Reliability of the Bolton tooth-size analysis when applied to crowded dentitions

W. Craig Shellhart, DDS, MS; D. William Lange, DMD; G. Thomas Kluemper, DMD, MS; E. Preston Hicks, DDS, MS, MSD; Alan L. Kaplan, PhD

In order for the maxillary teeth to fit well with the mandibular teeth, there must be a definite proportionality of tooth size (the sum of the widths of the mandibular teeth must be somewhat smaller than the sum of the widths of the maxillary teeth because the mandibular teeth are aligned along an arc that is smaller than that of the maxillary teeth). The presence of disproportionately-sized teeth in either arch can make it difficult to obtain an occlusion with good alignment, ideal overjet, ideal overbite, and a Class I molar relationship. Occasionally, disproportionately-sized teeth are easily recognizable. However, significant discrepancies can occur between the overall sizes of the maxillary and the mandibular teeth that are difficult to detect by inspection alone. If a discrepancy goes undetected initially, it may be only as the treatment of a malocclusion is nearing completion and difficulties in achieving an ideal occlusion occur that the discrepancy is discovered. Discovery during the finishing stages could lead to embarrassing delays in the completion of treatment, or even worse, to a compromised result as options that may have been possible at one time are no longer viable. Therefore, the ability to analyze the proportionality of the maxillary and mandibular teeth is an important diagnostic tool and one that would be best used at the initial diagnostic stage.

One of the most common malocclusions is den-

Abstract

The Bolton tooth-size analysis is widely taught and used in orthodontics. However, its reliability has not been documented. The purpose of this study was to evaluate the reliability of the analysis when performed with needle-pointed dividers and a Boley gauge. Four clinicians measured the teeth on 15 sets of casts with two instruments at two sessions. The measurements were used to calculate tooth-size excess. To evaluate the measurement error, the difference between the two analyses made by the same investigator on the same set of casts was calculated. More of the same-investigator analyses were significantly correlated when the Boley gauge was used than when the needle-pointed dividers were used. Between-investigator analyses revealed significant correlations for each measurement session with both instruments. Every investigator was found to have at least one measurement error for each analysis and with each instrument that was as large as a clinically significant result of a Bolton analysis. The results of this study demonstrate that clinically significant measurement errors can occur when the Bolton tooth-size analysis is performed on casts with at least 3 mm of crowding. The Boley gauge demonstrated a higher frequency of significantly correlated repeated measures and thus may be somewhat more reliable for this analysis than the needle-pointed dividers.

Key Words

Bolton • Tooth-size analysis • Needle-pointed dividers • Boley gauge • Reliability

Submitted: April 1994 **Revised and accepted:** January 1995 Angle Orthod 1995;65(5):327-334.

Table 1
Overall tooth-size discrepancies when measured with a Boley gauge (a negative number indicates a mandibular excess, a positive number indicates a maxillary excess).

	Investigator/session							
	1/1	1/2	2/1	2/2	3/1	3/2	4/1	4/2
Subject								
01	+1.8	+1.2	-2.3	+3.7	+1.5	+2.8	+1.8	+1.8
02	+0.5	+1.4	+0.4	+1.8	+2.2	+2.3	+2.0	+1.3
03	-0.2	+2.4	+2.2	+3.1	+3.6	+3.5	+1.5	+1.7
04	+0.9	+0.1	+1.9	+2.0	+2.3	+1.3	-0.2	+1.6
05	+1.3	+1.6	+2.8	+2.0	+2.9	+1.7	+1.6	+1.2
06	+0.9	+0.9	+3.3	+1.5	+2.8	+3.1	+2.2	+2.8
07	+1.7	+3.1	+1.6	+2.0	+2.8	+3.1	+2.6	+2.6
08	-1.6	-1.0	+0.8	+0.5	+0.0	+0.3	-1.3	-0.8
09	-0.5	-0.4	-2.2	+3.0	-1.0	-1.0	-1.8	-1.2
10	+1.9	+3.0	+5.7	-0.2	+1.4	+1.9	+1.9	+1.6
11	+1.1	+0.7	+1.2	+1.1	+2.9	+2.9	-1.7	+2.2
12	-0.4	-0.4	-1.1	-0.1	-0.8	-0.6	+0.4	+0.5
13	-0.7	-0.I	-1.8	-1.2	+0.5	+0.2	-1.7	+0.7
14	+0.7	+1.2	+2.9	+0.9	+2.0	+1.5	+1.0	-2.2
15	+1.5	+1.1	+0.8	+0.5	+1.0	+3.1	+0.1	-0.3
Mean	+0.6	+1.0	+1.1	+1.4	+1.6	+1.7	+0.6	+0.9

