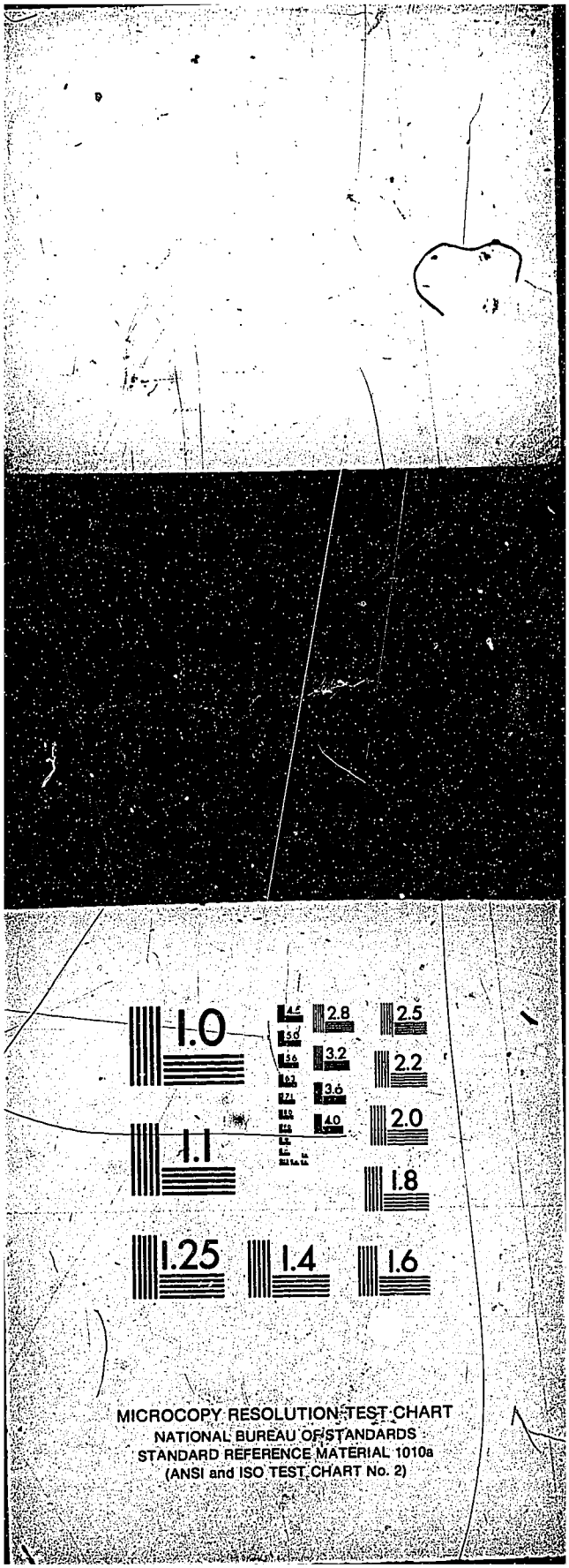
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS
STANDARD REFERENCE MATERIAL 1010a
(ANSI and ISO TEST CHART No. 2)

DOCUMENT RESUME

ED 236 215 TM 830 733

AUTHOR Wilkinson, Cheryl Yelich; Oakland, Thomas
TITLE Stability of the SOMPA's Health History Inventory.
PUB DATE Aug 83
NOTE 32p.; Paper presented at the Annual Meeting of the American Psychological Association, (Anaheim, CA, August 26-30, 1983). For related document, see TM 830 734.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Anglo Americans; Biographical Inventories; Black Students; Elementary Education; Health; *Longitudinal Studies; Medical Case Histories; Mexican Americans; *Scores; Screening Tests; *Test Reliability
IDENTIFIERS *System of Multicultural Pluralistic Assessment; *Test Retest Reliability; Test Score Variance

