Patterns of Individual Differences in Conceptual Understanding and Arithmetical Skill: A

Meta-Analysis

Camilla K. Gilmore

Learning Sciences Research Institute, University of Nottingham

Marietta Papadatou-Pastou

Department of Experimental Psychology, University of Oxford

Abstract

Some theories from cognitive psychology and mathematics education suggest that children's understanding of mathematical concepts develops together with their knowledge of mathematical procedures. However, previous research into children's understanding of the inverse relationship between addition and subtraction suggests that there are individual differences in the way this concept develops. To determine whether these differences are reliable and reflect alternative paths of development, we examined data from 14 studies of children's understanding of inversion. Cluster analyses and meta-analytic techniques were used to quantify the size of the inversion effect and examine factors influencing its size and to test the stability of patterns of individual differences across the studies. Evidence was found for reliable patterns of individual differences, which have implications for current theories of concept development.

Patterns of Individual Differences in Conceptual Understanding and Arithmetical Skill: A

Meta-Analysis

An enduring question in mathematical cognition concerns how children develop

understanding of mathematical concepts. This has attracted attention from researchers in both

cognitive psychology and mathematics education. One of the key questions for both fields is

how children's understanding of mathematical concepts develops in relation to their ability to

perform mathematical procedures.

Within cognitive psychology, there has been a move from earlier debates over whether

mathematical concepts or mathematical skills develop first (e.g., Briars & Siegler, 1984;

Riley, Greeno, & Heller, 1983) to the proposal that mathematical concepts and skills develop

iteratively, each building on the other (e.g., Baroody & Ginsburg, 1986; Hiebert & Wearne,

1996) or develop simultaneously (Baroody, 1992, 2003; Rittle-Johnson & Siegler, 1998).

These iterative- or simultaneous-development theories propose that conceptual and

procedural knowledge are closely related, with developments in procedural knowledge

leading to improvement in conceptual knowledge and vice-versa.

Similarly, process-object theories from mathematics education research have proposed a

close relationship between the development of mathematics procedures and concepts (e.g.,

Cottrill, Dubinsky, Nichols, Schwingendorf, Thomas, & Vidakovic, 1996; Gray & Tall,

1994; Sfard, 1991). According to these theories, mathematics concepts, or objects, are

developed through the encapsulation or reification of a process, or procedure. This can take

place at a number of levels, so that procedures performed on existing concepts may lead to

the development of new concepts. Thus, procedural knowledge plays an integral role in the

development of conceptual understanding. In this way, individuals can develop flexible

thinking in which processes and concepts are tied together (proceptual thinking; Gray & Tall,

1994).

Many cognitive psychology and mathematics education theories therefore suggest that knowledge of procedures plays an important role in the development of conceptual understanding. Although the details of these theories differ somewhat, they emphasise that procedural knowledge is necessary for the development of a full conceptual understanding (e.g., Baroody, Feil, & Johnson, 2007). Thus to develop flexible understanding of concepts, children also develop proficient procedural skills (Gray & Tall, 1994). It is important, however, to examine alternative ways in which children may learn about arithmetic concepts. Although some theorists have suggested that concept development may follow different routes in different mathematical domains (Rittle-Johnson & Siegler, 1998), there has been little examination of the ways in which concept development differs across individuals.

Recent work has begun to examine individual differences in the development of mathematical understanding (see Clements & Sarama, 2004; Dowker, 2005, for reviews). This work has attempted to measure, and in a few cases account for, individual differences in children's arithmetic performance (Hecht, Close, & Santisi, 2003; Swanson & Kim, 2007), conceptual understanding (Canobi, 2004, 2005; Canobi, Reeve, & Pattison, 1998; Dowker, 1998) and strategy use (Imbo & Vandierendonck, 2007; Kerkman & Siegler, 1997). These studies reveal the importance of considering the variety of ways in which children develop understanding of arithmetic.

Children's understanding of the relationship between addition and subtraction provides an example of these individual differences. Understanding that addition and subtraction are inversely related is an important step in children's development of arithmetic. Previous research has examined the development of this concept and highlighted three types of individual differences relating to children's understanding and use of this principle. First, there is wide variation in the age at which children may use the inverse principle in problem solving (e.g., Bisanz & LeFevre, 1990). If children understand that addition and subtraction

are inversely related, they can solve problems involving addition and subtraction of the same quantity ($a + b - b$) without using calculation. This is demonstrated through higher accuracy for problems involving inversion (e.g., $9 + 7 - 7$) than for those that require calculation (e.g., $8 + 6 - 3$) – the inversion effect. When inverse and control problems are matched for computational difficulty, higher accuracy for inversion problems can indicate use of a conceptually-based shortcut. Use of this shortcut reveals that children have some understanding of the inverse nature of the transformations involved. Use of this shortcut does not, however, necessarily imply conscious awareness of this relationship, or full understanding of the abstract logical principle. Although some children appear to show evidence of an inversion effect during the preschool period, other children fail to use this principle by the end of primary school. For example, around three-quarters of preschoolers (age 4) in a study by Rasmussen, Ho, and Bisanz (2003) had response and behaviour patterns that showed better performance for inverse problems such as $4 + 2 - 2$ than matched control problems. In contrast, less than half of the 7- and 9-year-olds tested by Bisanz and LeFevre (1990) and 60% of the 10-year-olds tested by Stern (1992) showed an advantage for inverse problems. Although differences in methodology and problem-size may account in part for the different findings of these studies, there remains a wide variation in the age at which children seem able to use understanding of inverse relations when solving problems.

