Some of the reported benefits connect directly to procedural skills. For example, conceptual knowledge has been shown to help people evaluate which procedure is appropriate in a given situation (e.g., Brownell, 1945; Byrnes & Wasik, 1991; Carr, Alexander, & Folds-Bennett, 1994; Garofalo & Lester, 1985; Greeno, 1978; Schneider & Stern, 2012). Conceptual knowledge also allows for more flexible problem solving, in that people who understand the conceptual underpinnings of a procedure are more likely to successfully generalize it to novel problems (e.g., Baroody & Dowker, 2003; Baroody, Feil, & Johnson, 2007; Blote, Klein, & Beishuizen, 2000; National Council of Teachers of Mathematics, 2000; Rittle-Johnson, Siegler, & Alibali, 2001). Once a problem has been solved, conceptual knowledge can also be used to check whether the solution is reasonable (e.g., Brownell, 1945; Carr et al., 1994; Garofalo & Lester, 1985).

It has also been suggested that conceptual knowledge provides more general benefits. The Common Core State Standards (National Governors Association Center for Best Practices and Council of Chief State School Officers, 2010), for example, explicitly mention that teaching conceptual knowledge in addition to procedures is a way to instill deeper and longer-lasting mathematical understanding. Thus, there is a widely held belief that conceptual knowledge plays an important role in mathematics learning.

Despite a clear movement in both research and educational practice toward emphasizing conceptual knowledge in addition to procedural knowledge, there are several obstacles standing in the way of a comprehensive understanding of conceptual knowledge. One major hurdle for researchers is that there does not appear to be a clear consensus in the literature as to what exactly conceptual knowledge is and how best to measure it. The term “conceptual knowledge” has come to denote a wide array of constructs, making it difficult to understand the major findings in the field, the ways in which conceptual knowledge relates to procedural knowledge, and the most effective ways to utilize current research to guide instructional practices (e.g., Baroody et al., 2007; Star, 2005). In particular, the diverse ways in which conceptual knowledge has been defined theoretically and the diverse ways in which it has been measured have created a wide-ranging literature in which a consistent “bigger picture” is hard to find.

[1] Some conceptualizations of mathematical knowledge include additional knowledge types (e.g., Rittle-Johnson & Koedinger, 2005) or divide knowledge into slightly different categories (e.g., Reason, 2003). The conceptual/procedural distinction, however, remains the dominant framework in the literature.

We believe that it is important to identify and adequately measure conceptual knowledge, as distinct from procedural knowledge – even though the "wilderness that the procedural/ conceptual distinction has created" (as one reviewer of an early version of this paper termed it) appears both dense and tangled. Taming this "wilderness" could have both practical and theoretical benefits. In practical terms, educators sometimes wish to assess whether students have appropriate conceptual knowledge, or whether different instructional practices lead learners to acquire or construct appropriate conceptual knowledge. For such purposes, tools that measure conceptual knowledge with high validity and reliability are needed.

From a theoretical perspective, understanding knowledge change is at the heart of understanding learning, development, and instruction. Any complete, mechanistic account of mathematics learning should be able to specify the nature of the knowledge that people activate and apply in mathematical reasoning, and how aspects of that knowledge change over time and with experience. Of course, it may be difficult to distinguish conceptual from procedural knowledge at some points in development, because the two forms of knowledge are deeply intertwined (see, e.g., Baroody et al., 2007). Nevertheless, it seems likely that, at certain points in development, people may have only one form of knowledge or the other, or they may apply only one form of knowledge, for example, when interpreting mathematical situations or solving problems. Conceptual and procedural knowledge may also develop along different paths in different circumstances (e.g., with different sorts of mathematical experience). Such phenomena are important for theories of cognitive development and learning to acknowledge and to explain. A deeper understanding of conceptual knowledge will be of value for researchers espousing a range of theoretical perspectives who are interested in mathematical thinking, learning, and instruction.

The current review: overview

To investigate these issues, we conducted a literature review characterizing the state of research on conceptual knowledge of mathematics. This work incorporates both (1) a general survey of the psychological literature on mathematical thinking, and (2) focused reviews of specific mathematical domains. We sought to identify how the term "conceptual knowledge" is used in the mathematical thinking literature, both in general, and by examining research in three mathematical domains: equivalence, cardinality, and inversion. Each of these domains has generated a substantial amount of research, much of which has focused on characterizing and measuring conceptual knowledge.

Note that the purpose of the general review was to identify definitions of conceptual knowledge in the literature "at large." Because of the wide range of content domains covered in the general review, it was not practical to also consider how conceptual knowledge was operationalized and measured in the general review; however, this was a central aim of the domain-specific reviews. For each of the three target domains, we consider not only the definitions of conceptual knowledge that researchers offer, but also how researchers measure conceptual knowledge, and how these measurement approaches align with the provided definitions.

General mathematics literature

For the general review, we performed a search in PsycINFO using the terms "concept*" and "math*" to identify a preliminary set of sources ($n = 14{,}897$).[2] From this set, we excluded sources using an initial set of broad criteria. Specifically, we excluded unpublished dissertations, book reviews, sources whose primary focus was a domain other than math (e.g., physics), experiments using animal subjects, neuroimaging studies, and sources unavailable in English. This narrowed our list to 4282 possible sources. Next we reviewed the abstracts of these sources to further narrow the search to those that focused on conceptual knowledge in some way ($n = 367$), excluding sources focused only on knowledge of a specific mathematical concept or on academic achievement in general. Our process for identifying sources from this set to include for the general review and the domain specific reviews is depicted in Fig. 1.

[2] Search current as of 9:30 am, March 21, 2014.

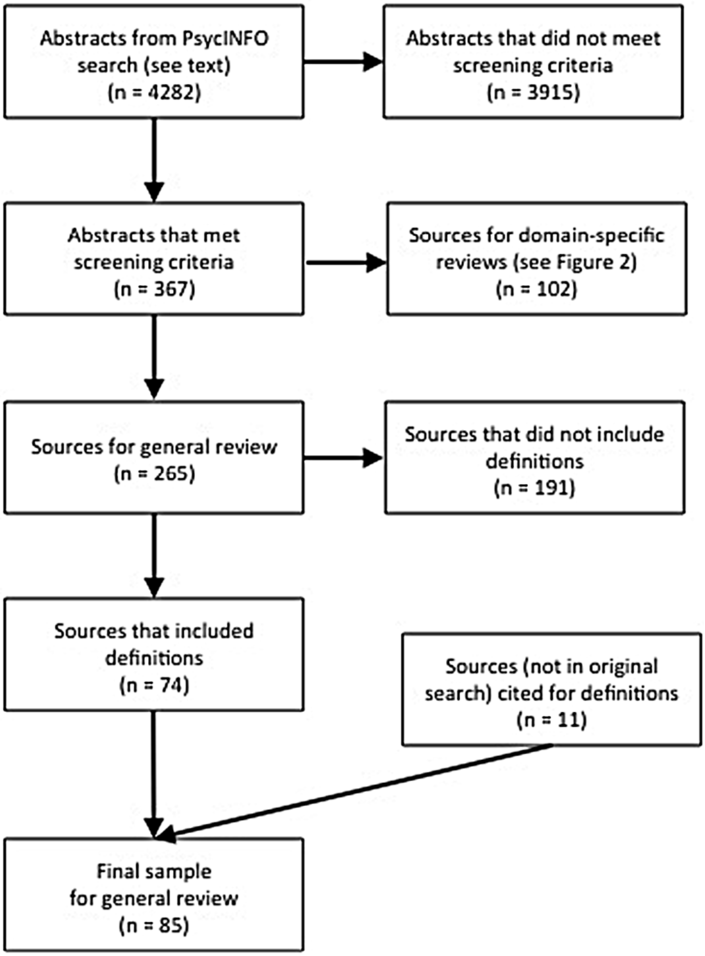

Fig. 1. The selection process for sources in the general literature review.

Within this set, those that focused on the topics of our domain specific reviews (i.e., equivalence, cardinality, and inversion) were set aside for those reviews ($n = 102$), so that there would not be overlap between the general review and the domain specific reviews. We then read each of the remaining general sources ($n = 265$) and identified those that explicitly defined conceptual knowledge. In this way, we generated a sample of 74 sources (articles, chapters, or books) for the general review. These sources spanned a range of topics, from general conceptual knowledge to knowledge in specific domains (e.g., proportional reasoning, decimals, and so forth).

For each source, we identified all given definitions of conceptual knowledge, allowing us to assess the types of definitions provided and the frequency with which different definition types occurred. We also noted instances in which specific definitions were cited from other sources, so that we could evaluate the frequency of unique definitions and determine which definitions were most frequently cited or influential. In doing so, we identified a small number of sources ($n = 11$) that were cited by other authors as the basis for their definitions of conceptual knowledge, but that did not appear in our original PsycINFO search. We added these sources to the sample for the general review, yielding a final sample of 85 sources. The definition types identified in the general review were later used to categorize definitions in the subsequent domain-specific reviews.

Table 1
Summary of conceptual knowledge definition types.

Definition type	Explanation	Example
Connection Knowledge	Relationships within a domain	"…knowledge that is rich in relationships. It can be thought of as a connected web of knowledge, a network in which the linking relationships are as prominent as the discrete pieces of information." (Hiebert & Lefevre, 1986, pp. 3–4)
General Principle Knowledge	General rules, facts, and definitions	"…static knowledge about facts, concepts, and principles that apply within a certain domain." (de Jong & Ferguson-Hessler, 1996, p. 107)
Knowledge of Principles Underlying Procedures	The basis for procedures	"…conceptual underpinning of the subject specific procedures" (Pardhan & Mohammad, 2005, p. 7)
Category Knowledge	The categories that organize information	"Subtypes of concepts include taxonomic categories…" (Byrnes, 1992, p. 236)
Symbol Knowledge	Symbol meanings	"…conceptual knowledge can be defined as the awareness of what mathematical symbols mean…" (Ploger & Hecht, 2009, p. 268)
Domain Structure Knowledge	The organization of mathematics	"…understanding of the underlying structures of mathematics" (Robinson & Dube, 2009a, p. 193)

Definitions of conceptual knowledge

This general literature review revealed six main types of definitions for conceptual knowledge: connection knowledge, general principle knowledge, knowledge of principles underlying procedures, category knowledge, symbol knowledge, and domain structure knowledge (see Tables 1 and 2). Some definitions fell into multiple categories and were coded as such. For example, "implicit or explicit understanding of the principles that govern a domain and of the interrelations between pieces of knowledge in a domain" (Rittle-Johnson & Alibali, 1999, p. 175) was considered both a general principle knowledge definition and a connections definition. Definitions varied in specificity and in the extent to which they distinguished conceptual and procedural knowledge.

Connection knowledge

The most common definition type, found in 65% ($n = 55$) of the sources, characterized conceptual knowledge as understanding of relationships and connections within a domain. The pieces being linked by these relationships varied across definitions, with some focusing on connections among ideas (e.g., Hiebert & Lefevre, 1986) and others on connections among things like symbols and numbers (e.g., Dixon & Moore, 1996).

General principle knowledge

Fifty-six percent ($n = 48$) of the sources defined conceptual knowledge as understanding of the principles that govern a domain. Some of these definitions noted that such knowledge is typically abstract, or that principle knowledge does not have to be verbalizable (e.g., Rittle-Johnson & Alibali, 1999; see below for discussion).

Knowledge of principles underlying procedures

Twenty-eight percent ($n = 24$) of the sources defined conceptual knowledge as understanding the basis for procedures, or knowing *why* a procedure works. Within this set, some articles discussed understanding the conceptual basis for entire procedures (e.g., Bolden & Newton, 2008), whereas others focused on knowing the rationale for individual steps within a procedure (e.g., Baroody et al., 2007).

Category knowledge

Definitions that focused on conceptual knowledge as category knowledge were found in 4% ($n = 3$) of the sources. These definitions noted that conceptual knowledge guides the formation of categories that can be used to organize knowledge within a domain. These definitions align more closely with general psychological research on concepts and categories (e.g., Medin, 1989; Rogers & McClelland, 2005) than with the bulk of the research on conceptual knowledge in mathematics.

Symbol knowledge

Definitions that focused on the importance of understanding and attaching meaning to mathematical symbols were found in 2% ($n = 2$) of the sources. Ploger and Hecht (2009), for example, stated that "conceptual knowledge can be defined as the awareness of what mathematical symbols mean…" (p. 268).

Domain structure knowledge

Definitions that focused on understanding the structure of mathematics as a whole or of a particular domain within mathematics were offered in 2% ($n = 2$) of the sources. For example, Richland, Stigler, and Holyoak (2012) defined conceptual knowledge as "attainment of an expertlike fluency with the conceptual structure of a domain" (p. 190).

Recurring definitions

There were two definitions that were cited frequently (i.e., more than five times) throughout the general review. Hiebert and Lefevre's (1986) assertion that conceptual knowledge "can be thought of as a connected web of knowledge, a network in which the linking relationships are as prominent as the discrete pieces of information" (pp. 3–4), was cited by 16 other articles. Additionally, Kilpatrick, Swafford, and Findell's (2001) definition of conceptual knowledge as "comprehension of mathematical concepts, operations, and relations" (p. 5) was cited in six articles.

Forms of conceptual knowledge

In addition to defining conceptual knowledge, several of the reviewed sources also discussed the ways in which such knowledge could be known or expressed. Most of these sources considered the question of whether conceptual knowledge must be explicit and verbalizable, or whether it can exist in more implicit forms. Although not all sources discussed this issue, of those that did ($n = 14$), the majority ($n = 12$) suggested that conceptual knowledge can be either explicit or implicit. For example, Hatano and Inagaki (1986) noted the existence of both "implicit and explicit forms of understanding" (p. 263).

Some literature suggests that there is a shift over time or development from more implicit knowledge to more explicit knowledge (e.g., Siegler & Stern, 1998). There may be change over time in the depth or amount of conceptual knowledge that learners possess, with implicit knowledge as a less advanced precursor to more explicit knowledge. In line with this view, a number of studies suggest that tasks requiring explicit verbalization may underestimate conceptual knowledge, as participants may have some conceptual knowledge that is not sufficiently advanced for them to be able to put it into words (e.g., Greeno, 1993). On the other hand, some of the evidence suggesting a shift from implicit to explicit forms of conceptual knowledge may be an artifact of the tasks used with different age groups. Young children, for example, often have difficulty on the more "explicit" knowledge tasks commonly used with older participants, but much of this difficulty may have to do with task demands (e.g., language skills) that are separate from knowledge of the mathematical concept itself. Broadly speaking, however, the consensus in the current literature is that conceptual knowledge can be either implicit or explicit.

Summary

This general review reveals that conceptual knowledge is defined in a range of ways in research on mathematical thinking. One point of agreement, however, is that conceptual knowledge can exist either in an explicit, verbalizable form or in a more implicit form. Next, we turned to detailed reviews of studies of conceptual knowledge in the three specific domains of equivalence, cardinality, and inversion. These reviews allowed us to explore possible differences within and across domains in how researchers define and measure conceptual knowledge.

