Simple Mental Addition in Children with and without Mild Mental Retardation

Rianne Janssen, Paul De Boeck, Mieke Viaene, and Lies Vallaey

University of Leuven, Leuven, Belgium

The speeded performance on simple mental addition problems of 6- and 7-year-old children with and without mild mental retardation is modeled from a person perspective and an item perspective. On the person side, it was found that a single cognitive dimension spanned the performance differences between the two ability groups. However, a discontinuity, or “jump,” was observed in the performance of the normal ability group on the easier items. On the item side, the addition problems were almost perfectly ordered in difficulty according to their problem size. Differences in difficulty were explained by factors related to the difficulty of executing nonretrieval strategies. All findings were interpreted within the framework of Siegler’s (e.g., R. S. Siegler & C. Shipley, 1995) model of children’s strategy choices in arithmetic. Models from item response theory were used to test the hypotheses. © 1999 Academic Press

Key Words: mental arithmetic; cognitive development; problem-size effect; mild mental retardation; item response theory.

Accurate and effortless performance on simple mental addition problems is an important but difficult developmental task for children in the first grades of primary education. In this paper we focus on this developmental task from two complementary perspectives. First, the nature of the development of addition skill is investigated by looking at the differences in performance of 6- and 7-year-old children with and without mild mental retardation. This approach can be described as modeling simple mental addition from the *person perspective*. Second, the nature of the simple mental addition task itself is investigated by looking at the differences in difficulty among addition problems. This approach can be described as modeling simple mental addition from the *item perspective*. Both perspectives on modeling simple addition can be inferred from the work of Siegler on a general model of children’s strategy choices in arithmetic (e.g.,

This paper was prepared while the first author was a postdoctoral researcher (PDM/97/042) funded by the University of Leuven (Belgium) and a postdoctoral Fellow of the Fund for Scientific Research–Flanders (Belgium). We thank Irwin Leenen and Wim Fias as well as the special issue editor for their comments on a previous version.

Address correspondence and reprint requests to Rianne Janssen, Department of Psychology, University of Leuven, Tiensestraat 102, B-3000 Leuven, Belgium. E-mail: rianne.janssen@ped.kuleuven.ac.be.

Siegler & Jenkins, 1989; Siegler & Shipley, 1995). In the present paper we show that both perspectives can be studied simultaneously using models from item response theory (IRT). In the following, we first present Siegler's developmental account of the automatization of addition facts. We then summarize previous research on modeling simple mental addition from the person perspective and from the item perspective. Finally, the design, IRT modeling, and hypotheses of the present study are described.

A DEVELOPMENTAL ACCOUNT OF THE AUTOMATIZATION OF ADDITION FACTS

Many studies have shown that children rely on several strategies to solve mental addition problems (Carpenter & Moser, 1984; Goldman, Mertz, & Pellegrino, 1989; Groen & Parkman, 1972; Groen & Resnick, 1977; Siegler, 1987; Siegler & Shrager, 1984). *Counting* strategies may involve counting fingers or verbal counting. The counting strategy itself may consist in counting both addends (the sum or counting-all strategy), counting the second addend starting from the first (the max or counting-on strategy), or counting the smaller addend from the larger (the min or counting-on-from-larger strategy). *Retrieval* strategies are used in the direct retrieval of number facts from memory. Retrieval may also be used in a decomposition strategy, in which a problem is split into simpler problems. For example, $6 + 7$ can be solved by decomposing it into $6 + 6 + 1$ (Svenson, Hedenborg, & Lingman, 1976; Svenson & Sjoburg, 1983). All strategies other than direct retrieval are commonly labeled as backup strategies.

Children show a transition from counting strategies to strategies relying on some form of retrieval (e.g., Cooney, Swanson, & Ladd, 1988; Geary, Bow-Thomas, & Yao, 1992; Geary & Brown, 1991; Geary, Brown, & Samaranayake, 1991). However, the developmental progression of addition strategies cannot be characterized by the simple succession of more mature addition strategies (Ashcraft, 1982). Rather, "development involves changes in the mix of existing strategies as well as the construction of new ones and the abandonment of old ones" (Siegler & Jenkins, 1989, p. 27). "In all, the normal development of skill in addition should reflect an increase in the use of the memory-retrieval strategy, with fewer retrieval errors, and a decrease in the use of counting and other backup strategies" (Geary et al., 1991, p. 788). These authors call this a "developmental shift in the distribution of strategy choices" (p.788).

Siegler's account of mathematical skill development in children (1988a; Siegler & Jenkins, 1989; Siegler & Shipley, 1995; Siegler & Shrager, 1984) provides a parsimonious framework for the above findings on how children learn to solve simple mental addition problems. The framework consists of two interrelated parts: a distribution-of-associations model and a model for adaptive strategy choices. According to the *distribution-of-associations model* for arithmetic knowledge, problems are associated with both correct and incorrect answers. Due to extensive practice, the shape of the distribution of associations

changes from a flat distribution to a distribution that is “peaked,” so that the association between a problem and the correct answer becomes markedly stronger than associations with incorrect answers. The more peaked the distribution of associations becomes, the more quickly and accurately an answer can be given. The peakedness of the distribution plays an important role in the *adaptive strategy choice* for a problem. The more peaked the distribution, the more likely a retrieval strategy will be used to solve the problem as compared to a backup strategy. The frequent and correct use of counting strategies increases the strength of association between a problem and its correct answer, thus leading to the eventual extinction of the strategies by ensuring that memory retrieval becomes the dominant problem-solving strategy (Siegler & Jenkins, 1989). According to Siegler’s model, the actual strategy chosen for solving a problem also depends on a child’s confidence criterion and search length criterion. The latter variables are of less importance for the present study.

Looking at simple mental addition problems from the person perspective, Siegler’s theoretical framework suggests that there is one developmental path to the acquisition of automated performance. This developmental path is based on the interplay between the distribution of associations and the adaptive strategy choice for a problem: Practice with a variety of backup strategies of increasing complexity not only changes the distribution of associations, but also makes a retrieval strategy more likely. Hence, given this one developmental path, individual differences among children in their performance on simple additions should be interpretable within this framework. On the other hand, when looking at simple mental addition problems from the item perspective, Siegler’s model states that, at any moment in development, a child may rely on a variety of strategies to solve addition problems, but the probability that a certain strategy is used may differ from problem to problem. Hence, Siegler’s model may also be used to study differences among simple mental addition problems. In the following, previous research findings regarding the person perspective and item perspective on modeling simple mental addition are discussed.