ABSTRACT

Interest in and use of the System of Multicultural Pluralistic Assessment's (SOMPA's) Health History Inventories (HHI) has been intensive despite the unavailability of psychometric data estimating its stability. This paper reports the results of a longitudinal study designed to provide data on the stability of the HHI over a four year period. The 44 elementary school children studied were drawn from middle and lower class Anglo, Black, and Mexican-American families. Variance in scores is observed between the two administrations of HHI. Means change somewhat, test-retest correlations generally are moderate and lines of best fit for the two sets of data suggest change in scores. Although some variance may be attributable to health incidents which occurred between the two administrations, some variance also is attributable to respondent inconsistency as noted on questions that should receive the same answer across administrations. Despite this variance, "at risk" and "not at risk" classifications made from each Health History Inventory generally are stable. (Author)

Stability of the SOMPA's Health History Inventory

Cheryl Yelich Wilkinson

and

Thomas Oakland

The University of Texas at Austin

Paper presented at the Annual Meeting of the American Psychological Association, Anaheim, California, 1983

SUMMARY: Interest in and use of the System of Multicultural Pluralistic Assessment's Health History Inventories (HHI) has been intensive despite the unavailability of psychometric data estimating its stability. This paper reports the results of a longitudinal study designed to provide data on the stability of the HHI over a four year period. The children studied were drawn from middle and lower class Anglo, black and Mexican-American families.

Variance in scores is observed between the two administrations of HHI. Means change somewhat, test-retest correlations generally are moderate and lines of best fit for the two sets of data suggest change in scores. Although some variance may be attributable to health incidents which occurred between the two administrations, some variance also is attributable to respondent inconsistency as noted on questions that should receive the same answer across administrations. Despite this variance, "at risk" and "not at risk" classifications made from each Health History Inventory generally are stable.

Stability of the SOMPA's Health History Inventory

Issues regarding the psychoeducational assessment of children have been of major concern to school psychology since its inception. Recent litigation (Bersoff, 1981), legislation (Bersoff, 1980), and political activities (Oakland, 1977) have made the public more aware of problems associated with the assessment of persons from different social classes, cultures, and racial-ethnic groups. Our profession has made numerous attempts over the last sixty years to devise more reliable and valid assessment techniques to use with different groups (Oakland, 1980). One of the most recent, the System of Multicultural Pluralistic Assessment (SOMPA; Mercer and Lewis, 1977), is an example of many which are introduced, often with much fanfare, but lacking adequate research prior to their widespread use. The SOMPA's Health History Inventories (HHI) are designed to obtain information about a child's health history in order to identify children who are in need of further medical attention or specialized educational resources (Mercer and Lewis, 1977, p.60).

Basic to any measure's effectiveness is its consistency, reliability or stability (Anastasi, 1982). In brief, a measure must have a high degree of stability or reliability in order to use a scale with confidence. Even though the HHI was introduced six years ago, data regarding its stability over time have not been available. This paper reports the results of a longitudinal study designed to provide data on the stability of HHI over a four year period.

Methodology

Subjects. The 44 children for whom HHI data are reported were part of a larger study involving approximately 450 children ages six through 12 on whom various

psychological, social, medical and educational data were acquired (Oakland, 1979, 1980). These 44 children were entering grades one through three when they were tested originally in 1976 (T_1) and were in grades four through six when they were retested in 1980 (T_2). Approximately one-third of the children in this follow-up study are from each of three major racial-ethnic groups (Anglo, black, and Mexican American), and approximately one-half are from each of two social classes (SES; lower and middle). Finally, approximately half of the children in the follow-up sample are male and half are female.

Instrument. The HHI is a part of the SOMPA's Parent Interview section. It consists of 45 questions presented in five subscales: (1) the Prenatal/ Postnatal Inventory, (2) the Trauma Inventory, (3) the Disease and Illness Inventory, (4) the Vision Inventory, and (5) the Hearing Inventory. Each inventory is scored separately, and the score on each inventory is used to classify a child as either "at risk" (and in need of further medical screening) or "not at risk" (and not in need of further screening).

