Chapter 6

The BSID-II-NL for assessing children with motor impairment

Abstract

Most of the developmental instruments that measure cognitive development in young children rely heavily on fine motor skills, especially for young children whose language skills are not yet well developed. This will typically result in low scores for motor-impaired infants. To be able to adjust guidance, intervention and education to the developmental level of a motor-impaired young child, it is necessary to gain insight into the cognitive development of a motor-impaired child, based on an assessment with a standardized and specifically adapted diagnostic tool. For this reason, we have adapted the procedures, item instructions and play material of a standardized instrument, the Bayley Scales of Infant Development –Second Edition (BSID-II). Results of validity studies support the assumption that low motor adaptations only compensate for the child’s motor impairment; item content and level of difficulty of the items did not change significantly. The Standard and Low motor versions were administered to twenty children experiencing normal development and 45 children with motor impairment within a period of two weeks. Results of a validity study support the value of the use of the Low-motor version with children with motor disabilities. Low-motor children score significantly higher on the Low motor version than on the Standard version. No significant difference was found for the group of “normal” children between test results on the standard and low motor version of the BSID-II-NL. This supports the assumption that item content and difficulty have remained unchanged. Because of this, the norm scores of the standard BSID-II-NL, can still be used with the low-motor version of the test.

4 This chapter has been submitted for publication, slightly adjusted and was co-authored by Bieuwe F. van der Meulen, Han Nakken and Henk C. Lutje Spelberg.
6.1 Introduction

If a child has a developmental delay or disability, early identification and intervention can be crucial in order to minimize lasting effects (Kenny & Culbertson, 1993; Conlon, 2002) and to promote a child’s successful adaptation to the environment (Yeates & Taylor, 1998). In his hypothetical model for explaining cognitive delay in young children with motor impairment, Vriesema (1990. p. 37) mentions several factors that increase the risk of cognitive delay in motor-impaired infants: lack of ability to explore the world, perceptual disorders, inadequate stimulation by the child’s social environment and lack of intrinsic motivation to explore the world and perform on tasks. To facilitate adequate decision making (e.g. school type), effective interventions and follow-up for low-motor children it is essential that appropriate and fair assessment is possible. An adapted diagnostic tool will enable the child to perform on a test without being hindered by its motor impairment, taking into account the child’s reduced fine motor skills, possibly less developed eye-hand coordination and motor-planning skills and delayed information processing. At the same time, an adapted and standardized instrument enables better research on the specific development of low motor children. 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.

Almost all developmental instruments that measure cognitive development in young children have a strong motor component. Especially with young children (when language skills are not yet well developed) performance on cognitive tasks highly relies on the ability to explore and manipulate play material. Worldwide, only a few diagnostic developmental tests are available that measure young children’s cognitive development without appealing strongly to the motor abilities of a child. Well known instruments are: The Kent Infant Developmental Scale (KIDS, Reuter & Brickett, 1985), the Fagan Test of Intelligence (Fagan & Shepherd, 1991) and the Snijders Oomen Non verbal intelligence test 2½ to 7 years (SON-R 2½-7; Tellegen, Winkel, Wijnberg-Williams, & Laros, 1998).
These scales all cover a large part of the aspects of a child’s cognitive development, but do not specifically take motor disabilities into account. The KIDS evaluates psychomotor development based on information from parents on 252 behavioral items grouped into five domains: cognitive, language, motor, self-help, and social. Although it is not a traditional instrument where a child needs to perform on tasks to show its cognitive skills, Stancin, Reuter, Dunn, & Bickett (1984) concluded from their research that developmental age estimates from the KIDS are highly correlated and similar to results of a traditional test like the Bayley Scales (BSID). Major drawbacks of the test are, however, the limited age range (1 to 14 months) and the fact that data are entirely dependent on informant report; this increases the possibility of bias. The Fagan Test of Infant Intelligence (FTII) is a screening device for testing visual recognition (novelty preference). In examining the predictive validity of the FTII, Tasbihsazan, Nettelbeck, & Kirby (2003) concluded that correlations between the BSID-II and the FTII were low and non-significant and that clinical utility is limited as a result of low sensitivity. Meta-analyses have found average correlations between results on the FTII and later scores on intelligence tests (e.g. Bornstein, Slater, Brown, Roberts, & Barrett, 1997). Apart from low demands placed on motor skills, the limited age range and non-conclusive validity data, make the FTII less appropriate as an alternative to the BSID-II-NL for low motor children. The SON-R is a non-verbal intelligence test for young children that does not ask much of the fine motor skills of a child, but the manual does not give standard low-motor adaptions for the test procedure and item instruction. It is left to the clinician how and to what degree adapted procedures and/or item instructions are necessary in individual cases. (Tellegen, Winkel, Wijnberg-Williams & Laros, 1998).

