CHAPTER 8

General Discussion
Chapter 8

Introduction

The general aim of this thesis was to investigate the contribution of dynamic testing procedures, in particular those with a graduated prompts approach, to assessment procedures focused on interventions and classroom recommendations. In addition, teachers’ values regarding the application of dynamic testing outcomes in formulating educational plans were investigated. In the context of assessment for intervention and the emphasis on adapting instructions to the needs of individual children we sought answers to the two general research questions. First, whether there is evidence for the use of dynamic testing procedures to identify individual differences in children’s need for instruction, which could contribute to needs based assessment. Second, how useful are outcomes of dynamic testing procedures for teachers, regarding educational interventions and formulating educational plans.

In this final chapter the general findings regarding these two main research questions will be discussed in reference with the literature. First, the findings of our studies regarding the identification of individual differences in need for instruction will be discussed (chapters 2-4). Subsequently, the applicability of the dynamic testing procedures in Dutch educational settings, from teachers’ perspectives (chapters 5-7) will be discussed. Next a critical discussion regarding the overall findings is presented, followed by practical implications of our findings and directions for future research.

Differences in need for instruction

To answer our first research question, we conducted a series of studies to investigate the specific information obtained with a dynamic test procedures regarding children’s potential for learning and need for instruction.

In the first study (chapter 2) presented in this thesis we showed that the need for instruction of a group of young children with moderate to mild learning disabilities in
special education classes could be measured with a dynamic test consisting of analogy problems, including a one session training based on a structured protocol of prompts and feedback, the Analogical Reasoning Learning Test (ARLT, Buchel & Schlatter, 2000; Hessels-Schlatter, 2002b). Based on the posttest scores of this dynamic test, different levels of need for instruction could be distinguished. Children could be categorized as being either high responsive, medium responsive or low responsive to the prompts and feedback provided during the training. These categories of responsiveness to instruction show resemblance to those reported in earlier research (e.g., Budoff, 1970; Büchel, Schlatter & Scharnhorst, 1997; Hessels-Schlatter, 2002a).

We, however, sought to find more refined information regarding children’s individual need for instruction during dynamic testing. In the subsequent studies therefore a dynamic testing model was followed, based on the concepts of the graduated prompts techniques employed by, for example, Campione & Brown (1987), Fabio (2005) and Resing, (1990, 1993, 2000). In this model potential for learning is defined by the minimum number and type of prompts children need to solve problems independently.

In the study described in chapter 3, the need for instructions among a group of 10-12 year old children with intellectual disabilities (IQ < 70) was investigated with the Children’s Conceptual and Perceptual Analogies Modifiability test (CCPAM, Tzuriel & Galinka, 2002). A structured training based on graduated prompts techniques was developed, including metacognitive, cognitive and modeling prompts, instead of the original mediation phase. Findings demonstrated, as expected on the basis of earlier research, that trained children achieved higher scores than untrained children on the posttest compared to the pretest (see for earlier findings e.g., Campione & Brown, 1987; Hessels-Schlatter, 2002; Resing, 1990, 1993, 2000; Resing, Tunteler, De Jong & Bosma, 2009; Resing, De Jong, Bosma & Tunteler, 2009; Resing & Elliott, in press). Despite of the verbal nature of the structured prompts-protocols that were followed, the results demonstrate the applicability of the training based on graduated prompts techniques regarding children with intellectual disabilities, which, so far, has been investigated only incidentally.
An important claim of researchers promoting dynamic testing in favor of, or parallel to, standard tests, is the assumption that dynamic testing scores will provide a different picture of the cognitive abilities of a child than standard test scores. The results of both studies described above support outcomes of previous studies on dynamic testing in which the levels of children’s potential for learning did not necessarily correspond to those of standard intelligence scores (e.g., Campione & Brown, 1987; Hessels, 1997; Resing, 1990; 1997; 2000). Children with a high responsiveness to training, and who had profited from instruction during training, varied, for example, from far below average to average intelligence test scores; similar ranges and variations in IQ’s were found for children with moderate or low responsiveness. Children’s need for instruction during training in the second study (chapter 3) varied not only between children with comparable pretest scores but also between children with comparable IQ-scores. Some children with a low IQ needed a similar amount and type of assistance as children with a moderate IQ and vice versa. These findings showed as well the additional value of dynamic testing procedures in identifying children’s need for instruction compared to regularly administrated intelligence tests. In particular the often underestimated potential for learning among children with intellectual disabilities was demonstrated once again (e.g. Campione, Brown & Ferrara, 1982; Hessels-Schlatter, 2002; Tzuriel & Klein, 1985).