Table 2
Anterior tooth-size discrepancies when measured with a Boley gauge (a negative number indicates a mandibular excess, a positive number indicates a maxillary excess).

	Investigator/session							
	1/1	1/2	2/1	2/2	3/1	3/2	4/1	4/2
Subject								
01	+3.1	+1.6	+4.2	+4.8	+2.9	+2.9	+3.1	+2.9
02	-0.1	+0.6	-0.1	+0.5	+1.0	+0.3	+0.7	+0.5
03	-0.8	+0.6	-0.2	+0.5	+0.1	-0.2	+0.3	-0.1
04	+0.0	-0.5	-0.3	-0.2	+0.8	-0.3	-0.2	+1.1
05	-0.8	-0.8	-0.8	-0.8	-0.1	-0.6	-0.4	-0.6
06	-0.2	-0.2	-0.2	-0.3	+0.7	-1.3	+0.5	+1.4
07	+0.9	+1.2	+1.6	+1.4	+1.3	-0.6	+1.3	+1.3
08	-0.3	-0.6	+0.8	-1.0	+0.4	+0.4	-0.5	-0.1
09	-1.3	-0.7	-2.2	+0.6	-0.7	-0.6	-0.5	-0.5
10	+1.2	+1.7	+1.6	-0.7	+0.9	+1.4	+1.0	-2.1
11	-1.3	-1.5	-1.1	-0.9	-0.7	-1.3	-1.5	-0.7
12	+0.2	+0.1	-0.7	+0.6	-0.4	-1.0	-0.5	-0.2
13	+0.3	+0.5	+0.6	-0.5	-0.7	-0.7	-0.5	+0.3
14	-0.3	+0.5	+1.7	+0.2	+0.4	-0.9	-0.1	-0.8
15	+0.4	+0.8	+0.5	-0.9	-0.3	-0.3	+0.7	+0.3
Mean	+0.1	+0.2	+0.4	+0.2	+0.4	-0.2	+0.2	+0.2

tal crowding. The presence of crowding may alter the relationship of the teeth in such a way that the ideal interproximal contact points must sometimes be estimated. This estimation may introduce variations in the measurement of the teeth such that analyses of tooth-size proportion conducted on crowded casts may have limited reliability.

Very few studies of the proportionality of the teeth have been published[1-4] and none are as clinically useful or as well accepted as the study published in 1958 by Bolton on the relationship of tooth-size disharmony to the treatment of malocclusion.[1] In this study, 55 cases with excellent occlusions were evaluated. The mesiodistal dimensions of the maxillary and mandibular teeth were measured. Ratios were calculated to produce a percentage relationship of mandibular size to maxillary size. This calculation was done for both the anterior teeth (canine to canine) and for the overall dentition (first molar to first molar). A statistical analysis of these percentage relationships resulted in an average for each grouping (overall = 91.3%; anterior = 77.2%) and a standard deviation (overall = 1.91%; anterior = 1.65%). The Bolton analysis has gained wide acceptance and is advocated in a popular orthodontic textbook.[5]

In spite of its prevalent use, the reliability of this analysis has not been studied extensively. There are at least two different instruments that could be used to measure the teeth. Bolton used needle-pointed dividers (compass, bow dividers) when he performed his study.[1] The Boley gauge is another instrument commonly used for the measurement of casts,[5] and it could also be used for this analysis. The reliability of the analysis when performed with either instrument has not been reported in the literature. The purpose of this study was to evaluate the reliability (intra-investigator and inter-investigator) of the Bolton analysis when performed with both the needle-pointed divider and the Boley gauge.