The second type of individual differences concerns the route through which children may develop understanding of the relationship between addition and subtraction. Canobi (2005) gave 5- to 7-year-old children two tasks to assess use of inversion and the related complement principle: three-term inversion and control problems of the type described above (e.g., $5 + 8 - 8 = ?$); and pairs of related addition and subtraction complement problems (e.g., $3 + 5 = 8$, $8 - 5 = ?$). Children's use of the inverse or complement principles rather than calculation to solve these problems was recorded. Using cluster analysis, Canobi (2005)

revealed that some children use the relationship between addition and subtraction first in terms of the complementary relationship between separate addition and subtraction problems, whereas other children use this relationship first in terms of the inverse addition and subtraction of the same number. This finding suggests that children take different routes to understanding the relationship between addition and subtraction.

Finally, a third type of individual differences concerns the relationship between children's understanding of inversion and their procedural calculation skills. A number of studies have found that children's performance on inversion tasks is not related to their numerical or calculation skills with similar sized quantities (Bryant, Christie, & Rendu, 1999; Rasmussen et al., 2003; Sherman & Bisanz, 2007). For example, children's ability to add 9 is not related to their use of a computational shortcut on problems involving $+ 9 - 9$. These findings seem to suggest that development of inversion understanding is not closely tied to development in calculation skills with similar-sized items. Recent research indicates that, in fact, there may be differences among children in the relationship between inversion understanding and calculation skills. Using cluster analysis to examine subgroups in a sample of 6- and 8-year-olds, Gilmore and Bryant (2006) found three subgroups of children. One group demonstrated good inversion performance and good calculation skills; a second group had poor understanding of inversion and poor calculation skills; the final group, in contrast, had good understanding of inversion despite poor calculation skills. These children demonstrated understanding of the relationship between addition and subtraction that far outstripped their proficiency with these operations.

The pattern of individual differences found by Gilmore and Bryant (2006) appears to suggest that children differ in the relationship between their conceptual understanding and calculation skill. The implications of this finding are difficult to judge, however, since there are three possible empirical interpretations. First, this may be a replicable pattern of

individual differences that is found at all stages of children's development of inversion understanding. In this case, some children are able to show conceptual understanding in the absence of proficient calculation skills with similar-sized control items. Second, this may be a replicable pattern of individual differences but that only applies to children at a particular age, or stage of learning about inversion. In this case, some children may go through a period of showing more advanced understanding of inversion than calculation skills, but before and after this their conceptual understanding may be more closely tied to their procedural skills. Finally, the pattern of individual differences may simply reflect chance variation in a particular sample, or a particular methodology. In this case the pattern will not be replicable in other studies.

In summary, there appears to be evidence for differences in the age at which children may understand inversion, the context in which they first show this understanding, and the relationship between understanding of inversion and calculation skills. If found to be reliable, these differences suggest that there may be multiple routes in which children acquire this concept. In particular, the role of procedural knowledge in concept development may be called into question.

In this paper we examine evidence for differences in the way children learn about inversion by bringing together data from a range of studies on inversion understanding. These published and unpublished studies include a range of age groups, different methodologies and have been carried out in diverse geographical locations. To examine the effect of age and context in children's use of inversion, a meta-analysis was used to quantify the size of the inversion effect (the difference in children's scores on inverse and control problems) and reveal what factors influence the size of the effect. To examine individual differences in the relationship between understanding inversion and calculation skills and to discover whether the subgroups found by Gilmore and Bryant (2006) are replicable in other studies, cluster

analyses were performed on each data set and a meta-analysis of each cluster was used to quantify the differences in performance across the clusters. These analyses help to establish the robustness and basis for patterns of individual differences, with implications for theories of concept development.

<div align="center">Method</div>

The analysis consisted of three phases: (a) a meta-analysis of studies to quantify the inversion effect and examine the effect of moderating variables; (b) cluster analysis for each data set to reveal subgroups of children; (c) meta-analyses of subgroups to reveal differences between the clusters in the size of the inversion effect.

Meta-analysis of Inversion Effect

Locating studies. The studies that entered the meta-analysis were located through the following procedure: first the computerized reference database "Web of Science" was searched using the search terms "invers* AND addition AND subtraction AND children" and "invers* AND arithmetic AND children". The reference list for all studies eligible for inclusion was scanned and citation searches were further performed for all selected studies. All relevant studies in a doctoral thesis conducted in this area were included (Gilmore, 2005). Authors of published papers were contacted to request information on any unpublished studies as well as to request raw data. Data collection ended on 17th October 2006.

Study selection. The following criteria were set for inclusion of an individual study in the meta-analysis: (a) participants under the age of 18, (b) studies using a method involving the presentation of two types of arithmetical problems: inverse problems that included an inverse transformation (e.g., $a + b - b = a$) and equivalent matched control problems that did not include an inverse transformation (e.g., $a + b - c = d$), (c) means and their standard deviations of accuracy scores reported for both types of problems, and (d) papers in English. A few studies included two or more samples from different age groups or using different

presentation types; in those cases the data sets were treated as separate (Bryant et al., 1999; Canobi, 2005; Gilmore, 2005; Gilmore & Bryant, 2006; Rasmussen et al., 2003; Robinson, Ninowski, & Gray, 2006; Rose, 2002).

