Developmental Review xxx (2009) xxx–xxx

Contents lists available at ScienceDirect

Developmental Review

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

The development of arithmetic principle knowledge: How do we know what learners know?

Richard W. Prather *,[1], Martha W. Alibali [2]

University of Wisconsin-Madison, Madison, WI 53706, United States

ARTICLE INFO

Article history:
Received 4 June 2008
Revised 11 September 2009
Available online xxxx

Keywords:
Cognitive development
Arithmetic
Number

ABSTRACT

This paper reviews research on learners' knowledge of three arithmetic principles: *Commutativity*, *Relation to Operands*, and *Inversion*. Studies of arithmetic principle knowledge vary along several dimensions, including the age of the participants, the context in which the arithmetic is presented, and most importantly, the type of knowledge assessment (e.g., application of procedures, evaluation of examples). The vast majority of studies utilize single-faceted knowledge assessments, which can lead to incomplete or misleading views of learners' knowledge. Both context and type of knowledge assessment can influence conclusions about learners' arithmetic principle knowledge. However, relatively few studies directly address the possible effects of context or type of knowledge assessment on their results. To move the field forward, research that utilizes multifaceted knowledge assessments is needed.

> Understanding of arithmetic essentially involves understanding of relationships—relationships among sets of numbers and relationships between operators.
> – Camilla Gilmore (2006)

* Corresponding author.
E-mail addresses: rwprather@wisc.edu, rwprathe@indiana.edu (R.W. Prather), mwalibali@wisc.edu (M.W. Alibali).
[1] Present address: Department of Psychological and Brain Sciences, Indiana University, 1101 10th Street, Bloomington, IN 47408, United States.
[2] Present address: Department of Psychology, University of Wisconsin-Madison, 1202 W. Johnson Street, Madison, WI 53705, United States.

0273-2297/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.dr.2009.09.001

The sum of two natural numbers must be greater than either of the numbers. Adding a number and then subtracting that same number leaves the original value unchanged. Regularities such as these are *principles* that apply in the domain of arithmetic. This review focuses on learners' acquisition of arithmetic principles, with an emphasis on two broad questions: what do learners know at various points in development? And, how does their knowledge change? In addressing these questions, we consider conceptual and methodological issues that arise in making sense of the literature, with a focus on generalization across contexts and the types of evidence used to assess knowledge. We also seek to identify potential gaps in the literature.

What is a principle?

Principles can be defined as fundamental laws or regularities that apply within a problem domain. For example, one principle that applies in the domain of mixture problems is that the concentration of the final solution must be in between the concentrations of the two initial solutions. Problem solvers have been shown to use principles in a variety of problem domains, including counting (e.g., Gelman & Gallistel, 1978), proportional reasoning (e.g., Dixon & Moore, 1996), and physics (e.g., Chi, Glaser, & Rees, 1982). A deep, flexible understanding of any domain presumably involves an understanding of principles.

Mathematical principles are fundamental properties about the functioning of the system of mathematics. Such principles represent basic truths about the parts of the system to which they apply. For example, the associative law for addition holds that, for all real numbers x, y and z, $(x + y) + z = x + (y + z)$. This is a fundamental truth about real numbers.

Principles are related to *concepts*, which can be defined as abstract or generic ideas that learners infer or derive based on specific instances. Principles are inherent aspects of a domain, whereas concepts are mental representations constructed by the learner. Knowledge of mathematical principles may be inferred from experience in the domain, or it may be explicitly taught. In arithmetic, some commonly taught principles include associativity for addition and multiplication, commutativity for addition and multiplication, and the distributive property of multiplication. However, being taught or formally recognized is not essential to being a principle. Principles are regularities that may be used by problem solvers regardless of whether or not they are formally recognized.

This review addresses three principles in the domain of arithmetic: *commutativity, relation to operands*, and *inversion*. These three principles describe important properties of mathematical operations, and all three have been extensively studied. Unless specifically noted, these principles are defined and investigated as they relate to natural numbers, which are positive whole numbers not including zero.

Commutativity pertains to the syntax of arithmetic operations; the order of the operands is irrelevant for operations that are commutative. Addition and multiplication are both order-irrelevant operations; $a + b = b + a$ and $a \times b = b \times a$. Subtraction and division are both order-relevant operations; $a - b \neq b - a$ and $a \div b \neq b \div a$.

Relation to operands describes the relation between the operands in a given arithmetic equation and the result of the operation; thus, it reflects knowledge about the expected outcomes of operations. Relation to operands is a family of principles, with the exact relation depending on the particular operation and the types of numbers being considered. When operating on natural numbers, in a simple addition equation ($a + b = c$), the sum (c) must be greater than both addends (a and b). In a simple subtraction equation ($a - b = c$), the difference (c) must be less than the minuend (a), however it may have any relation with the subtrahend (b). In a simple multiplication equation ($a \times b = c$), the product (c) must be greater than both multiplicands (a and b). In a simple division equation ($a \div b = c$), the quotient (c) must be smaller than the dividend (a); however it may have any relation with the divisor (b). These statements are all examples of *relation to operands* principles.

Inversion describes relations between operations. Certain arithmetic operations can be viewed as opposites of one another. For example, addition and subtraction are inverse operations; subtraction "undoes" addition and vice versa. The inversion principle holds that inverse operations involving the same value result in no net change, thus, $a + b - b = a$, and $a \times b \div b = a$. The inversion principle applies regardless of the order of the operands (e.g., $b + a - b = a$).

Knowledge evaluation: how do we know what learners know?

How may investigators infer knowledge based on behavior? Any study that seeks to characterize learners' knowledge must utilize some method of assessing that knowledge. Since we have yet to develop a method of reading minds, experimenters must provide participants opportunities to display behavior that allows their knowledge to be characterized. Each individual study may employ a unique method of assessing knowledge; however, across studies, there are several general types of knowledge assessments used. These categories of knowledge assessments were discussed in previous work by Bisanz and LeFevre (1992).

For purposes of this review, we build on Bisanz and LeFevre's prior work to propose a slightly modified list of assessment types, which includes some distinctions that Bisanz and LeFevre did not draw. Our list is as follows.

Application of procedures

The learner uses a procedure that implies knowledge of the principle. Researchers may design a task specifically to elicit the procedure, or use a task that is commonly used in the domain. For example, some procedures for solving simple addition equations can be taken as evidence of commutativity knowledge. One such procedure is counting-on-larger. A learner who uses this procedure solves $a + b = ?$ equations by counting up from the larger addend. For example, 3 + 5 would be solved by counting 5...6, 7, 8. Use of the counting-on-larger procedure implies knowledge of commutativity because the procedure treats the order of the addends as irrelevant. Thus, learners who use this procedure when appropriate are inferred to have knowledge of commutativity, at least as revealed through application of the procedure.

Evaluation of procedures

The learner recognizes the validity of a procedure that implies knowledge of the principle. For example, a learner may state that the equation 3 + 7 can be solved by switching it to 7 + 3 and counting up from 7. The learner may also evaluate procedures used by another individual. For example, in some studies, learners evaluate whether a puppet can legitimately use particular procedures in solving arithmetic equations (e.g., Canobi, Reeve, & Pattison, 1998).

Justification of procedures

The learner provides a justification for the use of a procedure by him or herself or others. For example, a learner may explain that one can solve 3 + 7 by looking at the result of 7 + 3, because the order of the addends does not matter.

Evaluation of examples

The learner differentiates between examples in the domain that violate a principle and those that do not. In this type of assessment, learners do not evaluate the procedure *per se*, but rather they evaluate a statement or problem that includes a result or that expresses a relationship. For instance, participants may indicate that equations that violate a principle are "worse" or "more wrong" than equations that do not (Dixon, Deets, & Bangert, 2001). Dixon and colleagues showed learners sets of equations that contained either principle violations (e.g., 5 + 3 = 4, a violation of relation to operands, which holds that, for natural numbers, the sum must be greater than both addends) or incorrect equations that were consistent with the principle (e.g., 5 + 3 = 12). Learners who rated sets with violation equations as worse than sets with only non-violations were inferred to have knowledge of the relation to operands principle. In this case, learners did not evaluate any of the many possible procedures that could have been used to solve the arithmetic equations; they simply evaluated the solved equations.

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

Explicit recognition

The learner provides or recognizes a statement that is consistent with a principle. This type of evidence can be obtained by stating a principle in general terms and asking the learner whether the principle applies in certain situations. For example, learners can be shown statements such as "The result is always larger than the first number in the problem" and asked to which operations the statement applies (Dixon et al., 2001).

Every investigation of learners' knowledge uses at least one of these types of assessments. The categories are to some degree fuzzy; there may be more than one way to utilize an assessment type, and in some instances researchers may employ a blend of more than one assessment type.

As a case in point, there are multiple routes to infer knowledge via procedure application. Researchers use a variety of tasks that differ in terms of how central knowledge of the principle is to completing the task. For some tasks, the only reasonable way for learners to "pass" the task is to use a procedure that implies knowledge of the principle. Many researchers construct tasks of this sort for assessing learners' knowledge. For other tasks, there may be several plausible ways for learners to complete the task, only one of which implies principle knowledge. Thus principle knowledge may be "optional" in completing the task. Tasks used by researchers can be conceived as lying on a continuum that reflects how essential principle knowledge is to completing the task.

Every type of assessment involves some degree of recognition, production or evaluation by the learner. Thus, each may tap different aspects of learners' knowledge. In light of these different task demands, it cannot be assumed that learners will perform similarly on all types of knowledge assessments.

This general issue has been addressed in domains other than arithmetic, as knowledge assessment is crucial to much behavioral research. For example, research on counting principles has yielded different conclusions about what learners know at various ages, depending on the types of knowledge assessments that are used. In the counting literature, this issue is often framed in terms of competence vs. performance. The requirements of the tasks used in different studies are not exactly the same (LeCorre, Van de Walle, Brannon, & Carey, 2006). Some learners may know the counting principles (i.e., have competence), but be unable to produce the performance to show it, depending on the particular knowledge assessment being used. For example, if the task requires learners to make long counts, other factors such as memory limitations may influence performance. Thus, a particular knowledge assessment may lead to a misrepresentation or underestimation of learners' knowledge.

Some tasks have been criticized as not showing the learner's "true" knowledge because the demands on the learner are too great—the idea being that an "easier" task would more accurately tap that knowledge. However, is also possible that an "easier" task might misrepresent the learner's knowledge in a different way. If a learner can only produce principle-consistent behavior on a contrived task that has stripped away any need to generalize knowledge of a principle, then how well does the learner really "know" that principle? Knowledge that can be employed in only an extremely limited fashion may be qualitatively different from knowledge that can be more broadly applied.

Moreover, evidence for knowledge can be potentially ambiguous. For example, not all researchers agree that application of a procedure necessarily implies knowledge of the relevant principle (Baroody & Ginsburg, 1986). The issue is not that there is one perfect knowledge assessment for any given concept or principle. No knowledge assessment is "best" or "ideal"; instead, each provides a slightly different window onto the learner's knowledge.[3]

It seems likely that performance on different types of knowledge assessments will vary across development, because of developmental differences in the abilities required for each task. In addition, there are several other reasons to expect performance on different types of knowledge assessments to vary with development. First, there is evidence that learners sometimes understand correct procedures that they do not use (e.g., Siegler & Crowley, 1994), and even when learners know a procedure, they do not always use it consistently (Miller & Seier, 1994). Thus, learners might succeed on

[3] It should be noted that unreasonable task demands are a real issue. The task demands need to be thoroughly characterized in order for an accurate conclusion to be made about learners' performance on the task.

assessments that rely on *recognition* of correct procedures, before they succeed on assessments that require overt *use* of a correct procedure. Second, learners sometimes use correct procedures before they understand the principles that underpin those procedures (e.g., Briars & Siegler, 1984). Thus, learners might succeed on assessments that rely on *use* of a correct procedure before they succeed on assessments that require *justification* of that procedure. Finally, both theoretical arguments and empirical findings suggest that development often involves a progression from more implicit to more explicit knowledge (Goldin-Meadow, Alibali, & Church, 1993; Karmiloff-Smith, 1986, 1992; Siegler & Stern, 1998). For any particular knowledge assessment, it may be difficult to ascertain whether implicit or explicit knowledge is required for success on that assessment. However, broadly speaking, learners might be expected to succeed on assessments that do not require explicitly articulated knowledge before they succeed on assessments that require producing explicit justifications.