Procedures

Subject Selection. The norms for the HHI include children between the ages of 5 and 11. Therefore, although the original (T_1) sample included about 450 children, only about 150 of them were less than 12 years old and still eligible for testing at T_2. About 80 eligible children were located using information from original testing records, the schools, local phone directories, the post office, and other sources (e.g., neighbors). These families were contacted by telephone in order to describe the follow-up study and make an appointment for a home visit. Families who did not have a telephone were contacted directly by home visits. Every effort was made to locate and contact all eligible families living in the Central Texas area.

Data Collection. HHI data were acquired through parent interviews in children's homes by psychologists or social workers. All interviewers had previous experience in interviewing families and were trained in using the interviewing and scoring procedures for the HHI. For follow-up (T_2) interviews, families and interviewers were matched on the basis of racial-ethnic group.

Indices of the Stability of HHI. Since no one psychometrically accepted definition of the long term stability of an instrument currently exists, several stability indices were considered for the HHI. These included the stability of group means, the stability of test-retest correlations, the stability of the line of best fit for both testings (i.e., T_1 and T_2) and the stability of classification of children as "at risk" or "not at risk" across HHI administrations. Stability criteria for each of these indices are presented in the Results and Discussion section.

One additional index of stability, the equality of individual answers, was considered for the Prenatal/Postnatal Health History Inventory. Because this measure (and this measure only) concerns events which occurred previous to both administrations of the HHI, answers for this measure should have been the same for both administrations.

Results and Discussion

Stability of Group Means. Means for each HHI show evidence of stability if they do not differ significantly ($p > .05$) between HHI administrations. Mean scores are reported for each HHI subscale for the total follow-up sample as well as for racial-ethnic, gender, and SES groups (Table 1). The results of a two-tailed t test for matched samples between 1976 and 1980 mean raw scores for each group and subgroup also are reported.

PUT TABLE 1 ABOUT HERE

In general, 1980 mean scores for each inventory are about equal to or slightly higher than 1976 mean scores. It seems likely, therefore, that respondents reported some health events at both interviews, and that children may have had some, but not a great number, of major health incidents between HHI administrations. The only exception to this pattern occurrs for the Disease and Illness HHI. Scores on this inventory decrease between administrations for all groups except Mexican Americans and middle SES children. Since all diseases or illnesses reported at the first HHI administration should have been reported again during the second interview, results for this inventory suggest some respondent inconsistency.

Statistically significant differences ($p \leq .05$) are found mainly on the Vision HHI. Means on this inventory increase significantly for the total sample, blacks, Mexican Americans, females and both SES groups. Since 1976 mean scores on this inventory were generally low, a change in the answer to just one question for several members of a group (such as whether a child wore reading glasses) could account for these differences.

Stability of Test-Retest Correlations. Test-retest rank order (rho) and Pearson Product Moment correlations for the two HHI administrations show evidence of stability if they achieve statistical significance ($p \leq .05$).

Rank order (rho) correlations are reported for the total sample (Table 2). The magnitude of these correlations varies widely, ranging from -.03 to .81. Only correlations for the Prenatal/Postnatal and Trauma HHIs are significant.

PUT TABLE 2 ABOUT HERE

Pearson Product Moment correlations are reported for each HHI for the total sample and all subgroups (Table 3). These correlations also vary in magnitude. Correlations are generally high ($\geq .89$) for the Prenatal/Postnatal HHI Correlations are moderate (50s to 60s) for the Trauma and Disease and Illness Inventories and generally low ($\leq .53$) for the Vision and Hearing Inventories. The low correlations found for these last two inventories may stem from the small number of questions included for vision and hearing as well as from actual changes in children's health.

PUT TABLE 3 ABOUT HERE

Stability of Lines of Best Fit. Line of best fit results indicate stability if the line which best describes 1976 (T_1) HHI scores and predicted 1980 scores has a slope of one and an intercept of zero. Under these conditions, the mean of 1980 predicted scores for a given 1976 score equals that 1976 score (i.e., within measurement error, scores are expected to be equal across administrations).

Predicted scores and equations for lines of best fit were derived using a series of linear models (Program MODEL, Ward and Jennings, 1973, pp.137-327). Results for racial-ethnic groups on the Hearing HHI could not be computed because some groups had no variance in scores (i.e., all scores were zero) for one HHI administration.

