CHAPTER 18

DOMAIN-SPECIFIC VERSUS METACOGNITIVE KNOWLEDGE EFFECTS ON TEXT RECALL AND COMPREHENSION

Joachim KÖRKEL and Wolfgang SCHNEIDER
Max Planck Institute for Psychological Research, Munich
Federal Republic of Germany

INTRODUCTION

In numerous studies, it has been shown that domain-specific knowledge influences how much and what children recall (cf. Bjorklund, 1987; Chi & Ceci, 1987; Ornstein & Naus, 1985; Schneider & Pressley, 1989). The focus of this paper concerns the impact of domain-specific knowledge on text processing in highly articulated domains, a topic that has also been treated extensively in the literature. We already know from numerous studies using the expert-novice paradigm that experts in an area learn more when studying new information in their domain of expertise than do novices in that domain (cf. Voss, Fincher-Kiefer, Green, and Post, 1986; Körkel, 1987, for reviews). We also know from studies based on the expert-novice paradigm that domain-specific expertise can compensate for low overall aptitude on certain domain-related text processing tasks, regardless of age (cf. Recht & Leslie, 1988; Schneider, Körkel, & Weinert, 1989; Walker, 1987). Thus, there is no doubt that rich domain-specific knowledge strongly influences text processing in the domain of interest.

In this chapter, we focus on two questions that were rarely addressed in studies using the expert-novice paradigm. The first question concerns the issue of developmental differences in experts' text recall. While there is little doubt that older child experts recall more text information related to their designated domain than younger child experts, it is less clear whether older experts additionally differ qualitatively from younger experts in how their knowledge is represented in recall.

311

The second question addresses the problem of how to conceptualize the relationship between domain-specific knowledge and (general) metacognitive knowledge in influencing memory performance. So far, most studies based on the expert-novice paradigm have emphasized the crucial role of domain-specific knowledge in facilitating text recall, ignoring possible influences of metacognitive knowledge (cf. Voss et al., 1986; Walker, 1987). On the other hand, studies using representative samples in order to explore developmental differences in text processing have demonstrated the importance of metacognitive knowledge but usually neglected issues of domain-specific knowledge (cf. Forrest-Pressley & Waller, 1984; Garner, 1987). Thus one major goal of this discussion is to explore the *relative* effects of domain-specific knowledge and metacognitive knowledge on memory for text. The crucial question is whether recall of a story dealing with a specific domain is solely determined by the richness of domain-specific knowledge, or may be additionally influenced by both procedural and declarative metacognitive knowledge.

To answer the two questions described above, we reanalyzed some of our data that seemed suited to deal with these problems (cf. Körkel, 1987; Schneider, Körkel, & Weinert, 1989). In this study, the expert-novice paradigm was used to explore the influence of knowledge about soccer on recall of a story about a soccer game. Before discussing the major findings concerning the two questions mentioned above, we want to give more details concerning the study that our secondary analysis is based on.

SUBJECTS, MATERIALS, AND DESIGN

A total of 185 middle-class children (64 third, 67 fifth, and 54 seventh graders) participated in the study. Children were selected from two public schools located near Heidelberg, Germany. An approximately equal number of boys and girls was included at each age level.

A thirteen-item questionnaire was used to assess children's knowledge about soccer. Ten multiple-choice items tapped subjects' knowledge about soccer rules, whereas the remaining three items assessed knowledge about important soccer events. Each item on the questionnaire was given a score of 1 or 0 (maximum score of 13). For each age group, children with scores above the median were classified as soccer experts, and those with scores below the median were considered soccer novices.

All subjects were presented with a narrative text dealing with a soccer game. The story was about a young soccer player's experiences in an important match. After a short description of the young hero and his activities before the game, the course of action during the game was described in detail. The story ended with a description of the hero's physical and psychological condition after the game was over.

Precautions were taken to ensure that most parts of the story were easily understandable for novices. Amstad's (1978) version of the Flesch formula was used to assess text readability. This formula yields values ranging from 0 (low readability) to 100 (high readability). The average score obtained for our story ($x = 80$) indicated that the text was easily readable for the different age groups under study. In addition, a structural analysis of the text based on the grammar of Mandler and Johnson (1977; Johnson & Mandler, 1980) revealed that the story could be considered simple and well-structured according to the criteria of these authors.