These tests (KIDS, FTII and SON-R) do not specifically take motor impairment into account. Low-motor children will, solely due to their impairment, get lower scores on certain items. Test results will not give a clear picture of the child’s cognitive abilities.

To compensate for the lack of standardized and adapted diagnostic instruments and concerns about motor impairment bias when using standard procedures and materials with motor-impaired children, we have developed an adapted version of a well known, standardized assessment instrument, the Bayley Scales of Infant Development – Second edition (BSID-II: Bayley, 1993). The Bayley is a widely used instrument for assessing
young children’s cognitive and motor development, also known as the ‘golden standard’ (Aylward, 2002; Gauthier, Bauer, Messinger, & Closius, 1999). Our current study was based on the Dutch version of this test (BSID-II-NL; Van der Meulen, Ruiter, Lutje Spelberg, & Smrkovsky, 2002). The Dutch version consists of a translation of the BSID-II, a Dutch standardization and a discussion of validity and reliability studies for the BSID-II-NL (Ruiter, Lutje Spelberg, & Van der Meulen, 2005).

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 in format, response, environment, timing or scheduling 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 modification is used. If sufficient evidence shows that the adapted version of the BSID-II-NL involves only accommodations, it will not be necessary to conduct large-scale and time-consuming standardization research for this specific group of children. The use of a standardized low-motor version, with adapted test procedures, item instruction and play material, combined with the original norm tables will make it possible to compare the cognitive development of low motor children with that of their average peers.

A pilot study provided us with the answer to the question whether the experimental version of the BSID-II-NL Low Motor is an applicable instrument for measuring cognitive development in young children (12 – 42 months) with motor impairment. Administering the pilot version to 22 low-motor children and using the judgments of experts (n = 15) from the practical as well as scientific field, we were able to show positive results with regard to the applicability and usability of the low motor version of the BSID-II-NL. The adapted testing procedure, item instruction and play material met the requirements of applicability for this target group. Suggested improvements and changes of item procedures, instructions and/or play material were evaluated and, when determined as clear and usable improvements, assimilated in the
final version of the BSID-II-NL Low Motor. The final instrument consists of 108 items; items that were acceptable as they were (29), biased items that could not be changed to be acceptable (10), items that required only material adaptation (13), only instruction adaptation (9) and items that required instruction and material adaptation (47).

In this chapter we give an account of validity studies for the BSID-II-NL Low Motor. Preliminary results of these studies provide us with an indicative answer to the research question: *Is the BSID-II-NL Low Motor a valid adaptation for use with children with motor impairment that allows the use of the original norm tables?*

### 6.2 Method

**Subjects**

Sixty five children participated in this study. Twenty children represent the standard population: all children with average development and without any specific physical or mental problems (in this chapter referred to as the group ‘normal’ children). All children were between 12 and 42 months old and all come from the city of Groningen. This group already participated in the standardization study of the BSID-II-NL. Their parents were asked to participate in this study as well.