Materials and methods

Pretreatment casts for 15 patients were analyzed by four investigators on two separate occasions. The casts were selected on the basis of having pretreatment crowding of at least 3 mm in one arch. The other criteria for selection that limited the number of casts available were related to a simultaneous study on the repeatability of an analysis when crowded dentitions were compared to an analysis of the same teeth after they had been well aligned. Therefore, the casts used in this study had to have matched posttreat-

ment casts and detailed written treatment records with no indication of interproximal tooth reduction. The total number of casts available that met these criteria was 15. As a result of this pre- and posttreatment matching, the teeth that were extracted during treatment were not included in the measurements. Bolton[6] demonstrated the use of his analysis for planning treatment with extractions. All of the investigators had been trained in the use of the Bolton analysis, and their experience with its use ranged from limited (senior in dental school) to moderate (1 to 3 years on faculty) to extensive (having used and taught the analysis for approximately 20 years). The measurement sessions were at least 2 weeks apart. The casts were analyzed with both the needle-pointed dividers and the Boley gauge at each measurement session. The mesiodistal dimensions of the teeth (mesial to the second molars) were measured. When the needle-pointed dividers were used, the width of each tooth was measured with the divider, then that width was punched along a straight line into a card. The width of each adjoining tooth was punched along the same line using one of the holes of the previously punched tooth. The summed widths were then measured with a Boley gauge to the nearest 0.1 mm. Measurement of the casts with the Boley gauge was simply a summation of the individual tooth-width measures that had been recorded to the nearest 0.1 mm. The presence and magnitude of any tooth-size discrepancy was calculated for both overall and anterior proportions. The formulas derived by Bolton were used in the calculations. They are:

Overall analysis:

$$\frac{\text{Sum of the mandibular widths}}{\text{Sum of the maxillary widths}} = .913$$

Anterior analysis:

$$\frac{\text{Sum of the mandibular widths}}{\text{Sum of the maxillary widths}} = .772$$

If the value derived from measuring the casts was exactly the value published by Bolton (and listed above), a tooth-size discrepancy of 0 mm was recorded. If the value from the casts was greater than Bolton's values, the mandibular teeth were recorded as excessively wide relative to the maxillary. This mandibular excess was recorded as a negative (-) discrepancy. If the value from the casts was less than Bolton's value, the maxillary teeth were recorded as excessive in width, and this was noted by a positive (+) value. To determine the amount of excess, the formulas were simply rearranged. For example,

Table 3
Overall tooth-size discrepancies when measured with needle-pointed dividers (a negative number indicates a mandibular excess, a positive number indicates a maxillary excess).

	Investigator/session							
	1/1	1/2	2/1	2/2	3/1	3/2	4/1	4/2
Subject								
01	+2.2	+2.0	+1.6	+1.9	+1.6	+1.8	+2.7	+1.7
02	+1.7	+2.4	+3.4	+1.5	+1.7	+0.7	+1.1	+1.0
03	+1.8	+3.0	+1.3	+3.4	+2.6	+2.9	+3.2	+1.3
04	+0.7	+1.2	-0.8	+5.3	+2.8	+2.2	+0.7	+1.5
05	+1.1	+1.9	+0.8	+2.6	+1.6	+1.3	+2.1	+0.8
06	+2.5	+2.1	+1.1	+3.1	+0.9	+3.1	+1.4	+1.9
07	+2.8	+2.0	+1.4	+3.1	+2.1	+1.3	+1.3	+3.0
08	-0.1	-0.9	+0.8	-0.1	-1.4	+0.1	-2.0	+0.5
09	-2.2	-0.5	-1.1	+1.7	-0.9	-2.0	-0.9	-1.0
10	+2.2	+1.0	+1.6	+1.9	+0.2	+1.6	+0.9	+1.2
11	+1.7	+0.9	+1.9	+2.3	+1.8	+5.4	+1.8	+1.7
12	-0.5	-1.3	-1.2	-0.7	+1.8	+0.6	-0.1	+0.3
13	-0.1	-1.2	-0.6	-3.1	+0.6	+1.9	-0.3	+1.8
14	+0.8	+1.4	+1.2	+2.4	+2.6	+1.8	+0.5	-3.8
15	+1.4	-1.1	-0.9	+0.1	-0.1	+3.0	+0.8	+1.0
Mean	+1.1	+0.9	+0.8	+1.7	+1.2	+1.7	+0.9	+0.9

Table 4
Anterior tooth-size discrepancies when measured with needle-pointed dividers (a negative number indicates a mandibular excess, a positive number indicates a maxillary excess).

