Examiner Variability of Several Malocclusion Indices

PAUL C. HERMANSON, B.S.

JOHN M. GREWE, D.D.S., Ph.D.

Occlusion and its deviation from the norm may be classified in several ways. Traditionally, Angle's classification has been used to describe occlusion and its variability in morphological terms.[1] Recently several methods have been devised to classify occlusion according to the severity of the deviations.[2-6] With these newer methods it is conceivable that a Class I malocclusion, although morphologically different from a Class III malocclusion, could be classified as equally severe.

The need for accurate indices that classify occlusions according to severity is increasing. This is probably due to the fact that these indices may be used to establish priorities for receiving treatment and for disbursing insurance and governmental funds for orthodontic treatment. The indices may also be used for testing specific preventive orthodontic procedures or for epidemiological studies. However, before any index is used for establishing a treatment priority, for evaluation of a preventive orthodontic procedure, or for an epidemiological study, it should satisfactorily meet certain basic requirements for precision and bias.[7] The authors of these several indices[2-6] have used varying degrees of statistical analyses to design an index that is as precise and unbiased as possible. However, no one has compared these different indices using the same population and the same examiners.

The objective of this study was to

test the precision and, in part, the bias of the Indian Health Malocclusion Survey,[2] Malocclusion Severity Assessment,[3] Occlusal Evaluation Technique,[4] Occlusal Index[5] and Treatment Priority Index.[6] To accomplish this objective both relative and absolute interexaminer variabilities were determined for each index. The degree of correlation between the ranks obtained using each of the five indices and a standard ranking, obtained by a subjective treatment priority estimation by ten orthodontists, was determined. In addition, the examiner's subjective evaluations of the five indices was summarized.

MATERIALS AND METHODS

The sample consisted of pretreatment study casts of all the fourteen-year-old females with a permanent dentition who were treated in the Orthodontic Clinic at the University of Iowa College of Dentistry from 1965 to 1967. Although the casts were not selected by any criteria other than age, sex and year of treatment, all morphological types of occlusions were represented by the fourteen sets of casts. Each cast was examined by five first-year orthodontic graduate students using the Indian Health Malocclusion Survey, the Malocclusion Severity Assessment, the Occlusal Evaluation Technique, the Occlusal Index and the Treatment Priority Index. The graduate students were approximately equal in educational background with respect to orthodontics and none had used any malocclusion index prior to this study. The examiners were randomly assigned to a different index each day for five con-

Table I

Mean values and standard deviations obtained by five examiners for five malocclusion indices on 14 preorthodontic casts of 14-year-old females.

	Examiner				
Index	$\overline{X} \pm SD$	$\overline{X} \pm SD$	$\overline{X} \pm SD$	$\overline{X} \pm SD$	$\overline{X} \pm SD$
Indian Health Malocclusion Survey	31.29±15.67	34.14±11.02	39.57±14.34	29.86±11.05	32.79±13.65
Malocclusion Severity Assessment	19.93±9.55	26.93±6.63	24.71±9.51	21.29±7.88	26.00±6.97
Occlusal Evaluation Technique	15.36±6.02	21.50±7.42	19.86±9.17	18.64±10.37	19.86±8.29
Occlusal Index	11.28±3.13	11.11±4.56	11.21±4.14	9.81±3.61	10.16±5.37
Treatment Priority Index	7.91±3.98	8.89±3.67	8.34±3.91	9.08±3.02	9.10±4.69

secutive days and none of the examiners used the same index on the same day. They were instructed not to discuss any of the indices with any of the other examiners during this period. Prior to scoring the randomly arranged casts, each examiner was given written instructions on how to use the appropriate index. The instructions used were designed by the authors of each of the respective indices. The examiners were also provided with the essential measuring devices as were specified in the instructions of the various indices. Each examiner, after using an index, completed a questionnaire designed to elucidate his subjective evaluation of that particular index.

The same casts were also subjectively ranked by five second-year orthodontic graduate students and five faculty members in the Orthodontic Department of the University of Iowa College of Dentistry. The instructions for the subjective ranking stated: "Assume that limited funds are allocated for orthodontic treatment in this state and you are the orthodontist who decides which individuals are eligible for this treatment. Arrange the casts in order of whom you would treat first and whom you would treat last basing your decision on the individual's need for treatment as judged from casts only."

Interexaminer variability was determined for the five indices tested in order to assess precision. This was accomplished for each index by an analysis of variance of the absolute scores. In addition to the analysis of variance, the Spearman rank correlation coefficient was determined for each of the possible combinations of examiners for each index and then tested using the "t" test for significance. Subjective rankings were compared with the rankings obtained by the examiners using the several indices in order to partially test bias. The subjective rankings were analyzed using both an over-all analysis of variance of ranks and by determining the Spearman rank correlation coefficient for each of the possible combinations of the subjective rankers. In addition, the mean rank for each index was tested with the mean subjective rank using an analysis of variance. The subjective evaluations as determined from the questionnaires were tabulated and evaluated subjectively.