Moderator variables. Three variables were considered as potential moderators of the size of the inversion effect: (a) mean age in years; (b) school year: Coded as 0 = preschool children; 1 = children attending their 1st year at school; 2 = children in their 2nd school year; 3 = children in their 3rd school year; 4 = children in their 4th school year; 5 = children in their 5th school year and 6 = children in their 6th year and beyond; and (c) type of presentation (digits, word problems, pictures and concrete materials). All of the studies included in the meta-analysis presented children with inversion and control problems to solve, but they differed in terms of how these problems were presented: as symbolically-presented arithmetic problems; as verbally-presented word problems, using pictures and symbols to demonstrate actions; using concrete items such as counters or blocks to demonstrate actions.

Statistical analysis. Data were analysed using Comprehensive Meta-Analysis (v. 2) software package. In order to perform a meta-analysis, an effect size is first calculated for each study. Here, we used the standardized mean difference d in children's accuracy between inversion and control problems as the effect size, which is the difference between the two means divided by the within-group standard deviation. The meta-analytic procedure took the following steps. First the mean difference in accuracy between inversion and control problems was calculated and standardized for each data set with the corresponding 95% confidence intervals. Then, an average of the effect sizes across data sets was calculated and weighted according to sample size, using a fixed-effects model. The effect sizes were further tested to see if they come from a single population using two tests of homogeneity, the Q statistic and the I^2 index. In the case of heterogeneity between the studies, the overall effect

size was calculated again using a random effects model. Weighting the effect sizes can therefore follow two models, depending on the absence or presence of between-study heterogeneity: the fixed effects model and the random effects model respectively. The fixed effects model assumes that all the data sets come from a single population; it asks what the best estimate of the effect size is. The random effects model assumes that the data sets are drawn from a distribution of populations; this model asks what the range of the effect sizes of the populations studied is.

Whenever a subset is heterogeneous, the presence of moderator variables might be causing this heterogeneity. The possible moderating effect of the categorical moderator variable (i.e., presentation type) was examined by comparing the average effect sizes in the different subgroups that form the levels of the moderator. In examining the effects of interval moderator variables (i.e. age and school year) meta-regression was used. Meta-regression is similar in essence to regression: the moderator variable is predicted according to the weighted effect size of each data set.

Forest plots were used to depict all the information visually. The studies were also tested for possible ascertainment bias, which can exist when significant results are produced by non-random sampling of the data-sets, using Egger's t statistical test and the fail-safe N.

Cluster Analyses of Individual Studies

Following the initial meta-analysis, cluster analysis was used to test whether the pattern of individual differences found by Gilmore and Bryant (2006) was replicated across studies. Cluster analysis can be used to group individuals into clusters of relatively homogeneous cases. A hierarchical cluster analysis using Ward's method was performed on each data set and the three-cluster solution was interpreted. The analysis was carried out on overall scores for inverse and control problems, combined across different presentation formats. Performance from problems that deviated substantially from the $a + b - b = a$ problem type

was excluded (complex inverse and control problems in Gilmore & Bryant, in press;

decomposition problems in Bryant et al, 1999). Where children from multiple age groups

were included in a study, the cluster analysis was performed on the combined sample where

children all received the same problem set, but performed on the samples separated by age

where children received different problem sets.

Meta-analysis of Cluster Structure

A second set of meta-analyses were carried out for each of the three clusters separately,

in order to quantify the inversion effect in each cluster. The same statistical procedures as in

the meta-analysis of the overall inversion effect were followed to calculate the overall effect

size. The differences between the inversion effects for each cluster were tested for

significance using the Kruskal-Wallis and Mann-Whitney nonparametric tests.

Results

The results of the three phases of analysis are examined separately below.

Meta-analysis of Inversion Effect

A total of 25 data sets (from 14 separate studies) were included in the analysis, totaling

745 subjects (357 male, 388 female). Table 1 shows the details of the studies included in the

meta-analysis.

Overall effect estimate. All but one data set had a standardized mean difference

indicating better accuracy levels for inversion than control problems, 22 of these were

statistically significant. Fixed effects analysis revealed significant heterogeneity among the

data sets (Q (24) = 127.38, $p < 0.01$, I^2 = 81.16%), indicating that one or more variables

influenced the size of the inversion effect. A random effects model was therefore employed

which gave a weighted average of the standardized mean difference across all data sets of

1.24 with a confidence interval of *95%CI* = 1.02 - 1.46 (Figure 1). In other words, the range

of the effect sizes in the distribution of populations studied is 1.02 to 1.46. Thus, for all

populations included in the meta-analysis a randomly selected child would score at least one

standard deviation higher on inversion problems than control problems.

--Insert Figure 1 about here--

Ascertainment bias. The data sets that were included in the meta-analysis were tested for

ascertainment bias. No ascertainment bias was detected using Egger's Test *(t* (23) < 0.01, *p* =

0.50). The fail-safe *N* was also calculated, *N* = 3,938. The fail-safe *N* is the number of data

sets in which the odds ratio is zero that would be needed to increase the *p*-value of the meta-

analysis to above 0.05. These analyses indicate that there is no evidence for non-random

sampling which would distort the estimates of effect sizes produced.

Moderating variables. The possible moderating effects of age, school year and

presentation type were examined. Meta-regressions run for age (Q (1) = 0.12, p = 0.73) and

school year (Q (1) = 0.04, p = 0.83) did not show any significant relationship between the

size of the inversion effect and age or school year.

