

Acceptability and Reliability of the Bayley Scales of Infant Development III Cognitive and Motor Scales among Children in Makwanpur

Manandhar SR,¹ Dulal S,² Manandhar DS,² Saville N,³ Prost A³

¹Department of Pediatrics, Kathmandu Medical College Teaching Hospital, ²Mother and Infant Research Activities (MIRA), Kathmandu, Nepal, ³Centre for International Health and Development, UCL, London, UK.

ABSTRACT

Background: The Bayley Scales of Infant Development III (BSID III) is an instrument to measure the development of children aged 1-42 months. Our study sought to assess the feasibility and reliability of the BSID III's cognitive and motor sub-scales among children in rural Nepal.

Methods: For this study, translation and back translation in Nepali and English for cognitive and motor sub-scale of BSID III were done. Two testers assessed a total of 102 children aged 1-42 months and were video-recorded and rescored by the third tester. Raw scores were calculated for each assessment. Inter and intra-observer reliability of scores across the three testers was examined. Raw score was converted into scaled score to examine the mean score. The study received ethical clearance from NHRC.

Results: A total of 102 children were assessed. The inter-rater reliability of the BSID III among three testers using the Intraclass Correlation Coefficient by age group was 0.997 (95% CI: 0.996-0.998) for the cognitive scale, 0.997 (95% CI: 0.996- 0.998) for the gross motor scale, and 0.998 (95% CI: 0.997- 0.999) for the fine motor scale. All were statistically significant ($p < 0.0001$). The mean scaled cognitive, fine motor and gross motor development scores in this group of children were 8.3 (SD: 2.5), 8.5 (SD: 2.6) and 9.5 (3.2), respectively.

Conclusions: Assessing the cognitive and motor development of children under five using the BSID III was feasible in Makwanpur district, Nepal. The inter-rater reliability was highly comparable among the three testers.

Keywords: Bayley scale of infant development (BSID); inter and intra -observer reliability.

INTRODUCTION

An estimated 200 million children under the age of five years in low and middle-income countries are not reaching their full developmental potential because of poverty, nutritional deficiencies and a lack of learning opportunities.^{1,2} Many of these children live in South Asia, but there is relatively little research on interventions to improve early childhood development in this setting. One reason for this is that the tool to measure children's development are often developed in high-income settings but not validated cross-culturally, which makes it difficult to evaluate the impact of interventions in

middle and low-income settings. In Nepal, interventions to improve early childhood development are promoted by government and civil society organizations, but there is a lack of reliable, validated tools to measure it.

The Bayley Scales of Infant Development, version III (BSID III) is a commonly used psychometric tool for assessing the development of children aged between 1 and 42 months.³ It has been shown to be a valid diagnostic tool for identifying children with developmental delays at an early age, and is widely used in clinical settings.^{4,5,6} The BSID III scales can be used to assess infant and toddler development across five developmental domains

Correspondence: Dr Sunil Raja Manandhar, Department of Pediatrics, Kathmandu Medical College Teaching Hospital, Kathmandu, Nepal. Email: drsunilraja@gmail.com.

- cognition, language, social emotional, motor and adaptive behavior.⁷ The cognitive scale, language scale and motor scales are administered to individual children in the presence of a caregiver. Nepal is linguistically diverse, and adapting the language scale was deemed unfeasible in this study. We therefore chose to focus on the cognitive and motor scales, as these are likely to be used for clinical and research purposes. The length of the assessment varies between 20 minutes and two and half hours, depending on the child's age and ability. It involves manipulating objects, responding to verbal requests and simple physical activities. We sought to assess the acceptability of the BSID III, with caregivers and children in rural Nepal, and to measure its reliability to assess children's development.

METHODS

The study took place in Makwanpur, which is located in the mid-hill region of Nepal, and covers an area of 2426 km² with a population of around 458,000.⁸ Around half of the population is Tamang-speaking (45.2%), with a near-even split between Hindu (49.3%) and Buddhist (47.6%) religions.⁸ The overall literacy rate is 63.1%, and female literacy 53.7%.^{8,9} 82.5% of people in Makwanpur live in rural areas.⁹ Among 44 VDCs, each VDC has an average population of around 8900.¹⁰

