RESEARCH PAPER

Development and Validation of Language Evaluation Scale Trivandrum for Children Aged 0-3 years - LEST (0-3)

MKC NAIR , *GS HARIKUMARAN NAIR, # AO MINI, # S INDULEKHA, #S LETHA, AND †PS RUSSELL

*From the Department of Pediatrics & Clinical Epidemiology and Director, Child Development Centre, Government Medical College campus, Thiruvananthapuram, Kerala; * Clinical epidemiology Resource and Training Centre and Professor of Radiology, Government Medical College, Thiruvananthapuram, Kerala; #Child Development Centre, Government Medical College campus, Thiruvananthapuram, Kerala; and †Child and Adolescent Psychiatry Unit, Christian Medical College, Vellore 632 002, Tamil Nadu; India.*

Correspondence to: Prof MKC Nair, Director, Child Development Centre, Government Medical College Campus, Thiruvananthapuram, Kerala 695 011, India. cdcmkc@gmail.com

Received: February 21, 2012; Initial review: May 03, 2012; Accepted; October 10, 2012.

Objective: To develop and validate a simple screening tool which can be used in the Community to identify delay in language development among children of 0-3 years of age.

Methods: The normal range for the 33-items of "Language Evaluation Scale Trivandrum for 0-3years-LEST(0-3)" were carefully selected from various existing language development charts and scales, by experts keeping in mind the face validity and content validity. The criterion validity was assessed using a community sample of 643 children of 0 to 3 years of age, including 340 (52.9%) boys. LEST (0-3) was validated against Receptive Expressive Energent Language Scale, for screening delay in language development among children of 0-3 years.

Results: When one item delay was taken as 'LEST delay' (test positive), the sensitivity and specificity of LEST(0-3), was found to be 95.85% and 77.5%, respectively with a negative predictive value of 99.8% and LR (negative) of 0.05.When two item delay was taken as 'LEST delay' (test positive), the sensitivity and specificity of LEST(0-3), was found to be 66.7% and 94.8% respectively with a negative predictive value of 98.7% and LR (negative) of 0.35. The test-retest and inter-rater reliability were good and acceptable (Inter-class correlation of 0.69 for test-retest and 0.94 for inter-rater).

Conclusion: LEST (0-3) is a simple, reliable and valid screening tool for use in the community to identify children between 0-3 years with delay in language development, enabling early intervention practices.

Key words: *Development, India, Language, LEST, Screening tool, Validation.*

PII: S097475591200164

Language encompasses every means of communication in which thoughts and feelings are symbolised, so as to convey meaning to others. It includes such widely differing forms of communication as writing, speaking, sign language, facial expression, gesture and art [1,2]. Language development occurs in a sequential fashion and as age advances the child has more and more to communicate, first learning to listen and understand language before they learn to talk.

Language can be divided into two major components. Firstly, the receptive language where the child understands from verbal and non- verbal communication, and secondly, the expressive language where the child says or does convey, what he/she wants to communicate. Thus in short, receptive language (understanding) is the ability to take in information presented through speech and actions of others. Expressive language (talking) describes children's ability to tell their needs, thoughts, ideas and feelings through their own speech and actions.

Delay in acquiring language development is often an early and most sensitive indicator of intellectual disability, pervasive developmental disorder and specific learning disorders. Language delay or abnormalities in speech and language should be detected during the early stages of life itself, so that early intervention could be instituted. Most of the children with language delays and disorders need systematic assessment and training. In the west, many tools are available for the purpose; for example Early Language Milestone Scale for 0-3 year (by James Coplan), 3 DLAT (Three Dimensional Language Assessment Tool) and REELS (original REELS and Newer versions of REELS) etc. But these tools are generally not accepted for a community setting.