MODELING SIMPLE MENTAL ADDITION FROM THE PERSON PERSPECTIVE

Geary and his colleagues (Geary & Brown, 1991; Geary et al., 1991, 1992) have successfully applied Siegler’s framework to the study of individual differences in mathematical skill development in children from the early grades of normal primary education. Geary and Brown (1991) found that gifted, normal, and mathematically disabled children differed in the strategies used to solve the problems and in the rate of information processing. The gifted group was characterized by a very developmentally mature strategy mix, followed in turn by the normal group and the mathematically disabled group. The gifted group also showed a verbal counting rate that was already within the adult range and that took less than half of the time in comparison to the counting rate estimated for

the two nongifted groups. In a longitudinal study, Geary et al. (1991) showed that a group of normal children from first and second grade showed an increased reliance on memory retrieval and a decrease in the use of counting strategies, as well as an increase in the speed of counting and of retrieving addition facts from long-term memory. During the 10-month study, the mathematically disabled children did not show a change in their mix of addition strategies nor in their speed of executing counting or memory-retrieval strategies. Geary et al. (1991) also found differences between the normal and mathematically disabled in working memory resources and suggested that this factor may contribute to a mathematical disability (see also Geary, 1990). Finally, Geary et al. (1992) showed that the relatively poor computational skills in cognitive addition of mathematically disabled children from first grade are related to a developmental delay in the understanding of basic principles and properties of counting, as evidenced by their relatively poor skills at detecting certain forms of counting error.

To summarize, Geary and his colleagues found that children of different ability groups differed in their effectiveness of counting and addition strategies and in their speed of executing cognitive processes involved in those strategies. These differences in strategy use correlated with differences in the level of automatization of basic addition facts, which was inferred from the speed and reliability of memory retrieval. These findings suggest that children from different ability groups indeed follow the same developmental path, as predicted by Siegler's model. Geary and Brown (1991) even hypothesized that a single cognitive dimension spanned the performance differences in cognitive addition between gifted, normal, and mathematically disabled children of third and fourth grade. They interpreted this dimension as the maturity of the long-term memory organization of basic addition facts. In terms of Siegler's model, this memory organization refers to the peakedness of the underlying distributions of associations between a problem and its possible answers.

MODELING SIMPLE MENTAL ADDITION FROM THE ITEM PERSPECTIVE

A pervasive and robust phenomenon in mental arithmetic is that problems involving larger numbers (e.g., $7 + 8$) are solved more slowly than problems involving smaller numbers (e.g., $2 + 3$), both by children and by adults (see, e.g., Ashcraft, 1992; McCloskey, Harley, & Sokol, 1991; Miller, Perlmutter, & Keating, 1984; Siegler, 1988b; Widaman & Little, 1992). This phenomenon is called the problem-size effect. Tie problems (e.g., $4 + 4$) form an exception as they are more easily processed than the size of their sum would predict. Since the seminal work by Groen and Parkman (1972), numerous theoretical explanations have been proposed. Most explanations refer to the activation characteristics of a network of stored numerical and arithmetic facts. For example, Widaman and his colleagues (Widaman, Geary, Cormier, & Little, 1989; Widaman & Little,

1992; Widaman, Little, Geary, & Cormier, 1992) considered arithmetic knowledge to be stored in a tablelike, symmetric matrix, in which the correct answer to a problem is stored at the interconnection of two entry nodes. When retrieving an arithmetic fact, the spread of activation rate in the network is a function of the area in the network activated, which is equal to the product of the addends of the sum. More recently, Campbell (1995) proposed a network-interference theory. This theory explains the problem-size effect by the interfering, automatic activation of analog representations of the magnitudes associated with the integers.

LeFevre, Sadesky, and Bisanz (1996) applied Siegler's theory to account for the problem-size effect in adults. They showed that the problem-size effect is partly an epiphenomenon of the relationship between problem characteristics and the selection of solution strategies. For larger problems the probability that respondents select a nonretrieval strategy is larger than for smaller problems, resulting in a higher solution latency when averaging across respondents and across strategies. Hence, LeFevre et al. (1996) provided evidence against the implicit assumption in network models that adults always retrieve on simple arithmetic problems. However, on retrieval trials, LeFevre et al. (1996) still found a (reduced) problem-size effect. This finding may be explained by a network retrieval model. In line with Siegler (1988a) and Siegler and Shipley (1995), LeFevre et al. (1996) proposed that retrieval latencies may reflect the use of nonretrieval strategies in the course of development. When children learn addition facts, they primarily rely on counting procedures. Associations between problems and answers are strengthened when correct answers are generated. Counting errors can be assumed to be more likely for problems with larger addends, because these problems need more counting. Hence, the strengthening of the problem-answer associations is likely to be more delayed on larger problems than on smaller problems. Consequently, larger problems are solved more slowly than smaller problems because of their less peaked distribution of associations. In support of their explanation, LeFevre et al. (1996) showed that the percentage use of nonretrieval strategies for a given problem was almost as powerful in predicting response latencies for individual participants on retrieval trials as were structural measures of problem size (e.g., the product or sum of addends).

This account of the problem-size effect provides a nice example of the continuity of development from childhood to adulthood. Given this account, one can also predict that the problem-size effect can be eliminated with extensive practice. Once problem-answer associations have been developed for all simple arithmetic problems, retrieval speeds should be uniform across all simple arithmetic problems, regardless of the magnitude of the integers in the problem. To test this hypothesis, Geary (1996) studied kindergarten and elementary school children from mainland China and the United States, because children in East Asian countries receive much more exposure to basic arithmetic facts. Geary (1996) found a problem-size effect in all the North American samples, but the

Chinese children showed a problem-size effect only in kindergarten and at the beginning of the first grade. The problem-size effect had disappeared in the Chinese children at the end of the first grade, and by the end of the third grade the large-valued addition problems were even solved more quickly than the smaller valued problems. The latter finding was explained by the larger emphasis on larger valued problems in comparison with smaller valued problems in the Chinese workbooks for the first, second, and third grade. This contrasts with the previously found emphasis on smaller valued problems in early education in North America (Ashcraft & Christy, 1995; Hamann & Ashcraft, 1986).