Three testers (SRM, SD, and AP) underwent three weeks of training on the administration of the BSID III at the International Centre for Diarrheal Diseases Research, Bangladesh (ICDDR,B), Dhaka, from 3rd- 20th September 2012. Following this, SRM and SD translated the spoken requests in the cognitive and motor sub-scales from English to Nepali. These were then back-translated into English by a bilingual researcher (NS) to check the translation's accuracy. SRM and SD then conducted pilot assessments on 34 children (17 children each) aged between 1-42 months in 17 age groups (two children from each age group) in a field office in Makwanpur District. Each caregiver and child pair was invited to the MIRA field office for the assessment. SRM and SD sought written consent from the parents or guardians of each child. They then assessed the cognitive and motor skills of each child using the BSID III scales, gave them a raw score for each scale, and finally a full BSID score. Raw scores were converted into scaled scores to examine the mean scores in this population. The study received ethical clearance from the Nepal Health Research Council (NHRC). In the pilot with 34 children, SRM and SD counted how many children passed each item, how many correctly identified pictures, and identified the questions that more than half of children had problems understanding in order to help us identify items that

needed adaptation. In the final 68 assessments, SRM and SD examined 34 children each, all aged between 1-42 months (34 boys and 34 girls). The assessments were carried out using the same procedure as in the pilot, with only minor modifications to the images in the picture. Three children were invited for re-assessment within three days to rectify implausible anthropometric measurements.

All assessments were filmed and scored by one of the two testers not performing the assessment (either SRM, or SD or AP). We compared the raw scores given to each child between two testers to establish inter-observer reliability. To establish intra-observer reliability, each child was observed and given a score by the second tester, while being tested by the first tester. We analyzed scores on 102 children from all three testers. We calculated the Inter-rater reliability of BSID III for the cognitive and motor scales using the Intraclass Correlation Coefficient (ICC) by age group. We then converted mean raw scores into scaled scores using the BSID III manual.⁷ This is an observational experimental study. Study period was conducted from 2nd January 2014 - 3rd January 2015 with a period of one year. Analyses were conducted in SPSS version 18 and Stata 13. The study received ethical clearance from the Nepal Health Research Council (NHRC) in July 2013.

RESULTS

Table 1 describes the characteristics of mothers and their children. Over half of mothers (n=61) belonged to the Tamang caste. 75% (n=76) of them could read easily, 12% (n=13) with difficulty and 12% (n=13) of could not read. 47% (n=47) mothers had received secondary class education. By design, half of the children (n=51) were boys and the other half girls. 47% (n=48) of the children were under 12 months of age. All 102 mothers reported that their children were born at term, but we had no clinical records to verify this. 18% of children were stunted (HAZ <-2SD of 2006 WHO growth standards), 4% were wasted (WHZ<-2 SD), and 2% were underweight (WAZ <-2SD).

Table 1. Characteristics of mothers and their children

	(N)102	Percent
All mothers and children		
Maternal ethnic/caste group		
Tamang	61	60

Newar	5	5
Brahmin	23	22
Chettri	8	8
Pariyar	3	3
Other	2	2
Maternal education		
No education	10	10
Primary education	20	20
Secondary education	46	46
Intermediate	17	17
Bachelor's degree or above	8	8
Maternal literacy		
Can read easily	76	74
Can read with difficulty	13	13
Cannot read	13	13
Children's age groups (months)		
0 - 11.9	48	47
12 - 23.9	27	27
24 - 35.9	18	18
36 - 42	9	8
Born too early*		
Yes	3	
No	93	
Missing	6	

* Caregivers were asked whether they were ever told that their child had been born too early

Table 2. Intra and inter-rater reliability coefficient of the BSID III cognitive and motor scales scores (three testers)

	Coefficient	95% CI
Cognitive scale	0.997	(0.996 - 0.998)*
Fine motor sub-scale	0.998	(0.997 - 0.999)*
Gross motor sub-scale	0.997	(0.996 - 0.998)*

* p < 0.001

Table 2 shows the ICC for inter-rater reliability among the three testers for the cognitive and motor BSID III scales.

The ICC by age group was 0.997 (95% CI = 0.996-0.998) for the cognitive scale, 0.997 (95% CI = 0.996-0.998) for the gross motor scale, and 0.998 (95%CI = 0.997-0.999) for the fine motor scale. Results for all three scales were statistically significant (p <0.0001).

Table 3. Mean scaled cognitive, fine motor and gross motor scores (n=102)

	Mean	Standard deviation
Scaled cognitive scores	8.3	2.5
Scaled fine motor scores	8.5	2.6
Scaled gross motor scores	9.5	3.2

Table 3 describes the mean scaled scores for the cognitive, fine motor and gross motor scales. All were below 10, the mean for the standard BSID III scores. We did not perform further regressions to explore the association between stunting and BSID III scores due to the low prevalence of stunting in this population.