The observed 3.8% [3] prevalence of speech and language delay in western literature and 4.5% [4] in Indian

literature indicates the need for screening of all infants, toddlers and young children. But at the community level these children are usually not identified due to the lack of user-friendly, brief tools that can be used by community health workers. The present study describes the development of the screening tool 'Language Evaluation Scale Trivandrum for 0-3 years-LEST (0-3), and its validation against the original 1971 version of Receptive-Expressive Emergent Language Scale (REELS).

METHODS

Design of LEST (0-3): The Language Evaluation Scale Trivandrum for 0-3 years-LEST (0-3) was designed and developed at the Child Development Centre, Government Medical College Campus, Trivandrum. Thirty-three test items were care-fully chosen, from the item pool developed from pilot studies done earlier among children with 0-1 year, 1-2 year and 2-3 years of age and compared separately for each year against original REELS. These items were chosen from the item pool developed by a team of experts (Paediatric Neurologists, Developmental Paediatricians, Developmental Therapists, Speech therapists, Child Psychiatrists, Clinical Psychologists and Epidemiologists) to include items for language development milestones, adequately spread over the first 3 years of age. The items and the range for each test item (represented by horizontal dark line) were selected from different existing developmental/speech and language assessment scales, tools and guidelines like Denver Developmental Screening Test [5], Receptive-Expressive Emergent Language Scale (REELS) [6], Early Language Milestone Scale (ELM scale-2) [7], Hearing check list: New Zealand Government, The Rossetti Infant Toddler Language Scale [8], Hearing check list: Toronto Preschool Speech and Language Services etc. Item reduction was not deemed required as the items were finalised from the item pool by the experts using content retention approach. The aim of developing this measure to develop a culturally appropriate screening tool which is simple to understand as well as easy to be used by a health worker in the community to identify probable speech and language delay among 0-3 year old children, so as to enable the mother to initiate speech and language stimulation at home itself. Translation from English to local language (Malayalam) and back translation of the tool was done and found acceptable to the experts. The questions were asked in local language (being asked by operators with Malayalam translations with them).

Validation of LEST (0-3): The total sample for validation was calculated as follows. N (the number of positives for language delay) required was $(1.96)^2PQ/d^2$; where P is the sensitivity expected Q is (1-P) and d is the precision desired. Taking the expected sensitivity of the new tool as 95% and precision as 10 and with 95% confidence interval, the sample size calculated was 18 positive cases (of language delay). Assuming 3% prevalence for delay in language development in 0-3years of age, the total sample to be studied was calculated as 600. Considering possibility of non-co-operation or incomplete tests (either the new tool or the reference standard) as 5%, the final sample was calculated as 632 (rounded to 650). Thus, a total of 653 children between 0-3 years belonging to 11 Anganwadi areas from an urban ward, 20 from a rural Panchayat and 3 from tribal area, participated in the study. They were accompanied by a parent/primary care giver. Final valid sample (where the new tool and the reference standard were available) obtained for analysis was 643. Anganwadi worker recruited the mothers and children between 0-3 years of age in their area to the respective Anganwadi, on a specified date and time. The screening test and the 'Reference Standard' were administered by independent observers blinded to the results of the other tests. Administration procedures for the tests were standardised. Same raters assessed children in all the Anganwadis to maintain the uniformity in the assessment process after standardisation. Two raters administered LEST and another two assessors evaluated the children using REELS. They were trained in Child Development Centre (CDC) Thiruvananthapuram for applying the tools. LEST (0-3), was applied by trained persons having similar educational qualification as that of an ICDS Supervisor and REELS by Speech and language Therapists, after acquiring written consent from the parent and verbal assent (here only means 'seeking co-operation') from the child. All children who satisfied the selection criteria (with proper consent from parents or primary care givers) in the respective Anganwadi areas were included in the study. Exclusion criteria include those who were ill and uncooperative for testing.

