THE CRIB-O-GRAM AS AN AID IN CLINICAL DIAGNOSIS

George T. Mencher, Ph.D.
Director
Nova Scotia Hearing and Speech Clinic

ABSTRACT

The use of the crib-o-gram as a hearing screening device for newborns is examined. Fourteen infants with known auditory thresholds were tested using this procedure. Emphasis was placed upon both the ability of the crib-o-gram to identify a hearing loss in the severely impaired child as well as inter-scorer reliability for the test results. The results indicate that this procedure is a valid means of differentiating infants with a severe impairment from those with normal hearing. Scorer reliability was found to be good with no inter-judge disagreement great enough to have altered the test results for any subject.

The United States Joint Committee on Infant Hearing Screening has issued a number of well publicized statements re: Screening Methods for Identifying Hearing Loss in the Newborn (Northern and Downs, 1974). In 1975, that Committee adopted the recommendations of the Nova Scotia Conference on Early Identification of Hearing Loss in which it was stated that a behavioral screening protocol could be used as a supplement for the newborn high risk register (Mencher, 1976). The behavioral protocol outlined may be either a manual method in which a trained observer scores a response when there is a behavioral change associated with auditory stimulation, or it may be an automatic method in which a machine records some physical or behavioral change associated with a response to sound.

The history of automatic devices for screening newborns for hearing loss is quite limited. Based on concepts originally presented in the early 1930's and 1940's by Aldridge, only two experimental automated systems have been developed, and those only recently.

The Accelerometer Recording System (ARS) of Altmann, Shenhav and Schaudinischky (1976) consists of a sound source, a cradle, a vibration analyzer, and an ink recorder. The system is quite primitive, having reached that stage in development where refinement of instrumentation and massive field testing is required before the unit could even begin to be considered more than a prototype. Nevertheless, Altmann, Shenhav and Schaudinischky reported a study in which observers' visual notation of infant response to sound was compared to results recorded on the ARS. A distinct ARS recording was obtained from 393 newborns out of a sample of 400, while observational response was noted in only 367 of the cases. Unfortunately, there was no comparative follow-up reported re: the presence of actual hearing loss in any of the children. Furthermore, there was no information provided re: any particular child's response to either of the test procedures. No other studies have been reported using the ARS. In summary, the system is defined but primitive, and it lacks field testing and comparative data.

The "crib-o-gram" work of Simmons and Russ (1974) has been highly publicized and is well known in North America. The test unit consists of a motion sensitive transducer and associated amplifier, an automated timing device, a strip chart recorder, and a loudspeaker. At intervals preprogrammed by the examiner, the machine automatically turns itself on, records a minimum of 10 seconds of baseline crib activity, presents a 2

second 92 dB SPL-NBN stimulus (between 2000 and 4000 Hz, peaking at 3000 Hz) and, of course, continuously records crib activity through and after the stimulus presentation. At approximately 5 seconds post stimulus off-set the machine automatically shuts itself off. Test presentations can be as close together as 2 to 3 minutes or as far apart as 24 hours, depending on the test protocol desired. Jones and Simmons (1977) recommend 20 trials in 24 hours.

Scoring is the only non-mechanical link in the entire procedure. The strip chart generated during the testing period is removed from the machine, divided into segments representing each of the test cycles, and scored according to a system devised at Stanford University which accounts for rate and degree of change from the baseline behavior within a 2 second time period following the stimulus on-set. Here, the objectivity of the system is lost, as a human scorer must read the chart, make a judgment on what is seen, and subjectively interpret that judgment as a response (or lack of a response). Granted, there are established guidelines for making such a judgment. However, suffice it to say, some cases fall into a marginal category and in those cases, judgment tends to be purely subjective.

Stanford University now has a grant from the United States Department of Health, Education and Welfare to field test the efficiency of the crib-o-gram, establish a rigid scoring procedure (preferably by computer) and to assess the number of false positives and false negatives likely to result from a screening program based on the use of that instrument.