THE PRESENT STUDY

A Speeded-Response Task

In the present study, two-term addition problems with a maximum sum of 5 were administered to 6- and 7-year-old children with and without mild mental retardation. In contrast with the majority of studies on mental addition, response accuracy instead of response latency was chosen as the dependent variable. This allows for a collective administration of the task. To prevent ceiling effects and to increase the likelihood that answers would depend on retrieval rather than backup strategies, a speeded-performance task was used in which the children were given a maximum of 4 s to encode and solve each problem and to write down their answer (cf. Siegler & Robinson, 1982).¹

It is assumed that differences in the probability of a correct response on the speeded-response task reflect the combined and interrelated influence of the differences in the peakedness of the underlying distributions of associations between problems and answers and of the differences in the developmental maturity of the strategy mix. First, the probability of a correct response can be expected to be proportional to the peakedness of the distribution. Flat distributions should generally lead to low probabilities of a correct response, while peaked distributions should generally lead to high probabilities. Second, as Geary and Brown (1991, p. 400) stated, "the developmental maturity of the strategy mix can be indexed by the frequency with which answers can be correctly retrieved from long-term memory." Hence, it is assumed that children having a more mature strategy mix have a higher probability of correctly responding on the speeded-response task. In line with the previous research on modeling simple mental addition problems from the person and from the item perspective, one can expect that, on the one hand, the probability of a correct response will vary between the children for a given problem, reflecting the individual differences in development. On the other hand, the probability of a correct response can also be expected to vary between the problems within each child, reflecting the differences in difficulty of the mental addition problems themselves.

¹ Use of a time limit on responses does not preclude the use of backup strategies entirely, however, as demonstrated by Baroody (1999).

IRT Modeling

The probability of a correct response is modeled using IRT. This psychometric approach is most commonly known only as a measurement device, but several psychometricians have argued that it can also be a viable research tool in cognitive psychology (e.g., Embretson, 1983; Mislevy, 1996) and in developmental psychology (e.g., von Eye & Clogg, 1994). In the most simple IRT model, namely the one-parameter logistic model or Rasch model (see, e.g., Fischer & Molenaar, 1995), the probability of a correct response for a person j on an item i is modeled on the basis of a person ability parameter θ_j and an item difficulty parameter β_i , which are both estimated on the same continuous scale. More specifically, in the Rasch model, the probability of a correct response is a function of the difference between θ_j and β_i , which can be written formally as

$$P(X_{ij} = 1) = f(\theta_j - \beta_i). \quad (1)$$

The function f is the logistic function. This S-shaped function transforms any value on the real line into a value between 0 and 1. One can infer from Eq. 1 that the higher the ability, the higher the probability of success, and the higher the difficulty, the lower the probability of success. When θ_j equals β_i , the probability of success is .50. The person ability and item difficulty parameters refer to the same latent continuum. This latent continuum is thus defined in two ways: as the dimension on which the persons are located in increasing order of ability and as the dimension on which the items are located in increasing order of difficulty. Hence, individual differences in the probability of a correct response are reflected in the θ parameter and differences in the probability of a correct response among the items are reflected in the β parameter.

When used to model developmental differences, the Rasch model implies that development consists in a continuous progression along one common path that is represented by the underlying latent scale. Higher development corresponds to higher probability of success in all problems of the same domain. Hence, the Rasch model only allows modeling quantitative change. However, a variant of the Rasch model, the so-called Saltus model (Wilson, 1989; Mislevy & Wilson, 1996), allows modeling of possibly qualitative change in development as well, without sacrificing the idea of one common underlying continuum. Formally, when comparing two groups of persons, the Saltus model states that the difficulty parameter changes by a certain amount for a subset of items in one group,

$$P(X_{ij} = 1) = f(\theta_j - \beta_i + \tau), \quad (2)$$

in which τ denotes the change in difficulty, also called the saltus parameter, and f is the logistic distribution function. A positive value of τ implies that the items to which the saltus parameter pertains all become easier to the same extent in that group. For the other items in this group and for all the items in the other group, the difficulty remains unchanged, and Eq. 1 holds. Hence, when two groups of persons who are assumed to be at different levels of development are compared,

a positive and significant value of the τ parameter for a subset of items in the more developed group may reflect a discontinuity in development and, hence, possibly qualitative change. The change consists in part of the items being easier. It is situated on the same ability dimension, and, hence, supplements a progression along the same latent scale. The jump made by the group at the higher developmental level may, for example, stem from automatization progressing along the continuum, so that the easier items become even easier.

Hypotheses

The person perspective. We investigated the hypothesis that children with and without mild mental retardation follow the same developmental path in learning to solve mental addition problems as outlined by Siegler's account of mathematical skill development. Children with mental retardation may be characterized by a slower and possibly delayed development in mathematical skills, given their general lower level of intellectual functioning. However, it is predicted that their development is based on the same principles as the development of children without mental retardation; that is, peaked distributions of associations are built on the basis of counting and other backup strategies. In line with this prediction, Baroody (1996) showed that during 5-month practice on single-digit addition, mentally retarded children were capable of inventing shortcuts on addition strategies autonomously, and that some children began to use retrieval during the course of study.