On the other hand, presentation type (Q (3) = 8.74, p = 0.03) did moderate children's

performance, albeit marginally. Table 2 presents details for each level of the categorical

variable. The inversion effect was greatest with picture presentation, similar for word or

concrete presentation and smallest for digit presentation of problems.

--Insert Table 2 about here--

Cluster Analyses of Individual Studies

The results of the cluster-analyses performed on each study are given in Table 3. This

lists the proportion of variance in each data set accounted for by the three-cluster solution, the

size of the individual clusters and the mean scores on the inverse and control problems (given

as the proportion of correct responses).

--Insert Table 3 about here--

The percentage of variance in the data explained by the three-cluster solution ranged from 63.8% to 86.4% across the studies. Thus, it was acceptable to interpret the three-cluster solution in each of the studies. This indicates that, for each of the studies, the sample could meaningfully be split into three subgroups.

The raw mean scores suggest that there is, on the whole, a similar pattern of clusters across the studies. The first cluster is characterized by very high means for inversion problems (all except three are more than 0.85) and reasonably high, although more varied, means for control problems. The second cluster is characterized by low means for both inversion and control problems (all except one less than 0.5). The third cluster is characterized by high means for inversion problems (all but two more than 0.5) and low means for control problems (less than 0.5). The one exception are the data from Canobi (2005). The children in Cluster 3 from this study had a mean score of 0.33 on inversion problems and 0.67 on control problems. This is the only group across all of the studies that showed a pattern of higher scores on control than inversion problems. The three-cluster solution appears to be identifying different subsets of individuals for this study compared to the other studies. The reason for this is unclear, however, only a small number of inversion problems (3) was included in this study, which may affect the reliability of estimates of inversion understanding. This group was therefore omitted from further analyses.

Meta-analysis of Cluster Structure

To quantify the inversion effect for each cluster, three separate meta-analyses were carried out. This procedure is more sensitive than simply examining mean scores, as it accounts for the different sample sizes across the data sets as well as for the between-study heterogeneity. The meta-analysis on Cluster 1 was performed on 15 data sets, totalling 256 subjects. All data sets had standardized mean differences indicating better accuracy levels for inversion than control problems; all but one of these contrasts was statistically significant.

Fixed effects analysis revealed significant heterogeneity among the data sets $(Q\ (14) =$ 114.81, $p < 0.01$, $I^2 = 87.80\%$). A random effects model was therefore employed which gave a weighted average of the standardized mean difference of 2.65 with a confidence interval of *95%CI = 1.94 – 3.36.*

The meta-analysis on Cluster 2 was performed on14 data sets, totalling 178 subjects (Rose, 2002, was not included as it had a mean score of zero with a standard deviation of zero on the control problems). Eight data sets had a significant standardized mean difference favouring inversion problems and six data sets had no significant difference between scores on inversion and control problems. Fixed effects analysis revealed significant heterogeneity among the data sets $(Q\ (13) = 39.68$, $p < 0.01$, $I^2 = 67.24\%$). A random effects model was therefore employed which gave a weighted average of the standardized mean difference of 0.99 with a confidence interval of *95%CI = 0.56 – 1.40.*

The meta-analysis on Cluster 3 was performed on 13 data sets, totalling 208 subjects (Gilmore, 2005 Exp. 2 was not included as it had a mean score of zero with a standard deviation of zero on the control problems). All data sets had standardized mean differences indicating significantly higher accuracy for inversion problems than control problems. Fixed effects analysis revealed significant heterogeneity among the data sets $(Q\ (12) = 85.50$, $p <$ 0.01, $I^2 = 85.96\%$). A random effects model was therefore employed which gave a weighted average of the standardized mean difference of 4.80 with a confidence interval of *95%CI =* 3.80 – 5.80.

The standardized mean differences for each cluster were compared using the Kruskal-Wallis and Mann-Whitney tests. The size of the inversion effect was significantly different in each of the clusters ($\chi^2 = 23.87$, $p < .001$). The inversion effect was larger in Cluster 3 than Cluster 2 ($U = 0.00$, $p < 0.001$) or Cluster 1 ($U = 42.00$, $p = 0.011$), and the inversion effect was larger in Cluster 1 than Cluster 2 ($U = 32.00$, $p = 0.001$).

Discussion

We draw three main conclusions from the analyses. First, by bringing together, for the first time, data from a range of studies on children's understanding of inversion we were able to produce a reliable estimate of the size of the inversion effect and to reveal what factors affect this. Second, we have revealed evidence for a robust pattern of individual differences across a range of studies in the relationship between children's understanding of inversion and their calculation skills. Finally, we have shown a replicable subgroup of children with a performance profile that has implications for theories of concept development in general and inversion development in particular. Each of these issues will be considered in turn.

The initial meta-analysis revealed that there was a reliable inversion effect, such that children in these studies scored on average 1.24 standard deviations higher on inverse than control problems. In only one study was there a trend in the opposite direction, with higher performance on control problems than inverse problems. This indicates that, on the whole, children can identify and use the inverse relationship between addition and subtraction when solving problems. Examination of the moderating variables revealed that only presentation type affected the size of the inversion effect. Relative performance on inversion and control problems did not change significantly with age or school year. The lack of a significant moderating effect of age group suggests that although arithmetic ability improves with age, the inversion effect does not appear to change. However, since the size of problems given to younger and older children tend to differ, this represents only stability in terms of the relative advantage for inversion problems over control problems.