For the administration of LEST, children need not go through the all 33 items of the measure. To rate LEST, the chronological age of the child was assessed first. A vertical line was drawn by keeping a scale (or just kept the scale vertically) at the point corresponding chronological age in months given horizontally in the X axis. All items (which are shown in blocks) completed fully to the left side of the scale were expected to be done by the child. If not attained by the child for that age, that item delay is assumed for the child. Thus the tool is designed to be simple and no expertise is required, when compared with REELS. The prematurity corrections were not done here for calculating the chronological age of the child, as this would make the new tool complicated and we feared that such corrections would make it unsuitable for a health

worker to use it in the community. Though such corrections (for preterm babies) would have improved our positive predictive value and specificity (without affecting sensitivity) we did not incorporate that in our tool because of our inclination was for a simpler tool (compromising on positive predictive value and specificity for a simple tool). Hearing tests were not done as the study was done in a community setting and we felt that, excluding those with hearing problems (some of them having language delay due to hearing problem) would affect our generalizability. More over the tool (LEST) was intended for the health workers to perform the screening in the community and in reality, there will not be an exclusion of those with hearing problems during actual screening.

The rating bias was minimised by independent rating and standardisation of administration. The tool was applied by the operators, who have got training in performing the test. First preference was given for observation of the child and testing of the items and if not possible then for parental reporting was considered as valid for some of the items. For inter-rater reliability assessment the LEST was administered by two raters independently in a sample of 50. For test-retest reliability assessment the LEST was administered twice by the same rater in a sample of 50, with a gap of two weeks between the tests. The study was conducted after getting approval from the Human Ethical Committee of Child Development Centre, Medical College Campus, Thiruvananthapuram.

Data were analysed using statistical functions available in Microsoft excel (and using DAG_stat a Microsoft excel based statistical software for diagnostic test evaluation) and SPSS (version 17) statistical software. Sensitivity, specificity, Positive Predictive Values, Negative Predictive Values, Accuracy and Likelihood Ratios for LEST (0-3) against REELS taken as "reference standard" were calculated. Intra class correlations were also calculated to assess test-retest and inter-rater reliability.

RESULTS

A total of 643 children were included for the validation, where both results (LEST as the tool and REELS as the reference standard) were available, 194 children (45.4% boys) were below 12 months, 197 were between 13-24 month old (54.8% boys), and 152 were between 25-36 month old (57.1% boys). ***Web Table* I** shows the 33 items with the age range at which the items are to be achieved normally.

The test re-test reliability of the tool was done in a valid sample of 50 and it was found to be acceptable (intra-class correlation was 0.69, 95% CI: 0.46-0.83). The inter-rater reliability for the study was done in a valid sample of 50 and it was also found to be acceptable (intra-class correlation was 0.94 having a 95% CI: 0.9 – 0.97)

***Table* I** shows the cross tabulation of LEST with one item delay taken as test positive (LEST delay) against REELS.

***Table* II** shows the cross tabulation of LEST with two items delay taken as test positive (LEST delay) against REELS.

***Table* III** shows the comparison of tool characteristics in 2 situations of new test (tool) criteria with one item delay as 'LEST delay' and two items delay considered as 'LEST delay'.

***Web Table* II** shows the tool characteristics at different age ranges.

DISCUSSION

An Anganwadi based survey of developmental delay/ disability in one ICDS block had observed that the speech and language delay was the commonest among developmental problems [9]. However, language development is not represented adequately in most developmental assessment tools. It was the felt-need to have a tool for assessing language delay, which can be used by health workers. A screening tool for language delay should be simple, less time consuming and easily understood by the

TABLE I LEST (0-3 YEARS) AGAINST REELS (ONE ITEM DELAY AS 'LEST POSITIVE')

LEST	*REELS*		
	Abnormal	*Normal*	*Total*
Test positive (1 item delay)	23	139	162
Normal	1	480	481
Total	24	619	643

REELS- Receptive-Expressive Emergent language Scale.