Dynamic testing with a training based on graduated prompts techniques enabled us to describe children’s need for instruction during training in more detail. Numbers and types of prompts provided during training varied strongly between children. Prompts were required frequently by the majority of the children in solving the items. The types of prompts provided during training varied as well. Cognitive and modeling prompts were needed most often, but sometimes just the provision of metacognitive prompts was of sufficient assistance, while in other cases very specific prompts that included a full explanation were necessary.

Individual differences in the need for instruction were further explored in a third study (chapter 4). Children with arithmetic difficulties in the 2nd grade of elementary education were administered an adapted version of the Seria-Think Instrument (Resing,
De Jong et al., 2009; Resing, Tunteler et al., 2009; see for the original training Tzuriel, 1998; 2000b), a complex dynamic test in which children had to solve seriation problems by means of measurement activities and insertions of rods. A training consisting of six scenario-protocols was developed, to be able to guide a child through all the steps in the problem solving process with the help of a series of graduated prompts structured from metacognitive to cognitive to modeling prompts. Providing prompts according to these multiple protocols delivered, as expected, more detailed information about number and types of prompts and feedback children required. In addition, using multiple protocols gave us the possibility of measuring the number and type of prompts children needed at different times in the solution process. Prompts were most frequently provided to assist children in employing a specific solving strategy but as well to encourage children to work systematically, organized or accurately.

This study further revealed four clusters of children, each with a different profile of need for instruction. Besides a large group of children who needed a relatively small number of prompts in general, a smaller second group was identified by their need for prompts regarding measuring, a specific cognitive strategy to solve the task. A third and fourth group of children both needed a relatively high number of prompts, but one group required in particular prompts for accurate measuring, while the other group was identified by the need for prompts regarding organizing, planning and measuring. The clusters of children did, as expected, not differ in regards to age, gender, ethnicity, memory or general cognitive functioning. Children in the 3rd cluster however had lower arithmetic scores. The clusters of children further showed differences in observed behavior during the training, for example in their level of self-confidence and concentration.

**Reversal analogy task**

One extra source of information regarding individualized need for instruction was the use of a reversal or construction task in the studies in chapters 2 and 3. The reversal task in the first study was administered after the dynamic test procedure, to be able to answer the question if such a task would give additional information regarding the
solving processes of a child. Children were asked to construct analogy problems and to instruct the tester how to solve these problems. The rationale behind administering such a task is the expectation that, if the child would be able to solve such a task correctly, this would be an indication that the child is able to transfer the trained principles of solving analogy problems to a completely new task. In this reversal task, children have to take the role of a teacher in order to show their ability to construct analogy problems, their knowledge of solving these problems, and their ability to explain the solution. Findings (chapter 2) showed that in particular children’s ability to construct analogies and their ability to explain the essential elements of the solution were related to their levels of responsiveness to the training as measured with the posttest scores. Children with high responsiveness to instructions given in the training outperformed children with moderate or low responsiveness in constructing and explaining analogies, while the weaker two groups rather equaled each other. The findings show how the reversal task provided additional information regarding children’s potential for learning. Besides the additional qualification of children’s potential for learning, the observations during the reversal task provided information about individual learning characteristics. The amount of assistance children required during the construction task in the study in chapter 3, provided comparable information regarding children’s need for instruction. In a recent study with a reversal task it was further demonstrated how the correct construction of an analogy item distinguished trained children from untrained children (Stevenson, Heiser & Resing in prep.).