The assumption has been that the crib-o-gram is useful as a screening device for newborns. Because of a recent rubella epidemic in the Maritimes, staff at the Nova Scotia Hearing and Speech Clinic have been concerned, not only with newborns, but with very young children (under 6 months) who have failed the high risk register and/or behavioral hearing screening, and from whom we have not obtained a conclusive follow-up audiometric result. Lack of visible response by a three month old infant in a sound room does not mean deafness. The purpose of this study was to use the crib-o-gram to evaluate a group of children in hospital or at home, to assess response patterns, and to determine the efficiency of using the crib-o-gram as a supplementary clinical tool in investigating for deafness in very young infants.

In order to appropriately evaluate the instrument it was necessary to consider scorer reliability and the validity of those judgments. Reliability was determined by comparing the independent scoring of crib-o-gram print-outs by 5 judges. The validity was assessed by analysis of the pass/fail scores determined by the judges and a comparison of those scores to actual hearing level as reported at least 1 year after the crib-o-gram examination.

Procedures

Subjects

Fourteen infants ranging in age from 2 months 16 days to 8 months 5 days served as the subjects. At the time the 7 boys and 7 girls were tested, 12 were actually suspect for hearing loss and were in the process of audiometric evaluation by serial appointments. The two normals were included in this study as a small control sample. Of the 12 suspect cases, 11 were the products of a rubella pregnancy. The 12 cases had been referred by the Department of Otolaryngology at the Izaak Walton Killam Hospital for Children.

All of the children were tested in their own crib, either at home or in the hospital. The motion transducer associated with the crib-o-gram was placed on the mattress, covered with a plastic diaper and then a hospital sheet. The speaker unit was placed at the head of the bed so that it was approximately 12 inches from the infant's ear and a signal of from 90 to 92 dB would be heard at the ear. The chart speed of the Grass ink recorder was set at 5 mm per second. The sensitivity of the motion transducer was adjusted according to the instructions provided by Stanford University when the crib-o-gram was loaned to us. Thus, the sensitivity was identical to that which would be used in the newborn nursery. It should be noted that the subjects in this study were somewhat heavier than newborn infants would be.

The crib-o-gram switch and timer was placed adjacent to the baby's crib. In all cases, whether it was one of the 10 infants tested at home or one of the 4 tested in a semi-private room at the hospital, the unit was programmed to produce at least 17 trials over the 12 hour period from 9 p.m. until 9 a.m. Trials were always at least 30 minutes apart. Specific attempts were made to avoid trials where the child was likely to be awake for feeding. It was understood that 17 scoreable trials might be unobtainable if the child was either particularly cranky or removed from the crib during one of the test periods. In point of fact, a minimum number of 17 trials were obtained from 13 of the 14 children. There were 16 scoreable trials obtained from the 14th child.

Scorer Reliability

A standard scoring technique was employed for evaluating each of the trials presented to the babies. First, the complete test strip was divided into individual trial segments, each representing one pre-stimulus, stimulus, and post-stimulus cycle. So as to avoid order effects while scoring, the individual test strips were coded numerically in a random order before presentation to the scorers. Finally, 5 individuals, all employees of the Nova Scotia Hearing and Speech Clinic, were asked to evaluate the test packets for each of the 14 babies. The 5 scorers varied in audiological experience. Three were trained audiologists at the M.A. level, one was an audiometrist having been trained on the job but with over 2 years of work experience in audiology at the time of scoring. The fifth individual had no specific audiology experience, **per. se.,** but was a research assistant hired through Stanford University with a total and complete knowledge of the crib-o-gram and its operation. Instructions to the scorers were quite restricted, limited to: "Score a 'yes' if there is a difference between the recorded behavioral patterns up to the point of the stimulus on-set and that recorded thereafter. Score a 'no' if there is no change in the recording which you can associate with the stimulus on-set."

Results and Discussion

As indicated, the individual scorers were asked to make a judgment of "yes" or "no" as to an apparent behavioral change associated with the stimulus on-set on each of the test strips. Each of the scorers' evaluation forms were compared to the other scorers' forms. Because there were only 16 trials obtained from one infant, it was decided that to be consistent, only the first 16 responses obtained from each of the babies should be compared. Comparisons were based on one of three possible combinations:

1) All observers agreed.
2) Four of the five observers agreed.
3) Three of the observers agreed.