The significant moderating effect of presentation type indicates that children can gain additional benefit in identifying conceptual relations when problems are presented with more context. The inversion effect was largest when problems were presented using pictures to describe actions and smallest when problems were presented using digits. It has been

suggested that additional context can help improve children's problem-solving and

calculation skills (e.g., Hughes, 1981). The current analysis suggests that increased context

also has a benefit in allowing children to recognize conceptual relations.

It is difficult, however, to disentangle the effects of the moderating variables since they

are largely confounded. School year and age are closely related (but not identical) and

different presentation methods were used for younger and older children. Furthermore, when

examining the effects of moderating variables in meta-analytic data the crucial characteristic

is the number of data sets (25 in this case) and not the number of subjects (745). This can

result in a surprisingly low power of meta-analytic studies when it comes to examining

moderating effects, in spite of their big subject numbers (Hunter & Schmidt, 1990).

The cluster analyses and meta-analyses of each cluster established that there is a

replicable pattern of individual differences across the majority of the studies. In all cases the

three-group cluster analysis accounted for a substantial proportion of the variance in scores

and in most cases these clusters replicated the findings of Gilmore and Bryant (2006). The

same pattern of subgroups was found across studies involving children of a range of ages (5 -

13 years) and using somewhat differing methods. This reveals that the pattern of individual

differences observed by Gilmore and Bryant (2006) was not due to chance variation in a

particular sample, neither does this only apply to children at a particular age, or stage of

learning about inversion. Rather, the evidence from these studies suggests it may be a

replicable pattern of individual differences that is found at all stages of children's

development of inversion understanding.

We found evidence for three subgroups with different performance profiles. Two of these

groups match the developmental model suggested by Gray and Tall (1994). The children with

good understanding of inversion and good calculation skills match the *proceptual thinkers*

described by Gray and Tall (1994). These children have developed flexible thinking in which

procedural knowledge and conceptual understanding are used as appropriate in different problem situations. The children with poor understanding of inversion and poor calculation skills appear to demonstrate *procedural thinking* (Gray & Tall, 1994). They have not developed understanding of arithmetical concepts and their procedural knowledge is inflexible and unreliable. In addition to these groups, we found evidence for children with good understanding of inversion but poor calculation skills. With the exception of one study (Canobi, 2005), there was no evidence for a reliable group of children with good calculation skills but poor understanding of inversion. The lack of evidence for such a fourth group may in part be due to the nature of the methodology used in each study that contributed to this analysis. The children's performance on control problems serves both as a measure of their calculation skills, and is compared with performance on inversion problems to give a measure of conceptual understanding. This may have affected the profiles of performance that were revealed using cluster analysis. Indeed, recent research that examined children's understanding of inversion and used separate measures of calculation skills, found four subgroups of children (Watchorn et al., 2007). Thus, a key finding of the present analysis is the demonstration that children show different profiles of performance, these are not the only profiles of performance that may be found.

The existence of a replicable subgroup of children at all ages who have good understanding of inversion but poor calculation skills, has implications for theories of conceptual development in general and the development of inversion understanding in particular. As outlined above, theories from both cognitive psychology and mathematics education propose that procedural skill plays an important role in the development of conceptual understanding. Iterative development theories (e.g., Hiebert & Wearne, 1996; Rittle-Johnson & Siegler, 1998) suggest that procedural knowledge and conceptual knowledge build on each other iteratively with advances in one leading to developments in

the other. Similarly, process-object theories (e.g., Cottrill et al., 1996; Gray & Tall, 1994; Sfard, 1991) suggest that mathematical concepts are developed out of a related mathematical process. Although these proposed developmental models may hold for many children (i.e., those in Clusters 1 & 2), theories that propose that procedural knowledge plays an integral role in the development of conceptual understanding cannot easily account for the performance of children in Cluster 3. Theories of concept development need to account for ways in which children develop conceptual understanding in the absence of proficient procedural skill.

As with all research in this area, the nature of the relationships between conceptual and procedural knowledge found here are affected by the operationalisation of these terms. The studies reviewed here examined the relationship between children's use of a conceptually-based shortcut and their general computation skills. This approach can be distinguished from studies in which children's understanding of a concept is compared to their use of procedures that rely on understanding of this concept. Both approaches examine the relationship between aspects of children's knowledge of arithmetic, however, they do so at different levels of generality. To discover how all aspects of arithmetic are related across development, it would be informative to conduct studies that cut across this distinction.

In keeping with theories of concept development, some researchers (e.g., Baroody & Lai, 2007; Canobi, 2005) have proposed that children develop understanding of inversion from experience and expertise in solving addition and subtraction problems. For example, Canobi suggests that "children may reach an understanding of the inverse relation between addition and subtraction after they become sufficiently experienced in problem solving to be able to reflect on the outcomes of the additions and subtractions they have carried out" (Canobi, 2005; p. 22). However, the results of the present analysis show that in the vast majority of studies examined there was a group of children with good understanding of inversion despite

poor calculation skills. Therefore, some children can show inversion understanding with numerbers for which they cannot consistently perform calculations. As discussed next, further research is needed to establish whether these children require a lower level of procedural skill prior to developing inversion understanding.