**Teachers’ appreciation of dynamic testing outcomes**

To answer our second research question, we conducted a series of studies to investigate teachers’ opinions regarding the use of dynamic testing outcomes for educational interventions and for formulating educational plans. We choose to study teachers’ evaluations of psycho-diagnostic reports and recommendations regarding children in
their actual classroom, which involved actually testing the children. Reports consisted of either the outcomes of dynamic testing or standard testing results. Prior to the start of the dynamic testing procedure we obtained a baseline of teachers’ general instructional practice and teacher-child interactions measured with two observational rating scales (Lidz, 2003; 1991; Van der Aalsvoort, 1994). Teachers were also interviewed and asked to rate the learning potential of children. After the reports and recommendations were reported to teachers, the observations, interviews and ratings were repeated to examine possible changes. With these semi-diagnostic designs we investigated the responses of teachers regarding first grade typically developing children (chapter 5), second grade children with arithmetic difficulties (chapter 6). We also explored the responses of a small number of special education teachers (chapter 3).

The observed teaching practices and teacher-child interactions between regular education teachers and their typical developing children could be, prior to the assessment procedure, interpreted as intentional, task-regulating and infrequent positive supporting. Observed teacher-child interactions regarding children with arithmetic difficulties showed overall lower ratings, indicating that most behaviors were inconsistently recorded. Teachers in both studies were observed to challenge children within their zone of proximal development only incidentally, and children were not or hardly ever informed about their learning progress. Observations of special education teachers showed similar patterns and corresponded to observations of daycare teachers as well (Van der Aalsvoort, 1994), although the observed level of positive support (praise) and provision of challenging tasks of daycare teachers appeared to be higher.

Changes in the teaching practice and teacher-child interactions were observed as well. Compared to teachers in the control group, who had read standard reports and recommendations, teachers of dynamically tested (typically developing) children were observed to give more frequent task regulating activities in their general teaching, and to improve their amount of positive feedback as well as their level of challenging activities to individual children. For these teachers, the second round of observations of teacher-child interactions showed improvement, which may indicate that teachers appeared to respond in their teaching practice to the reports and recommendations. In
the general teaching practice of teachers of 2nd grade children with arithmetic weaknesses, we observed more intentional, task-regulating and contingent responses, as well as positive feedback. However, within this sample, teachers who received dynamic reports and recommendations hardly ever changed their behaviors for the better compared to teachers reading standard reports. They even showed less positive feedback and challenge. Comparably, no changes were observed in teacher-child interactions in special education. The observed responses of the teachers of the typical group of children may be influenced by the fact that results of observations were given in both dynamic as well as standard reports.

Another focus in the study regarding typical children were teachers’ ratings and estimations of potential for learning. Teachers rated the potential for learning directly during the interviews and filled in a school behavior checklist (SCHOBL-R; Bleichrodt, Resing & Zaal, 1993) to which we added a factor ‘learning potential’. Both estimations appeared to be related to the posttest scores as well as to the number of prompts children required. Learning potential estimations appeared not affected by the type of reports and recommendations teachers received.

Although teachers’ practices appeared not to be affected by the type of reports and recommendations they received, it was observed that teachers in general responded to the reported observations. Perhaps our expectations were too ambitious as teacher change is difficult to establish (Richardson & Placier, 2001), takes time and changes might even be resisted (Witt, 1986).

Teachers’ evaluations of the practical use of the reported dynamic testing outcomes were obtained by teachers’ responses to structured interviews. The majority of teachers of typical first grade students judged the dynamic testing reports as positive and supplemental compared to the standard reports they had read as well. Evaluations of the recommendations demonstrated that teachers valued a number of recommendations as applicable for their practice, in particular the information regarding the learning potential of the child and the type of assistance the child needed. Both these reported recommendations were based on dynamic testing outcomes. Further, teachers appreciated the feedback from observations and recommended interactions.
However, the variation in responses between teachers in both groups was high and no differences between teachers reading dynamic reports versus standard reports were found. Explanations of teachers demonstrated that information was often a confirmation of teachers’ own experience with the child. Moreover, it turned out that for some teachers this was the reason for rating the information as neutral, whereas others appreciated the additional value of the recommendations.

The evaluation of the reports and recommendations by teachers of children with arithmetic weaknesses confirmed the above findings. Teachers valued the dynamic reports as meaningful to very meaningful; however, the ratings of teachers reading standard reports were almost equal. Responses on a follow-up questionnaire and report tended to be higher for the recommendation sections of the dynamic reports, yet ratings in both groups varied highly and the average report ratings were comparable.