Agreement could, of course, be either "Yes, there was a response", or "No, there was no response".

At the outset, the reader should understand that failure on a crib-o-gram evaluation is not based on the number of trials in which the score is marked a failure, but rather, the number of trials on which a pass is scored. In the course of routine newborn evaluations, a pass on 3 trials constitutes a pass of the screening test — **even if the child fails all other trials.** This scoring is predicated on the assumption that the probability of a chance response occurring in direct relationship to the on-set of the stimulus on 3 trials with the same child within the same 24 hours is so slight, as to defy all laws of chance. Statistical analysis by Simmons and Russ (1974) has clearly established the adequacy of this scoring procedure.

Observer agreement is reflected in Table I.

It is significant to note that all 5 observers agreed to "No response" on all 16 trials obtained from 3 children (subjects 1, 3, and 5). Further, subsequent results have confirmed that those 3 children are severely hard of hearing, with no responses to stimuli below 90 dB in the speech range.

Results obtained from subjects 9, 11, and 14 are also most interesting.

Subject 9 would, by any scoring system, be scored a failure. It should be noted that 1 scorer passed him on 1 of the 16 trials. However, the other 4 scorers failed him on that trial, and all scorers failed him on all other trials. In the author's opinion, the positive score is a judgmental error. The trial is presented in Figure I for the reader's own assessment. The child has a hearing loss in the severe category, with no responses below 90 dB.

All scorers agreed subject 11 failed 14 trials. However, all 5 scorers also agreed that he passed on 1 trial. That trial is also presented in Figure I. Whether the child actually responsed to the stimulus or whether, by chance, he moved at the exact moment of the test is not known. Certainly, either choice is a strong possibility. Based on the report of Mencher et al (1977), chance response would be expected in 1 in 98 trials. Given 14 subjects with 16 trials each, or 224 trials, 2 or 3 false positive responses should be expected in the course of this study. The child seen as subject 11 has a hearing loss in excess of 90 dB in the speech range.

Nearly the same problem is evident with subject 14. For 13 of the 16 trials all scorers agreed to "No response". Further, although there is no single trial in which all scorers agreed to a response, there is 1 trial in which 4 of the scorers did feel they could score in the positive column. That is the only trial in which a majority of the scorers agreed to seeing a response. Once again, the question of true response versus chance movement comes to the fore. It should also be noted that 1 scorer did, in fact, mark a pass on 3 trials, or a sufficient number of trials to pass this child on the crib-o-gram test according to the standards previously discussed and established by Stanford University. The child has a 65 dB sensori-neural hearing loss in the speech range, which is less than the degree of hearing loss the crib-o-gram is designed to identify. Thus, the fact that the child failed at all, is a very positive result, and the fact that he passed a number of trials is not surprising.

The results for subjects 2 and 4 are also most interesting. Those children are the two normals included as "controls" in this study. Both were scored as unanimous passes on 13 of the 16 trials, both have failed 1 trial, and both have 2 trials which were scored as a "pass" by a majority, but not unanimously. The number of passes scored for these children is clearly and definitively much greater than that scored for any of the other subjects except number 10. Subject 10 is a baby born of a rubella pregnancy, but without

hearing loss or other anomaly. Audiology tests have repeatedly confirmed normal hearing.

We now come to a review of the scoring for subjects 6, 7, 8, 12 and 13. Using that scoring protocol, all 5 babies passed the test, even though there were as many as 9 unanimous failures by 1 child. In point of fact, subject 8 has normal hearing; subject 13 has a mild ($\leq$35 dB) hearing loss, which may be of middle ear origin; and subjects 6, 7, and 12 have normal hearing, but appear to have definite auditory perceptual disorders and/or to be mentally retarded.

Subject 2 was 6 months 16 days when tested, subject 10 was 6 months 9 days when tested, and subject 3 was 6 months 15 days when tested. Clearly age did not effect the results in these children, amongst the oldest in the group. The same argument holds true when subject 11 (4 months 10 days) and subject 4 (4 months 10 days) are compared.