How else might children develop understanding of inversion? Other researchers have proposed that children's quantitative understanding of this concept may arise out of earlier qualitative forms (Bisanz, Sherman, Rasmussen, & Ho, 2005; Bryant & Nunes, 2007; Rasmussen et al., 2003). If children use mental models to represent numbers of external objects and the transformations on them, they may notice the equivalence of sets before and after transformations with the same number. Social experience may play a role in this as children take part in give-and-take games with other children and adults (Klein & Bisanz, 2000). Children's experience with small quantities (e.g., 1, 2, 3) might allow children to infer localized informal knowledge of inversion, that becomes more general with experience of larger quantities (Baroody & Lai, 2007). Recent evidence also suggests that a system of approximate non-symbolic representation of number may play a role in children's developing understanding of quantitative inversion. Five-year-old children can identify and use inverse relationships for large numbers with approximate symbolic or nonsymbolic (dot array) quantities before they can do so for exact symbolic quantities (Gilmore & Spelke, 2008). Children can recognize inverse relationships with approximate representations of quantity (i.e., those that are not distinguishable as different). They may do this by building on a mechanism for representing, comparing and manipulating approximate quantities (see Dehaene, 1997). Thus, one route by which children may develop understanding of quantitative inversion relating to exact numbers is out of a pre-existing ability to recognize this relationship in terms of approximate numerosity.

In conclusion, when describing how children develop understanding of mathematical concepts, it is important to consider multiple ways in which children may come to this understanding. The challenge for theorists from both cognitive psychology and mathematics education is to account for the variety of ways in which children might discover and integrate conceptual knowledge with their knowledge of mathematical procedures.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Baroody, A. J. (1992). The development of preschoolers' counting skills and principles. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number* (pp. 99–126). Hillsdale, NJ: Erlbaum Associates.

Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds), *The development of arithmetic concepts and skills: Constructing adaptive expertise. (pp. 1-33).* Mahwah, NJ: Erlbaum.

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

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., & Lai, M. (2007). Preschoolers' understanding of the addition-subtraction inverse principle: A Taiwanese sample. *Mathematics Thinking and Learning, 9*, 131-171.

Bisanz, J., & LeFevre, J.-A. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 213-244). Hillsdale, NJ: Erlbaum.

Bisanz, J., Sherman, J. L., Rasmussen, C., & Ho, E. (2005). Development of arithmetic skills and knowledge in preschool children. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition*. New York: Psychology Press.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20*, 607-618.

*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.

Bryant, P., & Nunes, T. (2007, August). *Children's understanding and use of the relation between addition and subtraction.* Paper presented at the biennial meeting of the European Association for Research in Learning and Instruction, Budapest, Hungary.

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development, 19*, 81-93.

*Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology, 92*, 220-246.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34*, 882-891.

Clements, D. H., & Sarama, J. (2004). Learning trajectories in mathematics education. *Mathematical Thinking and Learning, 6,* 81-89.

Cottrill, J., Dubinsky, E., Nichols, D., Schwingendorf, K., Thomas, K., & Vidakovic, D. (1996). Understanding the limit concept: Beginning with a coordinated process scheme. *Journal of Mathematical Behavior, 15*, 167-192.

Dehaene, S. (1997). *The number sense*. Oxford, UK: Oxford University Press.

Dowker, A. (1998). Individual differences in normal arithmetical development. In C. Donlan (Ed.), *The development of mathematical skills*. Hove, UK: Psychology Press.

Dowker, A. (2005). *Individual differences in arithmetic: Implications for psychology, neuroscience and education*. Hove, UK: Psychology Press.

*Gilmore, C. K. (2005). Children's understanding of the inverse relationship between addition and subtraction: Doctoral dissertation, University of Oxford, UK.

*Gilmore, C. K. (2006). Investigating children's understanding of inversion using the missing number paradigm. *Cognitive Development, 21*, 301-316.

*Gilmore, C. K., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetical skill. *British Journal of Educational Psychology, 76*, 309-331.

*Gilmore, C. K., & Bryant, P. (in press). 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*.

Gilmore, C. K., & Spelke, E. S. (2008). Children's understanding of the relationship between addition and subtraction. *Cognition, 107,* 932-945.

Gray, E. M., & Tall, D. O. (1994). Duality, ambiguity, and flexibility: A "proceptual" view of simple arithmetic *Journal for Research in Mathematics Education, 25*, 116-140.

Hecht, S. A., Close, L., & Santisi, M. (2003). Sources of individual differences in fraction skills. *Journal of Experimental Child Psychology, 86*, 277-302.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding and skill in multidigit addition and subtraction. *Cognition and Instruction, 14*, 251-283.

Hughes, M. (1981). Can preschool children add and subtract? *Educational Psychology, 1*, 207-219.

Hunter, J. E., & Schmidt, F. L. (Eds.). (1990). *Methods of meta-analysis*. Newbury Park, Califonia: Sage.

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*, 284-309.

Kerkman, D., & Siegler, R. (1997). Measuring individual differences in children's addition strategy choices. *Learning and Individual Differences, 9*, 1-18.

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology-Revue Canadienne De Psychologie Experimentale, 54*, 105-116.

*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.

Riley, M., Greeno, J. G., & Heller, J. I. (1983). Development of children's problem solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). New York: Academic Press.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills. Studies in developmental psychology* (pp. 75-110). Hove, UK: Psychology Press.

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

*Rose, S. (2002). *The role of inversion, decomposition and the decimal structure in children's arithmetic*. Unpublished manuscript, University of Oxford, UK.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics, 22*, 1-36.

