Chapter 7

Discussion
7.1 Introduction

The Bayley Scales are commonly used to assess the general development of infants. It offers the opportunity to determine norm scores for cognitive and motor development in a standardized and (because of the attractive play materials) child-friendly way. In combination with the judgment of the behavior of the child, test results can be used to compare the early development of a child to a relevant reference group. The test is very suited for this purpose (e.g. Neisworth & Bagnato, 1996; Dunst, 1998), but has limited use with children with specific impairments. These children may be unable to (optimally) use the common (physical and sensorial) functions to perform on a standardized instrument. Without suitable adaptations, test results are often biased and do not reflect the true level of development. The purpose of this research was to develop adapted versions of the BSID-II-NL in such a way as to allow children with specific impairments, maximum opportunity to show their cognitive and motor skills and at the same time allow the use of the standard norm tables. The adaptations should only compensate for the child’s impairment; item content and difficulty should remain unchanged.

This dissertation is a description of the construction and psychometric qualities of a standardized instrument for determining the current developmental level of young children and also of adapted versions of this instrument that allow a better assessment of children with specific impairments. The main research question to be answered was: *Is the Dutch translation and adaptation of the BSID-II a valid instrument for individual developmental assessment in general and for children with specific impairments in particular?*

Our research, as described in previous chapters, shows positive results. The translation and adaptation of the BSID-II for use in the Netherlands proved to be valid en reliable and the adapted materials for applying the test to children with hearing and/or speech/language problems (BSID-II-NL non-verbal), children with visual impairments (BSID-II-NL Low Vision) and children with motor impairments (BSID-II-NL Low Motor) proved to be highly valued by practitioners and results from validity studies,
though limited, support the assumption that the adaptations can be characterized as accommodations (and not modifications) so that the standard norm tables still apply.

The findings are however subject to some limitations. The specific limitations to the five studies collected in this dissertation are described in the discussion at the end of each chapter. In this chapter (7) important limitations are discussed that concern all studies. First, the criticism concerning the use of the BSID-II in clinical situations is discussed. This criticism mainly focuses on applying the test procedure and basal and ceiling rules when administering the test to children with severe developmental delay. Second, the limited validity and reliability studies are discussed. Especially with regard to the adapted versions, extended validity studies are necessary to support the assumption that item content and difficulty has remain unchanged so that the standard norm tables still apply, whatever (combination of) version(s) of the BSID-II-NL that is used. Next to these limitations, the general purpose of developmental assessment is discussed and the necessity of applying standardized, norm-referenced, but adapted instruments to children with specific impairments.

7.2 Test procedure and basal and ceiling rules

Several years of use have led to some criticism about the use of the BSID-II. Next to remarks about materials and item instructions, an important subject of debate is the influence of the starting point in combination with the basal and ceiling rules on the test results of a child. In some cases it might be advisable not to start in the chronological age group, but in an age group directly above or below the chronological age. However, several researchers (e.g. Matula & Aylward, 1997; Gauthier, Bauer, Messinger & Closius, 1999) have shown that the scores on the BSID-II are influenced by the item group in which the test is started. The lower the item group that the test is started in, the lower the final score. These differences in scores can be substantial in individual cases and are most likely to occur in at-risk children with an uneven development (Ross & Lawson, 1997). To minimize the influence of the testing procedure on the end result, adapted ceiling and basal rules are used in the BSID-II-NL. The extended basal and ceiling levels offer a better possibility for administering the items from different age
groups, allowing us to get a broader picture of development. To ensure a standardized administration of the test and to make sure that it isn’t determined after the test that a child would have reached the basal level in higher age groups as well (and therefore would have reached a higher standard score) it is important to always take the item group for the chronological age (or as close to it as possible) as the starting point of the test. The standardization study was performed this way, and only an identical operating procedure allows for a valid use of the norm tables.

7.3 Adapted versions

To facilitate adequate decision making (e.g. school type), effective interventions and follow-up for children with specific disabilities, it is essential that appropriate and fair assessment is possible. As Davidson & Dolins (1992) mention in their discussion of assessment of the young child with visual impairment and multiple disabilities, test result from a core instrument (such as the Bayley Scales; see for the use of the BSID-II as a core instrument also Mccune, Kalmanson, Fleck, Glazewski, & Sillari, 1990) in combination with more specific instruments (such as a language tests) should enable the clinician to distinguish between global developmental delay or disability and deficits due to a specific disability. Putting the psychological test results into perspective, by combining the information from the physical, behavioral and environmental assessment results, will provide the clinician (and parents) with information concerning the physical, behavioral and psychological causes for the developmental delay.