However, there were a few exceptions to this rule. Occasionally, sentences were shortened; that is, important information was omitted that had to be inferred by the reader. Moreover, several contradictions were built into the text that could only be detected by careful reading. For example, the hero was first described as a fast soccer player, but later referred to as very slow and sluggish. While prior knowledge about soccer was important in order to draw correct inferences, it was not always necessary to detect the contradictions in the text. The story was taped and presented twice to the subjects.

Three different memory performance variables were assessed: First, subjects freely recalled the soccer story. The instruction was that the children should do their best to recall the story as accurately and comprehensively as possible. The recall protocols were analyzed according to the procedure developed by Mandler and Johnson, that is, in terms of "semantic" or idea units. At maximum, 36 different idea units could be reproduced.

A cloze test was used as a second measure of (supported) recall. All subjects were presented with a written version of the story that included 20 blanks, and were asked to fill the gaps as accurately as possible. One point was given for each correct completion of the text (maximum score of 20).

A story reconstruction test followed about two weeks after the first test session. Children were given an envelope containing all sentences of the soccer story in random order, with each sentence typed on a separate card. The subjects were instructed to reconstruct the original soccer story

as accurately as possible. The resulting ordering of sentences was then compared with the original sequence, and the number of reconstruction errors was used as a quantitative outcome variable. In addition, a clustering score (ARC Score) was used as a measure of quality of reconstruction. Note that the story reconstruction test was not given to the third graders because of time constraints.

In addition to the three performance measures, several indicators of strategic operations and metacognitive processes were available in this study. These indicators included subjects' importance ratings of the text, their performance prediction for the free recall task, their feeling-of-knowing judgments when completing the cloze test, and their declarative metacognitive knowledge about text recall.

The importance rating procedure is a very popular tool in the area of text memory and comprehension (see Brown, Bransford, Ferrara, & Campione, 1983, for a review). The version used in this study required the children to select and underline those ten sentences in the soccer text that they considered the most important and necessary for an efficient reproduction of the text. Children's importance ratings were then compared with an "ideal" importance rating of the text based on the judgments of 20 adult soccer experts (students and faculty members of different departments at the University of Heidelberg). One point was given when a sentence judged as important corresponded with an "objectively" important sentence selected by the adult experts (maximum score of 10).

Immediately after the importance rating procedure, subjects were asked to predict the number of sentences they could freely recall if asked to do so. Prediction accuracy was assessed by relating performance prediction and actual performance in the free recall task. More precisely, it was defined as the absolute value of the difference between the recall estimate and actual recall, divided by recall. According to this formula, smaller scores correspond with better prediction accuracy.

With regard to the feeling-of-knowing judgments, the relevant information was taken from the cloze test. When completing the blanks in the cloze test, subjects were asked to indicate whether the completion of the text was correct or incorrect in their opinion. Correct evaluation, that is, hits (i.e., both completion and evaluation correct) and correct rejections (i.e., completion incorrect and evaluation correct) were summed, yielding a maximum score of 20.

Children's declarative metacognitive knowledge (metamemory) about text processing and recall was assessed by using a 17-item questionnaire

modelled after the interviews developed by Kreutzer, Leonard, and Flavell (1975) and Myers and Paris (1978). One point was given for each item, thus yielding a maximum score of 17.

Finally, subtests of a German cognitive ability test (Heller, Gaedike, & Weinläder, 1985) and the Culture Fair Intelligence test (Cattell & Weiss, 1978) were used to assess children's verbal and nonverbal intelligence.

RESULTS OF SECONDARY DATA ANALYSIS

Developmental differences in experts' recall

The analysis of free recall data in the original study (Körkel 1987) had yielded significant main effects of grade and expertise. Experts outperformed novices at each grade level. As depicted in Table 1, a reversal of developmental trends was observed. Third grade experts recalled more idea units than both third and fifth grade novices.

TABLE 1

Mean Percentage of Idea Units Recalled as a Function of Grade and Expertise (data from Körkel, 1987)

Grade	Soccer Experts	Soccer Novices
3	54	32
5	52	33
7	61	42

Similarly, fifth grade experts outperformed seventh grade novices. Taken together, these results demonstrate how greatly domain-specific knowledge influences memory performance.

As can be seen from Table 1, recall differences among experts of different age groups were less impressive. A one-way ANOVA with recall as dependent variable and grade as independent variable yielded a significant effect of grade, $F(2,86)=4.43$, $p<.05$. Subsequent Student-Newman-Keuls tests revealed that seventh grade experts showed better recall than both third grade and fifth grade experts, who did not differ from each other.