Table II

Analysis of variance data for the assessment of occlusion by five examiners using five malocclusion indices (absolute scores).

Source of Variation	d.f.	Indian Health Malocclusion Survey		Malocclusion Severity Assessment		Occlusal Evaluation Technique		Occlusal Index		Treatment Priority Index	
		M.S.	F	M.S.	F	M.S.	F	M.S.	F	M.S.	F
Between Examiners	4	195.80	3.68*	140.50	8.02**	73.75	5.48**	7.25	1.86	3.80	0.91
Between Casts	13	705.10	13.30**	224.50	5.01**	297.38	22.03**	73.00	18.70**	60.46	14.50**
Residual	52	53.20	1.02	28.00	0.54	13.46	0.26	3.90	0.08	4.17	0.08

* Significant at 0.05 level
** Significant at 0.01 level

RESULTS

The mean values and the standard deviations were determined for each of the five malocclusion indices tested (Table I). To determine the interexaminer variability for each index an analysis of variance test was undertaken. The results of these tests are shown in Table II. Using these tests, only the Occlusal Index and the Treatment Priority Index showed nonsignificant interexaminer variability at the one per cent level. Examiner disagreement for both the Malocclusion Severity Assessment and the Occlusion Evaluation Technique was significant at the one per cent level. The examiner disagreement for the Indian Health Malocclusion Survey was significant at the five per cent level. As expected, the casts were significantly different for all five indices.

In addition to the analysis of variance for each index, the Spearman rank correlation coefficient was calculated for each of the possible combinations of examiners for each index and then tested using the "t" test for significance. Significant correlations (P<.01) between different examiners using the same index were observed in nine of ten possible combinations in the Occlusal Evaluation Technique and the Indian Health Malocclusion Survey, seven of ten possible combinations in the Occlusal Index and the Treatment Priority Index and six of ten in the Malocclusion Severity Assessment (Table III).

The subjective rankings were obtained for the sample to partially determine bias. This was followed by an analysis of variance of the subjective rankings of the ten examiners and the fourteen casts which showed no significant difference among the examiners ($F_{9,117} = 1.17$). The Spearman rank correlation coefficient was determined for each possible combination of ex-

TABLE III

Examiner variability within five malocclusion indices as determined by the
Spearman rank correlation coefficient and the "t" test

Index	Possible Interexaminer Combinations with an Index	Probability of Examiner Agreement According to Ranking		
		P<.01*	P<.05**	N.S.***
Indian Health Malocclusion Survey	10	9	0	1
Malocclusion Severity Assessment	10	6	2	2
Occlusal Evaluation Technique	10	9	1	0
Occlusal Index	10	7	3	0
Treatment Priority Index	10	7	1	2

　* Significant at 0.01 level
　** Significant at 0.05 level
*** Not significant

aminers who subjectively ranked the casts and then tested using the "t" test for significance. At the five per cent level there was no significant difference between any combination of any two examiners. However, at the one per cent level two of the forty-five possible combinations between the subjective rankings by the examiners did not significantly correlate. This is significantly more precise than the twelve of fifty possible combinations between the examiners that did not correlate (Table III) for the five indices tested ($x^2 = 8.84$ with 1 df). The mean rank for the five examiners for each index was then tested with the mean subjective rank using an analysis of variance of rank. No significant differences ($F_{4,52} = .33$) were observed between the indices. The mean ranks obtained by the five examiners for each of the malocclusion indices and the mean subjective ranks obtained by the ten subjective evaluators are shown in Table IV. Comparisons between the five examiners with the five malocclusion indices and the mean subjective rankings by ten different examiners were determined by the Spearman rank correlation coefficient as shown in Table V. Using this method

of analysis the Occlusal Index and the Occlusal Evaluation Technique were the most precise indices.

Subjectively, the five examiners did not favor or disfavor any one index. The examiners reported that all indices had at least one component that was inadequately defined. In addition, a minimum of one examiner for each index stated that at least one variable, essential for arriving at a reliable estimate of malocclusion severity, was lacking.

Three of the five examiners stated that the Treatment Priority Index was easy to understand and fast to score. Four of the examiners noted that there was no category for posterior openbite and believed that this omission resulted in an inaccurate portrayal of the severity of a case with a posterior openbite. In addition, two of the examiners believed that a category for spacing was also necessary. Two of the five examiners stated that the provision for quantitative measurement rather than subjective evaluation was the strongest feature of this index.