Context of arithmetic

A second key issue to be considered in addressing knowledge of arithmetic principles is the context in which the arithmetic is presented. Context has been shown to be important, not only in arithmetic, but in problem solving more generally. Previous studies have shown that context affects problem solving in many different domains. Examples include the Wason selection task (Ahn & Graham, 1999; Wason, 1966), the Tower of Hanoi problem (Evans, 1982; Kotovsky, Hayes, & Simon, 1985), and the river crossing problem (Jeffries, Polson, Razran, & Atwood, 1977), to name only a few. In each case, small changes to seemingly superficial details of the problem presentation lead to significantly different patterns of performance by learners. For example, some versions of the Tower of Hanoi problem can take on average 16 times longer to solve than others (Kotovsky et al., 1985).

Studies of arithmetic knowledge typically use one of several contexts: symbolic, verbal, object or abstract. Though the labels and the exact details of the contexts vary from study to study, the important point is that context can vary. The *symbolic* context generally involves a symbol only presentation of an arithmetic operation (e.g., $5 + 3$). The *verbal* context (also called story or word problem context) generally involves a written or oral story scenario, such as "John has five apples and Jane has three apples". The *object* context (often called the *non-verbal* context) involves sets of actual objects or visual presentations of objects that can be combined or otherwise manipulated to correspond with arithmetic operations. The *abstract* context involves conveying arithmetic operations with nonspecific values. This can be done with either symbols (e.g., $x + y$), words (e.g., more, less) or objects, where exact amounts are concealed. Every study presents arithmetic in a particular context, even if the effect of context is not the focus of the study.

The effects of context on arithmetic problem solving have been well documented. Several studies have reported variations in performance on arithmetic or algebraic problems across different contexts (e.g., Jordan, Huttenlocher, & Levine, 1992; Koedinger & Nathan, 2004; Koedinger, Alibali, & Nathan, 2008; Levine, Jordan, & Huttenlocher, 1992). For example, Levine et al. (1992) examined children's arithmetic performance in object, story and symbolic contexts, and found that 4-year-old participants performed better in the object context than in either of the other contexts. Koedinger and Nathan (2004) investigated high school students' performance on simple algebraic problems in symbolic and verbal contexts, and found that students performed more accurately in verbal problem contexts.

Resnick's (1992) theory of general arithmetic development postulates that arithmetic thinking progresses through several stages that correspond to thinking about number in different contexts. According to Resnick, learners first think in terms of proto-quantities (object context), then quantities (verbal context), then numbers (symbolic context) and finally operators (abstract context). Gradually, through experience, learners gain capabilities for thinking about number in these contexts. The stages are not all-or-nothing and the transition from stage to stage can be relatively slow. Other theories of development also postulate that children progress from more concrete to more symbolic or abstract thinking (e.g., Karmiloff-Smith, 1992; Piaget, 1952; Werner & Kaplan, 1963).

This general idea may apply, not only to proficiency and performance on typical arithmetic tasks, but also to knowledge of arithmetic principles (see Resnick, 1992). Indeed, given the ample evidence that context plays a role in problem solving, it seems highly likely that context affects principle

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

understanding. It seems plausible that learners might display knowledge of principles in object contexts before they do so in verbal or symbolic contexts, as Resnick's theory would predict. However, most studies of arithmetic principle knowledge do not bear on this issue, because they utilize only a single context. Older children are typically tested only in a symbolic or verbal context (e.g., Siegler & Stern, 1998, inversion in a symbolic context with 8-year-olds), while younger children are often tested only in an object context (e.g., Klein & Bisanz, 2000, inversion in an object context with 4-year-olds). Only a handful of studies compare understanding of the same children across multiple contexts. Given that different studies use different contexts, and context may affect performance, it is important to take context into consideration when attempting to draw conclusions about the development of arithmetic principle knowledge.

Knowledge profiles

In view of the complexity of evaluating learners' knowledge, we endorse the use of multifaceted knowledge assessments, or *knowledge profiles*, which summarize individuals' performance across multiple knowledge assessments. Bisanz and colleagues (Bisanz & LeFevre, 1992; Bisanz, Watchorn, Piatt, & Sherman, 2009) have argued that multifaceted assessments characterize learners' knowledge more accurately than single-faceted assessments. Along similar lines, other researchers have argued that knowledge is often best characterized as partial or graded rather than binary or all-or-none (e.g., McNeil & Alibali, 2005; Munakata, 2001); such knowledge is often displayed on one type of assessment but not another. Despite the advantages of knowledge profiles, however, they are rarely used in practice. For example, Bisanz and colleagues (2009) discuss the merits of a particular type of knowledge profile in characterizing knowledge of inversion, but such profiles have yet to be used in any empirical study of inversion. Across principles, most studies utilize single-faceted assessments that yield a single conclusion regarding whether a learner "has" or "does not have" the knowledge in question.

At face value, more narrow assessments of learners' knowledge seem less valuable than more comprehensive assessments. However, this is not to say that narrow, targeted knowledge assessments have no value. Rather, the value depends on the goals of the researcher. While a particular type of knowledge assessment may not fully characterize the state of the learner's knowledge, the targeted aspect of the learner's behavior is still valuable to understand.

This review illustrates the range of knowledge assessments in use in the literature. Participants' behavior varies as a function of the type of assessment and the context in which the arithmetic is presented. These variations imply that participants' knowledge is more complicated than "have" or "have not", and they underscore the importance of knowledge profiles. We return to this issue in the concluding section.

Examining arithmetic principles

In the following sections, we look in detail at three arithmetic principles that have been investigated in the literature: *commutativity*, *relation to operands* and *inversion*. This is by no means an exhaustive list of arithmetic principles; however, these three principles have each received a great deal of research attention. For each principle, we will review the relevant findings with the issues of assessment type and context in mind. Of course, these are not the only factors that affect whether children display knowledge of principles; other factors, such as problem size and problem format, may also come into play (e.g., Canobi & Bethune, 2008). However, assessment type and context are factors that cut across much of the literature, and that raise important considerations for theories about how principle knowledge is acquired, so we have chosen to focus on them here. Within each section, we address how understanding of the principle is manifested across age groups.

Commutativity

Commutativity is possibly the most extensively researched principle, most likely because it is often explicitly taught in formal mathematics classes. The studies considered here are those that seek to characterize learners' knowledge of *commutativity*. Studies of commutativity knowledge have used several types of knowledge assessments. Most common is application of procedures, though several studies have also utilized justification of procedures. Individual studies and their outcomes are summarized in Tables 1(a–d).

Commutativity knowledge based on application of procedures

The most popular type of assessment used in the literature as an indicator of commutativity knowledge is application of procedures that imply knowledge of the principle. Studies of commutativity that use applications of procedures fall in one of two categories of relevance to the previously discussed issue of how essential principle knowledge is to completing the task. Researchers either construct their own "new" tasks or use more domain-typical tasks. For domain-typical tasks, there are usually several plausible procedures that learners can use to complete the task, only one of which implies principle knowledge. For such tasks, there may be relatively little motivation for learners to use a procedure that suggests principle knowledge, as opposed to other procedures they may be familiar and successful with. For constructed tasks, learners generally have fewer options for procedures to use. The target procedure may be the only reasonable way for the learner to "pass" the task.

Constructed tasks

Many investigations of commutativity knowledge have used tasks specifically designed to elicit procedures consistent with commutativity. In one such task, participants were given a series of $a + b$ equations to solve by any method they wished (Baroody, Ginsburg, & Waxman, 1983). Participants' answers to previous equations were always displayed, giving them the opportunity to reference a previously solved equation to solve the current equation. This procedure, termed "looking back", is consistent with commutativity if the current equation $(a + b)$ is the commuted pair of the previous equation $(b + a)$. Participants who used this procedure when appropriate were inferred to have knowledge of commutativity. As a second example of a constructed task, Canobi (2002) showed participants colored boxes of candy, where the color corresponded to a certain amount of candy. Participants then

Table 1a
Studies of commutativity using application of procedures on constructed tasks as knowledge evidence.

Study	Age group (yrs)	Context	Operation	DV	Outcome (success)
Sophian et al. (1995)	3–4	Object	Addition	% of participants	57
Sophian et al. (1995)	5	Object	Addition	% of participants	64
Canobi et al. (2002)	4–5	Object	Addition	% of trials	70
Cowan and Renton (1996)	5	Object	Addition	% of participants	54
Cowan and Renton (1996)	5	Abstract	Addition	% of participants	54
Baroody and Gannon (1984)	5–6	Symbolic	Addition	% of participants	51
Canobi et al. (2002)	5–6	Object	Addition	% of trials	85
Canobi et al. (2002)	5–6	Object	Addition	% of trials	79
Wilkins et al. (2001)	6	Verbal	Addition	% of participants	77
Baroody et al. (1983)	6	Symbolic	Addition	% of trials	72
Cowan and Renton (1996)	6–9	Object	Addition	% of participants	92
Cowan and Renton (1996)	6–9	Symbolic	Addition	% of participants	96
Cowan and Renton (1996)	6–9	Abstract	Addition	% of participants	77
Baroody et al. (1983)	7	Symbolic	Addition	% of trials	83
Baroody et al. (1983)	8	Symbolic	Addition	% of trials	83
Canobi (2009)	8	Symbolic	Addition	% of trials	75[a]
Canobi (2005)	7–9	Symbolic	Addition	% of trials	94
Canobi (2005)	7–9	Object	Addition	% of trials	97

[a] Averaged across conditions at pretest.

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

Table 1b
Studies of commutativity using application of procedures on domain tasks as knowledge evidence.

Study	Age group (yrs)	Context	Operation	DV	Outcome (success)
Groen and Resnick (1977)	4–5	Symbolic	Addition	% of participants	50
Cowan and Renton (1996)	5	Symbolic	Addition	% of participants	46
Baroody and Gannon (1984)	5–6	Symbolic	Addition	% of participants	31
Canobi et al. (2003)	5–8	Symbolic	Addition	% of participants	76
Canobi et al. (2003)	5–8	Symbolic	Addition	% of trials	25
Carpenter and Moser (1984)	6	Symbolic	Addition	% of participants	10
Canobi et al. (2002)	6	Symbolic	Addition	% of participants	33
Canobi et al. (2002)	6	Symbolic	Addition	% of trials	5
Canobi et al. (1998)	6–8	Symbolic	Addition	% of participants	96
Canobi et al. (1998)	6–8	Symbolic	Addition	% of trials	53
Cowan and Renton (1996)	6–9	Symbolic	Addition	% of participants	83
Carpenter and Moser (1984)	7	Symbolic	Addition	% of participants	70
Carpenter and Moser (1984)	8	Symbolic	Addition	% of participants	55
Squire et al. (2004)	9	Verbal	Multiplication	% of participants	52
Squire et al. (2004)	10	Verbal	Multiplication	% of participants	92
Campbell (1999)	Adults	Symbolic	Multiplication	Reaction Time	Yes
Rickard and Bourne (1996)	Adults	Symbolic	Multiplication	Reaction Time	Yes
Rickard and Bourne (1996)	Adults	Verbal	Multiplication	Reaction Time	No

Table 1c
Studies of commutativity using evaluation of procedures as knowledge evidence.

Study	Age group (yrs)	Context	Operation	DV	Outcome (success)
Canobi et al. (2003)	5–8	Object	Addition	% of trials	70
Canobi et al. (2003)	5–8	Symbolic	Addition	% of trials	72
Canobi et al. (2003)	5–8	Abstract	Addition	% of trials	70
Canobi et al. (1998)	6–8	Symbolic	Addition	% of trials	87

Table 1d
Studies of commutativity using justification of procedures as knowledge evidence.