The question of major interest concerned possible *qualitative* differences in text recall. We used the proportions of idea units recalled for various importance levels as a dependent variable. The mean

proportion of idea units correctly recalled as a function of importance level is illustrated in Figure 1. Importance level 1 indicates units judged to be most important, whereas importance level 4 contains units judged to be least important by soccer experts. A 3 (grade) x 4 (importance level) mixed analysis of variance yielded a significant main effect of importance level, $F(3,252)=22.13$, $p < .01$.
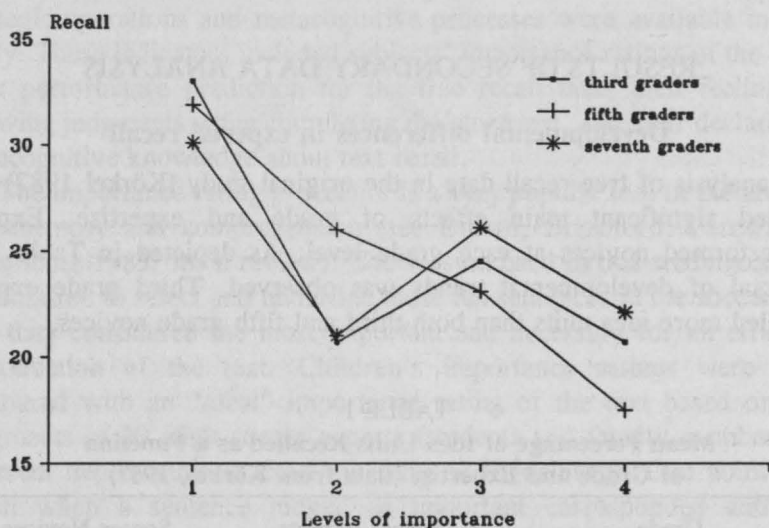

FIGURE 1

Mean Proportion of Text Units Recalled by Soccer Experts as a Function of Age and Importance Level

The importance level x grade interaction was just short of statistical reliability, $F(6,252)=2.09$, $p < .06$. The visual impression from Figure 1 suggests that seventh graders' recall patterns differed from the patterns of the two younger expert groups. To test this assumption, contrasts were specified that compared seventh graders' recall patterns with the combined results for third and fifth graders. This analysis yielded a significant importance level x group interaction, $F(3,252)= 3.13$, $p < .05$, thus confirming the hypothesis that the older experts' recall pattern differed significantly from that obtained for the two younger expert groups.

As a main result, these analyses suggest that there are no major qualitative differences in third and fifth grade soccer experts' recall patterns. No major developmental differences in the way these experts

memorized the text could be detected. Seventh grade soccer experts' recall patterns differed from those obtained for the two younger groups in that the mean proportion of recall decreased almost linearly with decreasing importance level, a result to be expected for adult soccer experts. Altogether, the structure of experts' recall patterns clearly differed from that of soccer novices who recalled as much important as unimportant text information, regardless of age.

RELATIVE EFFECTS OF DOMAIN SPECIFIC KNOWLEDGE AND METACOGNITIVE KNOWLEDGE ON TEXT RECALL

Given the fact that domain-specific knowledge has such a strong effect on text processing in the designated domain, it is by no means obvious whether other knowledge components (e.g., declarative and procedural metacognitive knowledge) additionally contribute to experts' memory performance.

In order to explore this issue more systematically, we decided to use structural equation modelling (SEM) procedures based on a latent variable approach. Their major advantages -- as compared to traditional regression analysis -- are that (1) a verbal theory has to be translated in a mathematical model that can be estimated; (2) structural/causal relationships are estimated at the level of latent variables or theoretical constructs and not on the basis of fallible observed variables; (3) the distinction between a measurement model describing the relationships among observed variables and latent factors on the one hand and a structural model describing interrelationships among theoretical constructs on the other hand allows for a separate estimation of measurement errors in the observable and specification errors in the structural part of the model; and (4) several so-called goodness-of-fit tests exist that detect the degree of fit between the causal model and the data set to which it is applied (cf. Bentler, 1987; Jöreskog & Sörbom, 1984; Schneider, 1989, for more details on SEM procedures).