To evaluate the hypothesis of a single developmental path, the performance of children with and without mild mental retardation will be modeled according to the Saltus model. When this model fits, the differences in performance in the two groups are spanned by a single cognitive dimension, as was hypothesized by Geary and Brown (1991) for children at the higher end of the intelligence continuum. As it is supposed that the speeded-response task taps the distributions of associations, this single cognitive dimension refers to the level of peakedness of these distributions and indirectly to the maturity of the strategy mix of the children, as was argued above. The significance of the saltus parameter in the Saltus model would indicate that qualitative differences, as well as quantitative differences, may exist between the two groups. Geary et al. (1991) alluded to a developmental shift in the distribution of strategy choices when memory retrieval becomes the most dominant strategy, presumably paralleling an increase in the peakedness of the distribution of associations between problems and answers. This developmental shift can be interpreted as qualitative change. For the simple mental addition problems, one can expect that the children within the normal intellectual ability range may already have shifted to retrieval for the easier problems. Hence, we hypothesize that the easier addition problems become easier for the children from normal education in comparison with the difficulties found for the children with mild mental retardation. Therefore, a τ parameter is expected that is positive and significantly different from zero for the easier addition problems in the group from normal education. For the more difficult

problems, the difficulties in the two groups of children are expected to be equal. When the Saltus model fits, but with a nonsignificant value for the saltus parameter, then the Saltus model reduces to a common Rasch model in the two groups of children. This finding would support a view of purely quantitative change along one dimension.

The item perspective. Given the pervasiveness of the problem-size effect, we expect that the difficulty of the problem in the speeded-response task correlates highly with the size of the problem. As an operationalization of problem size, the product of the addends of the sum will be used. However, in line with Siegler and Shipley (1995, p. 52) and in line with the explanation of LeFevre et al. (1996) for the problem-size effect, we hypothesize that the difficulty of executing backup strategies on a problem is a more fundamental determinant of problem difficulty than is the size of the problem as such. For example, although their problem size is the same, $4 + 1$ may be easier than $1 + 4$, both when using a counting-on strategy (add 1 versus add 4) and when using a counting-on-from-larger strategy (because of the cognitive reversal in $1 + 4$, but not in $4 + 1$). These differences in difficulty of executing a backup strategy on a problem should be reflected in the distributions of associations for the two problems in children at an early stage of mathematical skill development. When it is more difficult to apply a counting strategy successfully to a particular problem, the strength of association between the problem and its correct answer will grow more slowly. Consequently, use of retrieval should be delayed for problems to which counting strategies are difficult to apply.

In a study by Viaene (1992) with normal 6- and 7-year-old children, it was shown that the degree of difficulty of mental addition problems with a maximum sum of 10 was almost perfectly predicted ($R = .97$) by five factors that were related to the difficulty of counting strategies. First, tie problems were easier than nontie problems. Second and third, the inclusion of a 1 or of a 2 as a term in the sum made an addition problem easier, which can be related to the smaller amount of counting involved in these two kinds of sums. Fourth, the inclusion of a 5 as an addend made an addition problem easier as well, but to a lesser degree. This factor may refer to the use of the number 5 (like the number 10) as a reference number in counting, possibly originating from strategies involving counting by fingers. Finally, the size of the second term, when it was larger than the first term, made a problem more difficult. This factor may reflect the reversal operation needed in a counting-on-from-larger strategy, and also that the difficulty of reversal increases with the size of the second addend.

In the present study, the predictive value of the difficulty factors of Viaene (1992) that apply to sums up to 5 were investigated. These difficulty factors are whether or not the problem is a tie, the inclusion of a 1 or a 2 as an addend (in order to account for differences in the min and sum strategies), and the size of the second addend when it is larger than the first (referring to the cognitive reversal in the min strategy). Each item gets a score on each difficulty factor using dummy

coding (using 0 and 1 for the first three factors and 0, 2, 3, and 4 for the fourth). We tested whether the four difficulty factors referring to backup strategies can account for the item difficulty of the 10 mental addition problems using the linear logistic test model (LLTM) of Fischer (e.g., 1983, 1996). In the LLTM, the Rasch item difficulty parameters β_i of each individual item are replaced with a weighted sum of K common difficulty factors:

$$\beta_i = \sum_{k=1}^K \eta_k W_{ik} + c. \quad (3)$$

In this equation, w_{ik} represents the value of item i on the difficulty factor k and η_k is the estimated weight of the difficulty factor k . The parameter c is a normalization constant, which is relative to the origin of the latent scale but with no relevant interpretation. The LLTM looks like a regression of the β_i on the W_{ik} variables with the η_k as an equivalent to a regression coefficient and c to the intercept. However, unlike a regression model, the LLTM does not have an error term and the estimated weight parameters η_k are estimated directly, without first estimating the β_i . In comparison with a regression analysis, an LLTM analysis is also more powerful because (a) the analysis is based on all the data and not just on the proportions correct, summary statistics of which there are only as many values as there are items, and (b) the model is tested against full recovery of the item difficulties; that is, the LLTM (with K item difficulty factors) is tested against the Rasch model (with one difficulty parameter per item) to determine whether it can account for item difficulty as well as the Rasch model can. A descriptive index of fit of the LLTM can be obtained by calculating the correlation between the item difficulties obtained under the Rasch model with the item difficulties as predicted by the LLTM (cf. Eq. 2). The squared correlation coefficient tells how much of the variation in the item difficulties in the Rasch model can be accounted for by the item characteristics W_{ik} . Methods for testing the goodness of fit of the LLTM are described below.

METHOD

Respondents

All children in the study came from elementary schools in the Flemish part of Belgium. The sample of children from normal education consisted of 150 first-grade children from three elementary schools. Exactly half of the children were boys. The mean age was 6 years 7 months, with a standard deviation of 3.5 months.

The sample of children with mild mental retardation consisted of 120 children from five schools for mentally disabled children. A child is allowed admission in these schools only after extensive testing by a school psychologist, with an important criterion in educational practice being a score between 60 and 80 on the Dutch edition of the Wechsler Intelligence Scale for Children-Revised

(WISC-R Projectgroep, 1986). The addition task was administered at the moment during the school year that the individual children were judged by their teachers to be functioning at the cognitive level of children from normal education at the middle of the first grade. The actual age of the children in the sample varied between 6 years 10 months and 7 years 9 months, with a mean of 7 years 3 months. Exactly half of the children were girls.

Items

Normal ability group. The addition problems comprised all possible pairings of the digits 1 to 9 with a resulting sum smaller than or equal to 10. The 45 items were divided into two lists, consisting of the same nontie sums but in reversed order and of three ties in List 1 and two ties in List 2. The order of presentation of the items was randomized in the two lists, with the restriction that the number of times a digit appeared in the same position in an adjacent problem was kept to a minimum. Only the responses to the problems with a maximum sum of 5 were analyzed in the present study.