Study	Age group (yrs)	Context	Operation	DV	Outcome (success)
Canobi et al. (2002)	5	Object	Addition	% of trials	>83% of correct trials[a]
Baroody and Gannon (1984)	5–6	Symbolic	Addition	% of participants	51
Canobi et al. (2003)	5–8	Object	Addition	% of participants	35
Canobi et al. (2003)	5–8	Symbolic	Addition	% of participants	37
Canobi et al. (2003)	5–8	Abstract	Addition	% of participants	36
Canobi et al. (2002)	6	Object	Addition	% of trials	>89% of correct trials[a]
Canobi et al. (2002)	6	Object	Addition	% of trials	>87% of correct trials[a]
Langford (1981)	6	Object	Addition	% of participants	3
Langford (1981)	6	Object	Multiplication	% of participants	6
Canobi et al. (1998)	6–8	Symbolic	Addition	% of trials	67
Canobi (2009)	8	Symbolic	Addition	% of trials	59[b]
Canobi (2005)	7–9	Symbolic	Addition	% of trials	90
Canobi (2005)	7–9	Object	Addition	% of trials	76

Each row corresponds to one reported result within a study. DV is the dependent variable as reported by the author.

[a] Canobi et al. (2002) report justification data for trials with correct judgments only, and the justification categories are not mutually exclusive, making it impossible to ascertain the exact percentage of children who provided correct justifications.

[b] Averaged across conditions at pretest.

viewed boxes being passed out to two puppets, and were asked if the puppets had the same number of candies: "Bill gets a box of reds, then he gets four greens. Kate gets four greens, then she gets a box of reds. Do Bill and Kate have the same number of Smarties?" Participants who responded that both puppets had the same number were inferred to have knowledge of commutativity.

Several versions of constructed tasks that employ this general framework have been used. In each version it is advantageous, if not required, for the participant to note that two expressions are equivalent. This type of paradigm has been employed in an object context (Canobi et al., 2002; Sophian, Harley, & Martin, 1995), a verbal context (Wilkins, Baroody, & Tiilikainen, 2001), a symbolic context (Canobi, 2009), and in multiple contexts (Cowan & Renton, 1996).

Domain-typical tasks

A different method involves examining use of principle-consistent procedures in typical domain tasks that are not expressly designed to elicit such procedures. The range of procedures used by children in solving simple arithmetic problems has been well cataloged (Baroody, 1987; Carpenter & Moser, 1984; Siegler, 1987; Siegler & Jenkins, 1989). Two such procedures used in addition are counting-on-larger (COL) and counting-all-larger (CAL). A learner using the COL procedure would solve both 3 + 5 and 5 + 3 by starting from five, and counting up three (5...6, 7, 8). A learner using CAL would do the same, but would start at one and count up to five, rather than start from five. Both of these procedures imply knowledge of commutativity because they treat the order of the addends as irrelevant. Thus, use of COL or CAL can be interpreted as evidence of commutativity knowledge, on the assumption that learners would use one of these procedures only if they thought that the order of the addends did not influence the sum.

Many studies measure use of COL or CAL (Baroody & Gannon, 1984; Canobi, Reeve, & Pattison, 2003; Canobi et al., 1998, 2002; Carpenter & Moser, 1984; Cowan & Renton, 1996; Groen & Resnick, 1977), which we interpret as possible evidence of commutativity knowledge, and therefore relevant to this review. However, not all researchers who report use of COL or CAL accept it as evidence of commutativity knowledge (see Baroody, Wilkins, & Tiilikainen, 2003, for discussion). For example, children might count on from the larger addend because they have learned that procedure in school, not because they understand the principle. Further, children who use COL or CAL may not realize that starting with the larger number *must* yield the same solution as starting with the smaller number; they may believe that starting with the larger number yields *a* possible solution, but not necessarily *the same* solution. Thus, use of COL or CAL may not always indicate understanding of commutativity. Use of COL or CAL is generally determined by self-report, though it is also possible to analyze participants' reaction times.

In some studies, differences in performance when commutativity is relevant vs. not relevant are interpreted as evidence for knowledge of the principle. For example, learners who are given $63 \div 7$ to solve will respond relatively quickly if they solved the same equation very recently, because the earlier equation primes the later equation (Fendrich, Healy, & Bourne, 1993; Rickard & Bourne, 1996). Campbell (1999) investigated whether multiplication equations are primed by their commuted pair—for example, is 6×9 primed by 9×6?

A similar approach was used by Squire, Davies, and Bryant (2004). In this case, participants were given multiplication word problems, and for some of the problems, knowledge of commutativity was useful in determining the correct answer. For example, "*Christopher has 33 bags of coins, each with 18 coins in them. Altogether he has 594 coins. James has 18 bags, each with 33 coins in them. How many coins does James have?*" Participants who performed better on commutativity-relevant word problems than on problems on which commutativity was not relevant were inferred to have knowledge of the principle.

Commutativity knowledge based on the evaluation of procedures

Principle knowledge can also be assessed by learners' evaluation of procedures. Rather than measuring use of a procedure, learners are presented with a procedure and are asked to evaluate whether the procedure is "right" (Canobi et al., 1998, 2003). For example, a learner may be told that a puppet solved 3 + 5 by looking at the result of 5 + 3. A learner who indicates that this is a valid procedure might be thought to understand the principle that this procedure is based on, commutativity.

Commutativity knowledge based on justification of procedures

Principle knowledge can also be assessed via learners' justification of procedures. This type of evidence involves participants reasoning about procedures used by themselves or by others. As noted previously, some studies utilize constructed tasks in which it is possible to solve a simple equation ($a + b$) by referencing a previously solved equation ($b + a$). Participants may be asked to provide a justification for this procedure, either when they produce it themselves (Baroody & Gannon, 1984; Canobi et al., 2002, 2003; Langford, 1981) or when it is shown to them via a puppet display (Canobi, 2005, 2009; Canobi et al., 1998, 2003).

Comparing assessment types

There are many routes to evaluating knowledge of commutativity. Application of procedures is the most common type of assessment used in the literature. Studies utilizing application of procedures are split between constructed tasks (e.g., solving $a + b$ problems in a setting in which answers to previously solved problems, including commuted pairs, are displayed) and more typical domain tasks (e.g., solving addition problems under ordinary circumstances). Fortunately, there have been a few within-study contrasts between these task types (Baroody & Gannon, 1984; Canobi et al., 1998, 2002, 2003). These studies should be highlighted because the use of multiple assessment types within a single study is potentially more informative than cross-study comparisons. Several studies have reported higher estimates of knowledge for constructed tasks than for domain-typical tasks (Baroody & Gannon, 1984; Canobi et al., 2002, 2003; see Table 1). The opposite is the case for one study (Canobi et al., 1998). Thus, learners generally show more knowledge on constructed tasks.

Whether or not learners apply a particular procedure depends to some degree on what options they have. In domain-typical tasks there are often many procedures that learners may use; the relevant procedure is not the only possible procedure. In contrast, in constructed tasks, a procedure based on commutativity knowledge may be by far the easiest way for the learner to navigate the task. Constructed tasks are generally designed so that the principle-consistent procedure is either required or advantageous, while in domain-typical tasks the principle-consistent procedure may not be any better or easier than other procedures the learner may know. Thus, there may be more motivation to use a commutativity-consistent procedure on a constructed task. This is not to say that constructed tasks are better knowledge assessments than domain typical tasks. However, differences in the demands of the tasks need to be kept in mind when comparing results.

The use of COL and CAL procedures in simple arithmetic is an important procedural advancement by young learners. Inferring commutativity knowledge based on use of COL or CAL in simple arithmetic may lead to lower knowledge estimates than inferring it based on performance on constructed tasks; however, this does not imply that domain-typical tasks are a less useful knowledge assessment than constructed tasks. It is just as important to characterize how knowledge is used than simply to claim its presence or absence.

Several studies have reported that participants justify and evaluate commutativity-consistent procedures as well as they apply them (Canobi, 2005; Canobi et al., 2002, 2003). These studies looked primarily at learners' justifications of their own procedure use (see Table 1c). Participants' justifications of others' procedures (Canobi et al., 1998, 2003) seem to be more elaborate than justifications of their own procedures (Baroody & Gannon, 1984; Canobi et al., 2002).

Commutativity knowledge across contexts

The vast majority of commutativity studies are conducted in the symbolic context only. A few studies use other contexts, such as verbal or object; however differences in ages of participants and other methodological particulars make it difficult to draw any firm conclusions about the effects of context. The best evidence regarding effects of context comes from within-study comparisons, of which there are three: Canobi et al. (2003), Canobi (2005), and Cowan and Renton (1996).

Canobi and colleagues (2003) investigated 5–8 year olds' knowledge of commutativity using evaluation of procedures and justification of procedures. This study compared participants' knowledge in object, symbolic and abstract contexts. For both assessment types, participants' performance did not vary across contexts. For evaluation of procedures, participants' behavior revealed principle knowledge on 70% of trials in the object context, 73% in the symbolic context, and 70% in the abstract context. For justification of procedures, participants displayed behavior consistent with the principle on 35% of trials in the object context, 37% in the symbolic context and 36% in the abstract context.

Similar results were found by Canobi (2005) with 7–9 year olds. This study compared participants' knowledge in applying a procedure (i.e., judging whether a puppet could work out a problem by "looking back" at a previously solved problem) and in justifying that procedure. Participants' judgments revealed knowledge of commutativity on 94% of trials in a symbolic context, and 97% of trials in an object context. Participants' justifications revealed slightly better performance in the object context, with 89.5% of justifications invoking commutativity in the object context, compared to 76% in the symbolic context.

Cowan and Renton (1996) compared 6–9 year olds' knowledge of commutativity using application of procedures in a constructed task. Most participants displayed knowledge of commutativity in an object context (91%) and in a symbolic context (95%); significantly fewer participants displayed knowledge of the principle in an abstract context (77%).

Recall that Resnick's (1992) theory implies that understanding should emerge first in an object context, then in a verbal context, then in a symbolic context, and finally in an abstract context. The findings just reviewed provide partial support for this view, in that children displayed less knowledge in a symbolic context than in an object context (Canobi, 2005), and less knowledge in an abstract context than in an object or symbolic context (Cowan & Renton, 1996). Moreover, across-study comparisons at a given age are broadly compatible with this theory. For example, studies of 6-year-old children reveal slightly better performance in an object context (85% and 79% of trials in two experiments; Canobi et al., 2002) than in a verbal context (77% of trials; Wilkins et al., 2001) or a symbolic context (72% of trials, Baroody et al., 1983). However, more rigorous within-study tests of these developmental orderings are still needed, as no study to date has clearly demonstrated the full developmental trajectory.

When does knowledge of commutativity emerge?

There is no specific answer to the question of when learners "have" knowledge of commutativity. Generally, older participants show more knowledge than younger participants, and participants who are further along in formal schooling show more knowledge than participants who are less far along. There are a few positive results for commutativity for addition with learners as young as four and five (Canobi et al., 2002; Sophian et al., 1995), though there are some negative results with this age group as well (Canobi et al., 2002; Cowan & Renton, 1996; Groen & Resnick, 1977). It seems that not until at least age 6 do most learners consistently show knowledge of commutativity for addition in any of the assessment types or contexts (Baroody et al., 1983; Canobi et al., 2002; Wilkins et al., 2001). At age 6, most learners display knowledge in constructed tasks; fewer do so in domain-typical tasks (Canobi et al., 2002; Carpenter & Moser, 1984).

Studies of commutativity for multiplication are rare, but the existing evidence suggests that children have knowledge of commutativity for multiplication as young as age 9, as revealed by application of procedures in a verbal context (Squire et al., 2004).

Directions for future research

The literature on commutativity knowledge has several gaps. Application of principle-consistent procedures is by far the most commonly used means of assessing knowledge of commutativity. There are no studies of commutativity using explicit recognition or evaluation of examples as assessments of knowledge. This greatly limits comparisons of commutativity knowledge across assessment types. Although it may seem difficult, it would be possible to investigate commutativity using evaluation of examples. Learners could be asked to evaluate the performance of a hypothetical student on a certain equation given that the student knows certain other equations. For example, if a student previously correctly solved $6 + 3$, then participants may find it unexpected if the student incorrectly solves $3 + 6$. In this case incorrectly solving $3 + 6$ would be evaluated as worse if the student had

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

previously solved the commuted pair 6 + 3 as opposed to an unrelated equation such as 5 + 2. An investigation of commutativity using evaluation of examples may reveal knowledge not shown by other assessment types.