TABLE II LEST (0-3 YEARS) AGAINST REELS (TWO ITEM DELAY AS 'LEST POSITIVE')

LEST	*REELS*		
	Abnormal	*Normal*	*Total*
Test positive (2 item delay)	16	32	48
Normal	8	587	595
Total	24	619	643

REELS- Receptive-Expressive Emergent language Scale.

TABLE III TEST CHARACTERISTICS WITH TWO DIFFERENT CRITERIA

Criterion for Test positivity	*One item delay in LEST as (tool positive)*	*Two items delay in LEST (tool positive)*
Sensitivity (%)	95.8 (95% CI: 78.9 – 99.9)	66.7 (95% CI: 44.7 – 84.4)
Specificity (%)	77.5 (95% CI: 74 – 80.8)	94.8 (95% CI: 92.8 – 96.4)
Positive Predictive Value (%)	14.2 (95% CI: 9.2 – 20.5)	33.3 (95% CI: 20.4 – 48.4)
Negative Predictive Value (%)	99.8 (95% CI: 98.9 - 100)	98.7 (95% CI: 97.4 – 99.4)
LR+(Likelihood Ratio positive)	4.3 (95% CI: 3.6 - 5.1)	12.9 (95% CI: 8.3 - 20.01)
LR-(Likelihood Ratio negative)	0.05 (95% CI: 0.008 - 0.37)	0.35 (95% CI: 0.2 - 0.6)
Accuracy (%)	78.2 (95% CI: 74.8 – 81.4)	93.8 (95% CI: 91.6 – 95.5)
Prevalence and Bias Adjusted Kappa (PABK)	0.57	0.88
Prevalence by the Gold Standard test	3.73(95% CI: 2.4 – 5.5)	

community health worker and the parents. LEST (0-3) was designed to meet these requirements. It was important to validate LEST against the 'best available' assessment tool. Here we used the original version of REELS. A conscious decision was taken by us to include only the REELS criteria "delay" to be taken as the reference standard positive to have a better specificity of the reference standard test. Problems with REELS in the community setting are that it is a time consuming test and difficult to administer in a community setting; and it can be administered only by speech and language pathologists. LEST, on the other hand can be administered by any person with minimal training. LEST is easy to administer, items are simple to perform and is in Chart form, which is easier than the former one. REELS is in compound and complex language which is difficult to understand and apply in the community, but LEST is in a simple language.

Every possible effort was taken to avoid bias in this study. The Anganwadi workers recruited all children in the Anganwadi area and 'all consented' were participants to the study. Thus 'selection bias' was minimised. Tests were done by equally trained persons (with educational qualifications similar to ICDS supervisor) not familiar with the children in the area (where as Anganwadi workers and ICDS workers were familiar with the children). All were given clear instructions to administer the tests and record the result systematically. Uniform hands-on training were given to all investigators. The observer who administered and interpreted the REELS did not know the screening test results and vice versa. Thus measurement bias was minimised.

Shifting the test positivity (tool positivity) from one item delay to two items delay (LEST positive), resulted in a drop of the sensitivity from 95.8% to 66.7% though test specificity showed an increase (from 77.5% to 94.8%). For using LEST as a screening tool for delay in Language, we selected two item delay as test positive accepting a lesser (out of the two options) sensitivity of 66.7% because of the relatively higher Positive Predictive Value (PPV of 14.2% and 33.3%, respectively for one item delay and two item delay as the tool criteria positive) and lesser false positives in the screened sample. This option gives an excellent Negative Predictive Value (NPV) of 98.7% also, which is desirable for a screening tool. The choice between one-item or two-item delay in LEST depends on the need in the community, whether to have a highly sensitive test with a very low PPV or to have a lesser sensitivity with a relatively higher and acceptable PPV (but still low due to the low prevalence of the condition to be screened).Unlike sensitivity and specificity, the PPV and NPV are more influenced by the prevalence of the disease. Among these measures, both the Likelihood ratios are least dependent on prevalence. Here we are getting a LR positive of 12.9 (95% CI: 8.3 - 20.01) with two- item delay as tool positive criteria.