Domain specific reviews

At the outset of this study, we selected three domains – equivalence, cardinality, and inversion – for more targeted investigations of how researchers define, operationalize, and measure conceptual knowledge. We chose these domains because there is extensive work on conceptual knowledge in each. Note that sources on topics in these three domains were not included in the general review, described above.

For the domain-specific reviews, we identified sources using three methods. First, we used the list of sources on conceptual knowledge in each domain that had been gathered in the larger literature review ($n = 102$). Second, we included sources that were cited by other authors within the domain as the basis for their definitions of conceptual knowledge, but that did not appear in the original search; this method added one source for the cardinality review. Third, for the domains of equivalence and inversion, we screened the sources cited in previous literature reviews (McNeil, 2014, for equivalence; Prather & Alibali, 2009, for inversion) to identify additional, potentially relevant sources.

These potential sources were then screened for whether they included explicit discussion of conceptual knowledge or measures of conceptual knowledge. The final set included 41 sources about equivalence, 31 sources about cardinality, and 32 sources about inversion (see Fig. 2).

For each source, we coded all definitions of conceptual knowledge in terms of the categories identified in the general review. Next, we examined the range of tasks used to assess conceptual knowledge in each domain.[3] Finally, we evaluated the alignment between definitions and tasks within each domain.

Equivalence

Mathematical equivalence is the idea that the two sides of an equation represent the same quantity. Procedural understanding of equivalence is typically measured by having participants solve equivalence problems, which are equations that have operands on both sides (e.g., $3 + 4 + 5 = 3 + \_\_$). Elementary school students in the US generally perform poorly on equivalence problems (e.g., McNeil, 2007), with some studies showing up to 86% of elementary school students failing to solve such problems correctly (Alibali, 1999). Because students often have the prerequisite arithmetic skills necessary to solve such problems, much of the research on equivalence has focused on deficits in conceptual knowledge as a possible explanation for their poor performance. Understanding mathematical equivalence is foundational for more advanced mathematics, particularly algebra (e.g., Knuth, Alibali, McNeil, Weinberg, & Stephens, 2005).

Defining conceptual knowledge

Only about one-quarter of the equivalence sources ($n = 11$) provided an explicit definition of the term "conceptual knowledge" (see Table 2). Among the studies that did include explicit definitions, four of the previously identified definition types were found: connection knowledge (73% of the given definitions), general principle knowledge (73%), category knowledge (9%), and symbol knowledge (9%). Note that definitions could fall into multiple categories.

Measuring conceptual knowledge

Twelve different tasks were used to measure various aspects of conceptual knowledge of equivalence. The tasks differ, not only in the specific mathematical concept of interest, but also in the form

[3] There is disagreement about whether using specific tasks to assess conceptual knowledge is appropriate (e.g., Faulkenberry, 2013) and also about whether conceptual knowledge can be measured separately from procedural knowledge. For the review portion of this paper, however, we focused on the authors' labeling of the tasks as measuring conceptual knowledge or not. We discuss this general issue in the final section of the paper.

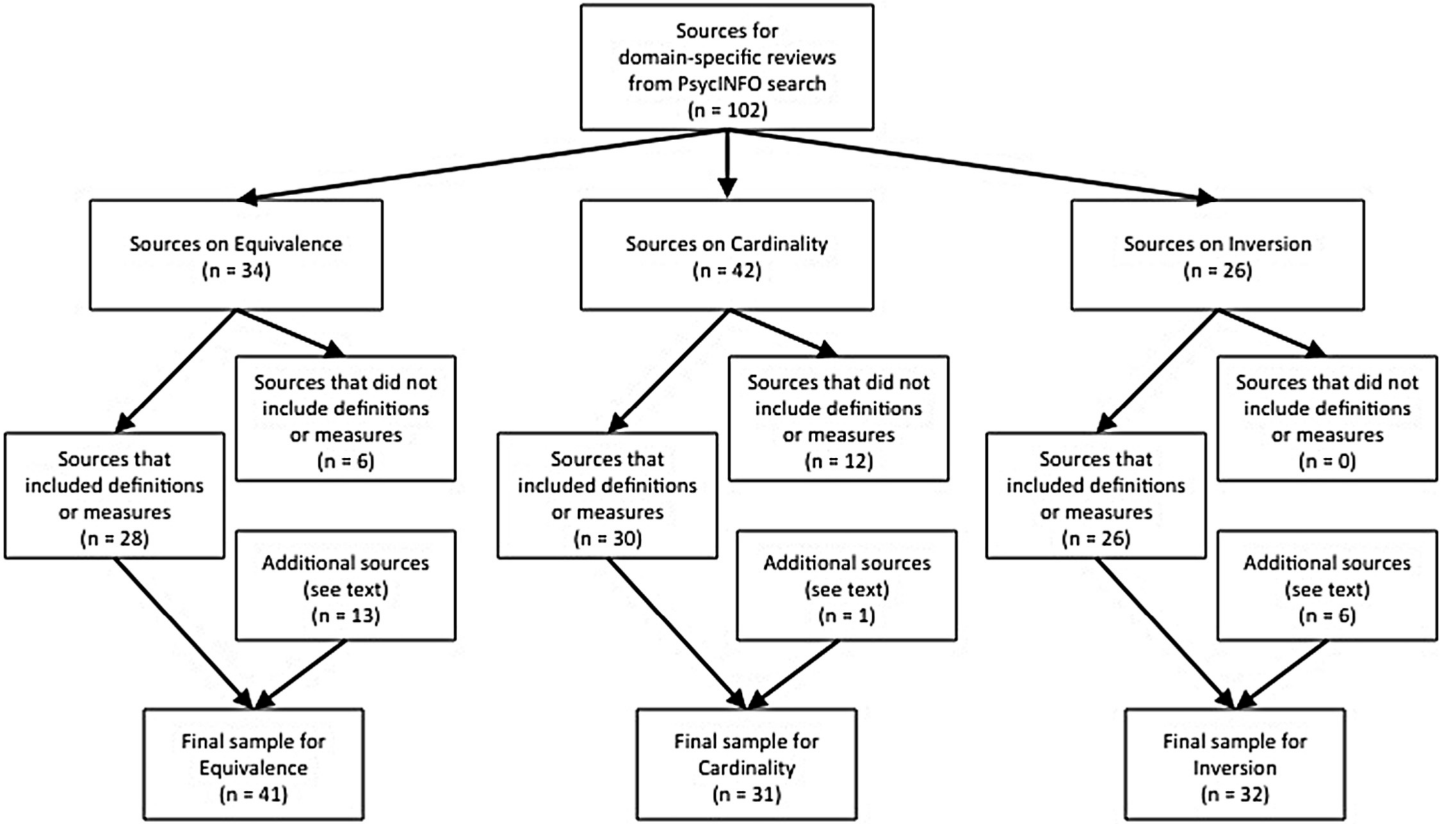

Fig. 2. The selection process for sources in the domain specific reviews.

Table 2
Conceptual knowledge definitions by type and domain.

Definition type	General	Equivalence	Cardinality	Inversion
Connection Knowledge	Alibali (2005) Ashlock (2001) Bisanz and LeFevre (1990) Blote et al. (2000) Bosse, Ries, and Chandler (2013) Burns (2011) Byrnes (1992) Byrnes and Wasik (1991) Capraro and Joffrion (2006) Carpenter (1986) Culatta (1993) Dixon and Moore (1996) Eisenhart, Borko, Underhill, and Brown (1993) English (1997) Gau Bartell, Webel, Bowen, and Dyson (2013) Goldman and Hasselbring (1997) Gutstein and Romberg (1995) Hallett, Nunes, and Bryant (2010) Hallett, Nunes, Bryant, and Thorpe (2012) Hiebert and Carpenter (1992) Hiebert and Lefevre (1986) Hiebert and Wearne (1986) Jitendra, DiPipi, and Perron-Jones (2002) Kadijevich and Haapasalo (2001) Kajander (2010) Kieren (1993) Kilpatrick et al. (2001) Kolloffel, Eysink, and de Jong (2011) Lachance and Confrey (2002) Lobato, Clarke, and Ellis (2005) Mabbot and Bisanz (2003) Mabbot and Bisanz (2008) Malloy (2009) Mann (2006) McCormick (1997) Middleton and Goepfert (1996) Miller and Hudson (2007) Mullins, Rummel, and Spada (2011) Ohlsson and Rees (1991) Ploger and Hecht (2009) Rittle-Johnson and Koedinger (2005) Rittle-Johnson and Siegler (1998) Rittle-Johnson and Star (2007) Rittle-Johnson and Star (2009) Rittle-Johnson et al. (2001) Rittle-Johnson, Star, and Durkin (2009) Rittle-Johnson, Star, and Durkin (2012) Saxe, Gearhart, and Suad Nasir (2001) Schneider and Stern (2010) Simon (1993) Simon (2006) Stump (2001) Tennyson and Cocchiarella (1986) Thanheiser (2012) Zell (2011)	Hattikudur and Alibali (2010) Matthews and Rittle-Johnson (2009) McEldoon, Durkin, and Rittle-Johnson (2013) McNeil and Alibali (2005a) Rittle-Johnson (2006) Rittle-Johnson and Alibali (1999) Saenz-Ludlow and Walgamuth (1998) Steinberg et al. (1991)	NA	Baroody, Lai, Li, and Baroody (2009); Baroody, Torbeyns, and Verschaffel (2009) Canobi (2005) Schneider and Stern (2009)

(continued on next page)

Table 2 (*continued*)

Definition type	General	Equivalence	Cardinality	Inversion
General Principle Knowledge	Alibali (2005) Baroody et al. (2007) Baroody, Wilkins, and Tiilikainen (2003) Berthold and Renkl (2009) Bisanz and LeFevre (1990) Burns (2011) Byrnes (1992) Byrnes and Wasik (1991) de Jong and Ferguson-Hessler (1996) Dixon and Moore (1996) English (1997) Faulkenberry (2013) Gau Bartell et al. (2013) Greeno (1993) Gutstein and Romberg (1995) Hecht and Vagi (2012) Hiebert and Wearne (1986) Jitendra et al. (2002) Kadijevich and Haapasalo (2001) Kaufmann, Handl, and Thony (2003) Kieran and Guzman (2006) Kilpatrick et al. (2001) Kolloffel et al. (2011) Lampert (1986) Lobato et al. (2005) Mabbot and Bisanz (2003) Mabbot and Bisanz (2008) Malloy (2009) Mullins et al. (2011) Niemi (1996) Ohlsson and Rees (1991) Patel and Canobi (2010) Peled and Segalis (2005) Rittle-Johnson and Koedinger (2005) Rittle-Johnson and Koedinger (2009) Rittle-Johnson and Siegler (1998) Rittle-Johnson and Star (2007) Rittle-Johnson and Star (2009) Rittle-Johnson et al. (2001) Rittle-Johnson et al. (2009) Rittle-Johnson et al. (2012) Schneider and Stern (2010) Schneider et al. (2011) Simon (1993) Simon (2006) Squire, Davies, and Bryant (2004) Stump (2001) Thanheiser (2012) Zell (2011)	De Cortre and Verschaffel (1981) DeCaro and Rittle-Johnson (2012) Fyfe, Rittle-Johnson, and DeCaro (2012) Hattikudur and Alibali (2010) Matthews and Rittle-Johnson (2009) McEldoon et al. (2013) Rittle-Johnson (2006) Rittle-Johnson and Alibali (1999)	Ferrara and Turner (1993) Frye, Braisby, Lowe, Maroudas, and Nicholls (1989) Gelman and Meck (1986) Greeno, Riley, and Gelman (1984) Nye, Fluck, and Buckley (2001) Sarnecka and Wright (2013) Siegler (1991) Sophian (1992)	Baroody and Lai (2007) Baroody et al. (2009a, 2009b) Bisanz et al. (2009) Canobi (2009) Robinson and LeFevre (2012) Stern (1992)

Table 2 (*continued*)

Definition type	General	Equivalence	Cardinality	Inversion
Knowledge of Principles Underlying Procedures	Baroody et al. (2003) Berthold and Renkl (2009) Bolden and Newton (2008) Bosse et al. (2013) Burns (2011) Eisenhart et al. (1993) Hatano (1988) Hatano and Inagaki (1986) Kajander (2010) Kieren (1993) Kilpatrick et al. (2001) Lampert (1986) Le Roux (2008) Lee (1998) Leinhardt (1988) Lithner and Palm (2010) Lobato et al. (2005) Malloy (2009) Montague (1998) Ohlsson and Rees (1991) Perry, VanderStoep, and Yu (1993) Resnick and Omanson (1987) Rittle-Johnson and Siegler (1998) Schmittau (2004) Semenza (2002)	NA	Briars and Siegler (1984) Freeman et al. (2000) Gelman and Meck (1986) Kamawar et al. (2010) LeFevre et al. (2006) Sarnecka and Carey (2008) Stock et al. (2009)	Baroody et al. (2009a, 2009b) Bisanz et al. (2009) Gilmore (2006) Gilmore and Bryant (2008)
Category Knowledge	Byrnes (1992) Byrnes and Wasik (1991) Miller and Hudson (2007)	DeCaro and Rittle-Johnson (2012)	NA	NA
Symbol Knowledge	Hecht (1998) Lachance and Confrey (2002) Ploger and Hecht (2009)	Saenz-Ludlow and Walgamuth (1998)	NA	NA
Domain Structure Knowledge	Bosse et al. (2013) Richland et al. (2012)	NA	NA	Canobi (2009) Robinson and Dube (2009a)

of conceptual knowledge being measured (i.e., implicit vs. explicit). Of the 12 tasks employed in the reviewed sources, some are used extensively throughout the literature, whereas others are used infrequently. Here we briefly describe the tasks that were used in at least 10% of the equivalence sources.

Equal sign definition

In this task, participants are asked to provide an explicit definition of the equal sign. This was the most common conceptual task in the equivalence literature and appeared in 56% of the sources ($n = 23$). Results from this task suggest that elementary and middle school students are generally poor at providing relational definitions for the equal sign (i.e., stating that the equal sign indicates that the two sides of an equation are the same/have the same value).

Sample equation evaluation

Forty-one percent of the sources ($n = 17$) included a task in which participants are asked to decide whether or not sample equations are correct or “make sense.” In these tasks, sample equations are typically presented in some combination of the following formats: traditional (e.g., $2 + 2 = 4$), non-traditional (e.g., $4 = 2 + 2$), equivalence/operations on both sides (e.g., $2 + 2 = 1 + 3$), and identity (e.g., $2 = 2$). Much of the data on such tasks is qualitative data from interviews (e.g., Baroody & Ginsburg, 1983; Falkner, Levi, & Carpenter, 1999). The typical finding is that, when faced with problems that do not follow a traditional format, many students assert that the equations are incorrect or that the experimenter must have made a mistake and needs to “fix” the equation to fit a more traditional problem format. These observations are supported by quantitative assessments of students’ abilities to evaluate sample equations, which demonstrate that students often have difficulty accurately evaluating equations with operations on both sides (e.g., Rittle-Johnson & Alibali, 1999).