Forty five children with motor impairment (in this chapter referred to as the group low-motor children) represent children with mild to severe motor impairment as a result of many different syndromes, anomalies or diseases. All low motor children are known to one of the centers for assessment and intervention for motor-impaired children in the Netherlands or to child-day-care centers for children with serious developmental delay. Children were recruited through the educational psychologists working in the centers. A clear description of this group, based on the diagnosis, is not possible because of the diversity in diagnoses and heterogeneity within the group of children with the same diagnose. Children with the same diagnosis may experience very different limitations in motor abilities, to a different degree, and additionally, often in combination with sensory problems. We decided to describe the target group by formulating inclusion criteria. These criteria describe the minimal necessary abilities of a child to be able to perform on the test items.
The low motor children met the following inclusion criteria:

- The child is able to sit up in a (wheel) chair to work at a table
- The child has sufficient visual abilities
- The child can use at least one hand

The (developmental) age of the low motor children is estimated by educational psychologists working with the children, between 12 and 42 months.

Table 1 provides an overview of the number of children that participated in this research, divided in the clinical group (low-motor children) and the children from the standard population (normal children).

Table 1
Relevant numbers on the low-motor children and normal children and the number of single Low motor administrations and the number of administrations of the Low motor in combination with the Standard version.

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Boy</th>
<th>Girl</th>
<th>Mean age (months)</th>
<th>Age range (months)</th>
<th>Single Low motor administration</th>
<th>Low motor and Standard administration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low motor children</td>
<td>45</td>
<td>29</td>
<td>16</td>
<td>57.1</td>
<td>15-300</td>
<td>23</td>
<td>18</td>
</tr>
<tr>
<td>Normal children</td>
<td>20</td>
<td>6</td>
<td>14</td>
<td>29.5</td>
<td>12-42</td>
<td>-</td>
<td>20</td>
</tr>
</tbody>
</table>

Experts of 24 centers for rehabilitation (centers for assessment, intervention, and education for motor-impaired children) and eight child day-care centers (centers for children with severe developmental problems) agreed to fill out surveys after administering the low-motor version of the BSID-II-NL to a low-motor child. All experts are qualified educational psychologists, who have ample experience in the assessment of young low-motor children.
Instruments

The BSID-II-NL. The BSID-II-NL consists of three scales, the Mental, Motor and Behavior Rating scale. Most of the items in the Mental and Motor scale are task items. The test administrator must try to produce wanted behavior in the child by presenting stimuli. The Mental scale consists of 178 items that measure the child’s cognitive skills. The Mental scale consists of items related to the processing of visual and auditory information, eye-hand coordination, imitation, language development, memory and problem solving. After administrating the Mental scale a raw score is calculated (RS<sub>MS</sub>) and a mental development index (DI<sub>MS</sub>) is determined. A 90% reliability interval is also provided, and it is possible to determine the classification of the score: seriously delayed, delayed, normal or accelerated development.

The Motor scale (MR) consists of 111 items that measure skills related to gross and fine motor control, including movements like rolling, crawling, standing, walking, running and jumping. This scale also tests fine motor manipulations, such as eye-hand coordination, adjusted use of writing materials and imitation of hand gestures. For the Motor scale a raw score (RS<sub>MR</sub>), a motor development index (DI<sub>MR</sub>), a reliability interval and a classification are also determined (as in the Mental scale).

The Behavior Rating scale (BRS) contains questions that are answered by the administrator after administering the test. The questions allow the administrator to judge the behavior of the child during the test. Raw scores can be converted to percentile scores. (A comprehensive description is provided in Van der Meulen, Ruiter, Lutje Spelberg & Smrkovsky, 2002 (Dutch version) and Bayley, 1993 (original version)).

The BSID-II-NL Low Motor. With the positive results of the pilot study, the final version of the BSID-II-NL Low Motor was developed. The general test procedure is adapted to the situation of testing a child with motor impairment. For instance, it is allowed to support the arm of the child by the elbow, to enable the child to control its arm/hand. In reviewing each item on the Mental scale, not only the content was addressed, but also the procedure for item instructions for the administrator and the response possibilities for the child. The item instructions of the Low Motor version include e.g. elimination of the time limits so that children can be given sufficient time to
explore the material and perform on the task. In addition, children are allowed to use eye-pointing (glance direction) in stead of finger pointing. Finally, adaptations to play material include for instance enlarged materials (e.g. pegs and peg board, puzzle board, and cubes), and separated pictures in the stimulus booklet in combination with the see-through-frame. This is a perspex framework that stands on the table between the child and administrator on which pictures can be attached (well apart) so that eye pointing of the child can clearly be observed by the administrator.