PUT TABLE 4 ABOUT HERE

Only scores on the Hearing HHI for the total sample, for both gender groups and for both SES groups meet the criterion for stability described above. For all other inventories, predicted 1980 (T_2) scores differed from 1976 scores, suggesting that scores differed across administrations.

Equations for predicted scores are presented in Table 4. Equations produce the most likely second administration score (Y) for a person with a given first administration score (X). For example, the equation for Anglo children for the Trauma HHI is $Y = 1.53 + 0.74X$. The predicted second administration score for an Anglo child with a first administration score of 3 is therefore 4 ($Y = 1.53 + (0.74 \times 3) = 3.75$, which rounds to 4).

Stability of At Risk-Not at Risk Classification. Two stability criteria are established for HHI classification. For the Prenatal/Postnatal Inventory, a criterion of 100% consistency of classification across administrations is used. Since scores on this inventory should have been equal across administrations, classifications should also have been perfectly consistent. For other inventories, a criterion of 75% consistency of classification is used. This less strigent criterion is intended to allow for the fact that health incidents which may occur between administrations can be expected to change the classification of some children. Such change does not reflect HHI instability.

PUT TABLES 5 THROUGH 9 HERE

The number and percentage of children who are and are not not classified consistently by both administrations are reported for each HHI (Tables 5-9). Results are presented for the total sample and for all subgroups. The criterion of 100% consistency is not reached for the total sample or any subgroup on the

Prenatal/Postnatal Inventory. Percentages of consistent classification for this inventory range between 83% and 96%. The criterion of 75% consistency is reached for all groups on all inventories except the Vision HHI. For that inventory, classifications are consistent for no more than 68% of the total sample, blacks, Mexican Americans, males, females and middle SES children.

PUT TABLE 10 ABOUT HERE

Stability of Answers to Prenatal/Postnatal HHI Questions. As has been explained above, scores on the Prenatal/Postnatal HHI can be expected to be the same across administrations. The number and percentage of cases for which scores on this inventory are in fact equal in 1976 and 1980 are lower than expected, with percentages of equal scores ranging from 18% for female children to 50% for black children (Table 10). Because the number of children for whom scores were equal falls so far below expectation, answers to items from the Prenatal/Postnatal HHI were examined individually. Answers changed most frequently for questions which concern the child's birth weight and whether there was anything unusual about of "wrong with" the child at birth. Answers changed least frequently for an item which asks whether the mother had measles or any similar childhood disease while carrying the child.

Summary. In assessing the overall stability of the HHI, various limitations associated with this study's sample should be considered. In addition to its relatively small size, the sample includes children who lived in the Austin, Texas area between 1976 and 1980, whose family residence could be traced in some way through school district, postal and/or telephone listings, and whose mothers agreed to participate in this study. Stability of residence may some-

how interact with childrens' actual health histories or their assessment.

PUT TABLE II ABOUT HERE

Results for all stability criteria for each HHI are summarized in Table II. HHI scores do vary between administractions. Score changes tend to be smallest on the Hearing HHI and greatest on the Vision HHI. Although some variance in scores probably is attributable to actual health incidents which occurred between the two administrations, respondent inconsistency also contributes to variance in scores. Consulting more permanent records pertaining to childrens' health histories (e.g. those kept in such places as baby books or scrap books) should be encouraged so as to improve the reliability of the SOMPA's Health History Inventories.

References

Anastasi, A. Psychological testing. New York: Macmillan, 1982.

Bersoff, D. Nonbiased assessment: Legal principles. In T. Oakland (Ed.) Nonbiased assessment. Minneopolis: University of Minnesota's National School Psychology Inservice Training Network, 1980.

Bersoff, D. Testing and the law. American Psychologist, 1981, 10, 1047-1056.

Mercer, J. & Lewis, J. System of Multicultural Pluralistic Assessment. New York: The Psychological Corporation, 1977.