One problem with using SEM procedures in the behavioural sciences is that verbal theories are not particularly well developed. In order to make sure that our preferred theoretical model fits the data better than alternative model specification, three alternative models were specified and tested. Our favourite model (model 1) specified that both verbal and nonverbal intelligence should influence the three knowledge components (i.e., domain-specific knowledge, declarative and procedural meta-cognitive knowledge). As declarative and procedural metacognitive

knowledge is conceived of as rather independent (cf. Brown et al., 1983), no relationship between these two concepts was specified. Similarly, there was no reason to pre-specify an interrelation between domain-specific knowledge and declarative metacognitive knowledge, as the latter concept addressed more general, domain-unspecific knowledge about text processing. On the other hand, procedural metacognitive knowledge is typically closely linked to the designated domain. That is, the quality of feeling-of-knowing judgments or importance ratings depends on the familiarity with item materials. Thus we assumed that domain-specific knowledge should influence procedural metacognitive knowledge, which in turn should affect memory performance. Finally, declarative metacognitive knowledge as well as intelligence should directly affect memory performance.

The only difference between model 1 and the three alternative models concerned the relationship between domain-specific knowledge and procedural metacognitive knowledge. In model 2, a correlation between both knowledge components was specified. In model 3, both knowledge components served as independent, unrelated factors which both directly influenced memory performance. Finally, in model 4, the relationship between both knowledge components specified in model 1 was just reversed: According to this "nonsense" model, procedural metacognitive knowledge should influence domain-specific knowledge.

Before turning to the results, we need to address a few technical problems. The most elegant solution to the problem of estimating the same structural equation model for all grade levels would have been to analyze the data from the three groups simultaneously, which is one of the special advantages of the computer program LISREL VI (Jöreskog & Sörbom, 1984) that we chose for our analyses. However, this approach turned out to be problematic in our case for two reasons. First, as indicated above, sample sizes per grade level ranged between 54 and 67. Multiple group comparisons based on such small sample sizes are probably strongly biased (cf. Tanaka, 1987). Second, we had to cope with the problem that some of the measures obtained for the third graders differed from those obtained for the two older age groups. For example, different intelligence subtests were given to the third graders, and the text reconstruction test was not available for this group. Due to this lack of correspondence of measures across age groups, multiple group comparisons seemed meaningless because it was impossible to test the degree of equality across covariance matrices of the observed variables.

Given these problems, we decided (1) to drop the third graders' data from the analyses, (2) to merge the data sets for the fifth and seventh graders, and (3) to introduce chronological age as an exogenous variable in the model. The LISREL solution for our theoretically preferred causal model (model 1) is depicted in Figure 2. Only the causal links (i.e., structural coefficients) are given for the sake of clarity. As can be seen from Figure 2, the measure of goodness-of-fit indicates that model 1 fits the data. This was not true for the three alternative models. While model 2 (intercorrelation of domain-specific knowledge and procedural metacognitive knowledge) turned out to be the best alternative model (yielding a Chi-square of 125.54 with 93 degrees of freedom), it did not fit the data (p < .05). In comparison, model 3 (independence of domain-specific knowledge and procedural declarative knowledge) yielded a chi-square of 169.7 (df=94), indicating that the data fit obtained for model 3 was significantly worse than that obtained for model 2. Not surprisingly, our "nonsense" model (procedural metacognitive knowledge influences domain-specific knowledge) did not fit the data at all (Chi-square=215.2, df=93).