Group with mild mental retardation. Only the addition problems with a maximum sum of 5 were administered to the children from special education. These 10 problems were presented in one list.

Procedure

Normal ability group. The addition task was administered in a collective classroom session in the morning during the last week of January (the school year starts in September). All teachers confirmed that the instruction in arithmetic in January was concentrated on the automatization of sums with a maximum total of 10. The addition problems were read aloud and at the same time written on the blackboard. The children were given 4 s to write down their answer. They then had to lay down their pencil, fold their (numbered) response sheet, and put it in a little container in front of them. After the completion of the first list, there was a short break. The order of presentation of the items of the two lists was reversed in half of the sessions. The time limit of 4 s was determined on the basis of pretesting of some other children with different time limits. Siegler and Robinson (1982) used the same time limit to encourage children to retrieve.

Group with mild mental retardation. The administration of the 10 addition problems to the children from special education was the same as for the children from normal education, but the groups were smaller and the instructions were repeated several times before the problems were given. The order of presentation of the 10 problems was reversed in half of the sample.

Estimation of the IRT Models

All psychometric analyses were done using the computer program CONQUEST (Wu, Adams, & Wilson, 1997), which allows estimation of both the Rasch model and its extensions to a Saltus model and an LLTM. The estimation

procedure used in CONQUEST is based on marginal maximum likelihood (MML). In this procedure, the item parameters are estimated first by aggregating the person ability parameter out of the likelihood using a hypothesized normal distribution for this parameter. Given the item parameter estimates, the person ability parameters are then estimable. For the Rasch model to be identified, the origin of the scale has to be fixed. Such an identification restriction has no influence on the interpretation of the scale, as differences in difficulty among the items and differences in ability among the persons remain the same, whichever identification restriction is imposed. In the analyses with CONQUEST, the origin of the scale was fixed at the mean of the technical distribution for the person ability parameter used in the MML procedure. Because the MML procedure implies that only one distribution of persons is used, between-group differences between the children with and without mild mental retardation were estimated by allowing the item difficulties to differ in the two samples. Between-group differences in the item difficulty parameters are equivalent to between-group differences in the ability parameter: When one group of children is on average more proficient than the other group, then the difficulty parameters in the former group should be lower (implying that the items are easier). The between-group differences will be modeled with a group-difference parameter, also called group parameter.

Testing of the IRT Models

The goodness of fit of the IRT models is tested in three ways. First, CONQUEST derives Wu's (1997) *parameter specific goodness-of-fit tests*. These statistics are approximate normal deviates that are based on transformation of a weighted average of squared standardized residuals. Each residual compares the observed contribution of a person to the estimator of a parameter with the corresponding expected value under the model. Taken together, the statistics give an indication of the overall goodness of fit of an IRT model. Second, more restricted models are tested against less restrictive models using the *likelihood-ratio test*. This test statistic provides a method for testing the relative goodness of fit of nested models. For example, the LLTM restricts the difficulties of a Rasch model to be linear combinations of item characteristics, whereas the item difficulties are free of restrictions under the Rasch model. When the LLTM gives an adequate account of the item difficulties of a Rasch model, the decrease in likelihood due to the smaller number of parameters in the LLTM (only $K \eta_k$ parameters) in comparison with the Rasch model (as many difficulty parameters as items) must not be significant. Finally, the necessity of including a saltus parameter in the Saltus model and of including a certain predictor variable in the LLTM can be tested by looking at *the statistical significance of the corresponding parameter*. A significance test can be performed by calculating the ratio of the estimated parameter value to its estimated standard error. This ratio corresponds to a t statistic and, hence, is asymptotically distributed as a standard

FIG. 1. Scatter plot of the logits of the mean proportion of success in the two samples of children.

normal distribution. As a rule of thumb, ratios that are greater than 2 in absolute value are significant at about the 5% level.

RESULTS

Descriptive Summary

The proportion of correct answers on the 10 addition problems with a sum smaller than or equal to 5 ranged from .79 to .97 (with a mean of .91) in the sample of normal children and from .24 to .65 (with a mean of .45) in the sample of children with mild mental retardation. Cronbach's alpha coefficient of internal consistency was .85 in the normal sample and .91 in the mentally retarded sample.

Figure 1 presents a plot of the logits of the proportion of correct answers for the addition problems in the two samples of children. The logit is a common transformation of proportions and probabilities to the real line (Andrich & Nelson, 1984). Technically, the logit of a proportion equals the natural logarithm of the ratio of the proportion to one minus the proportion. The difference in the mean proportion of correct answers in the two groups of children is inferable from the different range of the logits on each axis. Figure 1 gives an indication of a discontinuity in the mastery of the problems for the children from normal education. If one would fit a line through the positions of the problems of the five more difficult problems and through the positions of the five easier problems,

TABLE 1
Overview of the Estimated IRT Models

Description	Label	Parameters ^a	Number of parameters ^a
Unrestricted Rasch Model	Model 0	10 × 2 item difficulty	21
Saltus model	Model 1	10 item difficulty Saltus, group, constant <i>c</i>	14
LLTM with product	Model 2	Product Saltus, group, constant <i>c</i>	5
LLTM with strategy difficulty factors	Model 3	“One,” tie, reversal, “1 + 2” Saltus, group, constant <i>c</i>	8

^a In all models, the variance of the assumed latent distribution of person ability parameters is also estimated.

then one would notice a discontinuity, with the line for the easy items in a higher position than the line for the difficult items. In the normal ability sample, the problems 2 + 2 and 3 + 1 are about equally as difficult as the problem 2 + 1, and they are clearly easier than the problems 1 + 3 and 4 + 1. For the children with mild mental retardation, the transition between the logits of the more difficult items and the logits of the easier items is more gradual. The significance of the saltus parameter in the Saltus model will provide a statistical test for this phenomenon.

Psychometric Analyses

Table 1 presents an overview of the different IRT models estimated. These models are now discussed consecutively.