	Investigator/session							
	1/1	1/2	2/1	2/2	3/1	3/2	4/1	4/2
Subject								
01	+1.4	+1.8	+1.6	+1.6	-1.2	+0.3	+1.6	+1.6
02	-0.5	+0.6	+1.3	+0.2	+0.1	-0.5	-0.1	-0.2
03	-0.1	+0.2	-1.0	-0.3	+0.0	-0.4	-0.5	-0.8
04	+0.1	-0.3	-2.7	+0.4	+0.5	-0.8	-0.2	-0.3
05	+1.0	-1.3	-1.7	-1.2	-1.3	-1.9	-0.8	-0.8
06	+0.2	+0.0	-1.5	+0.3	-0.5	-0.8	-0.5	-0.4
07	+0.7	+0.5	-0.7	+1.1	+0.0	+0.1	+0.4	+0.9
08	-1.6	-1.2	-1.1	-0.7	-0.5	-1.3	-1.2	-0.4
09	-1.7	-0.7	-1.8	-0.4	-1.3	-2.1	-1.0	-1.5
10	+1.1	+0.3	-0.9	-0.4	-0.2	+0.3	+0.4	+1.4
11	-3.3	-1.8	-1.8	-1.5	-1.9	-2.8	-1.4	-2.3
12	+0.9	-1.4	-1.3	-1.3	-0.5	-1.1	-1.4	-0.5
13	-0.7	-0.7	-0.5	-1.2	-0.7	-0.5	-0.8	-0.2
14	-0.6	-0.3	-1.8	-1.0	-0.5	-1.2	-0.5	-2.2
15	-0.6	-1.5	-1.0	-0.4	-0.7	+0.1	-0.1	-0.1
Mean	-0.2	-0.4	-1.0	-0.3	-0.6	-0.8	-0.4	-0.4

Table 5
Difference between two analyses of overall tooth-size discrepancy (molar-to-molar) measured with a Boley gauge by the same investigator.

	Investigator			
	1	2	3	4
Subject				
01	0.6	6.0	1.3	0.0
02	0.9	1.4	0.1	0.7
03	2.6	0.9	0.1	0.2
04	0.8	0.1	1.0	1.8
05	0.3	0.8	1.2	0.2
06	0.0	1.8	0.3	0.6
07	1.4	0.4	0.3	0.0
08	0.6	0.3	0.3	0.5
09	0.1	5.2	0.0	0.6
10	1.1	5.9	0.5	0.3
11	0.4	0.1	0.0	3.9
12	0.0	1.0	0.2	0.1
13	0.6	0.6	0.3	2.4
14	0.5	2.0	0.5	3.2
15	0.4	0.3	2.1	0.4
Mean	0.67	1.79	0.55	1.01
Std. dev.	0.66	2.11	0.60	1.23

Table 6
Difference between the two analyses of anterior tooth-size discrepancy (canine-to-canine) measured with a Boley gauge by the same investigator.

	Investigator			
	1	2	3	4
Subject				
01	1.5	0.6	0.0	0.3
02	0.7	0.6	0.7	0.2
03	1.4	0.7	0.3	0.4
04	0.5	0.1	1.1	1.3
05	0.0	0.0	0.5	0.2
06	0.0	0.1	2.0	0.9
07	0.3	0.2	1.9	0.0
08	0.3	1.8	0.0	0.4
09	0.6	2.8	0.1	0.0
10	0.5	2.3	0.5	3.1
11	0.2	0.2	0.6	0.8
12	0.1	1.3	0.6	0.3
13	0.2	1.1	0.0	0.8
14	0.8	1.5	1.3	0.7
15	0.4	1.4	0.0	0.4
Mean	0.50	0.98	0.64	0.65
Std. dev.	0.45	0.86	0.66	0.77

Table 7
Difference between two analyses of overall tooth-size discrepancy (molar-to-molar) measured with needle-pointed dividers by the same investigator.