Problem reconstruction

Problem reconstruction tasks were included in 31% of the sources ($n = 13$). In these tasks, participants view an equation for a set amount of time, and then the equation is hidden from view and they write down exactly what they saw. Performance on these tasks is typically poor, with many students incorrectly reproducing the structure of the target equations (e.g., McNeil, Fyfe, Petersen, Dunwiddie, & Brletic-Shipley, 2011).

Equal sign definition ratings

In this task, participants rate given definitions of the equal sign. About one quarter of the sources ($n = 11$) included a definition rating task. Participants are typically asked to rate three types of definitions: relational (e.g. “the same as”), operational (e.g., “the total”), and distractor (e.g., “the end of the problem”). Higher ratings of relational definitions, compared with operational or distractor definitions, are taken as evidence of conceptual knowledge. Data from these tasks provide varying estimates of students’ knowledge, although performance is typically better than for the *equal sign definition* task.

Equivalent equations

Equivalent equation tasks were found in 22% of the sources ($n = 9$). These tasks included five different types of problems in which participants are required to provide some information about equations. In type 1, participants are shown two equations with variables (e.g., $n + 3 = 7$ and $n + 3 - 5 = 7 - 5$) and asked to indicate whether the value of n is the same in both. In type 2, participants are shown two equations with variables and asked if knowing the value of one of those variables would assist them in solving the other equation (e.g., In the equation __ + 18 = 35, 17 goes in the blank. Can you use that knowledge to figure out what goes in the blank in __ + 18 + 27 = 35 + 27?). In type 3, participants are asked to explain whether transforming an equation in a specific way would maintain its value (e.g., does $a + b = (a + 1) + (b - 1)$?). In type 4, participants are asked to explain why performing the same operation on both sides maintains the equivalence relationship in an equation. Finally, in type 5, participants are asked to insert the number or symbol that would make two incomplete expressions equivalent. Performance on these tasks varies widely by type, with knowledge estimates based on type 1 exceeding estimates based on the other equivalent equations tasks. Additionally, the developmental trajectory of performance differs depending on the specific type of equivalent equations task used.

Problem solving

Although almost all of the studies required participants to solve equivalence problems, in only 17% ($n = 7$) of the sources was this activity characterized as conceptual in nature. In four of these sources, participants’ solutions were used to infer conceptual knowledge, whereas in the other three, participants were required to explain their problem-solving procedures.

Other, infrequently used tasks

Six additional, specific tasks were identified in the equivalence literature, but were used in fewer than 10% of sources. The *identifying sides* task, in which participants are asked to identify the two sides of an equation, was used in three sources (Matthews & Rittle-Johnson, 2009; Rittle-Johnson & Alibali, 1999; Rittle-Johnson, Matthews, Taylor, & McEldoon, 2011). Two sources included a *problem recognition* task in which participants view an equation for a set amount of time and then, once the equation has been hidden, select the equation they saw from of a set of samples (Alibali, Phillips, & Fischer, 2009; McNeil & Alibali, 2005b). In four sources, participants were asked to decide whether two given numbers were equal (*number equality* task; e.g., Rittle-Johnson et al., 2011). In one source participants were asked to define explicitly what it meant for two sets to be “equal” (*generic equal definition;* Rittle-Johnson & Alibali, 1999). One source included a *procedure assessment* task, in which participants decide whether a given procedure makes sense (Rittle-Johnson & Alibali, 1999). Finally, two sources included a *symbol sort* task, in which participants are presented with a set of cards depicting symbols

– three operators (e.g., +, –), three relational symbols (e.g., <, =), and two numbers – and asked to place the cards into three groups (Hattikudur & Alibali, 2010; Rittle-Johnson, 2006).

Linking definitions and measurement of conceptual knowledge

As demonstrated in this review, a variety of tasks have been used to assess conceptual knowledge in the domain of equivalence. These tasks vary widely in their performance requirements and in how outcomes are measured. For many of these tasks, however, it is unclear how they align with the broader definitions of conceptual knowledge espoused in the equivalence literature.

Of the five types of conceptual knowledge definitions explicitly mentioned, only symbol knowledge was clearly measured in a majority of sources. In 23 sources, participants were required to provide an explicit verbal definition for a mathematical symbol (*equal sign definition*) and in 11, participants were asked to rate given definitions of the symbol (*equal sign definition ratings*). Additionally, two sources included a *symbol sort* task in which participants were asked to group a variety of mathematical symbols. The frequency of these symbol-based tasks is notable, considering that the idea of conceptual knowledge as symbol knowledge was provided in only one source about equivalence (Saenz-Ludlow & Walgamuth, 1998). The study reported in that article did not use a task that specifically assessed symbol knowledge, but instead analyzed qualitative data from classroom interactions, and noted a change in students' interpretations of the equal sign symbol over time.

The other theoretical definitions provided in the literature were less clearly tied to the measures used in the studies of equivalence. The idea of conceptual knowledge as general principle knowledge links clearly to only two tasks, both of which were used infrequently: the *number equality* task and the *generic equal definition* task. The *number equality* task measures general principle knowledge by having participants demonstrate that they can compare two values and decide if they are the same, an ability that requires a general understanding of the concept of equality. The *generic equal definition* task requires participants to provide a general explanation of what it means for two things to be equivalent. Although other tasks get at more specific notions of equivalence, this task is one of few that differentiate general principle knowledge from symbol knowledge.

Although none of the equivalence sources defined conceptual knowledge as understanding principles underlying procedures, this type of knowledge was measured in a few ways. The clearest measure of this is the *procedure assessment* task, found in only one study, in which participants evaluate a procedure as making sense or not (Rittle-Johnson & Alibali, 1999). One would expect that participants would rate procedures adhering to principles they know more highly than procedures violating those principles. Some versions of the *equivalent equations* tasks could also be seen as measuring understanding of the conceptual basis for a procedure – specifically, these tasks measure knowledge of problem-solving shortcuts (e.g., if the same thing is done to both sides, computation is unnecessary), which is taken as demonstrating conceptual knowledge in other domains (discussed in more detail below, in the Inversion section). Seven sources used data from students' problem solving to infer conceptual knowledge of this type.

The most common definition of conceptual knowledge provided in the literature on equivalence, connection knowledge, was the least clearly measured. Although some tasks could be seen as requiring participants to connect pieces of knowledge (e.g., for the *equal sign definition* task, participants must connect their general understanding of equality to a specific symbol), none of the tasks directly measure the nature or quality of these connections. Only one task, the *symbol sort* task, requires participants to relate their knowledge of the equal sign to their understanding of both arithmetic operations and relational symbols. It could be argued that this task allows one to assess whether participants believe the equal sign to be more strongly connected to operators than to other relational symbols or to numbers. At the same time, this task might also be viewed as tapping category knowledge, specifically knowledge that relational symbols form a category that is distinct from operation symbols. Regardless of how the *symbol sort* task is construed, it was used in only two sources (Hattikudur & Alibali, 2010; Rittle-Johnson, 2006), and in neither was it explicitly linked to either a connection knowledge definition of conceptual knowledge or a category knowledge definition.

Many of the tasks had no clear ties to any given theoretical definitions. The *problem reconstruction* and *problem recognition* tasks, for example, are used to measure students' ability to encode mathematical equations without conceptual errors, even though conceptual knowledge was never defined as the ability to accurately encode perceptual information.

Overall, despite the large variety of tasks designed to measure conceptual knowledge of equivalence, it does not appear that all of the tasks have solid theoretical foundations. It is striking that the most commonly provided definition of conceptual knowledge in the equivalence literature (connection knowledge) was the least frequently measured in any direct way, whereas the most commonly measured aspect of conceptual knowledge (symbol knowledge) was only mentioned once as being a defining feature of conceptual knowledge. Additionally, there were some definitions identified in the general review that were never used in the equivalence literature, but that were nonetheless tapped by some of the tasks. The *symbol sort* task, for example, requires participants to put symbols into categories, but the category knowledge definition was never explicitly provided in this domain.

Cardinality

Research on children's ability to count has focused on knowledge of a set of "counting principles" that govern the ways in which sets can be enumerated and the types of things that can be counted. One of the "how to count" principles that has received a great deal of research attention is the cardinal principle. The cardinal principle states that, when counting a set, the numerical tag given to the final item represents a property of the set as a whole: its total numerosity (i.e., its cardinal value) (Gelman & Gallistel, 1978). Children's ability to count is foundational in mathematics, and relates to their learning of later arithmetic as well as to many real-world activities (e.g., Stock, Desoete, & Roeyers, 2009).

Defining conceptual knowledge

Almost 50% of the reviewed sources on cardinality provided a definition of the term conceptual knowledge ($n = 14$; see Table 2). These definitions fell into two categories: general principle knowledge (57% of the given definitions) and knowledge of principles underlying procedures (50%).

Measuring conceptual knowledge

A variety of tasks have been designed to assess children's understanding of cardinality, with most taking advantage of children's ability to enact or assess a counting procedure. The eight tasks identified in the current review varied in their performance requirements, with some requiring verbal responses and others relying on action. Measures occurring in 10% or more of the reviewed sources on cardinality are described here.

How many

Found in almost two-thirds of the sources ($n = 20$), *how many* tasks require children to count a set and then provide the cardinal value by answering the question "how many?" The specifics of the task vary by study, with some varying in the constraints placed on children's counting (e.g., start counting with this object), some varying in the types of items children are asked to enumerate (e.g., objects, sounds, etc.), and some varying in set size. One issue with these measures is that, in some cases, children may succeed on the tests without actually understanding the cardinal principle. Instead, they may simply be demonstrating "last word" responses, indicating that they have learned that the appropriate response to questions about numerosity is always the last word in the count (e.g., Bermejo, Morales, & Garcia deOsuna, 2004). Despite such concerns, data from these tasks suggest that understanding of cardinality improves with age and that performance is influenced by both problem context and set size.

Procedure assessment

Almost 40% of the sources ($n = 12$) included a task that required children to decide whether a counting procedure enacted by the experimenter was correct or incorrect. Importantly, most included not only correct and incorrect counts but also included "correct-but-unusual" or "pseudoerror" counts. These counts are technically correct, but do not follow a typical counting pattern (e.g., starting with the middle object of an array). Even fairly young children are generally very good at stating which counts are correct, but appropriately assessing incorrect and correct-but-unusual counts is harder.

Give a number

Almost 40% of the sources ($n = 12$) included a task in which children were asked to give the experimenter a certain subset of objects chosen from a larger set. For example, a participant might be given 15 objects and then be asked to give the experimenter six of them. Performance on these tasks is typically measured in one of two ways. The first way simply tallies success on each trial. The second way involves looking at the behaviors children use to solve the problem: those who count out the objects are considered to understand the link between counting and cardinal value, whereas those who simply grab a set of objects are not. As with the *how many* task, performance on *give a number* tasks improves with age and is affected by set size.

Other tasks

Five additional tasks occurred in fewer than 10% of the cardinality sources. Two of these tasks measured children's ability to predict the outcome of a specific count procedure; in one task, children simply predicted what would happen if a certain count sequence were enacted (*procedure outcome* task), whereas in the other, children decided whether two different count sequences would have the same result (called the *windows* task by the authors, because it involved counting windows; Cowan, Dowker, Christakis, & Bailey, 1996). An additional prediction task, found in two studies, required participants to use information about one enumerated set to predict the quantity of a different set (the *compare sets* task; Bialystok & Cobb, 1997; Sarnecka & Wright, 2013). In one source, children evaluated a puppet's statements about the result of a count sequence (called the *counterfactual* task because it involved rating some statements that were counterfactual; Freeman, Antonucci, & Lewis, 2000). Finally, two sources included a task designed to measure children's belief about numerosity as a defining feature of a set (called the *magic paradigm* by Gelman and Gallistel (1978) because it involved surreptitious changes to arrays as if by "magic"). In the *magic paradigm*, children decide whether or not a set stays fundamentally the same over a variety of changes (e.g., changes in color, numerosity, and so forth).

Linking definitions and measurement of conceptual knowledge

Although a variety of tasks are used to assess children's conceptual knowledge of the cardinal principle, the tasks are fairly similar. Many require participants to utilize their understanding of the counting procedure to determine the numerosity of a set. Although it may seem problematic that these tasks do not cleanly separate the procedural and conceptual aspects of cardinality knowledge, the tasks are actually quite relevant to the definitions of conceptual knowledge provided most frequently in the counting literature. In particular, the focus on conceptual knowledge as understanding of the principles underlying procedures connects well with the tasks used to measure cardinal principle understanding.

Although the tasks are generally well aligned with the definition of conceptual knowledge as knowledge of principles underlying procedures, different tasks measure slightly different facets of that understanding. Some tasks – the *how many* task, for example – require explicit implementation of a counting procedure, which may lead researchers to underestimate the knowledge of children who have some conceptual knowledge but who have not yet generated or mastered a correct procedure. Additionally, although those tasks can demonstrate that children are able to use a procedure that does not violate the cardinality principle, the tasks do not show that children know *why* the procedure is appropriate.

The *procedure outcome, windows, compare sets*, and *procedure assessment* tasks, on the other hand, allow for measurement of more implicit knowledge of the principles underlying procedures, since children are not required to count objects themselves. The goal of these measures is to disentangle children's ability to assess cardinality from their own ability to count. These tasks allow researchers to separate children's understanding of the core counting principles from knowledge of the standard counting procedure. For example, the inclusion of correct-but-unusual counts in the evaluation tasks requires children to know that a count can be correct even if it does not follow the conventional counting procedure. Overall, the tasks used in the cardinality literature appear to measure children's knowledge of the principles underlying the counting procedure, but the specific aspects of knowledge being tapped by different tasks vary.

One task, the *magic paradigm*, in which participants decide whether or not a set stays fundamentally the same over a variety of changes, aligns closely with the notion of conceptual knowledge as general principle knowledge. This task does not require children to count objects. Instead, it assesses the importance that children place on cardinal value in contrast to other characteristics of a set, and it provides information about what children believe to be the most important features of an array. Success on this task demonstrates that children understand that numerosity is a defining feature of a set, a general concept that will be useful for later mapping of the count words to specific cardinal values.

In sum, in the cardinality literature, commonly used conceptual knowledge tasks do map onto the theoretical definitions provided for conceptual knowledge. Many of the tasks used, however, make it hard to distinguish conceptual knowledge from superficial procedural knowledge (e.g., last word responding).

Inversion

The mathematical inversion principle holds that "inverse operations (e.g., addition and subtraction) involving the same value result in no net change" to the original quantity (i.e., $a + b - b = a$) (Prather & Alibali, 2009, p. 236). The inversion principle as it applies to addition and subtraction has received a great deal of research attention, and recent work has begun to address the principle as it applies to multiplication and division as well (e.g., Robinson & Dube, 2009a). Understanding the inverse relationships between operations is considered foundational to deep understanding of each operation on its own (i.e., one cannot fully understand addition until one understands how it relates to the other operations) (e.g., Bryant, Christie, & Rendu, 1999). Many researchers view understanding of inversion as an extension of the more general Piagetian notion of reversibility (e.g., Baroody & Lai, 2007).