As model 1 depicted in Figure 2 fits the data significantly better than the various alternative models, we decided to accept this model as the best description of the underlying causal process. The most interesting aspects of model 1 concern the interrelations among the three knowledge components and their influences on memory performances. As predicted, there was no link between declarative and procedural metacognitive knowledge on the one hand, and declarative metacognitive knowledge and domain-specific knowledge on the other hand. The direct path from declarative metacognitive knowledge to memory performance was reliable but not substantial, indicating that declarative metacognitive knowledge only played a modest role in predicting text recall.On the contrary, procedural metacognitive knowledge turned out to be an important predictor of recall. As can be seen from model 1, procedural metacognitive knowledge is strongly affected by domain-specific knowledge which shows substantial direct and indirect effects on text recall. Taken together, the results obtained for model 1 not only confirm the dominating role of domain-specific knowledge in predicting text recall but also indicate that metacognitive knowledge does have an additional effect. It appears, then, that metacognitive knowledge does make a difference even in cases where domain-specific knowledge is rich.
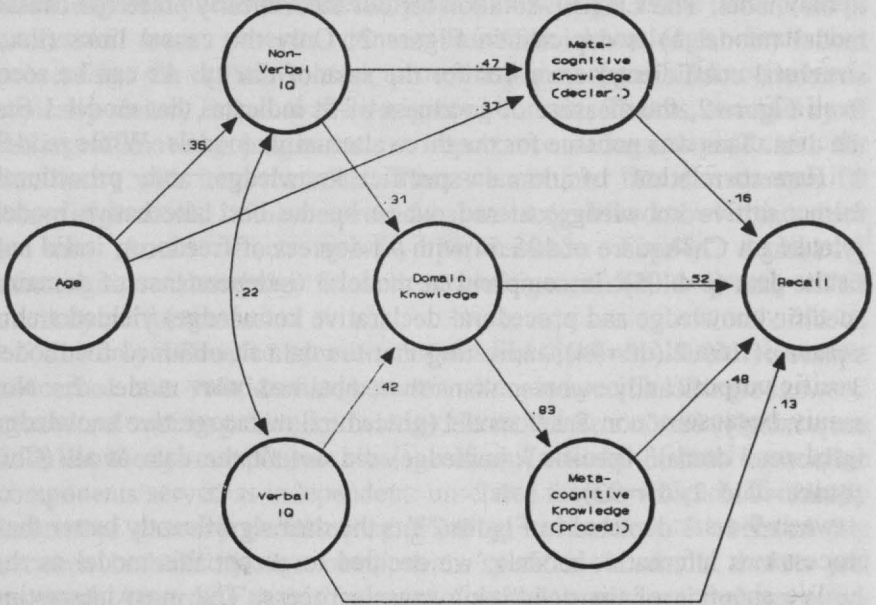
FIGURE 2
**Structural Equation Model Describing the Impact of Intelligence
and Various Knowledge Components on Text Recall**

CONCLUDING REMARKS

The secondary data analysis summarized in this article aimed at exploring
two issues rarely addressed in research using the expert-novice paradigm.
Regarding the first issue, namely, possible developmental differences in
experts' text recall, we did not find reliable qualitative differences in the
recall patterns of soccer experts from grade levels three and five.
However, seventh grade experts' recall patterns differed from that of the
two younger expert groups in that the proportions of recall decreased
continuously with decreasing level of importance. Overall, idea units
stemming from the most important text passages were best recalled,
regardless of age. This finding corresponds well with the experts'
importance ratings in that no age differences were found for the rating

procedure used in our study (cf. Körkel 1987). When asked to select those sentences in the soccer story that they considered most important, the child experts in our study tended to choose those sentences that already had been classified as most important for an adequate understanding of the story by our sample of adult soccer experts. Obviously, this knowledge is reflected in the structural pattern of text recall. The major difference between the older and the two younger expert groups concerns the recall of less important text units.

How to account for the finding that the two younger expert groups did not differ qualitatively in how their knowledge is represented? Is it because the story chosen was too difficult for the younger experts? Our results do not support such an explanation because even the recall rates for the youngest children were far away from bottom. Is it because our measures tapping qualitative differences in recall were too crude for this specific purpose? Again, we doubt that such an explanation is sufficient, mainly because a more complex qualitative analysis of recall patterns based on Mandler and Johnson's (1977) story grammar yielded similar results. Thus we are inclined to believe that soccer experts ranging between 9 and 11 years of age do not differ markedly in the way they process text information related to their domain of interest.

We should note that the finding of qualitative differences between younger and older experts' knowledge representations extends to other domains. Means and Voss (1985), for example, conducted a developmental study of expert and novice knowledge structures by using the domain of "Star Wars". Means and Voss, indeed, also found qualitative differences in the "Star War" representations of younger and older experts, recruited from grade levels 2, 3, 5, 7, 9, and college. The authors attributed the qualitative differences between younger and older experts to differential prior schematic knowledge, with older experts activating a more developed schema based on superior world knowledge. In our view, differential effects of world knowledge on experts' text recall are likely to occur even when the age range of subjects is not similarly broad as in the Means and Voss (1985) study.

Regarding our second issue, namely, the *relative* effects of domain-specific knowledge and metacognitive knowledge on memory for text closely related to the designated domain, the findings seem to be clearcut: While our results based on causal modelling procedures confirm the crucial role of domain-specific knowledge on text recall, they also indicate that domain-specific knowledge and procedural metacognitive knowledge are functionally related, and that procedural metacognitive

322   Körkel & Schneider

knowledge has an independent, reliable direct effect on text recall. Note, however, that the model estimates are based on relatively small sample sizes, and that we were not able to estimate this model for our third graders due to the technical problems mentioned above. Given these constraints, cross-validation studies testing this model by using larger sample sizes and a broader age range are badly needed in order to evaluate the generalizability of our findings.