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

Stern, E. (1992). Spontaneous use of conceptual mathematical knowledge in elementary

    school children. *Contemporary Educational Psychology, 17*, 266-277.

Swanson, L., & Kim, K. (2007). Working memory, short-term memory, and naming speed as

    predictors of children's mathematical performance. *Intelligence, 35*, 151-168.

Watchorn, R., Bisanz, J., Fast, L., LeFevre, J.-A., Smith-Chant, B., Skwarchuk, S.-L., et al.

    (2007, March). *Calculation skills and the use of conceptually advanced solution*

    procedures. Paper presented at the Poster presented at the Biennial Meeting of the

    Society for Research in Child Development, Boston.

Author Note

Footnote

[1] For one study it was not possible to obtain the raw data to perform this analysis, and thus this study was not included in the subsequent analyses.

Table 1

Studies Included in the Meta-analysis of Inversion Effect

Study	Source	*N*	Mean age(s) (years; months)	Location	Presentation	Notes
Bryant, Christie & Rendu (1999)	Journal of Experimental Child Psychology (Exp. 1)	38	5;8 & 6;9	UK	Concrete, word, abstract	Separated by age into 2 datasets
Bryant, Christie & Rendu (1999)	Journal of Experimental Child Psychology (Exp. 2)	55	6;10, 7;7 & 8;9	UK	Concrete	Separated by age into 3 datasets
Rose (2002)	Unpublished student study	62	8;4, 9;6, & 10;6	UK	Verbal	Separated by age into 3 datasets
Rasmussen, Ho & Bisanz (2003)	Journal of Experimental Child Psychology	48	4;7 & 6;10	Canada	Concrete	Separated by age into 2 datasets
Canobi (2005)	Journal of Experimental Child Psychology	60	5;8, 6;7, & 7;9	Australia	Digit	Separated by age into 3 datasets
Gilmore (2005)	Doctoral thesis (Exp. 2)	23	6;3	UK	Picture	
Gilmore (2005)	Doctoral thesis (Exp. 5)	53	8;11 & 9;10	UK	Digit	Separated by age into 2 datasets. Order of operators also varied

Gilmore (2006)	Cognitive Development (Exp. 1)	49	9;2	UK	Digit, Word	
Gilmore (2006)	Cognitive Development (Exp. 2)	51	9;5	UK	Digit	Order of elements also varied
Gilmore & Bryant (2006)	British Journal of Educational Psychology (Exp. 1)	66	8;11	UK	Digit, Word, Picture	
Gilmore & Bryant (2006)	British Journal of Educational Psychology (Exp. 2)	59	6;6 & 8;8	UK	Picture	Separated by age into 2 datasets
Robinson, Ninowski & Gray (2006)	Journal of Experimental Child Psychology	90	11;6 & 13;6	Canada	Digit	Separated by age into 2 datasets. Mean of small and large problems
Gilmore & Bryant (in press)	British Journal of Developmental Psychology	67	8;7	UK	Digit	Problems also varied order of elements; excluded complex problems.
Sherman & Bisanz (2007)	Journal of Cognition and Development	24	3;6	Canada	Concrete	

Table 2

Mean Standardized Difference in Accuracy between Inversion and Control Problems for all Levels of the Presentation Type Moderator Variable.

Presentation type	Number of data sets	Standardized mean difference	95% Confidence Interval
Concrete	4	1.42	1.18-1.66
Digit	8	0.91	0.56-1.25
Picture	3	1.61	1.26-1.97
Word	3	1.40	0.90-1.89

Table 3

Results of Cluster Analyses (Split by Age Group Only Where Separate Problem Sets Used)

Study	Variance explained (%)	Cluster 1			Cluster 2			Cluster 3		
		N	Inverse	Control	N	Inverse	Control	N	Inverse	Control
Bryant, Christie & Rendu (1999) Exp. 1	75.6	10	0.86 *(.23)*	0.76 *(.12)*	16	0.33 *(.22)*	0.05 *(.07)*	12	0.87 *(.10)*	0.24 *(.18)*
Bryant, Christie & Rendu (1999) Exp. 2	74.3	35	0.89 *(.11)*	0.43 *(.12)*	6	0.08 *(.10)*	0.04 *(.10)*	14	0.54 *(.17)*	0.16 *(.17)*
Rose (2002)	77.2	36	0.94 *(.11)*	0.63 *(.17)*	4	0.03 *(.06)*	0 *(0)*	21	0.93 *(.10)*	0.15 *(.12)*
Rasmussen, Ho & Bisanz (2003) Age 4	63.8	12	0.62 *(.08)*	0.44 *(.10)*	9	0.34 *(.11)*	0.38 *(.07)*	3	0.42 *(.01)*	0.14 *(.13)*
Rasmussen, Ho & Bisanz (2003) Age 6	83.6	4	0.98 *(.04)*	0.60 *(.08)*	8	0.31 *(.19)*	0.17 *(.11)*	12	0.92 *(.09)*	0.11 *(.10)*
Canobi (2005)	68.6	26	0.94 *(.13)*	0.77 *(.30)*	22	0.24 *(.29)*	0.05 *(.12)*	12	0.33 *(.28)*	0.67 *(.20)*
Gilmore (2005) Exp. 2	78.8	10	0.90 *(.13)*	0.38 *(.13)*	7	0.11 *(.13)*	0.11 *(.20)*	6	0.63 *(.21)*	0 *(0)*
Gilmore (2005) Exp. 5	81.3	17	0.91 *(.08)*	0.79 *(.13)*	17	0.25 *(.15)*	0.15 *(.13)*	19	0.81 *(.12)*	0.33 *(.18)*