Outspoken preferences for types of reports were not reported in other studies as well, for example, regarding behavior problems (Salvagno & Teglasi, 1987) and regarding children with special needs (Delclos et al., 1993). Reasons might have been that teachers were in particular attentive to information regarding a child’s general functioning (Delclos et al., 1993) or focused on practical hands-on recommendations (Salvagno & Teglasi, 1987; Haywood & Lidz, 2007). Regarding children with mathematics difficulties it was noticed that teachers appreciated in particular the information with regard to arithmetic level and working memory, which were conveyed in both dynamic and standard reports. Further, we found some indications that teacher’s age and experience may have influenced reports preferences, which was further examined in the next study (chapter 7).

In an internet survey we investigated to what extend elementary school teachers’ background variables affected their opinions regarding the usefulness of information and dynamic testing outcomes for educational planning. The overall findings showed that teachers appreciated information regarding a child’s learning process and learning potential as very useful for educational planning. These aspects are generally not included in standard intelligence reports and this appreciation merits the use of dynamic testing procedures. Inquiry of teachers’ opinions regarding specific examples of
dynamic testing outcomes (Freeman & Miller, 2000) showed additional justification for the use of dynamic testing procedures. The majority of teachers responded that inclusion of these examples would make the report “useful to very useful” for educational planning. On the other hand, teachers judged the information regarding a diagnosis or a child’s limitations also as important information for educational planning. This was as well found regarding the identification of learning disabilities (Machek & Nelson, 2010). It appeared that a combination of both dynamic and standard information provided the best starting point for formulating and developing individual educational plans.

Although expected, teachers’ sense of self efficacy did not relate to teachers’ preferences for information nor to teachers’ values regarding dynamic testing examples. Moreover, teachers’ preferences for information appeared to be influenced by teachers’ experience. Relatively inexperienced teachers tended to give less positive ratings than their older peers. These differences were found for the information about learning processes as well as for the majority of the specific examples of dynamic testing outcomes.

The contribution of dynamic testing in needs-based assessment revisited

The aim of this thesis is to investigate the contribution of dynamic testing to assessment procedures with a focus on recommendations and interventions. When we reflect on the various studies we can conclude that our dynamic test procedures enabled us to measure the need for instruction of elementary school children, across populations of children with severe to mild learning disabilities in special or regular education classes. In particular, the use of the graduated prompts technique provided us information regarding individual needs for instruction and feedback, helping elucidate a child’s learning process and what types of help the child profits from.

Both elements have been found lacking in the outcomes of intelligence tests (e.g. Elliott, 2003), but they are explicitly acknowledged as information necessary to conduct
assessment with the focus on recommendations and interventions (Haywood & Lidz, 2007; Pameijer, 2006). Reports and recommendations, formulated based on these outcomes as described in chapters 5 and 6, were valued by teachers in general as useful and meaningful for their practice and for writing educational plans; however, teacher evaluations were hardly different from teachers reading standard reports in which no information regarding a child’s potential for learning was provided.

Despite the absence of clear report preferences, we may conclude that teachers do appreciate the dynamic testing information. Support for this conclusion can be found in chapter 7, as responses on our survey showed how teachers considered a variety of dynamic testing aspects as valuable information for their teaching practice and for writing educational plans. The fact that standard information, for example IQ score or diagnostic information, was also valued explains perhaps the reasons for not finding differences in report preferences we sought in the studies in chapter 5 and 6. Whether unfamiliarity with the standard reports and information may have played a role has to be examined in future research. Apparently a combination of both standard and dynamic information would provide the best starting point for writing educational plans as seen from teacher’s perspectives. Yet, putting information into practice is a rather big step for teachers. Our expectations regarding changes in teachers’ practice in the short time-span succeeding reading reports and recommendations to indicate the value for teachers’ practice may have been too optimistic. Realizing such forms of change in classroom practice has been difficult to accomplish (e.g., Witt, 1986) and requires a lot of effort from teachers.