REFERENCES

Amstad, T. (1978). *Wie verständlich sind unsere Zeitungen?* Unpublished doctoral dissertation. Zürich.

Bentler, P.M. (1987). Drug use and personality in adolescence and young adulthood. Structural models with nonnormal variables. *Child Development*, 58, 65-79.

Bjorklund, D.F. (1987). How age changes in knowledge base contribute to the development of children's memory: An interpretive review. *Developmental Review*, 7, 93-130.

Brown, A.L., Bransford, J.D., Ferrara, R.A., & Campione, J.C. (1983). Learning, remembering, and understanding. In J.H. Flavell & E.M. Markman (Eds.), *Handbook of child psychology* (Vol. III, Cognitive development, pp. 77-166). New York: Wiley.

Cattell, R., & Weiss, R. (1978). *Grundintelligenztest, Skala 2.* Braunschweig: Westermann.

Chi, M.T.H., & Ceci, S.J. (1987). Content knowledge: Its role, representation, and restructuring in memory development. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20, pp. 91-142). Orlando: Academic Press.

Forrest-Pressley, D.L., & Waller, T. (1984). *Cognition, metacognition, and reading.* New York: Springer-Verlag.

Garner, R. (1987). *Metacognition and reading comprehension.* Norwood, NJ: Ablex Publishing Corp.

Heller, M.S., Gädike, A.K., & Weinläder, H. (1985). *Kognitiver Fähigkeitstest für 4. bis 13. Klassen (KFT 4-13)* (2nd ed.). Weinheim: Beltz.

Jöreskog, K.G., & Sörbom, D. (1984). *LISREL VI - Analysis of linear structural relationships by maximum likelihood instrumental variables, and least squares methods* (Users Guide). Mooresville: Scientific Software.

Johnson, N.S., & Mandler, J.M. (1980). A tale of two structures: Underlying surface forms in stories. *Poetics*, 9, 51-86.

Körkel, J. (1987). *Die Entwicklung von Gedächtnis- und Metagedächtnisleistungen in Abhängigkeit von bereichsspezifischen Vorkenntnissen.* Frankfurt/Main: Lang.

Kreutzer, M.A., Leonard, C., & Flavell, J.H. (1975). An interview study of children's knowledge about memory. *Monographs of the Society for Research in Child Development*, *40*, (1, Serial No. 159).

Mandler, J.M., & Johnson, N.S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111-151.

Means, M., & Voss, J. (1985). Star Wars: A developmental study of expert and novice knowledge structures. *Memory and Language*, *24*, 746-757.

Myers, M., & Paris, S.G. (1978). Children's metacognitive knowledge about reading. *Journal of Educational Psychology*, *70*, 680-690.

Ornstein, P.A., & Naus, M.J. (1985). Effects of the knowledge base on children's memory strategies. In H.W. Reese (Ed.), *Advances in child development and behavior* (Vol. 19, pp. 113-148). Orlando: Academic Press.

Recht, D.R., & Leslie, L. (1988). Effect of prior knowledge on good and poor readers' memory of text. *Journal of Educational Psychology*, *80*, 16-20.

Schneider, W. (1989). Problems of longitudinal studies with children: Practical, conceptual, and methodological issues. In M. Brambring, F. Lösel, & H. Skowronek (Eds.), *Children at risk: Assessment, longitudinal research, and intervention* (pp. 313-335). New York: De Gruyter.

Schneider, W., Körkel, J., & Weinert, F.E. (1989). Domain-specific knowledge and memory performance: A comparison of high- and low-aptitude children. *Journal of Educational Psychology*, *81*, 306-312.

Schneider, W., & Pressley, M. (1989). *Memory development between 2 and 20*. New York: Springer-Verlag.

Tanaka, J.S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, *58*, 134-146.

Voss, J.F., Fincher-Kiefer, R.H., Greene, T.R., & Post, T.A. (1986). Individual differences in performance: The contrastive approach to knowledge. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 3, pp. 297-334). Hillsdale, NJ: Erlbaum.

Walker, C.H. (1987). Relative importance of domain knowledge and overall aptitude on acquisition of domain-related information. *Cognition and Instruction*, *4*, 25-42.