An adapted diagnostic tool will enable the child to perform on a test without being hindered by its specific impairment, e.g. motor, visual or language impairment. Specifically adapted test procedures, item instructions and play material should take into account the child’s reduced sensory and/or physical possibilities and delayed information processing, but should not significantly alter the content or difficulty of the test items. At the same time, an adapted and standardized instrument enables better research on the general development of children with specific impairments. The use of more applicable instruments can provide us with information on how a specific impairment influences a child’s development compared normal development. Finally, with regard to the
motivation of the child to perform in test situations, adapted play material will increase the child’s intrinsic motivation. It allows the child to better explore and manipulate the material and therefore enhance the opportunity to show its skills in a test situation.

With the development of adapted versions of the BSID-II-NL, we provide diagnostic tools that are based on a well-known, psychometrically sound instrument that can be used to examine the cognitive and motor development of children with specific impairments in an adapted, but standardized way. Although each adaptation of the test procedure, item instruction and/or play material relevant to the standard instructions implies that the original norm tables no longer apply, the adaptations we propose are only meant to compensate for the child’s impairment. By leaving the item content and degree of difficulty essentially unchanged the original standardization will apply as before. Batshaw-Clair, Church, and Batshaw (2002) described these changes that do not alter in a significant way what the test measures as accommodations. In contrast, when changes in the assessment alter what the test is supposed to measure or the comparability of the scores, the term modifications is used. If sufficient evidence shows that the adapted versions of the BSID-II-NL involve only accommodations, it will be unnecessary to conduct large-scale and time-consuming standardization research for the specific groups of children mentioned in this dissertation. To obtain sufficient support for a valid use of the standard norm tables with adapted versions, extended validity studies should be conducted. The research described in this dissertation was meant primarily to construct adapted test procedures, item instructions and play material for adequate and appropriate use of the BSID-II-NL with children with specific impairments and to gain (indicative) information on the psychometric qualities of the adapted versions. For the non-verbal and the low motor version we performed a more extensive validity study. Results from these studies support the validity of the adapted versions. With regard to the Non-verbal version, however, more research is needed to examine the appropriateness and value of the adaptations for other clinical groups of children, e.g. children with hearing impairments and children with specific language difficulties. For the Low motor version, continued research is needed to extend the sample sizes. Despite the substantial number of centres that participated in the research, the amount of research data was lower than expected. With regard to the validity of the Low vision version only results from pilot
studies are incorporated in this dissertation. The results of these studies are positive, but too limited (as a result of sample size) to conclusively answer the research question whether the low vision adaptations to the standard version are applicable for measuring cognitive and motor development in young children with low vision (30-42 months). Meanwhile a follow-up study is being conducted with an adapted BSID-II-NL for children with low vision in the age range of 1 – 42 months. Incorporating the results from the pilot study described in chapter 5 with this new experimental version will facilitate large scale research on the applicability and psychometric properties of the adapted BSID-II-NL for assessing children with low vision.

The adaptations of the standard BSID-II-NL for using the test with prematurely born children differ from the adaptations made in the non-verbal, low vision and low motor versions. No adaptations were made to the item instructions and test material, except for a recommended alternative test procedure to take into account the developmental pattern of this group of children. Future research should provide us with the answer to the question until what age correction for prematurity is recommended and whether corrected or uncorrected test results reflect the child’s current level of functioning in a more realistic and valid way.

In discussing the purpose of assessment, some practitioners mention the decrease of generalizability of the test results when an adapted version of a standardized test is administered to a child with a specific impairment. Performance on tasks in a completely adapted situation does not provide practitioners with a realistic picture of a child’s functioning in the daily situation. The purpose of psychological assessment with the BSID-II-NL is however not to provide the user with practical advise for daily living situations. Besides the fact that the test has limited predictive value, the purpose of the test is only to determine the current developmental level of a child. This can then be used to decide eligibility for support, intervention and education (e.g. Black & Matula, 2000; Lichtenberger, 2005, Dunst, 1998). Besides this, information from the adapted tests can be used in practice by observing to what degree a child profits from the adaptations when performing the tasks. Information from an adapted test situation can be used to adapt the child’s learning and living environment to its abilities to maximize its functioning and development.
The question can be raised why we prefer to compare a child with specific impairments to children from the standard population and not, instead, develop separate norms. We have three reasons for doing so. The first reason concerns the necessity to adhere to formal guidelines. The Dutch government has developed guidelines with regard to formal indication of children for intervention programs and education. Results on standardized, norm-referenced tests are required and often the decisive factor in determining eligibility of a child for appropriate intervention and education. Developing separate norms would therefore be less relevant. The second reason has to do with the fact that the development of separate norms raises considerable practical and methodological difficulties. The limited number of children with specific impairments make it difficult to compose a representative sample that consists of a sufficient number of children. And in our opinion it would be unfair to compare the abilities of a severely motor-impaired child to those of a child with mild motor problems. A relevant number of the children is impaired in such a way that not every item in the adapted versions will enable them to show the required cognitive skill. Because of this, developing separate norms will be complicated and time consuming. The third reason concerns the generalizability of the adapted versions worldwide. By only compensating for the child’s impairment and leaving the item content and difficulty unchanged, the adapted versions of the BSID-II(-NL) can be used worldwide.