The adaptations in test procedure, item instruction and play material are an addition to the standard material of the BSID-II-NL. The Low Motor manual only provides adapted instructions for administering the BSID-II-NL Low Motor (12-42 months). The low-motor play materials are applied in combination with standard material. The materials are interchangeable; depending on the child’s degree of impairment and therefore need for adaptation, standard or adapted play material is used.

Procedure

The group of normal children were examined by master students orthopedagogy in the playroom of the orthopedagogy department in Groningen. The students were all experienced BSID-II-NL administrators and they received an intensive training in administering the Low Motor version. All normal children were examined within a period of two weeks in counterbalanced order, once with the Standard and once with the Low motor version. The low-motor children were examined by experts with ample experience in the assessment of low-motor children and in administering the BSID-II-NL. They received training in the Low Motor version as well. All children were examined at the centers.

Validity is examined in four ways:

(1) By comparing test results on the Standard (BSID-II-NL) and Low Motor version (BSID-II-NL Low Motor).

Within a period of two weeks, children from the standard group and the low motor group were tested twice; once with the standard version and once with the low motor version in counterbalanced order and under the same conditions. The expectation for the normal children was that test results on the Standard and Low motor version do not significantly
differ. For the low motor children the expectation was that applying the Low Motor version would maximize test results, and the test results on the Low Motor version would significantly surpass those on the Standard version. As a result of the adaptation, validity of the test for the group of low-motor children will increase.

(2) By comparing test results on item level.
Test results on the Low Motor and Standard version of the BSID-II-NL for both groups of children were compared on item level. Items were assigned to several categories. One category is formed by the non-verbal items (category 0), the other categories (1-5) are based on the type of adaptation (see figure 2) The expectation was that low-motor children will get significantly higher test scores on the items in the categories 0, 2, 3 and 4. Test results for normal children should not vary significantly between the low motor and standard version.

(3) By comparing test results on the Low Motor version and expert judgments.
All test results of motor-impaired children on the BSID-II-NL standard and low motor version are compared to the developmental level of the child estimated beforehand by the experts (administrators), based on file information. The child’s expected performance is rated on a scale 1 to 5: 1: very seriously delayed; 2: seriously delayed 3: delayed; 4: average; 5: above average. The expectation is that expert judgments correlate significantly higher with the test results on the Low Motor version than on the Standard version of the BSID-II-NL.

(4) By asking the experts to fill out surveys after applying the low motor version of the BSID-II-NL to a low-motor child about how appropriate the BSID-II-NL Low motor adaptations were for this specific child and are expected to be for the target group in general.
6.3 Results

The collected preliminary data allows us to present preliminary results on the comparison between test results on the Standard and Low Motor version for the normal children and the low-motor children and also to analyze expert judgments on the applicability of the BSID-II-NL Low Motor with regard to specific children within the group of low-motor children and in general. The results of the four approaches to examine validity (as mentioned under ‘method’) are discussed successively.

(1) With regard to the comparison between the test results on the Standard version and the Low Motor version of the BSID-II-NL, figure 1 shows differences between the test results on the Low motor and Standard version of the BSID-II-NL within the group of low motor children and within the group normal children. The raw scores in the figures are calculated as follows: all items until the basal-item plus all positive items belonging to a specific category within the age groups that are administered.

![Figure 1](image)

*Figure 1*

*Differences between the raw scores of low motor (n=18) and normal children (n=20) on the Low motor and the Standard version of the BSID-II-NL*
The results of performing the Wilcoxon signed-ranks test, in order to compare test results of the standard version of the BSID-II-NL and the low motor version in both the normal group and the low-motor group, are shown in table 2. Due to age (>42 months) or a raw score that was too low to convert to a developmental index, results are based on raw score and developmental age equivalents.