With respect to contexts, the vast majority of studies of commutativity knowledge utilize a symbolic context only (see Table 1), though a few use verbal (Squire et al., 2004; Wilkins et al., 2001) or object contexts (Canobi et al., 2002). In only three cases are multiple contexts used within the same study (Canobi, 2005; Canobi et al., 2003; Cowan & Renton, 1996). Furthermore, even these within-study comparisons do not allow firm conclusions to be drawn about the developmental ordering or emergence of commutativity knowledge across contexts.

In summary, the commutativity literature is largely focused on symbolic contexts and application of procedures. The literature generally ignores the knowledge learners may show in other contexts, especially young learners in an object context, and the knowledge learners might reveal using different types of knowledge assessments.

Relation to operands

Relation to operands principles describe the relative values of the operands and the result in any simple equation. In the case of adding natural numbers ($a + b = c$), the sum must be greater than both of the addends. Similarly, for multiplying natural numbers ($a \times b = c$), the product must be greater than or equal to both multiplicands. For subtraction ($a - b = c$), the difference must be less than the minuend. For division ($a \div b = c$), the quotient must be smaller than or equal to the dividend. Research relevant to these principles also includes studies of less formal versions of the principles, such as the ideas that "addition makes more" and "subtraction makes less." Individual studies and their outcomes are summarized in Tables 2(a–c).

Knowledge of relation to operands based on application of procedures

Several studies have investigated young learners' knowledge of relation to operands using application of procedures to assess knowledge. This is done by giving learners opportunities to indicate the magnitude of the result relative to the operands. Studies in both object contexts (Brush, 1978; Starkey, 1992) and verbal contexts (Sophian & McCorgray, 1994) have utilized this approach.

Table 2a
Studies of relationship to operands using application of procedures as knowledge evidence.

Study	Age group	Context	Operation	DV	Outcome (success)
Starkey (1992)	18 mos	Object	Addition/subtraction	% of trials	76
Starkey (1992)	24 mos	Object	Addition/subtraction	% of trials	82
Starkey (1992)	30 mos	Object	Addition/subtraction	% of trials	96
Starkey (1992)	36 mos	Object	Addition/subtraction	% of trials	96
Starkey (1992)	42 mos	Object	Addition/subtraction	% of trials	92
Starkey (1992)	48 mos	Object	Addition/subtraction	% of trials	97
Sophian and McCorgray (1994)	4 yrs	Verbal	Addition	% of trials	30
Sophian and McCorgray (1994)	4 yrs	Verbal	Subtraction	% of trials	55
Brush (1978), Exp. 2	4 yrs	Object	Addition	% of participants	95
Brush (1978), Exp. 2	4 yrs	Object	Subtraction	% of participants	100
Brush (1978), Exp. 1	4–6 yrs	Object	Addition	% of participants	96
Brush (1978), Exp. 1	4–6 yrs	Object	Subtraction	% of participants	96
Brush (1978), Exp. 2	5 yrs	Object	Addition	% of participants	100
Brush (1978), Exp. 2	5 yrs	Object	Subtraction	% of participants	100
Sophian and McCorgray (1994)	5 yrs	Verbal	Addition	% of trials	45
Sophian and McCorgray (1994)	5 yrs	Verbal	Subtraction	% of trials	65
Brush (1978), Exp. 2	6 yrs	Object	Addition	% of participants	100
Brush (1978), Exp. 2	6 yrs	Object	Subtraction	% of participants	100
Sophian and McCorgray (1994)	6 yrs	Verbal	Addition	% of trials	70
Sophian and McCorgray (1994)	6 yrs	Verbal	Subtraction	% of trials	80

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

Table 2b
Studies of relationship to operands using evaluation of examples.

Study	Age group	Context	Operation	DV	Outcome (success)
Wynn (1992)	5 mos	Object	Addition	Looking time	Yes
Wynn (1992)	5 mos	Object	Subtraction	Looking time	Yes
Simon et al. (1995)	5 mos	Object	Addition	Looking time	Yes
Cohen and Marks (2002)	5 mos	Object	Addition	Looking time	No
Feigenson et al. (2002)	7 mos	Object	Addition	Looking time	Yes
Chiang and Wynn (2000)	8 mos	Object	Addition	Looking time	Yes
Uller et al. (1999)	8 mos	Object	Addition	Looking time	No
McCrink and Wynn (2004)	9 mos	Object	Addition	Looking time	Yes
Prather and Alibali (2007)	7–8;5 yrs	Symbolic	Addition	Ratings	No
Prather and Alibali (2007)	7–8;5 yrs	Verbal	Addition	Ratings	No
Prather and Alibali (2007)	7–8;5 yrs	Symbolic	Subtraction	Ratings	No
Prather and Alibali (2007)	7–8;5 yrs	Verbal	Subtraction	Ratings	No
Dixon et al. (2001)	14 yrs	Symbolic	Addition	Rating	Yes
Dixon et al. (2001)	14 yrs	Symbolic	Subtraction	Rating	No
Dixon et al. (2001)	14 yrs	Symbolic	Multiplication	Rating	Yes
Dixon et al. (2001)	14 yrs	Symbolic	Division	Rating	No
Prather and Alibali (2007)	Adults	Verbal	Addition	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Addition	Rating	Yes
Prather and Alibali (2007)	Adults	Symbolic	Addition	Rating	Yes
Prather and Alibali (2008b)	Adults	Symbolic	Addition	Rating	Yes
Prather and Alibali (2008a)	Adults	Symbolic	Addition—negative numbers	Rating	No
Prather and Alibali (2007)	Adults	Verbal	Subtraction	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Subtraction	Rating	Yes
Prather and Alibali (2007)	Adults	Symbolic	Subtraction	Rating	No
Prather and Alibali (2008b)	Adults	Symbolic	Subtraction	Rating	No
Prather and Alibali (2008a)	Adults	Symbolic	Subtraction—negative numbers	Rating	No
Dixon et al. (2001)	Adults	Symbolic	Multiplication	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Division	Rating	No

Table 2c
Studies of relationship to operands using explicit knowledge assessments.

Study	Age group	Context	Operation	DV	Outcome (success)
Dixon et al. (2001)	Adults	Symbolic	Addition	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Subtraction	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Multiplication	Rating	Yes
Dixon et al. (2001)	Adults	Symbolic	Division	Rating	Yes

Each row corresponds to one reported result within a study. DV is the dependent variable as reported by the author. Outcome success is determined by a significant result involving the dependent variable that implies principle knowledge.

In one study, simple arithmetic equations were conveyed in an object context using marbles placed into containers (Brush, 1978). Participants were shown two containers that had the same number of marbles. The experimenter then removed or placed marbles into one of the containers to convey simple subtraction or addition. Participants with knowledge of relation to operands are able to indicate which container then had more marbles.

In another study, young children's arithmetic knowledge was tested in an object context using a reaching task (Starkey, 1992). Children were shown a set number of objects that had been placed into a container. The experimenter then performed a transformation, either placing objects into the container or taking objects out of the container. The participant was then asked to remove the objects from the container. Children's expectations about the number of objects were revealed by the number of times they reached. A child with knowledge of relation to operands should reach less often if objects were taken out and more often if objects were placed into the container.

Arithmetic transformations can also be conveyed in verbal contexts. For example, Sophian and McCorgray (1994) showed participants sets of toys that corresponded to story scenarios such as:

- *One morning Mickey's and Raggedy's bunnies decided to have a party. At first Raggedy's five bunnies were there, then Mickey's three bunnies came. How many bunnies were at the party after that?*

Young learners are often unable to provide exactly correct answers to these types of problems. However, some arithmetic knowledge may be inferred on the basis of their incorrect solutions. Learners who have knowledge of relation to operands should consistently indicate that addition results in more objects while subtraction results in fewer objects.

Knowledge of relation to operands based on evaluation of examples

Studies that use evaluation of examples often test very young participants, including infants. There is a great deal of interest in characterizing the numerical knowledge of infants (see Mix, Huttenlocher, & Levine, 2002, for a review), and some of this research investigates the possibility of simple arithmetic knowledge in infants. Perhaps the most well-known of these is Wynn's (1992) assessment of infants' arithmetic knowledge. In this paradigm infants viewed one Mickey Mouse doll that was then covered with an occluder. Another doll appeared and was placed behind the occluder. The occluder was then dropped to reveal either two dolls (correct) or one doll (magic). Infants' looking times were measured to evaluate their expectations about each event. The "magic" trials represent a violation of the relation to operands principle for addition, because adding objects to a set should result in more objects, not fewer or the same amount. This display was repeated in variations that corresponded to $1 + 1 = 2$, $1 + 1 = 1$, $2 - 1 = 1$ and $2 - 1 = 2$. This can be seen as a very basic test of relation to operands for addition and subtraction, using evaluation of examples as evidence.

A myriad of infant studies, most using a similar paradigm, have been conducted to respond to Wynn's original claim that 5-month-olds understand simple arithmetic. Many of these subsequent studies have also provided evidence that infants have knowledge of relation to operands (Chiang & Wynn, 2000; McCrink & Wynn, 2004; Simon, Hespos, & Rochat, 1995). However, as with many areas of infant research, there are differing perspectives on what sorts of cognitive abilities underlie infants' behavior. A major issue is that non-numerical factors make infants' behavior in these experiments difficult to interpret (see Mix et al., 2002). For example, Cohen and Marks (2002) report that familiarization with the display and objects used prior to test affects infants' behavior, suggesting that their reactions to the arithmetic display may be due to their familiarity with the objects, and not arithmetic knowledge. Feigenson, Carey, and Spelke (2002) suggest that infants' behavior is driven by violations in object expectations based on overall surface area, not number per se. The effect is also vulnerable to slight changes in methodology, such as the timing of the presentation of stimuli (Uller, Carey, Hunter-Fenner, & Klatt, 1999). Thus, the issue of whether arithmetic principle knowledge drives infants' looking behavior in this paradigm is still unresolved.

It should be noted that although infants' behavior in this paradigm may be consistent with relation to operands, the most informative comparisons are not made in most infant studies. Infants' reactions to correct equations are compared to incorrect equations, but the crucial comparison of violation and (incorrect) non-violation equations (e.g., $1 + 1 = 1$ vs. $1 + 1 = 3$) is usually not made. Thus, the evidence for infants' knowledge of the relation to operands principle is not as strong or as clear-cut as it could be.

Evaluation of examples assessments have also been used to investigate older learners' knowledge of relation to operands (Dixon et al., 2001; Prather & Alibali, 2008a, 2008b). All three of these studies utilized an equation set rating task to assess participants' knowledge. Participants were shown sets of symbolic equations that had been produced by hypothetical students, and were asked to rate each student's attempt at arithmetic. The equation sets were constructed so that all equations were incorrect; however, half of the sets contained one or more incorrect equations that violated the relation to operands principle. Participants who have knowledge of the principle should rate sets that contain violations lower than non-violation sets. Dixon and colleagues (2001) investigated eighth-grade students' (mean age = 14 years) and adults' knowledge in a symbolic context with addition, subtraction, division and multiplication. Prather and Alibali (2008a) expanded on this work to include symbolic equations that contain negative numbers, the idea being that principle knowledge for arithmetic with positive numbers may not generalize to arithmetic with negative numbers.

Knowledge of relation to operands based on explicit recognition

Participants' knowledge can also be evaluated using more explicit knowledge assessments. Dixon et al. (2001) gave participants statements about arithmetic operations and asked them to rate how likely it was that each statement referred to one of the four basic operations. The statements rated by the participants pertained to the relation of the operands and the result in arithmetic equations. For example, participants read the statement "The answer is always larger than the first number in the problem". This is true for addition and multiplication, but not for division or subtraction. Thus, participants who have knowledge of relation to operands for all four operations should respond "very likely" for addition and multiplication and "very unlikely" for division and subtraction.

Comparing assessment types

For the relation to operands family of principles, there is little overlap in the assessment types used with different age groups of learners. Studies of infants typically use evaluation of examples, with looking time as the outcome measure. Studies of younger children (18 months–6 years) tend to utilize application of procedures. Studies of older children are few and far between, but those that have been done utilize evaluation of examples, with ratings as the outcome measure. Studies of adult participants also largely use evaluation of examples, with ratings as the outcome measure.