Defining conceptual knowledge

Researchers who study inversion provided some explanation of the term "conceptual knowledge" in about one-third of sources ($n = 11$; see Table 2). The provided definitions, however, tended to be somewhat vague. Throughout the inversion literature, four of the definition types from the general review were found: connection knowledge (27% of the given definitions), general principle knowledge (55%), knowledge of principles underlying procedures (36%), and domain structure (18%).

Measuring conceptual knowledge

The inversion literature included only a small number of tasks assessing conceptual knowledge, with most tasks requiring participants to enact, explain, or evaluate a procedure. Of the four tasks identified in the inversion literature, only two occurred in more than 10% of the sources.

Inversion problem solving

The most common task for assessing conceptual knowledge of inversion, found in 84% of the sources ($n = 27$), was having participants solve inversion problems. Inversion problems follow a format that allows them to be solved using a shortcut. For example, the inversion problem $8 + 3 - 3 = ?$ can be solved in two main ways. A participant can calculate the answer, by first adding together 8 and 3 and then subtracting 3. Alternatively, a participant could use an "inversion shortcut," recognizing that no computation is needed because the operations "cancel" one another out. Although this measure may appear procedural, many researchers contend that conceptual knowledge of the inversion principle directly underlies use of the shortcut (e.g., Gilmore & Papadatou-Pastou, 2009).

Performance on inversion problems is typically assessed in one of three ways. The first way consists of simply measuring accuracy on inversion problems. The second way involves having participants solve both inversion problems and standard problems (e.g., $8 + 3 - 2 = ?$) and comparing accuracy across the two problem types. The third way again involves having participants solve both inversion problems and standard problems but comparing solution times (as opposed to accuracy) for the two problem types. The second and third methods rely on the assumption that accuracy will be higher and reaction time lower for inversion problems if one knows and uses the shortcut, since it does not require slow and potentially error-ridden calculations.

Inversion problems have been used in a variety of contexts (e.g., symbolic, concrete), with different pairs of operations (i.e., addition/subtraction, multiplication/division), and across different ages. Estimates of understanding vary depending on the specific form of the task used (e.g., Prather & Alibali, 2009).

Procedure assessment

In this task, participants are asked to observe two different methods for solving an inversion problem: full computation and the inversion shortcut. Participants are then asked to decide whether each procedure is acceptable, which procedure is best, and, in some cases, to explain their reasoning. Found in four sources, this task allows participants to demonstrate some understanding of the inversion shortcut before they begin to use it themselves (e.g., Robinson & Dube, 2009c). Some participants are able to clearly explain why one procedure is better, and other participants show a strong preference without being able to explain why (e.g., Robinson & Dube, 2009a). Thus, there is variability, not only in participants' preference for certain procedures, but also in their abilities to explicitly justify their preferences.

Other tasks

Two additional measures of conceptual knowledge were found in the inversion literature, but in fewer than 10% or in an unspecified number of the reviewed sources. One source included a *violation of expectation paradigm* in which participants were shown transformations of sets of physical objects and asked whether the resulting array was "normal" or not (Vilette, 2002). In other sources, participants were required to provide an explanation of their method for solving inversion problems. The frequency of this task is unclear, however, as many studies made mention of asking participants such questions but did not report any data from the responses.

Linking definitions and measurement of conceptual knowledge

Research on conceptual knowledge of inversion relies on a fairly small set of measures, allowing for easier comparison across studies than is possible in many mathematical domains. This has allowed researchers to draw a clear picture of development over time and across problem-solving contexts.

Most of the tasks used align clearly with the idea of conceptual knowledge as understanding of the principles underlying procedures. In fact, all but one of the tasks (the *violation of expectation paradigm*) assess participants' abilities to use, explain, or evaluate a procedure related to the inversion principle. However, despite the clear relationship between the definition and the tasks, only five of the sources actually provided this type of definition for conceptual knowledge. Additionally, these measures can be problematic because it is possible to use a procedure without understanding its conceptual basis, or to have some conceptual knowledge but not yet have generated a procedure based on it (Bisanz & LeFevre, 1992). Studies that ask participants to provide verbal justifications for their procedures or evaluations appear to avoid this issue, but such measures were rare in this literature, and they may also underestimate participants' implicit conceptual knowledge.

It could be argued that the *inversion problem solving* and *procedure assessment* tasks also align with two other definitions of conceptual knowledge. Specifically, understanding of the inversion shortcut might imply understanding of the relationship between operations, demonstrating both connection knowledge and understanding of domain structure. The link between the tasks and these definitions, however, is less direct, and it was not explicitly made by the researchers. Indeed, since one basic tenet of the inversion principle is that there is an inverse *relationship* between operations, all of the tasks end up providing some measure of conceptual knowledge as connection knowledge. This appears to be a characteristic of the domain, and is not necessarily based on theoretical notions of conceptual knowledge.

The *violation of expectation paradigm* does not strongly align with any of the definitions of conceptual knowledge found in the inversion literature, although a case could be made for its connection to any of those definitions. Children's ability to successfully assess an inverse transformation could be seen as an early form of understanding the connections between operations or as evidence for knowledge of the general principle of inversion. This task is also one of the few inversion tasks that provides a measure of implicit conceptual knowledge, in that it does not require participants to complete or explain any type of calculation.

In sum, although the majority of tasks used in the inversion literature do closely align with at least one of the theoretical definitions of conceptual knowledge, the specific form that the tasks take makes it hard to draw a clear line between responses based on conceptual knowledge and those based on procedural knowledge (e.g., Robinson & Dube, 2009b).

Conceptual knowledge across domains

A review of the literature on equivalence, cardinality, and inversion brings to light a few issues that cut across domains in research on conceptual knowledge. The first issue is that, despite the fact that all of the reviewed sources reported being "about" conceptual knowledge, only about 35% of the sources explicitly defined the term "conceptual knowledge." The lack of definitions makes it difficult to assess whether the methods used really tap the intended constructs. Additionally, the three domains differed in the frequency and clarity with which conceptual knowledge was defined. The cardinality literature most frequently included definitions, followed by the inversion literature. The definitions provided in the inversion literature, however, were among the least precise. The equivalence literature seldom offered explicit definitions, but those that were provided tended to be more descriptive.

The second major issue is that, across the sources that did explicitly define conceptual knowledge, there appeared to be little consensus about what this type of understanding actually entails (see Table 2). The use of multiple definitions presents a challenge for forming a coherent picture of the development of mathematical knowledge across content domains. The most commonly used definitions differed by mathematical domain, making it hard to compare across the domains. The definition used might affect the inferences one makes about how much children know at different points in developmental time, or about whether particular types of instruction have the intended effects.

For example, there has been much interest in understanding how procedural and conceptual knowledge connect and the order in which they develop (e.g., Ohlsson & Rees, 1991; Rittle-Johnson & Siegler, 1998). Results are mixed, however, partially due to lack of a clear definition of conceptual knowledge. This does not imply that using a standardized definition of conceptual knowledge would lead to a unified conclusion about the developmental time course of mathematical knowledge. It would, however, allow us to rule out simple differences in definitions as a possible reason for differences across domains and to focus on more important factors (e.g., amount of exposure to different domains; see Rittle-Johnson & Siegler, 1998).

There are at least two possible explanations for the plethora of definitions. First, researchers may not have clearly defined notions of the theoretical construct "conceptual knowledge." Alternatively, or in addition, the variety of definitions may reflect the existence of multiple facets of conceptual knowledge. The current review points toward a mixture of both accounts. Some sources do present clearly articulated explanations of conceptual knowledge; others simply imply that there is mathematical knowledge that is separate from procedures and that this type of knowledge is important, but do not specify what that knowledge actually is.

One piece of evidence for researchers' lack of clarity about conceptual knowledge as a theoretical construct comes from the third issue identified in this review: the frequent incongruity between the definitions offered and the tasks used to measure conceptual knowledge. The domains differ in the degree of alignment between definitions and tasks. The cardinality literature, for example, demonstrated fairly good alignment between theory and measurement. Research on equivalence, however, employed a great variety of tasks, very few of which clearly mapped onto definitions, when they were provided. Indeed, in each of the domains, there were a number of sources in which the measurements and definitions did not align. In fact, the current review probably overestimates the amount of alignment between definition and measurement, as definitions and tasks were aggregated across sources for discussion purposes.

It does appear, however, that the second possibility raised above also holds some merit: the tasks and definitions that did align were not all centered on a single notion of conceptual knowledge. This suggests that, in order to adequately characterize conceptual knowledge, it may be best to think about it as having multiple facets. Moreover, these distinct facets may be differentially important across mathematical domains. In order to assess this second possibility, it is helpful to compare the definitions and tasks used in each of the three reviewed domains.

Defining conceptual knowledge

Despite a general lack of consensus regarding the definition of conceptual knowledge, a comparison of the three reviewed domains illustrates the similarities and differences among the literatures.

Definitions found in all three of the reviewed domains

General principle knowledge

The idea of conceptual knowledge as general principle knowledge was explicitly stated in 21% of the sources ($n = 22$) across domains, and it occurred in at least six sources in each domain. This was the only definition type present in all three domains, suggesting that it may represent a central facet of conceptual knowledge. It appears to align clearly with many specific concepts of interest in the reviewed domains (e.g., the cardinality principle, the inversion principle). One issue with this type of definition is that researchers rarely (and in some domains, never) defined the general term "principle." Most did, however, provide an explanation of the specific principle of interest.

Definitions found in two of the reviewed domains

Knowledge of principles underlying procedures

This definition type, found in 23% ($n = 7$) of the cardinality sources and 13% ($n = 4$) of the inversion sources, defined conceptual knowledge as an understanding of the conceptual basis for a procedure or for steps in a procedure. Although this definition might be criticized as blurring the line between procedural and conceptual knowledge, its frequency across the literature and its clear alignment with many of the tasks suggest that it may represent another central facet of conceptual knowledge.

Connection knowledge

Conceptual knowledge as knowledge of relationships was found in 9% of the sources ($n = 3$) on inversion and 20% of the sources on equivalence ($n = 8$). The majority of these definitions were general statements about the importance of links between pieces of knowledge or concepts within a domain. One possible explanation for the imprecision of these definitions is that many were based on theoretical articles that did not focus on specific mathematical domains (e.g., Hiebert & Lefevre, 1986). Although unoperationalized definitions may be useful in considering mathematical cognition at a general level, they are less useful in work on mathematical thinking and development within specific domains. Despite not occurring in all the domains, the connection knowledge definition was among the most common, both in the general review and in the equivalence literature.

Some researchers have argued that well-connected knowledge is not indicative of the *type* of knowledge (e.g., conceptual versus procedural), but rather is an index of knowledge *quality* (Star, 2005). This seems reasonable in light of the fact that some elements of conceptual knowledge may exist as isolated pieces of knowledge. For example, asking a child to define the equal sign may reveal that the child knows the definition, but it does not show that the child's understanding of the symbol is well connected to other concepts.

Connection knowledge is a very broad notion, encompassing knowledge of a wide range of types of relationships, for example, among elements of domain structure, among category members, between symbols and their meanings, and so forth. As such, the construct of connection knowledge is potentially too broad to be theoretically useful as a specific form of conceptual knowledge on its own. However, as Star (2005) suggests, it may be productive to view connections as a way of indexing depth of understanding.

Definitions found in one of the reviewed domains

Domain structure

Two sources in the inversion literature defined conceptual knowledge as understanding of the structure of mathematics (Canobi, 2009; Robinson & Dube, 2009a). These definitions did not specify what aspects of structure were of interest, although in the domain of inversion, structural relationships among operations are presumably the most important. In addition, these definitions were not well connected to the specific tasks used to assess inversion understanding.

Symbol knowledge

Conceptual knowledge was defined as symbol knowledge in one equivalence article (Saenz-Ludlow & Walgamuth, 1998). Despite its infrequency, this definition aligned closely with one of the specific concepts discussed more broadly in the equivalence literature, relational understanding of the equal sign.

Category knowledge

One equivalence article mentioned the idea of category knowledge as a type of conceptual knowledge (DeCaro & Rittle-Johnson, 2012). Additionally, at least one of the tasks used in the equivalence domain (the *symbol sort* task) could be viewed as aligning with this definition, although it was not presented as such by the authors.

Measuring conceptual knowledge

In contrast to the limited number of definitions, there was no shortage of tasks assessing conceptual knowledge. In fact, across the three domains, researchers used 24 distinct types of tasks, and there were many minor variations, making the actual variety even greater.

There were substantial differences in the approach taken to measuring conceptual knowledge across studies and across domains, although the basis for these differences was not always clear. For example, some studies relied on one or more explicitly labeled "conceptual" measures that were analyzed separately from "procedural" measures (e.g., McNeil & Alibali, 2000). Others relied on characterizing the specific procedures used in a task to infer conceptual knowledge, as opposed to labeling the task itself as "conceptual" (e.g., Dube & Robinson, 2010a). This approach has been supported by researchers who have argued that tasks themselves are not procedural or conceptual in nature, but rather that the procedures used by participants to complete the tasks indicate which type of knowledge they are using (e.g., Faulkenberry, 2013). Finally, a few studies relied on patterns of performance across tasks to assess conceptual knowledge (e.g., Rittle-Johnson et al., 2011). It has been suggested that conceptual knowledge can only be inferred based on patterns of behavior across a variety of tasks, and not simply through the use of a single "conceptual" task (e.g., Bisanz, Watchorn, Piatt, & Sherman, 2009). Although each of these approaches was used in some sources, the choice of approach was rarely discussed, and it was even more rarely discussed how the chosen approach provides an adequate measure of conceptual knowledge. Thus it appears that some of the difficulty stems from trouble conceptualizing mathematical knowledge, and some from trouble measuring conceptual knowledge in a theoretically grounded way.

To characterize current practices in measuring conceptual knowledge, we classified tasks into a small set of categories (see Table 3), using a classification scheme based on others used in the mathematics cognition literature (e.g., Bisanz & LeFevre, 1992; Prather & Alibali, 2009). The categories we identified were (1) application and justification of procedures, (2) evaluation of procedures, (3) evaluation of examples, and (4) explanation of concepts. The classification of tasks into categories was based on the primary performance requirements of each task.

Table 3
Summary of types of measures of conceptual knowledge.