Oakland, T. Psychological and educational assessment of minority children. Larchmont, NY: Brunner/Mazel, 1977.

Oakland, T. Research on the ABIC and the estimated learning potential. School Psychology Digest, 1979, 8 (2), 63-70.

Oakland, T. An evaluation of the ABIC pluralistic norms and estimated learning potential. Journal of School Psychology, 1980, 18, 3-11.

Ward, J. & Jennings, E. Introduction to linear models: New Jersey: Prentice-Hall, Inc., 1973.

Table 1

Raw Score Means for 1976 and 1980 Administrations of the Health History Inventories

	Prenatal/Postnatal Inventory		Trauma Inventory		Disease and Illness Inventory		Vision Inventory		Hearing Inventory	
	1976	1980	1976	1980	1976	1980	1976	1980	1976	1980
Total Group (N=43)	4.2	4.2	2.8	3.6	2.0	1.8	0.3	1.0	.04	.02
Anglos (N=15)	3.7	3.7	3.9	4.5	2.6	1.2	0.2	0.4	.06	0.0
Blacks (N=12)	4.6	4.4	1.4	2.6	2.2	1.2	0.3	1.6**	.08	0.0
Mexican-Americans (N=16)	4.5	4.5	2.6	3.4	1.4	2.9*	0.5	1.3*	0.0	0.6
Males (N=22)	3.4	3.8	1.7	2.0	1.9	1.7	0.3	0.9	.04	0.0
Females (N=21)	5.1	4.6	3.8	5.2	2.1	2.0	0.3	1.2**	.04	.04
Low SES Family (N=19)	4.9	4.3	3.9	4.1	1.6	1.2	0.1	0.9**	0.1	.05
Middle SES Family (N=24)	3.7	4.2	1.8	3.2	2.3	2.4	0.5	1.2*	0.0	0.0

The probabilities are from t-tests for matched samples.

*$p \leq .05$

**$p \leq .01$

Table 2

Rank Order Correlations between 1976 and 1980 Health History Inventory Raw Scores for the Full Sample

Health History Inventory	N	RHO
Prenatal/Postnatal	43	.83*
Trauma	44	.51*
Disease and Illness	44	.31
Vision	44	.23
Hearing	44	-.03

* $p \leq .001$

Table 3: Pearson Product Moment Correlations Between 1976 and 1980 Health History Inventory Scores

	N	Prenatal/Postnatal Health History Inventory	Trauma Health History Inventory	Disease and Illness Health History Inventory	Vision Health History Inventory	Hearing Health History Inventory
Total Group	44	.91***	.65***	.51***	.34*	-.03
Anglos	16	.78***	.75***	.52*	.12	Uncomputable (1)
Blacks	12	.96***	.26	.72**	.38	Uncomputable (1)
Mexican Americans	16	.91***	.65**	.71***	.45*	Uncomputable (1)
Males	22	.93***	.23	.70***	.08	Uncomputable (1)
Females	22	.91***	.74***	.33	.53**	-.05
Low SES Family	20	.89***	.76***	.32	.04	-.08
Middle SES Family	24	.94***	.49**	.59***	.39*	Uncomputable (1)

* $p \leq .05$
** $p \leq .01$
*** $p \leq .001$

1. These correlations could not be computed because there was no variance in either 1976 or in 1980 scores for the group considered.

Table 4: Equations for Predicted Second Adminstration Health History Inventory Scores[1]