There has been only one within-study contrast of assessment types (Dixon et al., 2001), and it involved adult participants. In this study, learners displayed knowledge when assessed via explicit recognition; however, they did not display knowledge when assessed via evaluation of examples.

Knowledge of relation to operands across contexts

Only one study of relation to operands has compared multiple contexts (Prather & Alibali, 2007). In this study, adult participants showed more knowledge of relation to operands for subtraction in a verbal context than in a symbolic context. Children (ages 7–8;5 years) did not display knowledge of the principle in either the verbal or symbolic context.

Cross-study comparisons are difficult because different contexts tend to be used at different ages. Object contexts tend to be used more often with younger children, and symbolic contexts more often with older children. Brush (1978) reports that 4–6 year olds use principle-consistent procedures very consistently (over 90% of trials) in an object context. Sophian and McCorgray (1994) report less optimistic results with similar age groups (4-, 5-, and 6-year-olds) in a verbal context, with better performance at age 6 than at age 4. Taken together, these findings suggest that knowledge of relation to operands may emerge first in object contexts, then in verbal contexts, and lastly in symbolic contexts (see Table 2a). This provides some empirical support for Resnick's (1992) theory of general arithmetic development; however, within-study comparisons are still needed.

When does knowledge of relation to operands emerge?

Though a large number of studies have investigated relation to operands for addition and subtraction in infant learners, the conclusions that can be drawn about infants' knowledge are anything but clear. Based on participants' evaluation of examples using looking time, Wynn (1992) concluded that 5-month-old infants' expectations about adding and subtracting objects were consistent with relation to operands. Several subsequent studies supported this conclusion (Chiang & Wynn, 2000; Simon et al., 1995). However, other studies have suggested that infants' behavior is not driven by arithmetic principle knowledge (Cohen & Marks, 2002; Feigenson et al., 2002; Uller et al., 1999). As yet, there is no consensus about exactly what sort of knowledge drives infants' looking behavior.

A study utilizing application of procedures as evidence for knowledge has reported behavior consistent with relation to operands for addition and subtraction in learners as young as 18 months (Starkey, 1992). The children were shown arithmetic transformations using balls and a container, and their expectations about the number of objects were inferred by the number of times they reached. Starkey did not test learners younger than 18 months, so it is unclear whether slightly younger learners would be successful using this type of method.

Studies of older learners have tested knowledge of relation to operands across the four basic operations. These studies indicate that by 14 years, adolescents have knowledge of relation to operands for addition and multiplication, but not subtraction or division (Dixon et al., 2001). Evidence about adults' knowledge of relation to operands for subtraction is mixed, with some studies showing positive effects (Dixon et al., 2001) and others not showing such effects (Prather & Alibali, 2007, 2008a, 2008b). The general pattern suggests that experience with operations leads to knowledge of the relation to operands principle.

Directions for future research

The literature on relation to operands favors certain contexts and assessment types, making questions about the effects of context and assessment type difficult to answer. The majority of studies of relation to operands principles use evaluation of examples in an object context, due to the very young ages of the participants in many of the studies. Only one of the studies reviewed used multiple contexts (Prather & Alibali, 2007), and only a handful of studies (Dixon et al., 2001; Prather & Alibali, 2007, 2008a, 2008b; Sophian & McCorgray, 1994) used any context other than object. Only one study contrasted assessment types (Dixon et al., 2001).

As noted above, there is little overlap in the assessment types used with different age groups of learners. There is a lively debate about what knowledge underlies infants' performance on evaluation of examples tasks. However, it would also be informative to compare, for example, young children's performance on application of procedures, evaluation of examples and explicit recognition.

Inversion

The *inversion* principle is based on the inverse relation between operations. For example, addition and subtraction are inverse operations, as are multiplication and division. The inversion principle holds that inverse operations involving the same value result in no net change, $a + b - b = a$, and $a \times b \div b = a$. Individual studies of the inversion principle and their outcomes are summarized in Tables 3a and 3b.

Evidence of inversion knowledge based on application of procedures

The vast majority of studies of inversion use application of procedures as evidence of knowledge. Most utilize tasks in which knowledge of the inversion principle can be used as a shortcut. Consider two types of three-term arithmetic equations: $a + b - c = x$ and $a + b - b = x$ (where x is an unknown). Learners could solve either equation using a variety of procedures. The second equation can be solved using a "shortcut" procedure; the addition and subtraction of b can be "cancelled out", leaving $a = x$. For many learners (especially younger ones), this procedure can be used relatively quickly and accurately as compared to other procedures such as computation. A few studies utilize self-report to determine whether participants use the inversion shortcut (e.g., Robinson & Dubé, 2009a, 2009b). However, in most studies, knowledge of inversion is inferred if learners solve $a + b - b = x$ equations faster and/or more accurately than $a + b - c = x$ equations (see Table 3c).

Many studies rely on differences in accuracy on inversion and standard equations to determine whether participants use the shortcut procedure (Brush, 1978; Bryant, Christie, & Rendu, 1999; Gilmore & Bryant, 2006, 2008; Gilmore & Spelke, 2008; Nunes, Bryant, Hallett, Bell, & Evans, 2009;

Table 3a
Studies of inversion using application of procedures.

Study	Age group	Context	Operation	DV	Success
Sherman and Bisanz (2007)	3 yrs	Object	Addition/subtraction	Accuracy	Yes
Baroody, Lai, Li, and Baroody (2009)	3 yrs	Object	Addition/subtraction	Response pattern	0%
Klein and Bisanz (2000)	4 yrs	Object	Addition/subtraction	Accuracy	No
Klein and Bisanz (2000)	4 yrs (subset)	Object	Addition/subtraction	Reaction time	Yes
Rasmussen et al. (2003)	4 yrs	Object	Addition/subtraction	Accuracy	Yes
Canobi and Bethune (2008)	4 yrs	Object	Addition/subtraction	Accuracy	67%[a]
Baroody and Lai (2007)	4 yrs	Object	Addition/subtraction	Response pattern	6%
Baroody et al. (2009)	4 yrs	Object	Addition/subtraction	Response pattern	13%
Brush (1978), Exp. 2	4 yrs	Object	Addition/subtraction $a + b - b$	% of participants	35%
Brush (1978), Exp. 2	4 yrs	Object	Addition/subtraction $a - b + b$	% of participants	65%
Brush (1978), Exp. 1	4–6 yrs	Object	Addition/subtraction $a + b - b$	Accuracy	54%
Brush (1978), Exp. 1	4–6 yrs	Object	Addition/subtraction $a - b + b$	Accuracy	77%
Gilmore and Spelke (2008)	5 yrs	Object	Addition/subtraction	Accuracy	Yes
Gilmore and Spelke (2008)	5 yrs	Verbal	Addition/subtraction	Accuracy	Yes
Canobi and Bethune (2008)	5 yrs	Object	Addition/subtraction	Accuracy	76%[a]
Bryant et al. (1999)	5 yrs	Object	Addition/subtraction	Accuracy	Yes
Bryant et al. (1999)	5 yrs	Verbal	Addition/subtraction	Accuracy	Yes
Bryant et al. (1999)	5 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Baroody and Lai (2007)	5 yrs	Object	Addition/subtraction	Response pattern	25%
Baroody et al. (2009)	5 yrs	Object	Addition/subtraction	Response pattern	56%
Brush (1978), Exp. 2	5 yrs	Object	Addition/subtraction $a + b - b$	% of participants	70%
Brush (1978), Exp. 2	5 yrs	Object	Addition/subtraction $a - b + b$	% of participants	80%
Bryant et al. (1999)	6 yrs	Object	Addition/subtraction	Accuracy	Yes
Bryant et al. (1999)	6 yrs	Verbal	Addition/subtraction	Accuracy	Yes
Bryant et al. (1999)	6 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Baroody and Lai (2007)	6 yrs	Object	Addition/subtraction	Response pattern	37%
Baroody et al. (2009)	6 yrs	Object	Addition/subtraction	Response pattern	63%
Brush (1978), Exp. 2	6 yrs	Object	Addition/subtraction $a + b - b$	% of participants	90%
Brush (1978), Exp. 2	6 yrs	Object	Addition/subtraction $a - b + b$	% of participants	90%
Canobi (2005)	5–7yrs	Object	Addition/subtraction	Accuracy	Yes
Canobi (2005)	5–7yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Canobi (2005)	5–7yrs	Object	Addition/subtraction	Reaction time	Yes
Canobi (2005)	5–7yrs	Symbolic	Addition/subtraction	Reaction time	Yes
Gilmore and Bryant (2006), Exp. 2	6–7 yrs	Object	Addition/subtraction $x + b - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	6–7 yrs	Object	Addition/subtraction $a + x - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	6–7 yrs	Object	Addition/subtraction $a + b - x = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	6–7 yrs	Object	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Rasmussen et al. (2003)	6–7 yrs	Object	Addition/subtraction	Accuracy	Yes
Baroody et al. (2009)	7 yrs	Object	Addition/subtraction	Response pattern	100%
Siegler and Stern (1998)	8 yrs	Symbolic	Addition/subtraction	Verbal report—% of participants	21%[b]
Nunes et al. (2009)	8 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Gilmore and Bryant (2008)	8 yrs	Symbolic	Addition/subtraction $x + b - b = a$	Accuracy	Yes

(*continued on next page*)

Table 3a (*continued*)

Study	Age group	Context	Operation	DV	Success
Gilmore and Bryant (2008)	8 yrs	Symbolic	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Gilmore and Bryant (2008)	8 yrs	Symbolic	Addition/subtraction $b - b + x = a$	Accuracy	Yes
Gilmore and Bryant (2008)	8 yrs	Symbolic	Addition/subtraction $b - b + a = x$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	8–9 yrs	Object	Addition/subtraction $x + b - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	8–9 yrs	Object	Addition/subtraction $a + x - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	8–9 yrs	Object	Addition/subtraction $a + b - x = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 2	8–9 yrs	Object	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Object	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Object	Addition/subtraction $x + b - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Verbal	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Verbal	Addition/subtraction $x + b - b = a$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Symbolic	Addition/subtraction $a + b - b = x$	Accuracy	Yes
Gilmore and Bryant (2006), Exp. 1	8–9 yrs	Symbolic	Addition/subtraction $x + b - b = a$	Accuracy	Yes
Gilmore (2006)	9 yrs	Verbal	Addition/subtraction	Accuracy	Yes
Gilmore (2006)	9 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Robinson et al. (2006)	11 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Robinson et al. (2006)	11 yrs	Symbolic	Addition/subtraction	Reaction time	Yes
Robinson et al. (2006)	11 yrs	Symbolic	Addition/subtraction	Verbal report—% of trials	44%
Robinson et al. (2006)	11 yrs	Symbolic	Multiplication/division	Accuracy	Yes
Robinson et al. (2006)	11 yrs	Symbolic	Multiplication/division	Reaction time	Yes
Robinson et al. (2006)	11 yrs	Symbolic	Multiplication/division	Verbal report—% of trials	19%
Robinson and Dubé (2009a)	6th grade	Symbolic	Multiplication/division	Verbal report—% of trials	18%
Robinson and Dubé (2009a)	7th grade	Symbolic	Multiplication/division	Verbal report—% of trials	14%
Robinson and Dubé (2009a)	8th grade	Symbolic	Multiplication/division	Verbal report—% of trials	17%
Robinson et al. (2006)	13 yrs	Symbolic	Addition/subtraction	Accuracy	Yes
Robinson et al. (2006)	13 yrs	Symbolic	Addition/subtraction	Reaction time	Yes
Robinson et al. (2006)	13 yrs	Symbolic	Addition/subtraction	Verbal report—% of trials	60%
Robinson et al. (2006)	13 yrs	Symbolic	Multiplication/division	Accuracy	Yes
Robinson et al. (2006)	13 yrs	Symbolic	Multiplication/division	Reaction time	Yes
Robinson et al. (2006)	13 yrs	Symbolic	Multiplication/division	Verbal report—% of trials	39%
Robinson and Ninowski (2003)	Adult	Symbolic	Addition/subtraction	Reaction time	Yes
Robinson and Ninowski (2003)	Adult	Symbolic	Multiplication/division	Reaction time	Yes

[a] Averaged across abstract, large, and small.
[b] The value reported is the proportion of participants who used the shortcut procedure on at least one trial at the outset (session 1) of a multi-session study. These participants were excluded from the remainder of the study.