Recommendations were overall valued as useful by teachers in our studies, although the variation in responses was high. For some teachers a simple general comment regarding challenging a child or providing feedback was very informative, but other teachers expected more specific recommendations. Explanation and suggestions regarding the amount or type of instruction a child needed were in the latter case judged as not concrete and helpful enough. Possible differences in teachers’ need for information may have complicated our findings. Gauging teachers’ assumptions and expectations prior to assessment is of course an important part of a real clinical
assessment (Pameijer, 2006; Pameijer & van Beukering, 2007); however, in balancing our experiments between a diagnostic, and experimental controlled design, we could only include some teacher variables, but not write reports differently for each individual teacher. The intermediary role of school psychologist should certainly be investigated.

Regarding differences between teachers, we found limited indications that teachers’ age and experience influenced evaluations of recommendations. Although in a small sample (chapter 6) it appeared that older elementary school teachers were more critical in their evaluations, the responses to the survey (chapter 7) indicated that younger teachers appeared overall less positive in their responses towards the usefulness of information. Research in teacher education has shown how beginning teachers for example were found to have lower levels of sense of efficacy (Woolfolk-Hoy & Burk-Spero) while on the other hand teachers in their late career appeared to become less positive about their work (Klassen et al., 2010), which might also affect attitudes towards new ideas and change.
Implications for educational practice and questions for future research

The results of this thesis show that teachers see advantages in using dynamic testing outcomes, which provides information on the potential for learning and need for instructions. This poses certain challenges for practitioners, test-developers and teachers.

The availability of good and valid dynamic tests with reference norms to a broad population is essential to encourage and convince school psychologists to include dynamic tests in their test inventory. At this moment several dynamic tests have been developed, but tests differ in procedures, target groups and hardly have norms available.

However, the challenge of longer administration time and complication of scoring when administering dynamic tests should be taken into account, as these are difficult to match with the current school psychology practices. Luckily, recent studies regarding computerized testing have shown that computerized versions of dynamic tests are as effective as the traditional paper and pencil versions (e.g. Stevenson, Touw & Resing, 2010), with additional benefit of automatic scoring and report generation. Moreover, scaffolding procedures on an electronic console or a tangible tablet are currently developed (e.g., Henning, Verhaegh & Resing, in press; Resing & Elliott, in press). Certainly these applications will be helpful in supporting the administration and scoring of the provided prompts.

We stress once again that measuring and recording individual needs for instruction on cognitive tasks provides a very clear picture of a child’s potential for learning and this information may have implications for the education a child needs. Comparisons with IQ results have shown how the potential of a certain group of children is underestimated, which is another argument for frequent use and further development of dynamic tests, in particular for children with special needs (e.g. Hessels-Schlatter, 2002), with a different cultural background (e.g. Hessels, 2000; Resing, Tunteler, et al., 2009) or with low social economic status (Tzuriel, 2000).
Yet, the question remains whether identification of a child’s need for instruction is informative enough for providing recommendations to teachers concerning their teaching practice. On one hand the familiarity of teachers with ideas and theories regarding dynamic testing needs improvement and it will help teachers to understand the concept behind the recommendations. In a small study (Verdel, 2007) we found that teachers were more positive about recommendations after viewing a video of a dynamic testing process, and familiarity may help teachers as well in understanding reports better (e.g., Delclos, Burns & Kulewicz, 1987). Therefore it is suggested that principles of dynamic testing are taught in teacher education programs and in regular educational workshops.

Dynamic testing could serve very well as a screening instrument regarding children’s need for instruction and provide a starting point for designing a suitable educational setting, as well as learning goals and strategies a child profits from. Moreover the instrument, complemented with for example observations during a reversal task, could guide (remedial) teachers in individual training (e.g., Fuchs, Fuchs, Compton, Bouton, Caffrey & Hill, 2007), in classroom-based interventions (e.g., Hessels, Hessels-Schlatter, Bosson & Balli, 2009) or guide clinical practitioners working with children with (severe) disabilities (e.g., Hessels-Schlatter, 2010). Cooperation between practitioners, teachers and test developers should aim at a balance between measuring children’s potential, needs or strategies, and the request and use of the practice, to provide appropriate education and interventions.