	Prenatal/Postnatal Health History Inventory	Trauma Health History Inventory	Disease and Illness Health History Inventory	Vision Health History Inventory	Hearing Health History Inventory
Total Group	Y =0.89+0.79X	Y =1.53+0.74X	Y =1.02+0.40X	Y =0.88+0.52X	Y =0.00+1.00X
Anglos	Y =1.76+0.54X	Y =1.53+0.74X	Y =0.65+0.23X	Y =0.88+0.52X	Not computable (No variance in scores for some groups.)
Black	Y =0.13+0.94X	Y =1.53+0.74X	Y =0.44+0.37X	Y =0.88+0.52X	
Mexican Americans	Y =1.20+0.73X	Y =1.53+0.74X	Y =1.74+0.83X	Y =0.88+0.52X	
Males	Y =0.61+0.94X	Y =1.53+0.74X	Y =1.02+0.40X	Y =0.88+0.52X	Y=0.00+1.00X
Females	Y =1.24+0.66X	Y =1.53+0.74X	Y =1.02+0.40X	Y =0.88+0.52X	Y=0.00+1.00X
Low SES Children	Y =0.92+0.68X	Y =1.53+0.74X	Y =1.02+0.40X	Y =0.88+0.52X	Y=0.00+1.00X
Middle SES Children	Y =0.79+0.91X	Y =1.53+0.74X	Y =1.02+0.40X	Y =0.88+0.52X	Y=0.00+1.00X

1. Equations produce the most likely second administration score (Y) for a person with a given score on the first administration (X). For example, the predicted second administration Trauma HHI score for an Anglo child with a first administration score of 3 is 4. (Y=1.53+(0.74X3) = 3.75 which rounds to 4)

Table 5: Consistency of At-Risk/Not At Risk Classifications
Prenatal/Postnatal Health History Inventory

	Number and Percentage Not At Risk-Both Administrations	Number and Percentage At Risk-Both Adminstrations	Number and Percentage with Same Classification Both Administration	Number and Percentage with Discrepant Classifications
3)	30 (70%)	9 (21%)	39 (91%)	4 (9%)
15)	11 (73%)	3 (20%)	14 (93%)	1 (7%)
12)	8 (67%)	2 (17%)	10 (83%)	2 (17%)
ericans (N=16)	11 (69%)	4 (25%)	15 (94%)	1 (6%)
2)	16 (73%)	3 (14%)	19 (86%)	3 (14%)
=21)	14 (67%)	6 (29%)	20 (95%)	1 (5%)
=19)	11 (58%)	5 (26%)	16 (84%)	3 (16%)
(N=24)	19 (79%)	4 (17%)	23 (96%)	1 (4%)

Table 6: Consistency of At-Risk/Not At Risk Classifications Trauma Health History Inventory

	Number and Percentage Not At Risk-Both Administrations	Number and Percentage At Risk-Both Administrations	Number and Percentage with Same Classification-Both Administrations	Number and Percentage With Discrepant Classifications
)	35 (80%)	4 (9%)	39 (89%)	5 (11%)
6)	11 (69%)	3 (19%)	14 (88%)	2 (13%)
2)	10 (83%)	0 (0%)	10 (83%)	2 (17%)
ricans (N=16)	14 (88 %)	1 (6%)	15 (94%)	1 (6%)
)	20 (91%)	0 (0%)	20 (91%)	2 (9%)
22)	15 (68%)	4 (18%)	19 (86%)	3 (14%)
20)	14 (70%)	3 (15%)	17 (85%)	3 (15%)
(N=24)	21 (88%)	1 (4%)	22 (92%)	2 (8%)

Table 7: Consistency of At-Risk/Not At Risk Classifications-Disease and Illness Health History Inventory

	Number and Percentage Not At Risk-Both Administrations	Number and Percentage At Risk-Both Administrations	Number and Percentage With Same Classification -Both Administrations	Number and Percentage With Discrepant Classification
	35 (80%)	2 (5%)	37 (84%)	7 (16%)
	12 (75%)	0 (0%)	12 (75%)	4 (25%)
	11 (92%)	1 (8%)	12 (100%)	0 (0%)
icans (N=16)	12 (75%)	1 (6%)	13 (81%)	3 (19%)
	18 (82%)	1 (5%)	19 (86%)	3 (14%)
2)	17 (77%)	1 (5%)	18 (82%)	4 (18%)
0)	17 (85%)	0 (0%)	17 (85%)	3 (15%)
N=24)	18 (75%)	2 (8%)	20 (83%)	4 (17%)

Table 8: Consistency of At-Risk/Not At Risk Classifications
Vision Health History Inventory