Table 3b
Studies of inversion using evaluation of examples.

Study	Age group	Context	Operation	DV	Success
Vilette (2002)	2 yrs	Object	Addition/subtraction	Error rate	No
Vilette (2002)	3 yrs	Object	Addition/subtraction	Error rate	No
Vilette (2002)	4 yrs	Object	Addition/subtraction	Error rate	Yes

Table 3c
Studies of inversion using evaluation of procedures.

Study	Age group	Context	Operation	DV	Success
Robinson and Dubé (2009a)	6th grade	Symbolic	Multiplication/division	Approval of shortcut	83%
Robinson and Dubé (2009a)	7th grade	Symbolic	Multiplication/division	Approval of shortcut	73%
Robinson and Dubé (2009a)	8th grade	Symbolic	Multiplication/division	Approval of shortcut	93%

Each row corresponds to one reported result within a study. DV is the dependent variable as reported by the author. Outcome success is determined by a significant result involving the dependent variable that implies principle knowledge.

Rasmussen, Ho, & Bisanz, 2003; Sherman & Bisanz, 2007). Given the relatively young age of the participants in most of these studies, solving a three term equation is not trivial; thus, participants are motivated to apply any procedure that would decrease the difficulty of these equations. If participants have knowledge of inversion, they should use the shortcut procedure on the inversion equations, resulting in more accurate performance.

Other studies have included participants' reaction times in the analysis (Canobi, 2005; Klein & Bisanz, 2000; Robinson & Ninowski, 2003; Robinson, Ninowski, & Gray, 2006; Siegler & Stern, 1998). The idea is that learners who use the inversion shortcut should have faster reaction times on inversion equations as compared to standard equations, for which computation is required.

Inversion knowledge based on evaluation of procedures

In addition to measuring what procedures learners themselves use, it is also possible to have them evaluate procedures used by others. Robinson and Dubé (2009a) presented different types of procedures for solving three-term division and multiplication problems, and asked participants whether they approved of each procedure. Positive evaluations of the shortcut procedure were interpreted as evidence for knowledge of the inversion principle.

Inversion knowledge based on evaluation of examples

A less common method of evaluating inversion knowledge is evaluation of examples. In one study (Vilette, 2002), 2-, 3-, and 4-year-old children were shown three types of equations, $2 + 1$, $3 - 1$, and $2 + 1 - 1$ using a variation of the Mickey Mouse procedure (Wynn, 1992). Each equation was displayed either with the correct answer or with an incorrect answer. Children were asked to evaluate the answer by responding to the question, "Is that normal?". Children were not required to perform calculations to successfully answer the question. Knowledge of the inversion principle should allow children to perform accurately at evaluating the $2 + 1 - 1$ equations.

Comparing assessment types

There is little variation in the assessment types used in investigations of inversion knowledge. All but two of the studies reviewed used application of procedures as evidence for knowledge of inversion.

Both application of procedures and evaluation of examples have been used in research with 4-year-olds, allowing some comparison across assessment types in this age group. Using evaluation of examples in an object context, Villete (2002) found that the vast majority of 4-year-olds (20 of 22) correctly stated both that $2 + 1 - 1 = 2$ was "normal" and that $2 + 1 - 1 = 3$ was not. Two other studies used application of procedures as evidence with participants in the same age range (Klein & Bisanz, 2000; Rasmussen et al., 2003). The results of these studies were somewhat mixed. Rasmussen et al. (2003) reported that 4-year-olds performed significantly better on inversion equations than standard equations in an object context. Klein and Bisanz (2000) found that overall, 4-year-olds showed no difference in accuracy (41% and 49%) or response time on standard and inversion equations. However, for the subset of participants who had sufficient data to analyze (i.e., participants who did not use overt calculation on at least two problems of each type), participants solved inversion equations faster than standard equations, suggesting use of the shortcut procedure on inversion problems. Taken together, these studies suggest that 4-year-olds may have emerging knowledge of inversion. Further, this knowledge may be more readily displayed using evaluation of examples than application of procedures.

Comparing multiple contexts

A few studies have directly compared learners' knowledge of inversion across object, verbal and symbolic contexts (Bryant et al., 1999; Canobi, 2005; Gilmore, 2006; Gilmore & Bryant, 2006; Gilmore & Spelke, 2008). None of these studies report any differences in learners' inversion knowledge in the tested contexts. However, it should be noted that learners in these studies were older than the age at which inversion is suggested to first emerge (around 3–4 years; see below). Thus, it seems that these learners may have reached a point where their knowledge is similar across contexts. Of the remaining studies, there are no available comparisons that address whether learners at earlier points in development can generalize their knowledge of inversion across contexts, or whether knowledge of the principle emerges in one context before others.

Gilmore and Panadatou-Pastou (2009) conducted a meta-analytic review of studies of the inversion effect (i.e., the size of the advantage for inversion problems over control problems). They found that age did not moderate the size of the effect, but context did so. The inversion effect was greatest for pictorial presentation of the items, moderate for verbal and concrete presentation, and smallest for symbolic presentation. This pattern is in line with the context effects predicted by Resnick's model; however it does not directly address whether knowledge of inversion emerges in one context before others.

When does knowledge of inversion emerge?

The earliest evidence of knowledge of the inversion principle is in 3-year-olds in an object context (Sherman & Bisanz, 2007). In this study, preschool children (3;0 to 3;11 years) performed more accurately on inversion equations than on standard equations. At an individual level, 83% of the participants were more accurate on the inversion items than on the standard items. Thus, this study suggests that knowledge of inversion with addition and subtraction begins to emerge by around 3 years. However, other studies in which knowledge is assessed at an individual level have suggested that very few preschoolers demonstrate understanding of inversion (e.g., only 1 of 16 four-year-olds in the study by Baroody and Lai (2007); see Bisanz et al. (2009), for discussion of this issue). Differences across studies may hinge on the particular assessment items used, or on the stringency of the criteria for demonstrating understanding.

Evidence about inversion with multiplication and division is fairly sparse. The youngest participants who have been tested with multiplication and division are 11-year-olds, and children of this age do demonstrate knowledge of the principle, both in application of procedures assessments (Robinson et al., 2006) and evaluation of procedures (Robinson & Dubé, 2009a).

Directions for future research

The inversion literature utilizes almost exclusively application of procedures as evidence for knowledge of the principle. Only one study to date has used evaluation of procedures, and one has used evaluation of examples. No studies have utilized justification of procedures or explicit recognition. Since the vast majority of studies take the same approach to evaluating knowledge, this literature may present an incomplete characterization of learners' knowledge. It is unclear whether learners who show knowledge in application of procedure studies would do better, worse, or equally well on a different type of knowledge assessment.

In terms of context, there is some variation, including at least three within-study comparisons (Bryant et al., 1999; Gilmore, 2006; Gilmore & Bryant, 2006). The remaining studies are split between object context and symbolic context. The studies that utilize an object context typically include much younger participants than those that utilize a symbolic context. This makes comparing results in regards to possible effects of context difficult. Given that the earliest evidence for inversion knowledge is from 3-year-olds, learners of this age would be ideal for investigating of the effects of assessment type and context on knowledge.

Remaining issues

The studies reviewed in this paper do not follow any one specific standard for assessing knowledge. While most studies present well thought out methodologies, they vary in terms of knowledge assessment type and the context in which arithmetic is presented. Though this may be sufficient to draw conclusions in individual studies, it makes progress difficult for the field at large, because findings across studies are often not directly comparable. Without a clear view of the relations between different types of knowledge assessments and contexts, it is difficult to paint a comprehensive picture of the acquisition of principle knowledge.

This review has shown that assessment type and context both affect performance on tasks designed to assess principle knowledge, so they must be taken into account in any complete account of the acquisition of principle knowledge. It is certainly not realistic to expect every study to simultaneously take on all these issues. However, it seems that the effects of knowledge assessment type and context have largely been ignored in the arithmetic principle literature.

The evidence reviewed herein indicates that the type of assessment used can influence whether principle understanding is displayed. In general, tasks that involve evaluation of examples tend to reveal greater understanding than tasks that involve application of procedures. Further, among tasks that involve application of procedures, tasks constructed specifically to diagnose principle knowledge tend to reveal greater understanding than tasks typical of the domain. Thus, conclusions about what learners of various ages "know" hinge in part on the knowledge assessments used. Few studies utilize more than one measure of principle knowledge, and there are very few within-study comparisons of different assessment types.

The evidence reviewed herein also indicates that context can affect principle understanding and whether it is displayed. Some evaluations of context effects require integrating findings across studies (e.g., effects of context on commutativity knowledge in 6-year-olds, reviewed above). Unfortunately, such comparisons are often problematic because of differences in methods, criteria for success, and statistical summaries used across studies. There are a handful of within-study comparisons of different contexts. Some of these within-study comparisons show null effects (e.g., Canobi et al., 2003); this may occur because the participants are old enough to have mastered the principle in all contexts. Other within-study comparisons show significant effects of context (e.g., Cowan & Renton, 1996). Resnick's (1992) theory holds that arithmetic understanding should emerge first in an object context, then in a verbal context, then in a symbolic context, and finally in an abstract context. Existing data on the acquisition of principle knowledge across contexts are generally in line with this theory. However, because few studies have been designed to directly assess the effects of context on principle understanding, this is an important arena for future work.

The place of multifaceted knowledge assessments

One possible approach to addressing these issues is the use of multifaceted knowledge assessments, or *knowledge profiles*, which summarize individual participants' performance across multiple knowledge assessments (see Bisanz et al., 2009, for discussion). Knowledge profiles provide a more comprehensive picture of learners' knowledge and how it changes than can be obtained from a mix of single-faceted knowledge assessments drawn from different studies. Such profiles are valuable because principle knowledge is not something that learners either "have" or "don't have". Instead, learners' knowledge of arithmetic principles appears to be graded, in the sense that it is displayed in some contexts and through some types of knowledge assessments but not others. The value of knowledge profiles is underscored by evidence that knowledge is graded in domains other than arithmetic (e.g., Munakata & Yerys, 2001; Zelazo, Frye, & Rapus, 1996). For example, explicit assessments of learners' rule learning in a card-sorting task yield very different results than application-of-procedures assessments of that same rule.

Multi-faceted knowledge assessments are more common—though by no means common—in the literature on arithmetic problem solving, as compared to the literature on principle knowledge. As one example, Jordan et al. (1992) tested calculation abilities in middle and low-income children in problem solving in both an object context and in three different verbal contexts. They found that middle-income children outperformed low-income children in all of the verbal contexts, but that the two groups performed similarly in the object context. By comparing multi-faceted knowledge assessments across two different groups of learners, these investigators gained traction on the issue of how environmental input contributes to the development of calculation abilities.

It is possible that very different conclusions would be drawn about the acquisition of principle knowledge if multi-faceted assessments were the norm. For example, for relation to operands, different types of single-faceted knowledge assessments tend to be used at different ages, making it impossible to draw firm conclusions about when knowledge of the principle emerges. Very different conclusions are warranted if one considers only evaluation of examples or only application of procedures (as early as 1 year vs. as late as 7 years). If both types of assessments were used with the same participants, one might be less likely to conclude that younger children ‘‘have" knowledge of the principle, even if they ‘‘pass" the evaluation of examples tasks.

In the literature on arithmetic principles, there are a few notable exceptions to the preponderance of single-faceted assessments. One is Canobi's (2005) use of knowledge profiles in investigating knowledge of commutativity and inversion as well as other principles relevant to addition and subtraction. Canobi evaluated participants' principle knowledge using different knowledge assessments (application of procedures and evaluation of procedures) and in different contexts (object and symbolic), and she also examined participants' performance on other measures of arithmetic performance. This allowed for conclusions that go beyond ‘‘what 8-year-olds know" to begin to address the nature of principle knowledge and its relation to arithmetic knowledge in general. The findings revealed important individual differences in how children responded to concrete materials, with one subset of the children recognizing concepts with reference to objects, and another subset doing so with objects absent. These individual differences in *patterns* of knowledge could only be revealed with a multi-faceted knowledge assessment.