Five independent scorers had no difficulty in agreeing on the results they scored 84% of the time (187 of 224 trials). That percentage increased to 94% when agreement by 4 of the 5 scorers is considered (209 of 224 trials). Results of this study indicate that in those cases in which there is some disagreement, that disagreement would not affect the results of the test as there are a sufficient number of trials to obtain clear and evident results exclusive of those in which there is some confusion. Further, it would appear that agreement is relatively unanimous in the cases of normal hearing and in those cases with severely hearing impaired children (hearing loss $\geq$ 75 dB). Apparently, scoring is more difficult for children with a mild hearing loss or with other associated problems. Of the 14 children in this study evaluated by the crib-o-gram, the 6 with severe hearing loss clearly failed the test. Those with normal hearing and no associated middle ear or perceptual problems passed easily, and with high scores. Children with middle ear problems, mild sensori-neural hearing loss and/or auditory perceptual problems also passed, but with lower scores and a greater number of trials in which the scorers agreed that "No response" was evident.

The results speak well for the use of the crib-o-gram as a supplement to pediatric clinical audiology when trying to assess a very young child who has been unresponsive in the audiology suites, for hearing loss. The validity of the test result is excellent when compared to hearing levels determined 1 year later. That validity, predicated on utilizing crib-o-gram results based on independent scorer agreement, attests to the general reliability of the procedure as well. That is not to say that the procedure is perfect. However, it appears no more or less subjective and/or reliable and valid than other pediatric behavioral assessment tools. Further, the addition of a computerized scoring procedure (now at the prototype stage) will undoubtedly strengthen the reliability of the scoring. That is not to say that manual scoring is significantly less reliable than computer analysis. However, assuming that the computer is fed the proper scoring protocol, it is faster and obviously less likely to make a subjective judgmental error. We have found clinically, and results of this study reaffirm, that children up to 6 months of age, but relatively immobile in the crib, can be examined. Crib-o-gram evaluation provides at least primary information regarding presence and approximate severity of hearing loss. The fact that children with problems other than hearing loss yielded the most difficult results to score should not pass unnoticed. Further and more detailed assessment of those response patterns in a larger population sample is warranted. It is conceivable, that the greatest information obtained from the crib-o-gram with older children may lie in that group between those with severe sensori-neural hearing loss and those with normal hearing.

TABLE I. Scoring of Crib-o-gram Responses

Subject	1		2		3		4		5		6		7		8		9		10		11		12		13		14	
	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
5 Agree	0	16	13	0	0	16	13	1	0	16	5	3	3	8	5	6	0	15	13	2	1	14	4	9	4	7	0	13
4 Agree	0	0	1	1	0	0	2	0	0	0	1	2	3	0	0	2	0	1	1	0	0	0	2	1	3	0	1	1
3 Agree	0	0	1	0	0	0	0	0	0	0	3	2	1	1	3	0	0	0	0	0	0	1	0	0	1	1	0	1
Total	0	16	15	1	0	16	15	1	0	16	9	7	7	9	8	8	0	16	14	2	1	15	6	10	8	8	1	15
Pass or Fail		F	P			F	P			F	P		P		P			F	P			F	P		P			F
Age	4M23d		6M16d		6M15d		4m10d		4m14d		2m16d		4m13d		2m16d		4m24d		6m9d		4m10d		6m29d		6m22d		8m5d	
Hearing	90dB+		normal		90dB+		normal		90dB+		normal		normal		normal		90dB+		normal		90dB+		normal		35dB+		65dB+	