The use of standardised psychodiagnostic instruments with children with motor or visual impairment, hearing impairment and/or language/speech disabilities or children with multiple disabilities, strongly appeals to the experience of the test administrator with children with specific impairments, his or her knowledge of the physical and sensory possibilities (and impossibilities) of this specific child and secondary developmental problems (e.g. behavioural problems), and last but not least the know how of good and appropriate instruments and the extent to which the administrator is familiar with the test content and procedures. With two of the described studies in this dissertation (part of the) tests were administered by students orthopedagogics from the University of Groningen. With concern to the above mentioned demands for an optimal administration of the test, we preferred the clinicians working with these specific groups of children to administer the tests with the selected children. If, however, this would have been a necessary
condition for setting up a collaborative study, we would not have been successful at all in collecting data for the validity studies that were conducted with the Non-verbal and Low-motor version of the BSID-II-NL. Reason for this was simply lack of time. Orthopedagogues and psychologists often are very limited in their clinical time, let alone their possibilities to cooperate in research by administering tests for only research purposes. The communication with the researchers and coordinating the internal organisation of the test administrations (selection of the children, informing parents, making reservations for the testing rooms) already strongly appealed to their time schedule. For this reason only, students were also employed to administer the tests at the centers throughout the Netherlands. Important advantages of working with these students were that they all received the same (certified) training and they could be monitored more easily than clinicians working in the centers. To meet the demands concerning the experience with the clinical group of children and the specific problems of each child to be tested, ample attention was given to inform the students about the specific groups of children and for each child to be tested, the clinician sent individual information relevant to the test administration. In cases where this was found necessary (by the clinician) a teacher or care-taker was present during the test administration.

7.4 General purpose of assessment

Three important reasons for early assessment and the use of psycho-diagnostic instruments with children who are at risk for developmental delay are: (a) to aid the development of infants and toddlers with disabilities and to minimize lasting effects of the child’s difficulties (Kenny & Culbertson, 1993; Conlon, 2002), (b) to promote a child’s successful adaptation to the environment (Yeates & Taylor, 1998), and (c) to enhance the capacity of families to meet the needs of children with developmental problems (Culbertson & Willis, 1993). This last reason also implies that good and fair assessment will clarify the causes of the developmental problems. This can be of help to parents to first of all accept the fact that their child has a developmental delay or disability and second, to help them to adjust their ideas about the future of their child and the role they will play as a parent. All young children have the right to learn and live
under conditions that enables development to their full potential. Especially those children with physical problems (e.g. very premature infants) and general or specific impairments (e.g. visually or motor impaired children) need optimal support and possibilities to explore and understand the world around them. To offer a child and its family the kind of guidance and support that meet their specific needs, it is essential to gain insight into the developmental course of the child and its environment. The content of this dissertation is limited to research of a psycho-diagnostic instrument that plays an important role in the assessment of young children. However, it cannot be emphasized enough that such an instrument should never be the sole indicator for making diagnostic and classification decisions. To put test results into perspective and to get a broader picture of a child, an assessment procedure should always include instruments to gain insight into the daily learning and living situation of a child as well. This also accommodates the shift from a child-centered to a family-centered approach in early intervention. All children are part of a family, and both the family’s and child’s needs must be met so that both the child and the family’s circumstances are improved (Hauser-Cram, Upshur, Kraus & Shonkoff, 1988).

**End note**

Permission for publication in the Netherlands and internationally is a condition for implementation of the adapted versions. The non-verbal version was published in 2005 (Ruiter, Hoekstra, Van der Meulen & Lutje Spelberg, 2005), as part III of the Dutch manual (and published in Ruiter, Hoekstra, Van der Meulen, Lutje Spelberg, & Nakken, 2006) and is also used by German psychologists. Publishing the Low vision and Low motor version, however, is more complicated. The Dutch department of Harcourt, publisher of the BSID-II-NL, supports our research, but has to adhere to the strict licency policy of Harcourt USA. Most of the centres that were involved in our research continue to use the experimental versions in anticipation of the official publications. We hope the publisher will recognize the importance of allowing these versions to be published and used in practice.
References