Task type	Explanation	Example task
Application of Procedures	Participant enacts a procedure. May also include a justification of that procedure.	Inversion problems (e.g., Robinson & Dube, 2009c)
Evaluation of Procedures	Participant provides an explanation or rating of a procedure.	Windows task (Cowan et al., 1996)
Evaluation of Examples	Participant provides an explanation or rating of an example.	Number equality task (Rittle-Johnson & Alibali, 1999)
Explanation of Concepts	Participant provides an explicit statement or explanation of a concept.	Equal sign definition (e.g., McNeil et al., 2006)
Miscellaneous	Varied	Violation of expectation paradigm (Vilette, 2002)

Application and justification of procedures

In the first category of tasks, participants enact a procedure or solve a problem. In some cases, participants also justify or explain the procedure they use. In the cardinality literature, the *how many* and *give a number* tasks fall into this category. Both tasks require children to count sets in order to arrive at a solution. In the inversion literature, the *inversion problem* and *procedure justification* tasks fall into this category. These tasks require participants to solve problems or to provide verbal explanations of their own problem-solving procedures. In the equivalence literature, participants solve equivalence problems and explain how they obtained their solutions in most studies, and data from problem solutions and explanations have been taken as evidence of conceptual knowledge in a few cases (e.g., Capraro, Capraro, Ding, & Li, 2007). Most researchers in the domain of equivalence, however, label these tasks as procedural, although they are similar to tasks used to measure conceptual knowledge in other domains. Additionally, *equivalent equations* tasks could be seen as falling into this category; in these tasks, participants solve problems that require understanding of equivalence for success.

Despite some disagreement across domains about whether or not these tasks are conceptual in nature, they are among the most widely used tasks for measuring conceptual knowledge. These tasks rely on inferring knowledge based on the procedure a participant has selected to complete the task (e.g., Faulkenberry, 2013); thus, these tasks align with the idea that, although a task itself may not tap conceptual knowledge, a procedure can do so. Additionally, these tasks allow for the measurement of explicit conceptual knowledge, as they require participants to actively solve or explain problems.

Evaluation of procedures tasks

In a second category of tasks, occurring in all three domains, participants evaluate or rate a procedure or the outcome of a procedure. The equivalence literature includes one such task: the *procedure assessment* task, in which participants assess the validity of different problem solving procedures. In the cardinality literature, the *procedure outcome, windows, compare sets, counterfactual, and procedure assessment* tasks all fall into this category. In the *procedure outcome, windows,* and *counterfactual* tasks, children evaluate the outcome of a counting procedure, whereas in the *procedure evaluation* task children evaluate the procedure itself. The inversion literature also includes a *procedure assessment* task in which participants appraise the validity of multiple procedures and select which procedures are best. The *violation of expectation paradigm* also falls into this category, as it requires participants to evaluate the "normality" of transformations.

Note that, in evaluation of procedures tasks, participants may draw on aspects of conceptual knowledge that they are unable to enact or use in their own problem solving. Thus, tasks that require participants to evaluate procedures may tap implicit conceptual knowledge.

Evaluation of examples tasks

In a third category of tasks, occurring in the equivalence and cardinality literatures, participants evaluate, rate, or identify examples relevant to a target concept. Three tasks in the equivalence literature fall into this category. In the *sample equation* task, participants evaluate whether sample equations of varying formats "make sense" or not. In the *number equality* task, participants identify equivalent numbers from sets of possible pairs. In the *equal sign definition rating task*, participants rate the "smartness" of various definitions of the equal sign. Specific patterns of ratings (e.g., relational definitions rated more highly than operational definitions) are taken as evidence of conceptual knowledge. In addition, one task in the cardinality literature, the *magic paradigm*, falls into this category. In the *magic paradigm*, children evaluate whether an array of objects stays fundamentally the same as it undergoes changes in some of its features (e.g., numerosity, color).

Like evaluation of procedures tasks, tasks that require participants to recognize or evaluate examples may tap implicit forms of conceptual knowledge. Participants who are not able to generate correct examples or definitions may nevertheless be able to identify them.

Explanation of concepts tasks

In the fourth category of tasks, participants provide explicit explanations or definitions of concepts (similar to the category of *explicit recognition* tasks identified by Prather & Alibali, 2009). Equivalence was the only domain in which such tasks were included as measures of conceptual knowledge, and there were two such tasks. In the *equal sign definition* task, participants explain what the equal sign means, and in the *generic equal definition task,* participants explain what it means for two sets to be equal. Note that tasks in which participants must provide explanations require explicit, verbalizable knowledge about the relevant concepts.

Tasks of this sort were not used consistently across domains. In fact, there were no measures of this type in the domains of cardinality or inversion. However, some relevant qualitative data were mentioned for cardinality. Specifically, some studies reported that children sometimes offered unprompted explanations for their answers on various cardinality tasks (e.g., Gelman & Meck, 1983).

Miscellaneous tasks

Within the equivalence literature, there were a number of tasks that do not fit into any of the above categories. This is not surprising, as the number of tasks in the equivalence domain greatly outnumbered those in the other two reviewed domains; there were 12 tasks explicitly labeled as conceptual in the equivalence literature, compared with eight tasks in the cardinality literature, and four in the inversion literature. The *problem reconstruction, problem recognition*, and *identifying sides* tasks, for example, are included in many studies and measure the ability to represent problems correctly.

The *symbol sort* task, in which participants group mathematical symbols and numbers, stands alone among the reviewed tasks in tapping implicit knowledge of symbols, as well as knowledge about connections between symbols and the categories to which those symbols belong. This task also does not fit cleanly into one of the categories described above.

Measuring conceptual knowledge: summary

The lack of coherent and specific definitions of conceptual knowledge presents a challenge for researchers seeking to measure it. However, difficulty knowing exactly how to measure conceptual knowledge has not stopped researchers from trying, leading to a wide range of tasks. Not only do the tasks differ in their details, but they also differ in what they require of participants. Some tasks, for example, require participants to give explicit verbal statements, whereas other tasks involve ratings or recognition. As a result, different studies measure different forms of conceptual knowledge (e.g., Greeno, 1993). Since researchers use widely varying measures to quantify conceptual knowledge, it is difficult to integrate empirical findings, both within and across domains.

The factors that lead researchers to choose one task or another are not always explicitly stated. In some instances, the choice of tasks appears to be based on characteristics of the participant sample, such as the age of the participants. Thus, some task choices are dictated by the behaviors that are typical or possible in certain age groups. Young children, for example, are unable to perform many of the tasks that adults can. This is generally unproblematic, as researchers investigating conceptual knowledge seldom directly compare performance between widely disparate age groups.

In other cases, however, the choice of tasks seems not to be based on any clear practical concerns. In the equivalence literature, for example, there seems to be a consensus that the definition of the equal sign is an important concept. The measurement of the concept varies, however, between requiring participants to generate explicit verbal definitions and requiring them simply to rate given definitions. The age range in which equivalence is studied is relatively small (compared with inversion, for example) and it is not always clear why one task is chosen over the other. The variability in choice of tasks makes it difficult to compare performance across studies, despite the similarity of the participant populations.

Additionally, for many of the tasks, it is not clear how the measures derived from the tasks align with theoretical definitions of the nature of conceptual knowledge. Some of the tasks focus on the use and understanding of procedures, such as tasks that require applying, justifying, or evaluating alternative procedures. Other tasks focus on the concepts themselves, such as tasks that involve explaining concepts or definitions, and tasks that involve evaluating examples of concepts. In the following sections, we build on this distinction to propose a framework that we believe will prove useful in integrating theory and measurement.

Integrating theory and measurement: a framework

An increased focus on defining and measuring conceptual knowledge consistently, both within and across mathematical domains, might foster a deeper, more integrated understanding of conceptual knowledge at a general level. A first step in this direction would be increased consistency and improved theoretical grounding in measuring conceptual knowledge. Of course, some variation in measurement techniques is inevitable due to the different specific concepts being measured and the different participant populations and age groups that are typical in each domain. Moreover, we recognize that, in some cases, lack of precision in measuring conceptual knowledge may be due to the challenges associated with distinguishing conceptual from procedural knowledge, rather than from inattention to these measurement issues. Nevertheless, we believe that it would be possible to measure conceptual knowledge more consistently than is the case at present. By carefully considering the goals, theoretical bases, and task demands for each task, researchers may be able to avoid mixed results that may cloud or obscure the "bigger picture."

A second step would be to utilize a common framework based on the definitions and measures identified in this review. We propose that many of the general and specific concepts of interest can be subsumed under the two types of conceptual knowledge that are relevant across mathematical domains: *general principle knowledge* and *knowledge of principles underlying procedures*. General principles can be defined as "fundamental laws or regularities that apply within a problem domain" (Prather & Alibali, 2009, p. 222). Thus, this proposed framework divides conceptual knowledge into principle knowledge of two types: general and procedure-specific. This framework incorporates many of the constructs that are consistently considered key components of conceptual knowledge, while excluding those about which there is less agreement (e.g., problem representation).

General principle knowledge

General principles can be known without relation to specific procedures. They include things like rules (e.g., "the final number in a count sequence represents the set's total numerosity"), definitions (e.g., the equal sign means "the same as"), and aspects of domain structure (e.g., addition and subtraction are inverse operations). This category incorporates several of the definitions of conceptual knowledge that are currently used in the literature, including knowledge of domain structure and symbol knowledge. Because we define principles as including "regularities" within domains, knowledge of domain structure fits naturally in this category. Symbol meanings are also regularities that are not specific to a particular procedure. For example, the equal sign always represents an equivalence relation.

In some cases, understanding general principles requires or implies knowledge of connections. For example, understanding the inversion principle requires understanding how inverse operations relate to one another. Finally, general principles may also underlie category knowledge. For example, knowing that the equal sign "belongs" in a category with the greater-than and less-than symbols requires understanding the general principle that mathematical statements express relations between quantities.

Tasks for measuring general principle knowledge

Several tasks already in use lend themselves well to the assessment of general principle knowledge, as defined above. Here, we discuss how currently used tasks can measure general principle knowledge.

Explanation of concepts tasks

Tasks that require verbal explanations can be used to assess explicit knowledge of general principles. Although most current implementations of such tasks ask participants to provide definitions only for symbols, it would also be possible to have participants provide verbal explanations of elements of domain structure (e.g., what does it mean for one operation to be the inverse of another operation?) or rules (e.g., why do we use the last number in the count sequence to name a set?). Due to the high performance demands of explanation tasks, they are best suited for use with older children and adults. Using such tasks across ages can provide information about the emergence of explicit general principle knowledge.

Evaluation of examples tasks

Tasks that require participants to recognize examples, definitions, or statements of principles can also be used to assess knowledge of general principles. Such tasks may tap more implicit forms of general principle knowledge, and they may also be suitable for younger children. In the equivalence literature, for example, the *equal sign definition rating* task taps participants' general understanding of the meaning of the equal sign, without requiring them to generate a definition. In the cardinality literature, the *magic paradigm* assesses general understanding of the importance of numerosity in characterizing a set, which is important for the development of full cardinality understanding. Evaluation of examples tasks may allow for early detection of conceptual knowledge, because their performance demands are low, relative to explanation tasks.

Knowledge of principles underlying procedures

In contrast to general principles, principles underlying procedures capture the *why* of problem solving – in particular, knowing why certain procedures work for certain problems and knowing the purpose of each step in a procedure. This type of conceptual knowledge maps well onto current curricular standards, which state that it is not enough for students to simply be able to solve problems; they must also have conceptual knowledge of the problem-solving process (e.g., National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Two other definitions of conceptual knowledge are also relevant to knowledge of principles underlying procedures. First, conceptual knowledge as connection knowledge is relevant to knowledge of principles underlying procedures, in the sense that there are connections among the steps in a procedure and between individual steps and their specific conceptual underpinnings. For example, understanding the conceptual basis for the inversion shortcut requires connecting knowledge of the inversion principle to knowledge of a specific problem-solving procedure. The idea of conceptual knowledge as category knowledge can also be tied in, since correctly identifying a mathematics problem as belonging to a specific problem category has implications for selecting applicable procedures. For example, in order to utilize the inversion shortcut, one must first identify a problem as belonging to the category of problems for which the shortcut is appropriate (i.e., inversion problems).

Tasks for measuring knowledge of principles underlying procedures

Knowledge of principles underlying procedures can be assessed with two types of tasks that are already widely used by researchers.

Application and justification of procedures tasks

Requiring participants to solve problems is one way to assess their knowledge of principles underlying procedures. The way in which participants' procedures align (or do not align) with principles within a domain provides insight into their knowledge of those principles (e.g., Dixon & Moore, 1996). Although the specific data gathered from such tasks may vary (e.g., reaction time, use of a specific procedure), procedure-based measures are fairly straightforward in many mathematical domains. However, since there remains the possibility that participants could succeed on application of procedures tasks due simply to memorization of a correct procedure, such tasks should be used in conjunction with other, non-procedure-based tasks to provide a more comprehensive knowledge assessment.

Evaluation of procedures tasks

A second type of task that can be used to tap understanding of principles underlying procedures is having participants evaluate specific procedures. By asking participants to evaluate both correct and incorrect procedures, researchers can measure their abilities, not only to judge that certain procedures are appropriate in certain situations, but also to judge that other procedures are not. Thus, such tasks allow for assessing conceptual knowledge as distinct from the ability to enact procedures themselves.

Evaluation of procedures tasks can lead participants to consider procedures that they have not generated on their own. In the cardinality literature, for example, having children evaluate correct-but-unusual counts can distinguish those children who understand the conceptual basis for counting from those who have simply memorized the conventional count sequence. Such tasks can also be used to elicit justifications, for example, about why a solver might have taken a particular step in solving a problem, or about why a particular procedure is correct or incorrect.

Implementing change

The proposed framework for thinking about conceptual knowledge suggests specific steps that could be taken in future research that would lead to more systematic treatment of conceptual knowledge across mathematical domains. Some of these steps are general recommendations for research on conceptual knowledge; others are recommendations that apply in the specific subdomains that we have reviewed.

General recommendations for research on conceptual knowledge

At the broadest level, we encourage researchers to evaluate the current "state of the art" in thinking about and measuring conceptual knowledge in their particular area of expertise, and to situate their work on conceptual knowledge within the more general literature on this topic. This step would foster researchers' making connections across subdomains, which may be beneficial for the field at large, as it may spark exchanges and progress on issues of measurement, theory, and practical application.

Second, we encourage researchers to provide explicit definitions of conceptual knowledge, and to use tasks that align well with the chosen definitions. Care should be taken to consider the range of possible ways in which participants can solve or respond to specific tasks (see Faulkenberry, 2013), so as to insure – or at least increase the likelihood – that tasks intended to tap specific facets of conceptual knowledge actually do so. A focus on strategic aspects of behavior to infer knowledge may lead to more accurate assessments than a simple focus on scores on "conceptual" tasks.

Third, we encourage researchers to explicitly consider the specific principles that are relevant within each mathematical domain, as doing so may guide research in fruitful ways. As one example, Rittle-Johnson and Alibali (1999) distinguished between understanding the meaning of equal quantities and understanding the meaning of the equal sign symbol, and they selected or developed tasks to measure each of these concepts separately. Although some studies do attempt to identify and measure knowledge of specific principles or concepts (e.g., Rittle-Johnson et al., 2011), we believe it would be beneficial for the field if researchers would do so more systematically.