Y = Yes, a response seen

N = No, no response noted

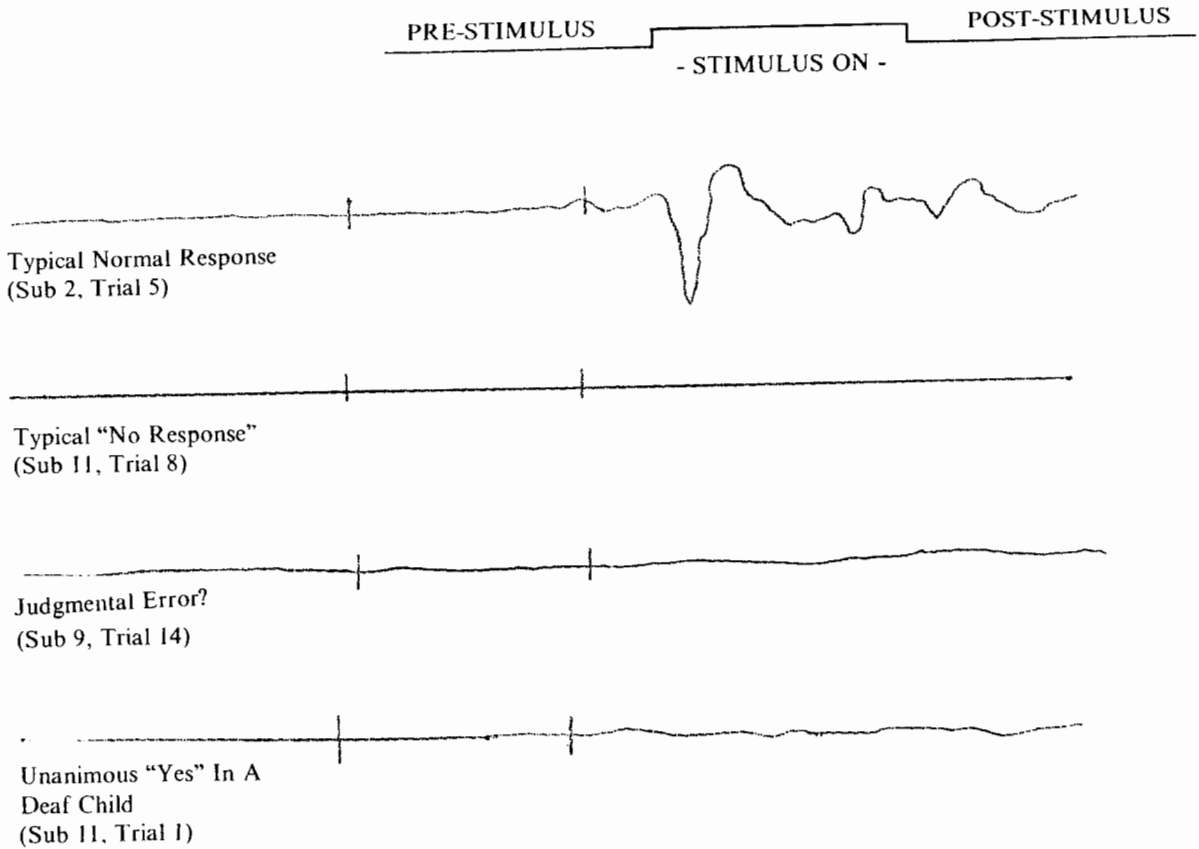

Fig. 1. Sample strip chart recordings from the crib-o-gram.

REFERENCES

Altmann, M.M., R. Shenhav, and L. Schaudinischky, "Semi-objective method for auditory mass screening of neonates," **Early Identification of Hearing Loss** (S. Karger: Basel, Switzerland), edited by G.T. Mencher, 1976.

Jones, F., and F.B. Simmons, "Early identification of significant hearing loss," **Hearing Instruments** 28:8-22, 1977.

Mencher, G.T., editor, **Early Identification of Hearing Loss** (S. Karger: Basel, Switzerland), 1976.

Mencher, G.T., B.J. McCulloch, A.J. Derbyshire, and R. Dethlefs, "Observer bias as a factor in neonatal hearing screening," **Journal of Speech and Hearing Research** 20:27-34, 1977.

Northern, J. and M. Downs, **Hearing In Children** (Williams and Wilkens and Co.: Baltimore), 1974.

Simmons, F.B. and F. Russ, "Automated newborn hearing screening, the crib-o-gram," **Archives of Otolaryngology** 100:1-7, 1974.