Abstract

In this doctoral thesis the Dutch standardization and validation of the Bayley Scales of Infant Development – Second Edition (BSID-II-NL) and its adapted versions for use with children with specific impairments are presented. The BSID-II-NL is a standardized and norm-referenced instrument to evaluate the general development of a child between the (developmental) age of 1 – 42 months. The test consists of a Mental scale, a Motor scale and Behavior Rating scale. To facilitate appropriate and fair assessment for children with specific disabilities, we developed adapted versions of the BSID-II-NL. By adapting test procedures, item instructions and/or play materials for specific groups of children, the test facilitates more adequate decision making in providing the right support, intervention or education. The research described in this thesis resulted in a Dutch translation and standardization of the BSID-II, adapted test procedures for using the test with prematurely born infants, a non-verbal version for children with hearing loss and/or speech and language problems, a low vision version for children with visual impairment and a low motor version for using the test with children with motor impairment. The main research question was: Is the Dutch translation and adaptation of the BSID-II a valid instrument for individual developmental assessment in general and for children with specific impairments in particular?

This question is answered, first, by analyzing results from validity and reliability studies for the BSID-II-NL for children from the standard population (norm population) and second, by analysing the value and validity of adapted versions of the BSID-II-NL for use with children with specific impairments. Account is given of research into the limitations of the BSID-II-NL for use with special groups of children and possible solutions in the form of constructing adapted versions of this instrument.

In chapter 2 the construction, standardization and psychometric characteristics of the BSID-II-NL are described. The BSID-II-NL contains of a translation of the original (US) materials and a Dutch standardization for the Mental, Motor and Behavior Rating Scale. The Dutch version was made in such a way as to preserve the original version as closely as possible. Only two important adaptations were made: a broadening of the basal and ceiling level and a different factor structure of the Behavior Rating Scale. The norms
of the three scales are based on data of 1909 Dutch children. Norm tables are constructed through a method that uses a fit procedure based on the score differentiation of all age groups together. This method produces a regression equation that enables the construction of the printed norm tables per month, but allows also for a more precise conversion of raw score into day norms (via a computerized program). Validity and reliability studies proved the BSID-II-NL to be an empirical sound instrument.

Chapter 3 discusses the influence of the BSID-II(-NL) administration procedure on the assessment of premature infants. Two suggestions for improvement are given. First the Dutch version of the BSID-II is discussed and how it may solve the problems resulting from the use of items sets and the adjusted basal and ceiling rules. Second, an alternative administration procedure is put forward, that allows the use of norm tables, but takes into account the deviating developmental pattern of premature infants.

The non-verbal version of the BSID-II-NL is presented in chapter 4. The non-verbal version contains a selection of the items of the original Mental scale and, where needed, adapted instructions are provided. These adaptations consist of extra emphasis on natural gestures or pantomime instructions. Separate norm tables were developed for the age range of 12 – 30 months. To further examine the validity of the BSID-II-NL NV, a comparative study was conducted between the standard and non-verbal version of the Mental scale of the BSID-II-NL with a group children with serious developmental delay. The results support the applicability and validity of the non-verbal version.

The adapted version of the BSID-II-NL for assessing children with visual impairment (BSID-II-NL Low Vision) is described in chapter 5. Test procedures, item instructions and play material of the Mental and Motor scale are adapted for a more appropriate use of the test with this specific group of children. The purpose was to preserve the original item content and degree of difficulty. By leaving the items essentially unchanged, the standard norms are still applicable to a ‘low vision’ administration in a valid way. This chapter describes the study that has produced an experimental version of the BSID-II-NL Low Vision for the age range of 30 – 42 months. A pilot study was conducted to find out how appropriate the adaptations are for children with low vision and to provide us with a first impression of the validity of the BSID-II-NL Low Vision.
Chapter 6 presents the BSID-II-NL Low Motor, an adaptation of the BSID-II-NL for a more appropriate use of the test with children with motor impairments. We adapted test procedures, item instructions and play material of the Mental scale for the age range of 12 – 42 months. Results from the validity study that are described in this chapter, support the assumption that low motor adaptations only compensate for the child’s motor impairment. Item content and degree of difficulty of the items did not change significantly, so that the standard norm tables are still applicable when administering a Low Motor version of the BSID-II-NL. The Standard and Low Motor version of the BSID-II-NL were administered to twenty children experiencing normal development and 45 children with low motor, within a period of two weeks. Children with low motor score significantly higher on the Low motor version compared to the Standard version. No significant difference was found for the group of normal children.

In chapter 7 the main research question is answered and general limitations, and implications for future research are discussed. Although more research is needed, results from the studies described in this thesis, provide us with empirical support for the conclusion that the Dutch translation and adaptation of the BSID-II is a valid instrument for individual developmental assessment in general and for children with specific impairments in particular.