Null model. First, an unrestricted Rasch model was estimated in which the item difficulty parameters were free to vary in the two groups of children. The model tests whether the children with and without mild mental retardation can be situated on one latent continuum per group. This unrestricted Rasch model was labeled Model 0, as it can be used as a reference model to test the other, more restrictive models using likelihood-ratio tests. Model 0 allows for group-specific dimensions, one in each group, and, hence, it allows for all kinds of qualitative differences, without imposing any kind of change along one common continuum for the two groups. The fit of Model 0 was reasonable. Wu’s (1997) parameter specific goodness-of-fit tests were significant for 2 of the 20 item difficulty parameters, but only at the 5% level. The mean item difficulty parameter for the 10 problems was −4.32 for the children from normal education and .46 for the children with mild mental retardation, meaning that the former group performed clearly better than the latter.

Modeling between-group differences. The Saltus model was estimated in order to test whether the differences between the two groups of children are

purely quantitative or whether a potentially qualitative difference (due to a saltus) can be found. This model is labeled Model 1. In line with Fig. 1, a saltus parameter was defined for the five easier items ($2 + 2$, $3 + 1$, $2 + 1$, $1 + 2$, and $1 + 1$), allowing for a saltus on those items in the normal ability group. In comparison with Model 0, equality restrictions were imposed on the item difficulty parameters of corresponding problems in the two groups of children, but the saltus parameter corrects for this equality when the five easy items are concerned. A group parameter was included to account for the differences in mean difficulty (and, hence, in mean ability) in the two samples.

Table 2 presents the results of this analysis. In this table, the items are ordered from easy to difficult on the basis of the estimated β_i parameter. The parameter specific goodness-of-fit tests showed that only for the problem $2 + 2$ a significant weighted standardized residual was obtained at the 1% significance level, meaning that the goodness of fit was poor for this item. A likelihood-ratio test of the Saltus model against Model 0 was not significant, $LR = 3.42$, $df = 7$, $p = .84$, meaning that the Saltus model explains the data equally as well as does the unrestricted null model (Model 0), although it uses seven fewer parameters. The results of the individual parameter estimates and their standard errors (see Table 2) show that the saltus parameter with a value of .48 is significantly different from zero (as its absolute value is more than twice its standard error). Its value implies for example that the item difficulty for the problem $3 + 1$ equals -1.50 for the children with mild mental retardation, but -1.98 for the children without mental retardation. Therefore, one may conclude that the change from one group to the other is not purely quantitative. The value of the group parameter, with a highly significant value, is of comparable size to the group difference found under Model 0. Its negative sign indicates that the problems are easier for the children from normal education; hence, children from normal education show on the average a higher proficiency parameter θ_j . In addition, given the saltus parameter, the easier problems are at the same time also easier for those children. The intercept parameter refers to the normalization constant in the model and does not have a relevant interpretation.

Investigating the problem-size effect. The estimated item difficulties in Model 1 increased in difficulty according to problem size, as can be seen in Table 2. The problem $4 + 1$ formed the only exception, as it was estimated to be easier than $2 + 2$ and $1 + 3$. Note that problem size also does not explain the variability among the problems with the same addends (like the differences between $1 + 2$ and $2 + 1$, or between $4 + 1$ and $1 + 4$). In order to investigate how well the problem-size effect can account for differences in difficulty, the product of the addends was used as the predictor of item difficulty in an LLTM, together with the saltus parameter and the group parameter as defined in Model 1. This model was labeled Model 2 and its results are given in Table 2 as well. The parameter specific goodness-of-fit tests did not reveal any lack of fit. A likelihood-ratio test

TABLE 2
Results of the Psychometric Models 1 to 3

Parameter	Estimate	Standard error	Standardized residual
Model 1			
1 + 1	-2.94	.25	-0.24
1 + 2	-2.70	.24	0.42
2 + 1	-2.08	.23	-1.35
3 + 1	-1.50	.22	-1.24
4 + 1	-1.50	.22	-1.01
2 + 2	-1.45	.22	-2.37
1 + 3	-1.22	.21	0.31
1 + 4	-0.69	.21	1.94
2 + 3	0.34	.20	-0.18
3 + 2	0.41	.20	-1.38
Group	-4.57	.12	0.18
Saltus	-0.48	.22	0.48
Constant <i>c</i>	1.79	.07	-0.20
Model 2			
Product	0.62	.05	0.28
Group	-4.45	.38	-0.03
Saltus	-0.73	.30	0.10
Constant <i>c</i>	-1.72	.61	0.44
Model 3			
One as addend	-1.65	.17	1.47
Reversal	0.15	.05	1.89
Tie	-1.41	.22	-0.07
1 + 2	-1.37	.27	0.73
Group	-4.46	.38	-1.17
Saltus	-0.69	.31	0.07
Constant <i>c</i>	1.85	.57	1.21

of Model 2 against the Saltus model without restricting the difficulties to be a linear function of the product (Model 1) was significant: $LR = 19.28, df = 9, p = .02$. Hence, the part of Model 2 that states the reduction of the 10 item difficulty parameters to the one product variable has to be rejected at the 5% level. Nevertheless, Model 2 still recovers the differences among the items of the unrestricted model rather well. As a descriptive goodness-of-fit index of Model 2, the correlation was calculated between the difficulty parameters obtained under Model 2 and those obtained under Model 0. Its value was .991, corresponding to 98.2% of the variance in the item difficulties in the unrestricted model being explained. The significant and positive value of the weight for the

product (see Table 2) indicates that the higher the product of its addends, the greater the difficulty of the problem. Note that in Model 2 the estimated value of the group parameter remains about equal in comparison with Model 1, but that the saltus parameter increased a little.