Table 2
Wilcoxon signed-ranks test

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Z</td>
</tr>
<tr>
<td>Low motor children</td>
<td>18</td>
<td>-2.85</td>
</tr>
<tr>
<td>Standard children</td>
<td>20</td>
<td>-1.23</td>
</tr>
</tbody>
</table>

**: significant at 1 % level

As expected, the Wilcoxon tests comparing the results on the standard and the low motor version reveal that significant differences occur in the low motor group. Raw scores and (thus) developmental age equivalents in this group are significantly higher on the low motor version than on the standard version of the BSID-II-NL. The results show no significant difference for the normal children when comparing results on both versions of the test.

(2) Both for the low motor children and the normal children results on item level are analyzed. As already mentioned, items (between the age of 12-42 months, nr. 71-178) are assigned to several categories. Figure 2 shows a comparison of the results on the standard and low motor version of the BSID-II-NL, both for low motor (n=18) and normal children (n=20).
Category 0: Non-verbal scale (selection of non-verbal items on the Mental scale)
Category 1: items are acceptable and do not require adaptation
Category 2: items with instruction adaptation
Category 3: items with material adaptation
Category 4: items with instruction and material adaptation
Category 5: items are biased but cannot be changed to be acceptable.

Figure 2

Raw score per item category for normal (n=20) and low motor (n=18) children
Based on the raw score, table 3 shows the results of the Wilcoxon signed ranks tests for every item categorie, compared between the Low Motor and Standard version of the BSID-II-NL.

Table 3

<table>
<thead>
<tr>
<th>Category</th>
<th>Low motor children</th>
<th>Normal children</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>0 non-verbal material</td>
<td>-2.61</td>
<td>-1.10</td>
</tr>
<tr>
<td>1 unchanged, acceptable</td>
<td>-1.77</td>
<td>-.55</td>
</tr>
<tr>
<td>2 material adaptation</td>
<td>-2.05</td>
<td>-.09</td>
</tr>
<tr>
<td>3 instruction adaptation</td>
<td>-1.41</td>
<td>-.54</td>
</tr>
<tr>
<td>4 material and instruction adaptation</td>
<td>-2.90</td>
<td>-1.96</td>
</tr>
<tr>
<td>5 unchanged, but biased</td>
<td>-1.41</td>
<td>-1.93</td>
</tr>
</tbody>
</table>

| Sig.                                  | .00**              | .27             |
|                                       | .076               | .58             |
|                                       | .04*               | .93             |
|                                       | .15                | .59             |
|                                       | .00**              | .05             |
|                                       | .16                | .05             |

* significant when p < .05
** significant when p < .01

Table 3 shows significant differences between test results on three item categories in the low-motor group on the Low Motor and Standard version. Low-motor children receive a significantly higher score on the non-verbal items on the low motor version than on the standard version, probably due to the high number of items where a child needs to manipulate material. As expected, children profit significantly from the material adaptation alone and from the material adaptation in combination with the instruction adaptation. For the items with just instruction adaptation no significant difference was found.
No significant differences were found between results on the low motor version of the BSID-II-NL and the standard version for the group of normal children, when taking into account a p value < .05.

(3) For the low-motor group we examined correlation between test results (low motor and standard version) and estimated developmental level by the expert before administering the test. Based on file information and clinical view, experts estimated the cognitive development on a scale of 1 to 5 (based on developmental indices): 1: very seriously delayed (<55); 2: seriously delayed (55-70); 3: delayed (70-85); 4: average (85-114); 5: above average (>114). Correlations between expert judgement and both versions were significantly high, but not significantly different: .58 for the standard version and .62 for the low motor version. Experts do not predict test results on the low motor better than results on the standard version.

(4) In general, experts mention the favorable effect of the possibility to decide on item level whether the item is administered according to the standard or low motor version of the BSID-II-NL. Due to the interchangeability of the item (instruction and material) an optimal test situation can be created for a child. Another remark mentioned often was that the adapted test materials increase the response possibilities of a child and therefore increase the child’s motivation to show its abilities. The adaptations proposed for the test procedures are found to be both adequate and appropriate. Particularly the optional use of personal aids and gestures are considered helpful. Several experts underline that a publication of the BSID-II-NL Low Motor will finally enable them to administer the test in a standardized way. Most experts present the test using their ‘home-made’ adaptations. The main part of the adapted item instructions met all criteria. Especially the elimination of time limits was positively evaluated. For some remaining items, experts mentioned that instructions became more complicated in the low motor version. The question referring to the appropriateness of the adapted play material was given the most attention by the experts. Probably as a result of the heterogeneity of the group children with motor-impairment, divergent comments were given concerning the same material. With regard to older children, the play material was found to be adequate and appropriate, but for the younger and smaller children (especially the ex-premature infants) most of the enlarged material was difficult to handle. Experts that were
experienced in using the see-through-frame commented positively on the use of this adaptation. Experts that were unfamiliar with this aid mentioned the advantage of enabling a child to use eye-pointing, but as a disadvantage they mentioned the time-consuming application of it.