Fourth, we encourage researchers to use multiple tasks when assessing conceptual knowledge. In many cases, conceptual knowledge may be best "diagnosed" using a battery or set of tasks, rather than a single task (e.g., Bisanz et al., 2009). Using a set of tasks would allow for assessment of different forms of conceptual knowledge (i.e., explicit and implicit knowledge) and for characterization of patterns of behavior that might indicate conceptual knowledge. Additionally, if a larger set of tasks were used in each study, there would be more room for direct comparison across studies. At present, it is often hard to compare performance across studies because each uses a different task to assess conceptual knowledge.

Fifth (and related to the previous points), we encourage researchers to utilize tasks that are specifically designed to tap each of the two forms of conceptual knowledge that we have identified: general principles and principles that underlie procedures. To do so effectively, researchers will need to explicitly identify the relevant principles (both general and procedure-specific) in their domain of study.

Finally, we encourage researchers to test specific predictions derived from our framework. Specifically, we encourage researchers to measure general principle knowledge and knowledge of principles underlying procedures separately, and then to use confirmatory factor analysis to test whether the data have a two-factor structure. On this point, however, we believe some notes of caution are in order. We expect that the two-factor structure will be readily observed only when knowledge of the principles underlying procedures is tested for procedures that are not transparently related to general domain principles. In the domain of equivalence, for example, it may be difficult to empirically distinguish knowledge of the meaning of the equal sign (i.e., a general principle) from knowledge of the principles underlying steps in the "make both sides equal" procedure (i.e., principles underlying procedures). It may be more straightforward to empirically distinguish knowledge of the meaning of the equal sign from knowledge of principles underlying steps in the "add–subtract" procedure (for a problem such as $3 + 6 = 5 + \_\_$, add the numbers on the left side of the equation and then subtract the number on right side), because the relation between the general principle and the add–subtract procedure is less direct.

Recommendations for research within the domains reviewed

Our framework suggests additional recommendations for research within each of the three domains reviewed in this paper. The reviewed domains displayed a variety of issues, leading to different recommendations in each area. Some suggestions are briefly outlined below, to serve as examples of potential directions for future research.

Equivalence

The literature on equivalence left the most to be desired in terms of linking definitions of conceptual knowledge to specific tasks. Therefore, one recommendation for this domain would be for researchers to focus on theoretically grounded measurement of conceptual knowledge. There has been a recent push in this area toward creating reliable and valid measurement instruments, leading to a rapid increase in the number of tasks being used to assess both conceptual and procedural knowledge (e.g., Rittle-Johnson et al., 2011). These efforts are a clear step in a positive direction, and they align well with the idea that assessing patterns of behavior over a variety of tasks may provide a better characterization of conceptual knowledge. In the future, it will be important to tie these carefully constructed measures to a more specific conceptualization of conceptual knowledge. Additionally, discussion of how the results from the battery of assessments provide insight into conceptual knowledge would be valuable.

A second recommendation is that researchers in the domain of equivalence should consider using data from application of procedures measures as evidence for conceptual knowledge, as is done in other domains. As mentioned previously, a majority of studies on equivalence require participants to solve problems and to provide procedural justifications. Many of these justifications include conceptual content and, moreover, the actual procedures used could be evaluated as being in line with domain principles or not, as is done in other domains. As noted by Faulkenberry (2013), these types of assessments might provide a particularly valuable picture of conceptual knowledge that does not rely on the specifics of the task, per se.

In this vein, the notion of a "shortcut" procedure as evidence for conceptual knowledge could be extended to the equivalence domain. Some equivalence problems lend themselves to being solved with a "shortcut" grouping procedure that reduces the need for computation (similar to use of the inversion shortcut on inversion problems). On problems where an addend is repeated on both sides of the equation (e.g., $3 + 4 + 5 = 3 + \_$), the grouping procedure takes advantage of the fact that some elements on each side are already equal (i.e., the 3s) and thus the solution can be reached by simply adding the other numbers (i.e., the 4 and 5). Recently collected data suggests that a majority of adult problem solvers utilize a grouping procedure on repeated-addend equivalence problems (Crooks & Alibali, 2013), much in the way that a majority of adults use the inversion shortcut on inversion problems. By recognizing that procedure-based tasks can tap aspects of conceptual knowledge, researchers in the equivalence domain could measure conceptual knowledge in a way that aligns with a common approach in other domains.

Cardinality

Of the three reviewed domains, cardinality demonstrated the best alignment between definitions and tasks. Few tasks within the cardinality literature assessed general principle knowledge, however, making it difficult to disentangle children's conceptual and procedural knowledge. Although completely separating conceptual and procedural knowledge may be impossible, it would be valuable to assess conceptual knowledge using tasks that cannot be solved with a simple procedural rule (i.e., last word responding; Bermejo et al., 2004). Although it may not be possible to cleanly separate conceptual and procedural knowledge in early counting, it seems prudent to use tasks that could at least differentiate last word responding from procedures with a stronger conceptual underpinning.

Future research could also include a wider variety of tasks, particularly tasks that allow children to demonstrate more general knowledge of the cardinality principle. Tasks that require children to provide explanations of outcomes, for example, have been used qualitatively in the counting literature (e.g., Gelman & Gallistel, 1978), but could also provide interesting quantitative data about the frequency with which children explicitly mention principles and how this tendency changes over development.

Inversion

The primary difficulty in the inversion literature was the absence of precise definitions of conceptual knowledge. In particular, there was a tendency to explain conceptual knowledge as understanding of principles, but then not define the term "principle." A general recommendation in this domain would be for researchers to more clearly define conceptual knowledge.

A second suggestion would be for researchers in this domain to utilize a wider range of tasks. Although behavioral measures of inversion problem solving provide important data, they could be supplemented with information about performance on other types of tasks. For example, one type of task that has appeared in the inversion literature with greater frequency in recent years is evaluation of procedures tasks. Researchers should continue to include this type of task, as it allows for assessment of implicit conceptual knowledge in participants who may not have yet developed a conceptually based procedure of their own. Another type of task that could be effectively adapted for use in this domain is justification of procedures tasks, in which participants are asked to explain why a certain procedure works or is better than another procedure. Participants could also be asked to explain the relationships among different operations or to group operations together based on different features (e.g., based on inverse relationships, or based on their effects on numerosity).

Conclusion

In the mathematical thinking literature, the term "conceptual knowledge" has come to denote a wide array of constructs, making it difficult to integrate major findings, to understand the ways in which conceptual knowledge relates to procedural knowledge, and to evaluate effective ways to use current research to guide classroom interventions (e.g., Baroody et al., 2007; Star, 2005). In particular, the ways in which conceptual knowledge has been defined theoretically and the ways in which it has been measured quantitatively have created a vast literature in which a "bigger picture" is hard to find. The current review highlights the lack of consistency in research on conceptual knowledge of mathematics. Specific issues include the fact that explicit definitions of conceptual knowledge are rare, that the given definitions are often vague or poorly operationalized, and that the tasks used to measure conceptual knowledge often fail to align with theoretical claims about the nature of such knowledge.

As one step toward addressing these problems, we have proposed a framework for thinking about conceptual knowledge that consolidates current notions into two specific types of conceptual knowledge: general principle knowledge and knowledge of principles underlying procedures. General principle knowledge involves understanding of mathematical ideas without relation to specific problems or procedures. Knowledge of principles underlying procedures, on the other hand, involves connecting concepts to specific procedures – for example, knowing why certain procedures work for certain problems or knowing the purpose of each step in a procedure. We do not claim that this framework will clear the "wilderness" of understanding mathematical knowledge, but we do believe that it represents a substantial step forward.

Considering these two types of knowledge could guide future efforts to measure conceptual knowledge across mathematical domains. By thinking about conceptual knowledge across domains in a consistent way, researchers may gain a clearer picture of how conceptual knowledge develops. Utilizing a common framework will also allow researchers to assess which experimental findings reflect true similarities or differences across domains and which are artifacts of the fractured state of the field. In turn, greater knowledge about the development of conceptual knowledge may contribute to generating and validating new methods for teaching conceptual knowledge in the classroom.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant #R305B090009 to the University of Wisconsin-Madison. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. We thank Chuck Kalish, Vanessa Simmering, and Percival Matthews for their feedback on earlier versions of this manuscript. Additionally, we thank Jackie Magee for assistance with the literature reviews.

References[4]

Alibali, M.W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*, 127–145.

Alibali, M. W. (2005). Mechanisms of change in the development of mathematical reasoning. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 33, pp. 79–123). San Diego, CA: Elsevier Academic Press.[G]

Alibali, M. W., Phillips, K. M. O., & Fischer, A. D. (2009). Learning new problem-solving strategies leads to changes in problem representation. *Cognitive Development*, *24*, 89–101.[E]

Ashlock, R. B. (2001). *Error patterns in computation: Using error patterns to improve instruction*. Columbus, OH: Merrill Prentice Hall.[G]

Baroody, A., Feil, Y., & Johnson, A. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education*, *38*(2), 115–131.[G]

Baroody, A. J., & Dowker, A. (2003). A. Schoenfeld (Ed.), *The development of arithmetic concepts and skills: Constructing adaptive expertise. "Studies in Mathematics Thinking and Learning" series*. Mahwah, NJ: Lawrence Erlbaum Associates.

Baroody, A. J., & Ginsburg, H. P. (1983). The effects of instruction on children's understanding of the "equals" sign. *The Elementary School Journal*, *84*(2), 198–212.[E]

Baroody, A.J., & Lai, M. (2007). Preschoolers' understanding of the addition-subtraction inversion principle: A Taiwanese study. *Mathematical Thinking and Learning*, *9*(2), 131–171, [I].

Baroody, A. J., Lai, M., Li, X., & Baroody, A. E. (2009a). Preschoolers' understanding of subtraction-related principles. *Mathematical Thinking and Learning*, *11*(1–2), 41–60.[I]

Baroody, A.J., Torbeyns, J., & Verschaffel, L. (2009b). Young children's understanding and application of subtraction-related principles. *Mathematical Thinking and Learning*, *11*(1–2), 2–9, [I].

Baroody, A. J., Wilkins, J. L. M., & Tiilikainen, S. (2003). The development of children's understanding of additive commutativity: from protoquantitative concept to general concept? In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp. 127–160). Mahwah, NJ: Erlbaum.[G]

Bermejo, V., Morales, S., & Garcia deOsuna, J. (2004). Supporting children's development of cardinality understanding. *Learning and Instruction*, *14*, 381–398.

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70–87.[G]

[4] Superscript annotations denote inclusion in the reported literature reviews. "G" indicates inclusion in the general review, "E" in the equivalence review, "C" in the cardinality review, and "I" in the inversion review.

Bialystok, E., & Cobb, J. (1997). Cardinal limits: evidence from language awareness and bilingualism for developing concepts of number. *Cognitive Development*, *12*, 85–106.[C]
Bisanz, J., & LeFevre, J. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 213–244). Hillsdale, NJ: Lawrence Erlbaum.[G]
Bisanz, J., & LeFevre, J. (1992). Understanding elementary mathematics. In J. Campbell (Ed.), *The nature and origins of mathematical skills* (pp. 113–136). Amsterdam, North Holland: Elsevier Science Publishers.[I]
Bisanz, J., Watchorn, R. P. D., Piatt, C., & Sherman, J. (2009). On "understanding" children's developing use of inversion. *Mathematical Thinking and Learning*, *11*, 10–24.[I]
Blote, A. W., Klein, A. S., & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction*, *10*, 221–247.[G]
Bolden, D. S., & Newton, L. D. (2008). Primary teachers' epistemological beliefs: some perceived barriers to investigative teaching in primary mathematics. *Educational Studies*, *34*(5), 419–432.[G]
Bosse, M. J., Ries, H., & Chandler, K. (2013). Putting the modern in algebra. *Primus*, *23*(1), 1–12.[G]
Briars, D., & Siegler, R. S. (1984). A featural analysis of preschooler's counting knowledge. *Developmental Psychology*, *20*(4), 607–618.[C]
Brownell, W. A. (1945). When is arithmetic meaningful? *The Journal of Educational Research*, *38*(7), 481–498.
Bryant, P., Christie, C., & Rendu, A. (1999). Children's understanding of the relation between addition and subtraction: inversion, identity, & decomposition. *Journal of Experimental Child Psychology*, *74*, 194–212.[I]
Burns, M. K. (2011). Matching math interventions to students' skill deficits: a preliminary investigation of a conceptual and procedural heuristic. *Assessment for Effective Intervention*, *36*(4), 210–218.[G]
Byrnes, J. P. (1992). The conceptual basis of procedural learning. *Cognitive Development*, *7*, 235–257.[G]
Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology*, *27*(5), 777–786.[G]
Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology*, *92*, 220–246.[I]
Canobi, K. H. (2009). Concept-procedure interactions in children's addition and subtraction. *Journal of Experimental Child Psychology*, *102*, 131–149.[I]
Capraro, M. M., & Joffrion, H. (2006). Algebraic equations: can middle-school students meaningfully translate from words to mathematical symbols? *Reading Psychology*, *27*, 147–164.[G]
Capraro, R. M., Capraro, M. M., Ding, M., & Li, X. (2007). Thirty years of research: interpretations of the equal sign in China and the USA. *Psychological Reports*, *101*, 784–786.[E]
Carpenter, T. P. (1986). Conceptual knowledge as a foundation for procedural knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 113–132). Hillsdale, NJ: Erlbaum.[G]
Carr, M., Alexander, J., & Folds-Bennett, T. (1994). Metacognition and mathematics strategy use. *Applied Cognitive Psychology*, *8*, 583–595.
Cowan, R., Dowker, A., Christakis, A., & Bailey, S. (1996). Even more precisely assessing children's understanding of the order-irrelevance principle. *Journal of Experimental Child Psychology*, *62*, 84–101.[C]
Crooks, N. M., & Alibali, M. W. (2013). Equivalence problem strategy data from undergraduate participants. Unpublished raw data.
Culatta, B. (1993). Building mathematics skills. In F. L. Rowley-Kelly & D. H. Reigel (Eds.), *Teaching the student with spina bifida* (pp. 145–170). Baltimore, MD: Paul H. Brookes Publishing.[G]
de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, *31*(2), 105–113.[G]
De Cortre, E., & Verschaffel, L. (1981). Children's solution processes in elementary arithmetic problems: analysis and improvement. *Journal of Educational Psychology*, *73*(6), 765–779.[E]
DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*, 552–568.[E]
Dixon, J. A., & Moore, C. F. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology*, *32*(2), 241–253.[G]
Dube, A. K., & Robinson, K. M. (2010a). The relationship between adults' conceptual understanding of inversion and associativity. *Canadian Journal of Experimental Psychology*, *64*(1), 60–66.[I]
Eisenhart, M., Borko, H., Underhill, R., & Brown, C. (1993). Conceptual knowledge falls through the cracks: complexities of learning to teach mathematics for understanding. *Journal for Research in Mathematics Education*, *24*(1), 8–40.[G]
English, L. D. (1997). Children's reasoning processes in classifying and solving computational word problems. In L. D. English (Ed.), *Mathematical reasoning: Analogies, metaphors, and images* (pp. 191–220). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.[G]
Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: a foundation for algebra. *Teaching Children Mathematics*, 232–236.[E]
Faulkenberry, T. J. (2013). The conceptual/procedural distinction belongs to strategies, not tasks: a comment on Gabriel et al. (2013). *Frontiers in Psychology*, *4*, 820.[G]
Ferrara, R. A., & Turner, T. (1993). The structure of early counting competence. *Bulletin of the Psychonomic Society*, *31*(4), 257–260.[C]
Freeman, N. H., Antonucci, C., & Lewis, C. (2000). Representation of the cardinality principle: early conception of error in a counterfactual test. *Cognition*, *74*, 71–89.[C]
Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development*, *60*(5), 1158–1171.[C]
Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: prior knowledge matters. *Journal of Educational Psychology*, *104*(4), 1094–1108.[E]
Garofalo, J., & Lester, F. K., Jr. (1985). Metacognition, cognitive monitoring, and mathematical performance. *Journal for Research in Mathematics Education*, *16*(3), 163–176.
Gau Bartell, T., Webel, C., Bowen, B., & Dyson, N. (2013). Prospective teacher learning: recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education*, *16*, 57–79.[G]