Modeling item difficulty using strategy difficulty factors. The hypothesis of whether the differences in item difficulty can be explained by referring to the difficulty of backup strategies was investigated in three LLTM analyses. In the first LLTM, the following four difficulty factors were used: being a tie problem, inclusion of a 1 or 2 as an addend, and the value of the second addend when it was larger than the first. In this model, the estimated weight for the factor whether a 2 was an addend was not significantly different from zero. However, when this factor was left out in the second LLTM, a large difference between the Rasch item difficulty and the predicted difficulty was found for the problem $1 + 2$. Therefore, a factor specific for the $1 + 2$ problem was included (defined as whether the problem is $1 + 2$ or not) in the third LLTM, which is labeled Model 3. The results in Table 2 show that the parameter specific goodness-of-fit tests were again nonsignificant. The likelihood-ratio test of Model 3 against Model 1 was nonsignificant ($LR = 10.49$, $df = 6$, $p = .11$), implying that the reduction of the 10 item difficulty parameters to the four strategy difficulty factors was successful. The descriptive index of goodness of fit consisting of the correlation between the item difficulty parameters obtained under Model 3 and under Model 0 was .995, corresponding to 98.9% of the variance in item difficulty under Model 0 explained. The estimated weight parameters for each of the difficulty factors were all significantly different from zero. The inclusion of 1 as an addend and the problem being a tie made an addition problem easier. The size of the second integer when it is larger than the first made a problem more difficult. Note that the estimated weight for the latter factor seems small in comparison with the other two factors, but the value of this parameter has to be multiplied with the value of the second, larger addend in order to derive the impact of this factor on item difficulty. The negative value for the weight of the $1 + 2$ factor indicates that this problem was relatively easier than the other addition problems when the other item difficulty factors are kept constant. A possible explanation for this factor is that the counting string association (1, 2, 3) facilitated learning of this problem (Siegler & Shipley, 1995).

Models 2 and 3 cannot be compared directly with a likelihood-ratio test, as these models are not nested. Both models accounted for a very high percentage of the variance in the item difficulty of the unrestricted Rasch model. However, Model 2 was rejected against Model 1, whereas Model 3 was not.

DISCUSSION

The present study provides evidence that children with and without mild mental retardation follow the same developmental path in acquiring automated performance on simple mental addition problems, as was inferred from Siegler's

account of mathematical skill development. This developmental path was studied from two complementary perspectives.

First, it was shown that the performance differences between children with and without mild mental retardation were spanned by a single cognitive dimension. However, the hypothesis of purely quantitative differences between the two groups had to be modified because of the significance of the saltus parameter. Children in normal education made a jump on the easier problems, presumably indicating a developmental shift toward using retrieval, as was predicted on the basis of Geary et al. (1991). It should be noted that a longitudinal study would be needed to confirm whether the saltus parameter presently used to study between-group differences can also be given a within-person interpretation. In a longitudinal perspective, the developmental shift in the distribution of strategy choices may be a more gradual process than the discontinuity in the performance between children of different ability levels suggests.

Second, it was shown that when the simple mental addition problems were modeled from the item perspective, the problem-size effect was found on the speeded-response task in both groups of children. Problem size, as operationalized by the product of the addends, explained the differences in the item difficulties about equally as well as did a model based on the difficulty of backup strategies. This finding is in line with the conjecture of Siegler and Shipley (1995) that the product is the best single predictor of problem difficulty because it reflects the combined influence of the most frequently used backup strategies (namely, the sum and min strategy) in shaping the strengths of problem-answer associations.

In the field of mental retardation there has already been a long controversy (e.g., Zigler & Balla, 1982) about whether mental retardation is best understood by the general principles of developmental psychology or whether it is necessary to invoke specific deficits over and above a generally lower rate and asymptote of cognitive development. Examples of such cognitive deficits are a slower speed of processing (e.g., Kail, 1992) and limited working memory resources (e.g., Ferretti & Cavalier, 1991). The present paper shows that the development versus deficit view may not be so incompatible with respect to mental addition skills. The deficits of children with mild mental retardation may account for slower development and for lagging behind that implies also qualitative differences, for example, because automatization has not yet progressed as much as in the group without mental retardation. However, the deficits may concern a general dimension, such as working memory resources, on which children within the normal ability range also differ from one another. In the same vein, Geary and his colleagues have argued that differences in performance on simple mental addition problems between children at the higher end of the intelligence continuum parallel differences in working memory resources and speed of counting. It is exactly on such cognitive measures that others have found deficits for children with mental retardation (cf. Hale & Borkowski, 1991).

All these findings suggest the following general picture of individual differences in mathematical skill development. It is a common finding that differences in general intelligence are related to differences in working memory resources (Carpenter, Just, & Shell, 1990; Embretson, 1995; Kyllonen & Christal, 1990) as well as in speed of processing (Jensen, 1998). The latter differences may even mediate the relation between working memory and intelligence (Fry & Hale, 1996). Working memory plays an important role in arithmetic problem solving in adults (e.g., Logie, Gilhooly, & Wynn, 1994), and differences in working memory resources may already induce differences in counting knowledge and counting strategies among young children (cf. Geary et al., 1991). Hence, at the beginning of primary education, children of higher intellectual ability may already be advanced in their use of counting strategies. This gives them a head start in learning basic addition facts because the frequent and correct use of counting strategies influences the distribution of associations between a problem and its possible answers, and it also forms a prerequisite for the later use of memory retrieval in mathematical problem solving. Children of higher cognitive abilities will therefore in general show a higher level of automatization of basic addition facts, as evidenced by a quicker and more reliable retrieval from memory.

REFERENCES

- Andrich, J. H., & Nelson, F. D. (1984). *Linear probability, logit, and probit models* (Series: Quantitative Applications in the Social Sciences No. 07-045). Beverly Hills, CA: Sage.
- Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review*, **2**, 213–236.
- Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, **44**, 75–106.
- Ashcraft, M. H., & Christy, K. S. (1995). The frequency of arithmetic facts in elementary texts: addition and multiplication in grades 1–6. *Journal for Research in Mathematics Education*, **26**, 396–421.
- Baroody, A. J. (1996). Self-invented addition strategies by children with mental retardation. *American Journal on Mental Retardation*, **101**, 72–89.
- Baroody, A. J. (1999). The roles of estimation and the commutativity principle in the development of third graders' mental multiplication. *Journal of Experimental Child Psychology*, **74**, 157–193.
- Campbell, J. I. D. (1995). Mechanisms of simple addition and multiplication: A modified network-interference theory and simulation. *Mathematical Cognition*, **1**, 121–164.
- Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review*, **97**, 404–431.
- Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, **15**, 179–202.
- Cooney, J. B., Swanson, H. L., & Ladd, S. F. (1988). Acquisition of mental multiplication skill: Evidence for the transition between counting and retrieval strategies. *Cognition and Instruction*, **5**, 323–345.
- Embretson, S. E. (1983). Construct validity: Construct validity in psychological tests. *Psychological Bulletin*, **93**, 179–197.
- Embretson, S. E. (1995). The role of working memory capacity and general control processes in intelligence. *Intelligence*, **20**, 169–189.
- Ferreti, R. P., & Cavalier, A. R. (1991). Constraints on the problem solving of persons with mental