A second study that utilizes a multifaceted knowledge assessment is Robinson and Dubé's (2009a) study of inversion with multiplication and division. In this study, data from two types of knowledge assessments allowed for a more nuanced view of participants' knowledge. A subset of the participants did not themselves use a procedure that suggested knowledge of the principle, but were able to note that the procedure was appropriate and explain why it could be used. This response pattern would not have been revealed if only one type of knowledge assessment had been used.

Thus, we endorse knowledge profiles as a means to gain leverage on variations in performance across contexts and knowledge assessment types. However, we are left with an outstanding issue: what do variations in performance tell us about the nature of learners' knowledge? The fact that learners' knowledge of principles varies as a function of context and type of assessment seems to imply that principle knowledge is not general or abstract. However, it also seems inaccurate to argue that principle knowledge is exemplar-based. We suggest that principle knowledge may initially be represented at an intermediate level of abstraction. Such knowledge is sufficiently abstract to generalize across problems *within a given context*, but not so abstract that it is generalized immediately *across contexts*. For example, a learner may have abstract, generalizable knowledge of arithmetic principles within an object context, but this knowledge may not generalize to a symbolic context, leading to different behavior on assessments that involve problems presented in a symbolic context.

It is worth noting that the principle knowledge used in a symbolic context is not necessarily more abstract than that used in a verbal or object context. Learners could show knowledge in a symbolic context, but fail to show it in other contexts (Lawler, 1981). However, given the studies reviewed in this paper, and in line with the theoretical accounts offered by Piaget, Resnick (1992) and others, we expect that in most cases, learners who use principle knowledge in a symbolic context will display that knowledge in verbal and concrete contexts as well. If there are learners for whom this is not the case, it would suggest that there may be multiple possible developmental trajectories in the acquisition of principle knowledge across contexts. Such a finding would compel further theoretical work on mechanisms of change in principle knowledge, because any proposed mechanisms would need to explain the full range of observed developmental trajectories.

We hypothesize that, at least initially, principle knowledge may be tied to the contexts in which it is learned. For example, if a learner has detected regularities about adding and subtracting objects, that learner may not immediately transfer this knowledge to adding and subtracting numbers—the relevant principles may need to be relearned in the symbolic context. Future research will need to address this possibility. Future studies could also examine whether there are "savings" in learning a principle in a new context, if that principle is already understood in another context.

We further suggest that principle knowledge may be represented more or less strongly, perhaps as a function of the amount of experience a learner has with the regularity in question. Different types of knowledge assessments may have different requirements regarding the strength of knowledge needed for success (see, e.g., Munakata, McClelland, Johnson, & Siegler, 1997). For example, a learner may have an internal representation of the principle that is sufficiently strong to support accurate performance on a recognition task, but that is too weak to support performance on an application-of-procedures task. Applying a procedure based on the knowledge may require a stronger internal representation because the response is more complex.

More broadly, what is needed to "show" knowledge on any particular assessment is never simply the knowledge in and of itself. Other factors come into play, such as memory limitations, task interpretations and the availability of competing procedures. It seems that knowledge assessments with fewer requirements might be more "pure" assessments than assessments with many requirements. However, there are important differences between having knowledge and using that knowledge. Knowledge assessments that also make other demands, such as demands on short-term memory, may be more accurate in predicting the practical use of that knowledge.

Issues in treatment of data

Integrating across studies is also difficult because different types of statistical summaries are used, even across studies that utilize the same assessment type. Many studies report the percentage of trials on which participants' behavior matches some predetermined criterion. Other studies report the percentage of participants who behave a certain way all the time, most of the time, or even only once. Thus, not only the type of knowledge assessment, but also the treatment of data can vary across studies, making it difficult to integrate findings. Examples from studies that do report multiple statistical summaries illustrate this point. In once such case (Canobi et al., 2002) the percentage of participants who used a procedure (96%) was very different from the percentage of trials on which it was used (53%).

Statistical summaries that aggregate across participants have the potential to detect small changes in overall behavior. If each participant's behavior changes only slightly, an overall effect may be detected only when aggregating across many participants. However, conclusions about the behavior of participants as a group may not accurately reflect each individual's performance. A subset of participants may drive the effect, leading to an overestimation of participants' overall knowledge (see Baroody & Lai, 2007, for discussion of this point). For example, in one study of 8-year-olds' understanding of inversion, group level analyses revealed that children performed better on inversion problems than on control problems, implying knowledge of inversion in this age group. However, at the individual level, fully 35% of the participants did not show this pattern (Nunes et al., 2009).

Broadly speaking, the appropriateness of any particular statistical summary depends on the nature of the specific research question being addressed. If the aim is to understand patterns of normative development, group level analyses may be most appropriate, but if the aim is to understand processes of change, individual level analyses will be required.

Mechanisms of change in the development of arithmetic principle knowledge

A key question in developmental research is, what do learners at various points in development know? As illustrated in this review, a good deal of arithmetic principle research has addressed this

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

question. An equally important question is, what mechanisms underlie learners' progression? In the case of arithmetic principles, the answer to this question is much less clear.

A small but growing number of studies have sought to explicitly address mechanisms of change in principle knowledge. These include three studies of inversion (Lai, Baroody, & Johnson, 2008; Robinson & Dubé, 2009b; Siegler & Stern, 1998), one of commutativity (Canobi, 2009), one of relation to operands (Prather & Alibali, 2008b), and one of the arithmetic principle direction of effect (Dixon & Bangert, 2005). All of these studies investigate changes in learners' knowledge as a function of experience with arithmetic operations (either in a symbolic context or with objects). In each case, particular sorts of experiences are hypothesized to lead to gains in principle knowledge.

Four of these studies have suggested that the *density of relevant experiences* may influence potential learning. When relevant experiences are presented closely together rather than distributed over time, participants show greater learning of the target principle (Dixon & Bangert, 2005; Robinson & Dubé, 2009; Prather & Alibali, 2008b; Siegler & Stern, 1998). For example, Siegler and Stern (1998) showed that participants began to apply the inversion principle more quickly when they encountered many inversion problems blocked together than when they encountered inversion problems interspersed with control problems (e.g., $a + b - c = ?$).

Another training study has shown that *variations in problem sequencing* can influence principle learning. Canobi (2009) provided some children with a set of practice arithmetic problems that included commuted pairs presented in sequence (e.g., 3 + 6 was followed by 6 + 3); other children received the same set of practice problems, presented in a random sequence. At posttest, children who received the conceptually sequenced practice problems were more likely than children who received randomly sequenced problems to generate accurate commutativity explanations in a puppet judgment task. Thus, highlighting conceptual relations through appropriate problem sequencing led some children to "discover" or to make more explicit their understanding of commutativity.

One study has suggested that principle-relevant *actions* may promote principle understanding. In a training study, Lai et al. (2008) showed some children addition or subtraction operations on collections of objects, and asked the children to "repair" those collections to their initial state. Thus, children needed to perform operations that inverted the operations that they had just viewed. The majority of participants who were unsuccessful at pretest and who received this action-based training showed progress in their understanding of inversion at posttest (71%, compared to 21% in a control condition).

Finally, one study has suggested that *exposure to principle violations* may also promote principle learning (Prather & Alibali, 2008b). In this study, participants who were exposed to a mixture of principle-consistent equations and equations that violated the principle showed greater gains in principle understanding than participants who were exposed solely to principle-consistent examples.

The types of experiences investigated in these studies may be effective because they help learners improve their encoding of arithmetic equations (Prather & Alibali, 2008b). It seems that both solving conceptually sequenced practice problems and contrasting principle-consistent examples and violations may serve to highlight key features of the problems, and consequently lead to more accurate encoding of those features. Future work is needed to test the role of problem encoding as a possible mechanism in acquiring principle understanding.

These studies have begun to address the mechanisms of arithmetic principle learning; however, the overall picture is still very much unclear. A more complete understanding of how learners acquire principle knowledge will require further research on mechanisms of learning, as well as research that provides "snap shots" of knowledge at various ages. Future work on mechanisms of knowledge change should also bear in mind the potential implications of knowledge assessment types and context for learners' performance.

Implications for cognitive development more broadly

The issues addressed in this review are not unique to the development of arithmetic principle knowledge, or even to the domain of mathematics. Researchers across areas of cognitive development must grapple with the implications of variations in children's performance across knowledge assessments and contexts. These variations make describing knowledge more difficult for the researcher; however, the use of multifaceted knowledge assessments yields more accurate characterizations of

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001

learners' knowledge and behavior, and ultimately may lead to more productive theorizing about mechanisms of developmental change.

Some illustrative examples can be drawn from research on the development of theory of mind. Clements and Perner (1994) utilized two measures of false belief knowledge—verbal prediction and direction of eye gaze—drawn from the same task. They presented children with a standard false belief task, in which a story character hid an object in one location and then went away. While the character was gone, the object was moved to another location. The experimenter then announced that the character was about to return, and said, "I wonder where he's going to look". Three-year-old children often *looked* to the correct location, but *verbally reported* the incorrect location. Thus, children's eye gaze revealed more knowledge than their verbal reports. Based on these findings, Clements and Perner argued that children acquire an implicit understanding of false belief before they acquire an explicit, verbally stateable understanding. This perspective sets constraints on possible mechanisms of change in the development of theory of mind, because any proposed mechanism must yield implicit knowledge before explicit knowledge.

Building on this work, Ruffman, Garnham, Import, and Connolly (2001) classified children in terms of their performance on both looking and verbal measures—a "knowledge profile" of sorts. They then went on to examine the performance of groups of children with different knowledge profiles on a "betting" task, in which children could bet plastic counters on locations where they thought the story character would go. The betting task was intended as a measure of the certainty of their responses. The data revealed that, among younger children, those who "passed" the eye gaze measure but failed the verbal report were more certain of (i.e., bet more on) the *incorrect* solution than those who "passed" the eye gaze measure and "passed" the verbal report. Thus, children who performed similarly on one measure (eye gaze) but differently on another measure (verbal response) showed different patterns of behavior on another task. Thus, multi-faceted assessment of children's knowledge yielded information that helped predict their behavior on another task.

Other research on the development of theory of mind has sought to develop a knowledge profile measure that can be used to evaluate individual children's acquisition of theory of mind. Wellman and Liu (2004) tested preschool children on a wide range of theory of mind tasks, and they identified a set of seven tasks that children pass in a predictable sequence. Children who passed tasks later in the sequence tended to pass all of the earlier tasks as well. This scale has been widely used to evaluate the development of theory of mind in various subgroups of children, including children growing up in different cultures (e.g., Kristen, Thoermer, Hofer, Aschersleben, & Sodian, 2006; Wellman, Fang, Liu, Zhu, & Liu, 2006), deaf children (Peterson, Wellman, & Liu, 2005), and children with autism (Peterson et al., 2005).

In the future, it may be possible to generate a comparable sort of scale for children's acquisition of arithmetic principles. It seems unlikely that all or most children will acquire arithmetic principles in a constant order. However, such a tool would allow investigations of how different factors, such as curricula, mathematical ability, and spatial ability affect children's acquisition of arithmetic principle knowledge.

Conclusion

This paper has reviewed research on learners' knowledge of the commutativity, relation to operands, and inversion principles in the domain of arithmetic. For all three principles, conclusions about learners' principle knowledge hinge both on the context in which the arithmetic is presented, and on the type of knowledge assessment used to evaluate learners' knowledge. However, relatively few studies directly address the possible effects of either of these factors. Systematic attention to these issues will encourage progress in understanding both the emergence of principle knowledge and possible mechanisms of change in that knowledge.

For all three principles, the vast majority of existing studies utilize single-faceted knowledge assessments, which can lead to incomplete or misleading views of learners' knowledge. In this paper, we have made a case for the importance of multifaceted knowledge assessments, or knowledge profiles, that characterize learners' knowledge across tasks, contexts, and assessment types. Such profiles require research designs that include multiple measures of principle knowledge for individual learners, and as such, they pose various challenges, both practical and statistical. Despite these challenges, we believe that research that utilizes multifaceted knowledge assessments is needed to spur both theoretical and empirical progress in understanding the development of principle understanding, not only in arithmetic but also in other domains. The potential payoff from such work should make addressing the challenges worthwhile.