6.4 Discussion

By developing the BSID-II-NL Low Motor, a diagnostic tool becomes available that takes into account a child’s motor impairment. In this chapter we described the BSID-II-NL Low Motor and preliminary results of validity studies to answer the research question whether the BSID-II-NL Low Motor is a valid adaptation for use with children with motor impairments that allows the use of the original norm tables.

The results as described in this chapter support our assumption that item content and difficulty remain essentially unchanged by the low-motor adaptations and that the norm tables can still be used when (parts of) the low motor version is (are) applied. The scores on the low-motor version for the group of children experiencing normal development do not differ significantly from the scores they achieve on the standard version. When items are grouped together based on the way in which they were adapted, their results on both versions still don’t differ significantly. For low-motor children, results indicate that they profit from the use of adapted test materials. Their scores on the low-motor version of the test are significantly higher. Item-level analysis shows that low-motor children score significantly higher on items that use adapted materials and items that use adapted materials in combination with adapted instructions. The significantly higher scores of low-motor children on the non-verbal items of the low motor version, show that low-motor children profit from the expanded response possibilities for those items that generally appeal to the fine motor skills of a child. For normal children, no such difference was found. These results support the validity of the BSID-II-NL Low Motor; item difficulty and content remain sufficiently unchanged to be able to apply the standard norm tables and the adapted response possibilities of the low motor version are a significant improvement for use with low-motor children.
The correlation between expert judgment and test results for the low motor version is not significantly higher than the correlation for the standard version. A possible explanation is that the experts had little experience using adapted test instruments and based their judgments on earlier experience with administering standard test material. A more statistical explanation could be that expert judgments were scored on more course scale than the test itself, preventing subtle differences from being statistically significant.

The findings in this chapter are subject to at least two important limitations. First, the tests were administered by many different administrators, who had not received additional guidance after the one-time training. This will lower the reliability of the test results. Second, results are preliminary and based on a small sample, a more extended sample is necessary to be able to answer the research question conclusively. Although all test administrators involved in this study were experienced users of the BSID-II-NL and were trained in administering the Low Motor version, it became apparent while processing data provided by clinicians, that regularly, standard procedures and item instructions were not applied correctly. From video tapings and notes and remarks on the score forms we deduced that mostly, mistakes were made in the way items were presented and the amount of assistance that was given. In some cases, the test administrator was forced to deviate from the instructions because of the specific limitations of the child being tested, but in other cases, the administrator clearly hadn’t sufficiently understood the instruction as described in the manual. In processing the results, it was possible to correct for the mistakes most of the times. In the cases where this wasn’t possible, the test was excluded from the analysis.

The results from this validity study support the assumption that the low motor version is a more valid instrument to examine cognitive abilities in young low motor children than the standard version of the BSID-II-NL, but to answer the research question conclusively, an extended study to supplement the sample, is recommended. An extended sample will also enable us to classify the heterogeneous sample of low motor children into more homogeneous groups, for example, according to Mittler (1974), into two main groups: children with motor impairment associated with or as a result of cerebral palsy and children with motor impairment that is not related to cerebral palsy. Children with cerebral palsy often experience more complex disabilities that affect their
intellectual, perceptual, sensory and/or speech and language abilities. Besides this, future research should provide insight in the question whether the test is sufficiently reliable to permit stable estimates of the cognitive development of young low-motor children.
References


*Mental Retardation and Developmental Disabilities Research Reviews, 8,* 234-240.