Gilmore (2006) Exp. 1	73.5	11	0.93 *(.08)*	0.66 *(.13)*	14	0.39 *(.20)*	0.12 *(.07)*	24	0.82 *(.11)*	0.31 *(.13)*
Gilmore (2006) Exp. 2	73.6	17	0.91 *(.08)*	0.63 *(.15)*	6	0.15 *(.09)*	0.03 *(.02)*	28	0.64 *(.19)*	0.25 *(.11)*
Gilmore & Bryant (2006) Exp. 1	70.9	30	0.92 *(.09)*	0.46 *(.14)*	26	0.40 *(.18)*	0.19 *(.13)*	10	0.84 *(.09)*	0.16 *(.06)*
Gilmore & Bryant (2006) Exp. 2 Age 8	75.8	7	0.77 *(.13)*	0.56 *(.17)*	4	0.21 *(.16)*	0.10 *(.08)*	18	0.96 *(.08)*	0.30 *(.14)*
Gilmore & Bryant (2006) Exp 2 Age 6	86.4	11	0.94 *(.07)*	0.19 *(.14)*	14	0.24 *(.12)*	0.04 *(.05)*	5	0.63 *(.10)*	0.02 *(.04)*
Gilmore & Bryant (in press)	68.4	26	0.95 *(.07)*	0.73 *(.12)*	21	0.58 *(.21)*	0.21 *(.20)*	20	0.94 *(.08)*	0.38 *(.11)*
Sherman & Bisanz (2007)	69.1	4	0.54 *(.05)*	0.29 *(.05)*	4	0.06 *(.08)*	0.17 *(.12)*	16	0.37 *(.08)*	0.10 *(.08)*

Figure Captions

Figure 1. Forest plot revealing the mean difference in accuracy scores between inversion and

control problems.

Study name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value		Std diff in means and 95% CI
Canobi (2005) Age 6	0.000	0.316	0.100	-0.620	0.620	0.000	1.000		
Canobi (2005) Age 5	0.157	0.317	0.100	-0.464	0.778	0.496	0.620		
Robinson, Ninowski & Gray (2006) Age 11	0.450	0.209	0.044	0.041	0.859	2.155	0.031		
Canobi (2005) Age 7	0.493	0.321	0.103	-0.136	1.122	1.536	0.125		
Gilmore (2005) Exp. 5 Age 9	0.650	0.269	0.073	0.122	1.178	2.413	0.016		
Robinson, Ninowski & Gray (2006) Age 13	0.682	0.222	0.049	0.247	1.117	3.073	0.002		
Rasmussen, Ho & Bisanz (2003) Age 4	0.734	0.298	0.089	0.150	1.319	2.462	0.014		
Bryant, Christie & Rendu (1999) Exp. 2	0.751	0.154	0.024	0.448	1.053	4.867	0.000		
Rose (2002) Age 8	1.061	0.315	0.099	0.444	1.678	3.369	0.001		
Bryant, Christie & Rendu (1999) Exp. 1	1.161	0.153	0.023	0.862	1.461	7.597	0.000		
Gilmore (2005) Exp. 5 Age 8	1.177	0.313	0.098	0.564	1.790	3.764	0.000		
Gilmore (2006) Exp. 2	1.247	0.216	0.047	0.823	1.671	5.762	0.000		
Gilmore (2005) Exp. 2	1.286	0.324	0.105	0.652	1.921	3.971	0.000		
Bryant, Christie & Rendu (1999) Exp. 2 Age 6	1.330	0.359	0.129	0.627	2.032	3.709	0.000		
Sherman & Bisanz (2007)	1.483	0.326	0.106	0.844	2.122	4.550	0.000		
Gilmore & Bryant (2006) Exp. 1	1.502	0.099	0.010	1.309	1.695	15.247	0.000		
Gilmore & Bryant (in press)	1.511	0.196	0.038	1.127	1.895	7.715	0.000		
Gilmore (2006) Exp. 1	1.530	0.133	0.018	1.270	1.790	11.536	0.000		
Bryant, Christie & Rendu (1999) Exp. 2 Age 7	1.836	0.387	0.150	1.078	2.595	4.747	0.000		
Gilmore & Bryant (2006) Exp. 2 Age 6	1.858	0.309	0.095	1.252	2.463	6.013	0.000		
Bryant. Christie & Rendu (1999) Exp. 2 Age 8	1.889	0.412	0.170	1.080	2.697	4.580	0.000		
Rasmussen, Ho & Bisanz (2003) Age 6	1.899	0.348	0.121	1.218	2.581	5.462	0.000		
Gilmore & Bryant (2006) Exp. 2 Age 8	1.979	0.321	0.103	1.351	2.608	6.175	0.000		
Rose (2002) Age 10	2.186	0.400	0.160	1.402	2.969	5.469	0.000		
Rose (2002) Age 9	3.407	0.508	0.258	2.412	4.403	6.708	0.000		
	1.230	0.045	0.002	1.142	1.319	27.187	0.000		

Std diff in means and 95% CI — scale: -4.00, -2.00, 0.00, 2.00, 4.00