The limitation of our study is the use of an imperfect reference standard, which we were forced to accept because of the non-availability of a Gold standard for language delay for the age group to be screened. In the case of language delay, it is difficult to get a gold standard and the original REELS was used for validation. In epidemiological studies, in the absence of a gold standard the researcher may be forced to take an 'imperfect' gold standard. In such situations researchers may utilise proximate measures of the 'gold standard' as the criterion to assess validity. In these situations the kappa statistics is commonly used to assess agreement for estimating "validity". Here we calculated the Prevalence and Bias Adjusted Kappa (PABK) of LEST with REELS and was acceptable (0.88) with two-item delay as tool positive criteria.

WHAT IS ALREADY KNOWN?

- Speech and language is one of the commonest development problem.

WHAT THIS STUDY ADDS?

- LEST (0-3) is a reliable scale to identify delay in Language development in children of 0-3 years in the community.

The observation that the prevalence of Language delay (as per the screening tool) coming down from 25.2% to 7.5% (due to drop in false positives) when shifting from one item delay to two items delay taken as the criteria for 'LEST delay', has implications in planning large scale community level language development assessments and for interventional programs.

Likelihood ratio positive and likelihood ratio negative, which are relatively independent of the prevalence of the condition, are also found to be good for LEST, and hence considered acceptable. The test-retest reliability here was a little low, but acceptable. This is explained by a recall effect (parents reporting the items better in the second time or due to the child performing the items better in the second time when examined by the same observer who have already established a rapport during the first examination, even though there was a gap of two weeks between measurements). The inter-rater ICC was 0.94, which was fairly high.

LEST (0-3 years) is a valid Indian tool for identifying children of 0-3 years with language delay in the community with an acceptable Sensitivity, Specificity, Positive Predictive Value and Likelihood Ratios. This is a simple tool also, which can be finished in 10 minutes, by a health worker, and requires only a pen/pencil/scale along with the tool and minimal training to apply the tool. Based on the test result, language stimulation by the mother can be easily done at home and also help in referring needy ones to an appropriate referral centre for intervention.

Acknowledgments: Dr Babu George, the Medical Superintendent Child Development Centre (CDC), Thiruvananthapuram,for his support during the study period. We thank Asokan N and Dr Leena Sumaraj, CDC, for their support. We also thank Dr. Abhiram, Dr. Princly Pappachan, Vinod S S and Bincymol K for their help as team members of the study project.

*Funding***:** National Rural Health Mission (NRHM) and Child Development Centre, Medical College, Thiruvananthapuram.

REFERENCES

1. Asher SR, SL Oden. Childrens' failure to communicate. An assessment of comparison and egocentrism explanations. Developmental Psychol. 1976;12:132-9.
2. Jakobson R. Verbal communication. Scientific American. 1972; 227:73-80.
3. Shriberg LD, Tomblin JB, McSweeny JL. Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. J Speech Hearing Res. 1999;42: 1461-81.
4. Nair MKC, George B, Philip E. Trivandrum Developmental Screening Chart. Indian Pediatr. 1991;28:869-72.
5. Frankenburg WK, Dobbs JB. The Denver Developmental Screening Test. J Pediat. 1967;71:181–91.
6. Bzoch KR, League R. Receptive-expressive Emergent Language Scale. 1st edition. Gainseville: The Tree of Life Press; 1971.
7. Coplan J. Early Language Milestone Scale. Austin, TX: Pro-Ed Inc; 1993.
8. Rossetti L. The Rossetti Infant-Toddler Language Scale: A Measure of Communication and Interaction. East Moline, IL: LinguiSystems;1990.
9. Nair MKC, George B, Padmamohan J, Sunitha RM, Resmi VR, Prasanna GL, *et al*. Developmental delay and disability among under-5 children in a rural ICDS block. Indian Pediatr. 2009; 46: S75-S78.