Acknowledgments

This work was supported by a predoctoral traineeship from the Institute of Education Sciences to Richard W. Prather, and by a Vilas Associate Award from the University of Wisconsin-Madison to Martha W. Alibali. The research was initiated as part of Richard W. Prather's preliminary examination for the doctoral requirements at the University of Wisconsin-Madison. We thank Mitchell Nathan, Colleen Moore, and Chuck Kalish for helpful discussion and comments on earlier drafts of this manuscript. We also thank Michael Koszewski for research assistance.

References

Ahn, W., & Graham, L. (1999). The impact of necessity and sufficiency in the Wason four-card selection task. *Psychological Science, 10*(3), 237–242.
Baroody, A. J. (1987). The development of counting strategies for single-digit addition. *Journal for Research in Mathematics Education, 19*, 141–157.
Baroody, A. J., & Gannon, K. (1984). The development of the commutativity principle and economical addition strategies. *Cognition & Instruction, 1*, 321–339.
Baroody, A. J., & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of arithmetic. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 75–112). Hillsdale, NJ: Erlbaum.
Baroody, A. J., Ginsburg, H., & Waxman, B. (1983). Children's use of mathematical structure. *Journal for Research in Mathematics Education, 14*, 156–168.
Baroody, A. J., & Lai, M. (2007). Preschoolers' understanding of the addition–subtraction inverse principle: A Taiwanese sample. *Mathematical Thinking & Learning, 9*, 131–171.
Baroody, A. J., Lai, M., Li, X., & Baroody, A. E. (2009). Preschoolers' understanding of subtraction-related principles. *Mathematical Thinking & Learning, 11*, 41–60.
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: The construction of adaptive expertise*. Mahwah, NJ: Erlbaum.
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.
Bisanz, J., Watchorn, R. P. D., Piatt, C., & Sherman, J. (2009). On "understanding" children's developing use of inversion. *Mathematical Thinking & Learning, 11*, 10–24.
Briars, D., & Siegler, R. S. (1984). A featural analysis of preschooler's counting knowledge. *Developmental Psychology, 20*, 607–618.
Brush, L. (1978). Preschool children's knowledge of addition and subtraction. *Journal for Research in Mathematics Education, 9*, 44–54.
Bryant, P., Christie, C., & Rendu, A. (1999). Children's understanding of the relation between addition and subtraction: Inversion, identity and decomposition. *Journal of Experimental Child Psychology, 74*, 194–212.
Campbell, J. (1999). Division by multiplication. *Memory and Cognition, 27*, 791–802.
Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology, 92*, 220–246.
Canobi, K. H. (2009). Concept–procedure interactions in children's addition and subtraction. *Journal of Experimental Child Psychology, 102*, 131–149.
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.
Canobi, K., Reeve, R., & Pattison, P. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34*(5), 882–891.
Canobi, K., Reeve, R., & Pattison, P. (2002). Young children's understanding of addition concepts. *Educational Psychology, 22*(5), 513–532.
Canobi, K., Reeve, R., & Pattison, P. (2003). Patterns of knowledge in children's addition. *Developmental Psychology, 39*(3), 521–534.
Carpenter, T., & Moser, J. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education, 15*, 179–202.
Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.). *Advances in the psychology of human intelligence* (Vol. 1, pp. 7–76). Hillsdale, NJ: Erlbaum.
Chiang, W., & Wynn, K. (2000). Infants' tracking of objects and collections. *Cognition, 77*(3), 169–195.
Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*, 377–395.
Cohen, L., & Marks, K. (2002). How infants process addition and subtraction events. *Developmental Science, 5*(2), 186–212.
Cowan, R., & Renton, M. (1996). Do they know what they are doing? Children's use of economical addition strategies and knowledge of commutativity. *Educational Psychology, 16*(4), 407–420.

Dixon, J., & Bangert, A. (2005). From regularities to concepts: The development of children's understanding of a mathematical relation. *Cognitive Development, 20*, 65–86.
Dixon, J., Deets, J., & Bangert, A. (2001). The representations of the arithmetic operations include functional relationships. *Memory & Cognition, 29*(3), 462–477.
Dixon, J., & Moore, C. (1996). The developmental role of intuitive principles in choosing mathematical strategies. *Developmental Psychology, 32*, 241–253.
Evans, J. (1982). *The psychology of deductive reasoning*. London: Routledge & Kegan Paul.
Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44*, 33–66.
Fendrich, D. W., Healy, A. F., & Bourne, L. E. Jr., (1993). Mental arithmetic: Training and retention of multiplication skill. In C. Izawa (Ed.), *Applied cognitive psychology: Applications of cognitive theories and concepts* (pp. 111–133). Hillsdale, NJ: Erlbaum.
Gelman, R., & Gallistel, C. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
Gilmore, C. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development, 21*, 301–316.
Gilmore, C., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetic skill. *British Journal of Educational Psychology, 76*, 309–331.
Gilmore, C. K., & Bryant, P. (2008). Can children construct inverse relations in arithmetic? Evidence for individual differences in the development of conceptual understanding and computational skill. *British Journal of Developmental Psychology, 26*, 301–316.
Gilmore, C., & Papadatou-Pastou, M. (2009). Patterns of individual differences in conceptual understanding and arithmetical skill: A meta-analysis. *Mathematical Thinking & Learning, 11*(1), 25–40.
Gilmore, C., & Spelke, E. (2008). Children's understanding of the relationship between addition and subtraction. *Cognition, 107*, 932–945.
Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review, 100*, 279–297.
Groen, G., & Resnick, L. (1977). Can preschool children invent addition algorithms? *Journal of Educational Psychology, 69*, 645–652.
Jeffries, R., Polson, P., Razran, L., & Atwood, M. (1977). A process model for Missionaries–Cannibals and other river-crossing problems. *Cognitive Psychology, 9*, 412–440.
Jordan, N., Huttenlocher, J., & Levine, S. (1992). Differential calculation abilities in young children from middle and low income families. *Developmental Psychology, 28*(4), 644–653.
Karmiloff-Smith, A. (1986). From meta-processes to conscious access: Evidence from children's metalinguistic and repair data. *Cognition, 23*, 95–147.
Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
Klein, J., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology, 54*(2), 105–115.
Koedinger, K. R., Alibali, M. W., & Nathan, M. J. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. *Cognitive Science, 32*, 366–397.
Koedinger, K., & Nathan, M. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences, 13*(2), 129–164.
Kotovsky, K., Hayes, J., & Simon, H. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology, 17*, 248–294.
Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G., & Sodian, B. (2006). Skalierung von 'Theory of Mind'-Aufgaben (validation of the 'Theory of Mind' scale). *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38*, 186–195.
Lai, M., Baroody, A. J., & Johnson, A. R. (2008). Fostering Taiwanese preschoolers' understanding of the addition–subtraction inverse principle. *Cognitive Development, 23*, 216–235.
Langford, P. (1981). A longitudinal study of children's understanding of logical laws in arithmetic and Boolean algebra. *Educational Psychology, 1*, 119–139.
Lawler, R. (1981). The progressive construction of mind. *Cognitive Science, 5*(1), 1–30.
LeCorre, M., Van de Walle, G., Brannon, E., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology, 52*(2), 130–169.
Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Differential calculation abilities in young children from middle- and low-income families. *Developmental Psychology, 28*, 644–653.
McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science, 15*(11), 776–781.
McNeil, N., & Alibali, M. (2005). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition & Development, 6*(2), 285–306.
Miller, P. H., & Seier, W. L. (1994). Strategy utilization deficiencies in children: When, where and why? *Advances in Child Development & Behavior, 25*, 107–156.
Mix, K., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin, 128*(2), 278–294.
Munakata, Y. (2001). Graded representations in behavioral dissociations. *Trends in Cognitive Sciences, 5*(7), 309–315.
Munakata, Y., McClelland, J. L., Johnson, M. J., & Siegler, R. S. (1997). Rethinking infant knowledge: Toward an adaptive process account of successes and failures in object permanence tasks. *Psychological Review, 104*(4), 686–713.
Munakata, Y., & Yerys, B. E. (2001). All together now: When dissociations between knowledge and action disappear. *Psychological Science, 12*(4), 335–337.
Nunes, T., Bryant, P., Hallett, D., Bell, D., & Evans, D. (2009). Teaching children about the inverse relation between addition and subtraction. *Mathematical Thinking & Learning, 11*, 61–78.
Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development, 76*, 502–517.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
Prather, R. W., & Alibali, M. W. (2008a). Understanding and using principles of arithmetic: Operations involving negative numbers. *Cognitive Science, 32*, 445–457.
Prather, R. W., & Alibali, M. W. (2007). Knowledge of an arithmetic principle in symbolic and verbal contexts: Do children know what adults know? Poster presented at the Biennial Meeting of the Society for Research in Child Development, Boston, MA.
Prather, R. W., & Alibali, M. W. (2008b). Implicit learning of arithmetic principles. In J. McClelland, & J. Weng (Eds.), *Proceedings of the 7th IEEE international conference on development and learning*.
Rasmussen, C., Ho, E., & Bisanz, J. (2003). Use of the mathematical principle of inversion in young children. *Journal of Experimental Child Psychology, 85*, 89–102.
Resnick, L. B. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In G. Leinhardt, R. T. Putnam, & R. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 373–429). Hillsdale, NJ: Erlbaum.
Rickard, T. C., & Bourne, L. E. (1996). Some tests of an identical elements model of basic arithmetic skills. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*(5), 1281–1295.
Robinson, K. M., & Dubé, A. K. (2009a). The multiplication and division inversion concept: Relationships between procedural, conceptual, and factual knowledge. *Cognitive Development*.
Robinson, K. M., & Dubé, A. K. (2009b). A microgenetic study of the multiplication and division inversion concept. *Canadian Journal of Experimental Psychology*.
Robinson, K., & Ninowski, J. (2003). Adults' understanding of inversion concepts. *Canadian Journal of Experimental Psychology, 57*(4), 321–330.
Robinson, K., Ninowski, J., & Gray, M. (2006). Children's understanding of the arithmetic concepts of inversion and associativity. *Journal of Experimental Child Psychology, 94*(4), 349–362.
Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology, 80*, 201–224.
Sherman, J., & Bisanz, J. (2007). Evidence for use of mathematical inversion by three-year-old children. *Journal of Cognition & Development, 8*, 333–344.
Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116*, 250–264.
Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology, 27*, 194–226.
Siegler, R., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
Siegler, R., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127*(4), 377–397.
Simon, T., Hespos, S., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn (1992). *Cognitive Development, 10*, 253–269.
Sophian, C., Harley, H., & Martin, C. S. M. (1995). Relational and representational aspects of early number development. *Cognition & Instruction, 13*, 253–268.
Sophian, C., & McCorgray, P. (1994). Part–whole knowledge and early arithmetic problem solving. *Cognition & Instruction, 12*(1), 3–33.
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.
Starkey, P. (1992). The early development of numerical reasoning. *Cognition, 43*, 93–126.
Uller, C., Carey, S., Hunter-Fenner, G., & Klatt, L. (1999). What representations might underlie infant numerical knowledge? *Child Development, 14*, 1–36.
Vilette, B. (2002). Do young children grasp the inverse relationship between addition and subtraction? *Cognitive Development, 17*, 1365–1383.
Wason, P. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology*. Harmondsworth, England: Penguin.
Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science, 17*, 1075–1081.
Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development, 75*, 523–541.
Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression thought*. New York: Wiley.
Wilkins, J., Baroody, A. J., & Tiilikainen, S. (2001). Kindergartners' understanding of additive commutativity within the context of word problems. *Journal of Experimental Child Psychology, 79*(1), 23–36.
Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749–750.
Zelazo, P. D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*(1), 37–63.

Please cite this article in press as: Prather, R. W., & Alibali, M. W. The development of arithmetic principle knowledge: How do we know what learners know?. *Developmental Review* (2009), doi:10.1016/j.dr.2009.09.001