DISCUSSION

A study done by Suzann et al conducted a study with 145 infants at 6 weeks of age and investigated the Inter-rater reliability between the Test of Infant Motor Performance (TIMP) and the Bayley Infant Scale III using an ICC (intra class correlation coefficient)¹¹. The ICC for TIMP was 0.79 (95% CI = 0.60-0.90), while for the Bayley Cognitive Scale it was 0.73 (95% CI=0.46-0.86), and for the Bayley Motor Scale it was 0.75 (95% CI = 0.46-0.88). Similarly, in this study we also found very high inter-rater reliability score by using ICC for BSID III among three testers.

In Nepal, there has been a strong push to strengthen Early Childhood Development (ECD) initiatives (http://www.unicef.org/evaldatabase/index_60110.html). However there have been few attempts to validate robust tools for measuring child development in Nepal and most of these have focused on children older than three years (e.g. Denver test). This study has provided evidence that the BSID III can be used to measure children's cognitive and motor development in Nepal, and can be further investigated for clinical use and the evaluation of ECD interventions.

Assessing the cognitive and motor development of children under-five using the BSID-III was feasible in Makwanpur district, Nepal. Inter-rater reliability was high among the three testers. The mean and standard

deviations for scaled cognitive, fine motor and gross motor scores in this population were all slightly below the BSID III standard means of 10 for each sub-scale. This study has provided the evidence that the BSID III can be used to measure children's cognitive and motor development in Nepal. Future work should focus on adapting the BSID III's language scale for Nepali, as well as testing the reliability of the BSID III's sub-scales against other age-appropriate measures of development.

CONCLUSIONS

Inter-rater reliability of the BSID III was highly comparable among three testers. In post-2015, there is a strong focus on going beyond child survival and focusing on strategies to enhance early child development, a significant determinant of educational attainment as well as subsequent health and productivity in adulthood. This study has provided the evidence that the BSID III test can be used to assess the cognitive and motor development of children aged 1-42 months in Nepal. This study also suggested that the BSID III cognitive and motor sub-scales are feasible and appropriate to use with rural Nepalese children and may be used in intervention studies to assess child development

ACKNOWLEDGMENT

We would like to acknowledge Prof Dr Anthony Costello, University College of London (UCL) for his continuous support in this study and the mothers of the children who have given written consent for the assessment of their children in this study.

REFERENCES

1. Grantham-McGregor S, Cheung YB, Cueto S, Glewwe P, Richter L, Strupp B. Developmental potential in the first 5 years for children in developing countries. *Lancet* 2006; 369: 60–70.
2. Walker SP, Wachs TD, Grantham-McGregor S, Black MM, Nelson CA, Huffman SL et al. Inequality in early childhood: risk and protective factors for early child development. *Lancet* 2011; 378: 1325-1338.
3. Bayley N: Bayley Scales of Infant and Toddler Development—Third Edition: Administration Manual. Harcourt Assessment: San Antonio; 2006.
4. Wu Y, Tsou K, Hsu C, Fang L, Yao G, Yao G et al. Taiwanese infants' mental and motor development 6–24 months. *J Pediatr Psychol* 2008, 33(1): 102–108.
5. Santos DCC, Gabbard C, Gonclaves VMG: Motor development during the first year: a comparative study. *J Genet Psychol* 2001, 162(2):143–153.
6. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle FW, The Victorian Infant Collaborative Group: Underestimation of developmental delay by the new Bayley-III scale. *Arch Pediatr Adolesc Med* 2010, 164(4):352–356.
7. Bayley N: Bayley Scales of Infant and Toddler Development—Third Edition: Technical manual. Harcourt Assessment: San Antonio, TX; 2006.
8. Sharma NK. Makwanpur. In: SR Joshi eds. District profile of Nepal 2007/2008, 1st edition, Kathmandu: Intensive Study & Research Centre, 16th September 2007: p 335-345, 346- 358.
9. Population census 2001, National Report. In: Nepal Government, National Planning Commission Secretariat, Central Bureau of Statistics in collaboration with UNFPA Nepal. June 2002.
10. Baseline survey MIRA (Mother and Infant Research Activities) Makwanpur 2005. (Unpublished).
11. Suzann KC, Laura Z, Kristin MR, Joseph CY, Nicole S, Zhuoying L and Rosemary WT. Concurrent Validity of the TIMP and the BSID III Scales at 6 Weeks Corrected Age. *Pediatr Phys Ther* 2013; 25(4): 395–401.