- retardation. In N. W. Bray (Ed.), *International review of research in mental retardation* (Vol. 17, pp. 153–192). San Diego: Academic Press.
- Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika*, **48**, 3–26.
- Fischer, G. H. (1996). The linear logistic test model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 225–243). New York: Springer-Verlag.
- Fischer, G. H., & Molenaar, I. W. (Eds.). (1995). *Rasch models: Foundations, recent developments, and applications*. New York: Springer-Verlag.
- Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science*, **7**, 237–241.
- Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, **49**, 363–383.
- Geary, D. C. (1996). The problem-size effect in mental addition: Developmental and cross-national trends. *Mathematical Cognition*, **2**, 63–93.
- Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, **54**, 372–391.
- Geary, D. C., & Brown, S. C. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal and mathematically disabled children. *Developmental Psychology*, **27**, 398–406.
- Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology*, **27**, 787–797.
- Goldman, S. R., Mertz, D. L., & Pellegrino, J. W. (1989). Individual differences in extended practice functions and solution strategies for basic addition facts. *Journal of Educational Psychology*, **81**, 481–496.
- Groen, G. J., & Parkman, I. H. (1972). A chronometric analysis of simple addition. *Psychological Review*, **79**, 329–343.
- Groen, G., & Resnick, L. B. (1977). Can preschool children invent addition algorithms? *Journal of Educational Psychology*, **69**, 645–652.
- Hale, C. A., & Borkowski, J. G. (1991). Attention, memory, and cognition. In J. L. Matson & J. A. Mulick (Eds.), *Handbook of mental retardation* (2nd ed., pp. 505–528). New York: Pergamon.
- Hamann, M. S., & Ashcraft, M. H. (1986). Textbook presentations of the basic addition facts. *Cognition and Instruction*, **3**, 173–192.
- Jensen, A. J. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kail, R. (1992). General slowing of information processing by persons with mental retardation. *American Journal on Mental Retardation*, **97**, 333–341.
- Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity. *Intelligence*, **14**, 389–433.
- Lefevre, J.-A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **22**, 216–230.
- Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory & Cognition*, **22**, 395–410.
- McCloskey, M., Harley, W., & Sokol, S. M. (1991). Models of arithmetic fact retrieval: An evaluation in light of findings from normal and brain-damaged subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **17**, 377–397.
- Miller, K. F., Perlmutter, M., & Keating, D. (1984). Cognitive arithmetic: Comparison of operations. *Journal of Experimental Psychology: General*, **115**, 307–330.
- Mislevy, R. J. (1996). Test theory reconceived. *Journal of Educational Measurement*, **33**, 379–416.
- Mislevy, R. J., & Wilson, M. (1996). Marginal maximum likelihood estimation for a psychometric model of discontinuous development. *Psychometrika*, **61**, 41–71.

- 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. (1988a). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, **59**, 833–851.
- Siegler, R. S. (1988b). Strategy choice procedures and the development of multiplication skill. *Journal of Experimental Psychology: General*, **117**, 258–275.
- Siegler, R. S., & Jenkins, E. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 16, pp. 241–312). New York: Academic Press.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.
- Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.
- Svenson, O., Hedenborg, M. L., & Lingman, L. (1976). On children's heuristics for solving simple additions. *Scandinavian Journal of Educational Research*, **20**, 161–173.
- Svenson, O., & Sjoburg, K. (1983). Evolution of cognitive processes for solving simple additions during the first three school years. *Scandinavian Journal of Psychology*, **24**, 117–124.
- Viaene, M. (1992). *Een cognitief psychometrische studie van het oplossen van eenvoudige reken-sommen bij zesjarigen* [A cognitive psychometric study of solving simple additions by six-year-old children]. Unpublished master's thesis, University of Leuven, Belgium.
- von Eye, A., & Clogg, C. C. (Eds.). (1994). *Latent variable analysis: Applications for developmental research*. Thousand Oaks, CA: Sage.
- Widaman, K. F., Geary, D. C., Cormier, P., & Little, T. D. (1989). A componential model for mental addition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **15**, 898–919.
- Widaman, K. F., & Little, T. D. (1992). The development of skill in mental arithmetic: An individual differences perspective. In J. I. D. Campbell (Ed.), *The nature and origin of mathematical skills* (pp. 189–253). Amsterdam: Elsevier.
- Widaman, K. F., Little, T. D., Geary, D. C., & Cormier, P. (1992). Individual differences in the development of skill in mental addition: Internal and external validation of chronometric models. *Learning and Individual Differences*, **4**, 167–213.
- Wilson, M. (1989). Saltus: A psychometric model of discontinuity in cognitive development. *Psychological Bulletin*, **105**, 276–289.
- WISC-R Projectgroep: P. P. van Haasen, E. E. J. De Bruyn, Y. J. Pijl, Y. H., Poortinga, H. C. Spelberg, G. Vander Steene, P. Coetsier, R. Spoelders-Claes, & J. Stinissen. (1986). *WISC-R, Wechsler Intelligence Scale for Children-Revised, Nederlandstalige uitgave*. Lisse, The Netherlands: Swets & Zeitlinger.
- Wu, M. L. (1997). *The development and application of a fit test for use with marginal maximum likelihood estimation and generalized item response models*. Unpublished master's dissertation, University of Melbourne, Australia.
- Wu, M. L., Adams, R. J., & Wilson, M. R. (1997). ConQuest: Generalised Item Response Modeling Software [Computer program and computer program manual]. Australian Council for Educational Research.
- Zigler, E., & Balla, D. (1982). *Mental retardation: The development-difference controversy*. Hillsdale, NJ: Erlbaum.