Four of the examiners criticized the Occlusal Index instruction booklet as

TABLE IV

Mean ranks obtained by five examiners for five malocclusion indices and mean subjective ranks obtained by ten examiners on 14 preorthodontic casts of 14-year-old females

Index

Indian Health Malocclusion Survey	Malocclusion Severity Assessment	Occlusal Evaluation Technique	Occlusal Index	Treatment Priority Index	Mean Subjective Rank
7	3.5	11	4	12.5	7
3	7	3.5	1	1	3
5	3.5	3.5	5	11	8
11	6	9	11	4	9
14	10	13	14	12.5	12
8	14	14	12	8	14
1	2	1	2	2	1
6	5	5	6	9	4
10	13	12	10	6	11
13.5	11	8	9	14	13
2	1	2	3	3	2
13.5	12	7	7	5	6
9	8	10	13	10	10
4	9	6	8	7	5

Table V

Comparison of five examiners with five malocclusion indices and the mean subjective ranking of ten different examiners as determined by the Spearman rank correlation coefficient

Index	Possible Combinations of Examiner vs Mean Subjective Ranking with an Index	Probability of Examiner Agreement with Subjective Ranking P<.01* P<.05** N.S.***		
Indian Health Malocclusion Survey	5	4	1	0
Malocclusion Severity Assessment	5	1	4	0
Occlusal Evaluation Technique	5	5	0	0
Occlusal Index	5	5	0	0
Treatment Priority Index	5	1	2	2

  * Significant at 0.01 level
 ** Significant at 0.05 level
*** Not significant

being too long and complex. However, the same four examiners stated that once they understood the instructions the index was the easiest to use. This index was generally agreed to be the most objective of the five indices. This objectivity was attributed to the detailed quantitative classification system. Two examiners stated that they liked the detailed system for classifying crossbites.

The strongest aspect of the Malocclusion Severity Assessment, in the opinion of four of the examiners, was that it was fast to use and required no measurements. One examiner stated that the directions were clear whereas two examiners regarded the directions as vague and too general. One examiner objected to the unquantitative subjective classification of overjet and overbite. Another examiner stated that the crowding and spacing categories were responsible for a disproportionate rise in score inconsistent with the overall malocclusion.

Four of the examiners stated that the Occlusal Evaluation Technique was very fast and easy to score. All five found the instructions easy to interpret. Two examiners believed that this index gave an accurate portrayal of the malocclusion scores. Two other examiners disagreed and noted that posterior openbite was not considered. Another examiner believed the "all or none" nature of the scoring method was unrealistic. Also several categories were noted to be unclear by one or more of the examiners.

Four examiners found the Indian Health Malocclusion Survey easy to understand and all five examiners agreed that it was one of the quickest to score. One examiner believed that this index had the most potential but that more categories such as posterior openbite should be included. Another examiner stated that it was difficult to score cross-

bite since only three classifications were possible.

DISCUSSION

Theoretically, a malocclusion index should be precise and unbiased. Precision in the present discussion can be regarded as the ability to obtain the same score or measurement when one or more examiners measure the same cast at the same or at a different time.[7] Bias, or systematic error, of an index or measuring process is the magnitude and direction of its tendency to measure something other than what was intended.[7] Hence, the results or scores of a precise index should be reproducible by the same examiner or by different examiners. Likewise, the scores of an unbiased index should accurately reflect the intended characteristic, namely, in this instance, severity of malocclusion and, concurrently, the priority for treatment based on need. It is notable that an index could be precise and at the same time be biased. In such a case, the scores would be reproducible but not an accurate portrayal of the occlusion.

From the data obtained from this study the precision in so far as interexaminer agreement is concerned is better for the Occlusal Index and the Treatment Priority Index than for the other three indices. This was determined by an analysis of variance of the absolute scores for each index. However, when considering the relative ranks of the casts using Spearman rank correlation coefficients for each of the possible combinations of the examiners, the Occlusal Evaluation Technique was found to be most precise. These results should not be considered as conflicting since it is theoretically possible for the ranking of casts to vary significantly within an index while the mean scores and variances of the same casts for the same index do not vary

significantly. Conversely, the means and variances may vary significantly but the ranks may not be significantly different. This information on where the variability occurs is important when an index is being considered for a specific use. For example, if absolute scores are used, the most precise and unbiased index from these data would be the Occlusal Index or the Treatment Priority Index. If relative ranks are used, either the Indian Health Malocclusion Survey or the Occlusal Evaluation Technique would be the most precise and unbiased index.

School of Dentistry
Univ. of Maryland
Baltimore, Md. 21201

REFERENCES

1. Angle, E. H.: Classification of Malocclusion, *Dent. Cosmos* 41:248-264, 1899.
2. Glauser, R. O.: Orthodontic Examination and Report PHS 2506-3, Section IV, Indian Health Manual TN No. 57, 1966, pp. 22-26.
3. Salzmann, J. A.: Handicapping Malocclusion Assessment to Establish Treatment Priority, *Am. J. Orthodont.* 54(10):749-765, 1968.
4. Stringfellow, R. D.: A Study of New Criteria for an Orthodontic Index in Epidemiological Surveys, *M.S. Thesis*, University of Tennessee, 1965, p. 96.
5. Summers, C. J.: A System for Identifying and Scoring Occlusal Disorders: The Occlusal Index, *Ph.D.. Thesis*, University of Michigan, 1966, p. 136.
6. Grainger, R. M.: Orthodontic Treatment Priority Index, Public Health Service Publication No. 1000-Series 2-No. 25, U. S. Government Printing Office, Washington, D.C., 1967, p. 49.
7. Eisenhart, C.: Expression of Uncertainties of Final Results, *Science*, 160(3833):1201-1204, 1968.