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.[C]
Gelman, R., & Meck, E. (1983). Preschoolers' counting: principles before skill. *Cognition*, *13*, 343–359.[C]
Gelman, R., & Meck, E. (1986). The notion of principle: the case of counting. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 29–57). Hillsdale, NJ: Erlbaum Associates.[C]
Gilmore, C. K. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development*, *21*, 301–316.[I]
Gilmore, C. K., & Bryant, P. (2008). Can children construct inverse relations on arithmetic? Evidence for individual differences in the development of conceptual understanding and computational skill. *British Journal of Developmental Psychology*, *26*, 301–316.[I]
Gilmore, C. K., & Papadatou-Pastou, M. (2009). Patterns of individual differences in conceptual understanding and arithmetic skill: a meta-analysis. *Mathematical Thinking and Learning*, *11*, 25–40.
Goldman, S. R., & Hasselbring, T. S. (1997). Achieving meaningful mathematics literacy for students with learning disabilities. Cognition and Technology Group at Vanderbilt. *Journal of Learning Disabilities*, *30*(2), 198–208.[G]
Greeno, J. G. (1978). Understanding and procedural knowledge in mathematics instruction. *Educational Psychologist*, *12*(3), 262–283.[G]
Greeno, J. G. (1993). For research to reform education and cognitive science. In L. A. Penner, G. M. Batsche, H. M. Knoff, & D. L. Nelson (Eds.), *The challenge in mathematics and science education: Psychology's response* (pp. 153–192). Washington, DC: American Psychological Association.[G]
Greeno, J.G., Riley, M.S., & Gelman, R. (1984). Conceptual competence and children's counting. *Cognitive Psychology*, *16*, 434–483, [C].
Gutstein, E., & Romberg, T. A. (1995). Teaching children to add and subtract. *Journal of Mathematical Behavior*, *14*, 283–324.[G]
Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology*, *102*(2), 395–406.[G]
Hallett, D., Nunes, T., Bryant, P., & Thorpe, C. M. (2012). Individual differences in conceptual and procedural fraction understanding: the role of abilities and school experience. *Journal of Experimental Child Psychology*, *113*, 469–486.[G]
Hatano, G. (1988). Social and motivational bases for mathematical understanding. *New Directions for Child Development*, *41*, 55–70.[G]
Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY: Freeman.[G]
Hattikudur, S., & Alibali, M. W. (2010). Learning about the equal sign: does comparing with inequality symbols help? *Journal of Experimental Child Psychology*, *107*, 15–30.[E]
Hecht, S. A. (1998). Toward an information-processing account of individual differences in fraction skills. *Journal of Educational Psychology*, *90*(3), 545–559.[G]
Hecht, S. A., & Vagi, K. J. (2012). Patterns of strengths and weaknesses in children's knowledge about fractions. *Journal of Experimental Child Psychology*, *111*, 212–229.[G]
Hiebert, J. (1986). *Conceptual and procedural knowledge: The case of mathematics*. Hillsdale, NJ: Erlbaum.
Hiebert, J., & Carpenter, T. P. (1992). Learning and teaching with understanding. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics* (pp. 65–97). New York, NY: Macmillan Publishing Co, Inc.[G]
Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: an introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Erlbaum.[G]
Hiebert, J., & Wearne, D. (1986). Procedures over concepts: the acquisition of decimal number knowledge. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 199–223). Hillsdale, NJ: Erlbaum.[G]
Jitendra, A., DiPipi, C. M., & Perron-Jones, N. (2002). An exploratory study of schema-based word-problem-solving instruction for middle school students with learning disabilities: an emphasis on conceptual and procedural understanding. *The Journal of Special Education*, *36*(1), 23–38.[G]
Kadijevich, D. J., & Haapasalo, L. (2001). Linking procedural and conceptual mathematical knowledge through CAL. *Journal of Computer Assisted Learning*, *17*, 156–165.[G]
Kajander, A. (2010). Elementary mathematics teacher preparation in an era of reform: the development and assessment of mathematics for teaching. *Canadian Journal of Education*, *33*(1), 228–255.[G]
Kamawar, D., LeFevre, J., Bisanz, J., Fast, L., Skwarchuk, S., Smith-Chant, B., et al. (2010). Knowledge of counting principles: how relevant is order irrelevance? *Journal of Experimental Child Psychology*, *105*, 138–145.[C]
Kaufmann, L., Handl, P., & Thony, B. (2003). Evaluation of a numeracy intervention program focusing on basic numerical knowledge and conceptual knowledge: a pilot study. *Journal of Learning Disabilities*, *36*(6), 564–573.[G]
Kieran, C., & Guzman, J. (2006). The number-theoretic experience of 12-15-year-olds in a calculator environment: the intertwining coemergence of technique and theory. In R. Zazkis & S. R. Campbell (Eds.), *Number theory in mathematics education: Perspectives and prospects* (pp. 173–200). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.[G]
Kieren, T. E. (1993). Rational and fractional numbers: from quotient fields to recursive understanding. In T. P. Carpenter, E. Fennema, & T. A. Romberg (Eds.), *Rational numbers: An integration of research* (pp. 49–84). Hillsdale, NJ: Erlbaum.[G]
Kilpatrick, J., Swafford, J. O., & Findell, B. (Eds.), (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.[G]
Knuth, E. J., Alibali, M. W., McNeil, N. M., Weinberg, A., & Stephens, A. C. (2005). Middle school students' understanding of core algebraic concepts: equality and variable. *Zentralblatt für Didaktik der Mathematik / International Journal on Mathematics Education*, *37*, 68–76.
Kolloffel, B., Eysink, T. H. S., & de Jong, T. (2011). Comparing the effects of representational tools in collaborative and individual inquiry learning. *Computer-Supported Collaborative Learning*, *6*, 223–251.[G]
Lachance, A., & Confrey, J. (2002). Helping students build a path of understanding from ratio and proportion to decimal notation. *Journal of Mathematical Behavior*, *20*, 503–526.[G]
Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction*, *3*(4), 305–342.[G]
Le Roux, K. (2008). A critical discourse analysis of a real-world problem in mathematics: looking for signs of change. *Language and Education*, *22*(5), 307–326.[G]

Lee, S. (1998). Mathematics learning and teaching in the school context: reflections from cross-cultural comparisons. In S. G. Paris & H. M. Wellman (Eds.), *Global prospects for education: Development, culture, and schooling* (pp. 45–77). Washington, DC: American Psychological Association.[G]

LeFevre, J., Smith-Chant, B., Fast, L., Skwarchuk, S., Sargla, E., et al. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through grade 2. *Journal of Experimental Child Psychology*, *93*(4), 285–303.[C]

Leinhardt, G. (1988). Getting to know: tracing students' mathematical knowledge from intuition to competence. *Educational Psychologist*, *23*(2), 119–144.[G]

Lithner, J., & Palm, T. (2010). Learning difficulties and mathematical reasoning. In B. Sriraman, C. Bergsten, S. Goodchild, G. Palsdottir, B. Dahl, & L. Haapasalo (Eds.), *The first sourcebook on Nordic research in mathematics education: Norway, Sweden, Iceland, Denmark, and contributions from Finland* (Vol. 10, pp. 283–298). Charlotte, NC: IAP Information Age Publishing.[G]

Lobato, J., Clarke, D., & Ellis, A. B. (2005). Initiating and eliciting in teaching: a reformation of telling. *Journal for Research in Mathematics Education*, *36*(2), 101–136.[G]

Mabbot, D. J., & Bisanz, J. (2003). Developmental change and individual differences in children's multiplication. *Child Development*, *74*(4), 1091–1107.[G]

Mabbot, D. J., & Bisanz, J. (2008). Computational skills, working memory, and conceptual knowledge in older children with mathematics learning disabilities. *Journal of Learning Disabilities*, *41*(1), 15–28.[G]

Malloy, C. E. (2009). Instructional strategies and dispositions of teachers who help African American students gain conceptual understanding. In B. Danny (Ed.), *Mathematics teaching, learning, and liberation in the lives of Black children* (pp. 88–122). New York, NY: Routledge/Taylor & Francis Group.[G]

Mann, E. L. (2006). Creativity: the essence of mathematics. *Journal for the Education of the Gifted*, *30*(2), 236–260.[G]

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: comparing self- explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, *104*, 1–21.[E]

McCormick, R. (1997). Conceptual and procedural knowledge. *International Journal of Technology and Design Education*, *7*, 141–159.[G]

McEldoon, K. L., Durkin, K. L., & Rittle-Johnson, B. (2013). Is self-explanation worth the time? A comparison to additional practice. *British Journal of Educational Psychology*, *83*, 615–632.[E]

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology*, *43*(3), 687–695.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: externally imposed goals influence what is learned. *Journal of Educational Psychology*, *92*(4), 734–744.[E]

McNeil, N. M., & Alibali, M. W. (2005a). Knowledge change as a function of mathematics experience: all contexts are not created equal. *Journal of Cognition and Development*, *6*(2), 285–306.[E]

McNeil, N. M., & Alibali, M. W. (2005b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*(4), 883–899.[E]

McNeil, N. M., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., & Brletic-Shipley, H. (2011). Benefits of practicing $4 = 2 + 2$: nontraditional problem formats facilitate children's understanding of mathematical equivalence. *Child Development*, *82*(5), 1620–1633.[E]

McNeil, N. M., Grandau, L., Knuth, E. J., Alibali, M. W., Stephens, A. C., Hattikudur, S., et al. (2006). Middle school students' understanding of the equal sign: the books they read can't help. *Cognition and Instruction*, *24*(3), 367–385.[E]

Medin, D. (1989). Concepts and conceptual structure. *American Psychologist*, *44*(12), 1469–1481.

Middleton, J. A., & Goepfert, P. (1996). *Inventive strategies for teaching mathematics: Implementing standards for reform*. Washington, DC: American Psychological Association.[G]

Miller, S. P., & Hudson, P. J. (2007). Using evidence-based practices to build mathematics competence related to conceptual, procedural, and declarative knowledge. *Learning Disabilities Research and Practice*, *22*(1), 47–57.[G]

Montague, M. (1998). Mathematics instruction in diverse classrooms. In K. R. Harris, S. Graham, & D. Deshler (Eds.), *Teaching every child every day: Learning in diverse schools and classrooms* (pp. 168–207). Cambridge, MA: Brookline Books.[G]

Mullins, D., Rummel, N., & Spada, H. (2011). Are two heads always better than one? Differential effects of collaboration on students' computer-supported learning in mathematics. *Computer-Supported Collaborative Learning*, *6*, 421–443.[G]

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. <http://www.nctm.org/standards/default.aspx?id=58>.

National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. <http://www.corestandards.org/assets/CCSSI_Math%20Standards.pdf>.

Niemi, D. (1996). Assessing conceptual knowledge in mathematics: representations, problem solutions, justifications, and explanations. *The Journal of Educational Research*, *89*(6), 351–363.[G]

Nye, J., Fluck, M., & Buckley, S. (2001). Counting and cardinal understanding in children with Down syndrome and typically developing children. *Down Syndrome Research and Practice*, *7*(2), 68–78.[C]

Ohlsson, S., & Rees, R. (1991). The function of conceptual understanding in the learning of arithmetic procedures. *Cognition and Instruction*, *8*(2), 103–179.[G]

Pardhan, H., & Mohammad, R. F. (2005). Teaching science and mathematics for conceptual understanding? A rising issue. *Eurasia Journal of Mathematics, Science and Technology Education*, *1*(1), 1–20.

Patel, P., & Canobi, K.H. (2010). The role of number words in preschoolers' addition concepts and problem-solving procedures. *Educational Psychology*, *30*(2), 107–124, [G].

Peled, I., & Segalis, B. (2005). It's not too late to conceptualize: constructing a generalized subtraction schema by abstracting and connecting procedures. *Mathematical Thinking and Learning*, *7*(3), 207–230.[G]

Perry, M., VanderStoep, S. W., & Yu, S. L. (1993). Asking questions in first-grade mathematics classes: potential influences on mathematical thought. *Journal of Educational Psychology*, *85*(1), 31–40.[G]

Ploger, D., & Hecht, S. (2009). Enhancing children's conceptual understanding of mathematics through Chartworld software. *Journal of Research in Childhood Education*, *23*(3), 267–277.[G]

Prather, R. W., & Alibali, M. W. (2009). The development of arithmetic principle knowledge: how do we know what learners know? *Developmental Review*, *29*, 221–248.