Group	Number and Percentage Not At Risk-Both Administrations	Number and Percentage At Risk - Both Administrations	Number and Percentage with Same Classification Both Administrations	Number and Percentage with Discrepant Classifications
Total (N=44)	28(64%)	1(2%)	29(66%)	15(34%)
Anglos (N=16)	13(81%)	0(0%)	13(81%)	3(19%)
Blacks (N=12)	7(58%)	0(0%)	7(58%)	5(42%)
Mexican Americans (N=16)	8(50%)	1(6%)	9(56%)	7(44%)
Males (N=22)	14(64%)	0(0%)	14(64%)	8(36%)
Females (N=22)	14(64%)	1(5%)	15(68%)	7(32%)
Low SES (N=20)	16(80%)	0(0%)	16(80%)	4(20%)
Middle SES (N=24)	12(50%)	1(4%)	13(54%)	11(46%)

Table 9: Consistency of At-Risk/Not At Risk Classifications
Hearing Health History Inventory

Group	Number and Percentage Not At Risk-Both Administrations	Number and Percentage At Risk-Both Administrations	Number and Percentage with Same Classfication Both Administrations	Number and Percentage with Discrepant Classifications
Total (N=44)	41 (93%)	0 (0%)	41 (93%)	3 (7%)
Anglos (N=16)	15 (94%)	0 (0%)	15 (94%)	1 (6%)
Blacks (N=12)	11 (92%)	0 (0%)	11 (92%)	1 (8%)
Mexican Americans (N=16)	15 (94%)	0 (0%)	15 (94%)	1 (6%)
Males (N=22)	21 (96%)	0 (0%)	21 (96%)	1 (5%)
Females (N=22)	20 (91%)	0 (0%)	20 (91%)	2 (9%)
Low SES (N=20)	17 (85%)	0 (0%)	17 (85%)	3 (15%)
Middle SES (N=24)	24(100%)	0 (0%)	24(100%)	0 (0%)

Table 10: Number and Percentage of Cases with Equal Scores on the Prenatal/Post Natal Health History Inventory

Group	Number with Equal Scores	Percentage of Group with Equal Scores
Total Sample[1] (N=43)	14	33%
Anglos (N=15)	4	27%
Blacks (N=12)	6	50%
Mexican Americans (N=16)	4	25%
Males (N=22)	10	45%
Females (N=21)	4	18%
Low SES family (N=19)	4	20%
Middle SES family (N=24)	10	42%

1. A 44th case involving an adopted child had scores for 1976 but not for 1980. If this case, for which the two interviewers differed in their perceptions of whether the mother had enough information to complete the inventory, is also considered unequal, scores were equal for 32% of the total sample.

Table II: Summary of Health History Inventory Stability Results (1)

	Prenatal/Postnatal Inventory	Trauma Inventory	Disease and Illness Inventory	Vision Inventory	Hearing Inventory
	LBF, CLASS	LBF	TR(R), LBF	GM, TR(R), LBF, CLASS	TR(R), LBF, CLASS
	LBF, CLASS	LBF	LBF	TR(P), LBF	
	LBF, CLASS	TR(P), LBF	LBF	GM, TR(P), LBF, CLASS	
ericans	LBF, CLASS	LBF	GM, LBF	GM, LBF, CLASS	
	LBF, CLASS	TR(P), LBF	LBF	TR(P), LBF, CLASS	
	LBF, CLASS	LBF	TR(P), LBF	GM, LBF, CLASS	TR(P)
nily	LBF, CLASS	LBF	TR(P), LBF	GM, TR(P), LBF,	TR(P)
Family	LBF, CLASS	LBF	LBF	GM, LBF, CLASS	

lity criteria were not met for groups and inventories indicated.

d for Stability Criteria: GM = Significant difference in group means between administrations.
TR(R) = Rank order test-retest correlation was not significant.
TR(P) = Pearson Product Moment test-retest correlation was not significant.
LBF = Line of best fit results suggest scores changed across administrations.
CLASS = At risk-not at risk classifications were not consistent across administrations.