Reason, M. (2003). Relational, instrumental and creative understanding. *Mathematics Teaching*, *184*, 5–7.
Resnick, L. B., & Omanson, S. F. (1987). Learning to understand arithmetic. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 3, pp. 41–95). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.[G]
Richland, L. E., Stigler, J. W., & Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educational Psychologist*, *47*(3), 189–203.[G]
Rittle-Johnson, B. (2006). Promoting transfer: effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1–15.[E]
Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: does one lead to the other? *Journal of Educational Psychology*, *91*(1), 175–189.[E]
Rittle-Johnson, B., & Koedinger, K. (2009). Iterating between lessons on concepts and procedures can improve mathematics knowledge. *British Journal of Educational Psychology*, *79*, 483–500.[G]
Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. *Cognition and Instruction*, *23*(3), 313–349.[G]
Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: a construct-modeling approach. *Journal of Educational Psychology*, *103*(1), 85–104.[E]
Rittle-Johnson, B., & Siegler, R. S. (1998). The relations between conceptual and procedural knowledge in learning mathematics: a review. In C. Donlan (Ed.), *The development of mathematical skill* (pp. 75–110). Hove, UK: Psychology Press.[G]
Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: an iterative process. *Journal of Educational Psychology*, *93*(2), 346–362.[G]
Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology*, *99*(3), 561–574.[G]
Rittle-Johnson, B., & Star, J. R. (2009). Compared with what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology*, *101*(3), 529–544.[G]
Rittle-Johnson, B., Star, J. R., & Durkin, K. (2009). The importance of prior knowledge when comparing examples: influences on conceptual and procedural knowledge of equation solving. *Journal of Educational Psychology*, *101*(4), 836–852.[G]
Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology*, *82*, 436–455.[G]
Robinson, K. M., & Dube, A. K. (2009a). A microgenetic study of the multiplication and division inversion concept. *Canadian Journal of Experimental Psychology*, *63*(3), 193–200.[I]
Robinson, K. M., & Dube, A. K. (2009b). Children's understanding of addition and subtraction concepts. *Journal of Experimental Child Psychology*, *103*, 532–545.[I]
Robinson, K. M., & Dube, A. K. (2009c). Children's understanding of the inverse relation between multiplication and division. *Cognitive Development*, *24*, 310–321.[I]
Robinson, K. M., & LeFevre, J. (2012). The inverse relation between multiplication and division: Concepts, procedures, and a cognitive framework. *Educational Studies in Mathematics*, *79*, 409–428.
Rogers, T.T., & McClelland, J.L. (2005). A parallel distributed processing approach to semantic cognition: applications to conceptual development. In L. Gershkoff-Stowe & D.H. Rakison (Eds.), *Building object categories in developmental time* (pp. 335–387). Hillsdale, NJ: Erlbaum.
Saenz-Ludlow, A., & Walgamuth, C. (1998). Third graders' interpretations of equality and the equal symbol. *Educational Studies in Mathematics*, *35*(2), 153–187.[E]
Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition*, *108*, 662–674.
Sarnecka, B.W., & Wright, C.E. (2013). The idea of an exact number: children's understanding of cardinality and equinumerosity. *Cognitive Science*, *37*, 1493–1506, [C].
Saxe, G. B., Gearhart, M., & Suad Nasir, N. (2001). Enhancing students' understanding of mathematics: a study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, *4*, 55–79.[G]
Schmittau, J. (2004). Vygotskian theory and mathematics education: resolving the conceptual-procedural dichotomy. *European Journal of Psychology of Education*, *1*, 19–43.[G]
Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, *47*(6), 1525–1538.
Schneider, M., & Stern, E. (2009). The inverse relation of addition and subtraction: a knowledge integration perspective. *Mathematical Thinking and Learning*, *11*(1–2), 92–101, [I].
Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, *46*(1), 178–192.
Schneider, M., & Stern, E. (2012). The developmental relations between conceptual and procedural knowledge: a multi-method approach. *Developmental Psychology*, *46*(1), 178–192, [G].
Semenza, C. (2002). Conceptual knowledge in arithmetic: the core of calculation skills. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *38*, 285–288.[G]
Siegler, R. S. (1991). In young children's counting, procedures precede principles. *Educational Psychology Review*, *3*(2), 127–135.[C]
Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: a microgenetic analysis. *Journal of Experimental Psychology. General*, *127*(4), 377–397.[I]
Simon, M. A. (1993). Prospective elementary teachers' knowledge of division. *Journal for Research in Mathematics Education*, *24*(3), 233–254.[G]
Simon, M. A. (2006). Key developmental understanding in mathematics: a direction for investigating and establishing learning goals. *Mathematical Thinking and Learning*, *8*(4), 359–371.[G]
Sophian, C. (1992). Learning about numbers: lessons for mathematics education from preschool number development. In J. Bideaud, C. Meljac, & J.-P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 19–40). Hillsdale, NJ: Lawrence Erlbaum Associates.[C]
Squire, S., Davies, C., & Bryant, P. (2004). Does the cue help? Children's understanding of multiplicative concepts in different problem contexts. *British Journal of Educational Psychology*, *74*, 515–532.[G]
Star, J. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education*, *36*(5), 404–411.

Steinberg, R. M., Sleeman, D. H., & Ktorza, D. (1991). Algebra students' knowledge of equivalence of equations. *Journal for Research in Mathematics Education*, *22*(2), 112–121.
Stern, E. (1992). Spontaneous use of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology*, *17*, 266–277, [I].
Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom*. New York, NY: Free Press.
Stock, P., Desoete, A., & Roeyers, H. (2009). Mastery of the counting principles in toddlers: a crucial step in the development of budding arithmetic abilities? *Learning and Individual Differences*, *19*, 419–422.
Stump, S. L. (2001). Developing preservice teachers' pedagogical content knowledge of slope. *Journal of Mathematical Behavior*, *20*, 207–227.[G]
Tennyson, R. D., & Cocchiarella, M. J. (1986). An empirically based instructional design theory for teaching concepts. *Review of Educational Research*, *56*(1), 40–71.[G]
Thanheiser, E. (2012). Understanding multidigit whole numbers: the role of knowledge components, connections, and context in understanding regrouping 3+ -digit numbers. *The Journal of Mathematical Behavior*, *31*, 220–234.[G]
Vilette, B. (2002). Do young children grasp the inverse relationship between addition and subtraction? Evidence against early arithmetic. *Cognitive Development*, *17*, 1365–1383.[I]
Zell, S. (2011). Mathematical literacy and how scientific experiments can promote that conception. In B. Sriraman & V. Freiman (Eds.), *Interdisciplinarity for the twenty-first century: Proceedings of the third international symposium on mathematics and its connections to arts and sciences* (Vol. 11, pp. 411–423). Charlotte, NC: IAP Information Age Publishing.[G]

Uncited references

Asquith, P., Stephens, A. C., Knuth, E. J., & Alibali, M. W. (2007). Middle school mathematics teachers' knowledge of students' understanding of core algebraic concepts: equal sign and variable. *Mathematical Thinking and Learning*, *9*(3), 249–272. doi:10.1080/10986060701360910.[E]
Baroody, A. J. (1986). Basic counting principles used by mentally retarded children. *Journal for Research in Mathematics Education*, *17*(5), 382–389.[C]
Baroody, A. J. (1993). The relationship between the order-irrelevance principle and counting skill. *Journal for Research in Mathematics Education*, *24*(5), 415–427.[C]
Behr, M., Erlwanger, S., & Nichols, E. (1980). How children view the equals sign. *Mathematics Teaching*, *92*, 13–15.[E]
Bermejo, V., & Lago, M. O. (1990). Developmental processes and stages in the acquisition of cardinality. *International Journal of Behavioral Development*, *13*(2), 231–250.[C]
Brush, L. (1978). Preschool children's knowledge of addition and subtraction. *Journal for Research in Mathematics Education*, *9*, 44–54.[I]
Canobi, K. H., & Bethune, N. E. (2008). Number words in young children's conceptual and procedural knowledge of addition, subtraction and inversion. *Cognition*, *108*, 675–686.[I]
Capraro, R. M., Capraro, M. M., Yetkiner, Z. E., Ozel, S., Gyu Kim, H., & Kucuk, A. (2010). An international comparison of grade 6 students' understanding of the equal sign. *Psychological Reports*, *106*(1), 49–53.[E]
Carpenter, T. P., Franke, M. L., & Levi, L. (2003). *Thinking mathematically: Integrating arithmetic and algebra in elementary school*. Portsmouth, NH: Heinemann.[E]
Chesney, D. L., McNeil, N. M., Brockmole, J. R., & Kelley, K. (2013). An eye for relations: eye-tracking indicates long-term negative effects of operational thinking on understanding of math equivalence. *Memory & Cognition*, *41*, 1079–1095.[E]
Chihara, T. (1985). Younger children's understanding of the counting principles. *Psychologia*, *28*, 203–215.[C]
Dube, A. K., & Robinson, K. M. (2010b). Accounting for individual variability in inversion shortcut use. *Learning and Individual Differences*, *20*, 687–693.[I]
Gilmore, C. K., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetical skill. *British Journal of Educational Psychology*, *76*, 309–331.[I]
Gilmore, C. K., & Spelke, E. S. (2008). Children's understanding of the relationship between addition and subtraction. *Cognition*, *107*, 932–945.[I]
Gregoire, J., & van Nieuwenhoven, C. (1995). Counting at nursery school and at primary school: toward an instrument for diagnostic assessment. *European Journal of Psychology of Education*, *10*(1), 61–75.[C]
Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, *38*(3), 258–288.[E]
Jones, I., Inglis, M., Gilmore, C., & Dowens, M. (2012). Substitution and sameness: two components of a relational conception of the equals sign. *Journal of Experimental Child Psychology*, *113*, 166–176.[E]
Jones, I., Inglis, M., Gilmore, C., & Evans, R. (2013). Teaching the substitutive conception of the equals sign. *Research in Mathematics Education*, *15*, 34–39. doi:10.1080/14794802.2012.756635;[E]
Jones, I., & Pratt, D. (2006). Connecting the equals sign. *International Journal of Computers for Mathematical Learning*, *11*(3), 301–325.[E]
Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: the concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology*, *54*, 105–115.[I]
Knuth, E. J., Stephens, A. C., McNeil, N. M., & Alibali, M. W. (2006). Does understanding the equal sign matter? Evidence from solving equations. *Journal for Research in Mathematics Education*, *37*(4), 297–312.[E]
Lai, M., Baroody, A. J., & Johnson, A. R. (2008). Fostering Taiwanese preschoolers' understanding of the addition-subtraction inverse principle. *Cognitive Development*, *23*, 216–235.[I]
Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*, 395–438.[C]
Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, *52*, 130–169.[C]

Li, X., Ding, M., Capraro, M. M., & Capraro, R. M. (2008). Sources of differences in children's understandings of mathematical equality: comparative analysis of teacher guides and student texts in China and the United States. *Cognition and Instruction*, *26*(2), 195–217.[E]

McCollister, T. S., Burts, D. C., Wright, V. L., & Hildreth, G. J. (1986). Effects of computer-assisted instruction and teacher-assisted instruction on arithmetic task achievement scores of kindergarten children. *Journal of Educational Research*, *80*(2), 121–125.[C]

McNeil, N. M. (2008). Limitations to teaching children 2 + 2 = 4: typical arithmetic problems can hinder learning of mathematical equivalence. *Child Development*, *79*(5), 1524–1537.[E]

McNeil, N. M. (2014). A change-resistance account of children's difficulties understanding mathematical equivalence. *Child Development Perspectives*, *8*(1), 42–47.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: students encode equations based on their knowledge of arithmetic. *Cognitive Science*, *28*, 451–466.[E]

McNeil, N. M., Chesney, D. L., Matthews, P. G., Fyfe, E. R., Petersen, L. A., Dunwiddie, A. E., et al. (2012). It pays to be organized: organizing arithmetic practice around equivalent values facilitates understanding of math equivalence. *Journal of Educational Psychology*, *104*(4), 1109–1121.[E]

Mix, K. S., Sandhofer, C. M., Moore, J. A., & Russell, C. (2012). Acquisition of the cardinal word principle: the role of input. *Early Childhood Research Quarterly*, *27*, 274–283.[C]

Molina, M., Castro, E., & Castro, E. (2009). Elementary students' understanding of the equal sign in number sentences. *Electronic Journal of Research in Educational Psychology*, *17*(7), 341–368.[E]

Munn, P., & Stephen, C. (1993). Children's understanding of number words. *British Journal of Educational Psychology*, *63*, 521–527.[C]

Nunes, T., Bryant, P., Burman, D., Bell, D., Evans, D., Hallett, D., et al. (2008). Deaf childen's understanding of inverse relations. In M. Marschark & P. C. Hauser (Eds.), *Deaf cognition: Foundations and outcomes* (pp. 201–225). New York, NY: Oxford University Press.[I]

Nunes, T., Bryant, P., Hallett, D., Bell, D., & Evans, D. (2009). Teaching children about the inverse relationship between addition and subtraction. *Mathematical Thinking and Learning*, *11*, 61–78.[I]

Ortega-Tudela, J. M., & Gomez-Ariza, C. J. (2006). Computer-assisted teaching and mathematical learning in Down Syndrome children. *Journal of Computer Assisted Learning*, *22*, 298–307.[C]

O'Rode, N. (2011). Latino/a students' understanding of equivalence: use of two standards-based curricula in eighth-grade algebra. In K. Tellez, J. Moschkovitch, & M. Civil (Eds.), *Latino/as and mathematics education: Research on learning and teaching in classrooms and communities* (pp. 19–36). Charlotte, NC: IAP Information Age Publishing.[E]

Perry, M. (1991). Learning and transfer: instructional conditions and conceptual change. *Cognitive Development*, *6*, 449–468.[E]

Perry, M., Breckinridge Church, R., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359–400.[E]

Petersen, L. A., & McNeil, N. M. (2013). Effects of perceptually rich manipulatives on preschoolers' counting performance: established knowledge counts. *Child Development*, *84*(3), 1020–1033.[C]

Rasmussen, C., Ho, E., & Bisanz, J. (2003). Use of the mathematical principle of inversion in young children. *Journal of Experimental Psychology*, *85*, 89–102.[I]

Renwick, E. (1932). Children's misconceptions concerning the symbols for mathematical equality. *British Journal of Educational Psychology*, *2*, 173–183.[E]

Robinson, K. M., & Dube, A. K. (2013). Children's additive concepts: promoting understanding and the role of inhibition. *Learning and Individual Differences*, *23*, 101–107.[I]

Robinson, K. M., & Ninowski, J. E. (2003). Adults' understanding of inversion concepts: how does performance on addition and subtraction inversion problems compare to performance on multiplication and division inversion problems? *Canadian Journal of Experimental Psychology*, *57*(4), 321–330.[I]

Robinson, K. M., Ninowski, J. E., & Gray, M. E. (2006). Children's understanding of the arithmetic concepts of inversion and associativity. *Journal of Experimental Child Psychology*, *94*, 349–362.[I]

Seo, K.-H., & Ginsburg, H. P. (2003). You've got to carefully read the math sentence. . .": classroom context and children's interpretations of the equals sign. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 161–187). Mahwah, NJ: Erlbaum.[E]

Sherman, J., & Bisanz, J. (2007). Evidence for use of mathematical inversion by three-year-old children. *Journal of Cognition and Development*, *8*(3), 333–344.[I]

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: benefits of solving problems with manipulatives. *Journal of Educational Psychology*, *101*(1), 88–100.[E]

Slusser, E. B., & Sarnecka, B. W. (2011). Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology*, *110*, 38–51.[C]

Stephens, A. C., Knuth, E. J., Blanton, M. L., Isler, I., Murphy Gardiner, A., & Marum, T. (2013). Equation structure and the meaning of the equal sign: the impact of task selection in eliciting elementary students' understandings. *The Journal of Mathematical Behavior*, *32*, 173–182.[E]

Stojanov, A. (2011). Children's understanding of quantitative and non-quantitative transformations of a set. *Review of Psychology*, *18*(2), 55–62.[C]

Zur, O., & Gelman, R. (2004). Young children can add and subtract by predicting and checking. *Early Childhood Research Quarterly*, *19*, 121–137.[C]