	Investigator			
	1	2	3	4
Subject				
01	0.2	0.3	0.2	1.0
02	0.7	1.9	1.0	0.1
03	1.2	2.1	0.3	1.9
04	0.5	6.1	0.6	0.8
05	0.8	1.8	0.3	1.3
06	0.4	2.0	2.2	0.5
07	0.8	1.7	0.8	1.7
08	0.8	0.9	1.5	2.5
09	1.7	2.8	1.1	0.1
10	1.2	0.3	1.4	0.3
11	0.8	0.4	4.6	0.1
12	0.8	0.5	1.2	0.4
13	1.1	2.5	1.3	2.1
14	0.6	1.2	0.8	4.3
15	2.5	1.0	3.1	0.2
Mean	0.94	1.70	1.29	1.15
Std. dev.	0.57	1.46	0.99	1.18

Table 8
Difference between two analyses of anterior tooth-size discrepancy (canine-to-canine) measured with needle-pointed dividers by the same investigator.

	Investigator			
	1	2	3	4
Subject				
01	0.4	0.0	1.5	0.0
02	1.1	1.1	0.6	0.1
03	0.3	0.7	0.4	0.3
04	0.4	3.1	1.3	0.1
05	2.3	0.5	0.6	0.0
06	0.2	1.8	0.3	0.1
07	0.2	1.8	0.1	0.5
08	0.4	0.4	0.8	0.8
09	1.0	1.4	0.8	0.5
10	0.8	0.5	0.5	1.0
11	1.5	0.3	0.9	0.9
12	2.3	0.0	0.6	0.9
13	0.0	0.7	0.2	0.6
14	0.3	0.8	0.7	1.7
15	0.9	0.6	0.8	0.0
Mean	0.81	0.91	0.67	0.50
Std. dev.	0.73	0.82	0.38	0.49

Table 9
The largest measurement errors made by each investigator and the average of the largest errors.

Investigator	1	2	3	4	Avg.
Boley gauge					
Overall analysis	2.6	6.0	2.1	3.9	3.7
Anterior analysis	1.5	2.8	2.0	3.1	2.4
Needle-pointed dividers					
Overall analysis	2.5	6.1	4.6	4.3	4.4
Anterior analysis	2.3	3.1	1.5	1.7	2.2

Table 10
Results of the t-test for related samples and probablility values for two analyses made by the same investigator at least two weeks apart (performed on data given in Tables 1-4).

Investigator	1	2	3	4
Boley gauge				
Overall analysis	1.67844 P=0.1154	0.40729 P=0.5342	0.63725 P=0.6900	0.83719 P=0.4165
Anterior analysis	0.88687 P=0.3901	-0.40991 P=0.8618	-2.94059 P=0.0107	-0.17737 P=0.6881
Needle-pointed dividers				
Overall analysis	-0.72447 P=0.4807	1.72775 P=0.1060	1.27805 P=0.2220	-0.04609 P=0.9639
Anterior analysis	-0.49298 P=0.6297	2.50363 P=0.0253	-1.35115 P=0.1981	0.10859 P=0.9151

when the overall excess was mandibular, the amount of excess was computed using the following formula in which X is the ideal mandibular width sum:

X = .913 x sum of maxillary widths.

Once the ideal width sum (X) was calculated, the difference between the actual widths and the ideal widths was determined and recorded.

To evaluate the size of the measurement error, the difference between the two analyses made by the same investigator on the same set of casts was calculated.

A t-test for repeated measures was used to evaluate the presence of a time-related trend in measurements when performed by the same individual at two sessions separated by at least 2 weeks.

Finally, an evaluation of the correlation of the measurements was performed, both of the measures performed by the same investigator (Pearson correlation coefficient), and of the measures performed by different investigators (interclass correlation coefficient).

Results

The results of the Bolton analyses are given in Tables 1-4. The majority of the repeated measures on the same casts yielded different results.

The difference between the two analyses performed by the same investigators at least 2 weeks apart are given in Tables 5-8. The magnitudes of the measurement errors range from 0.0 mm to 6.1 mm.

Table 9 lists the highest measurement errors made by each investigator and the average of these errors. These average highest errors range from 2.2 mm (needle-pointed dividers, anterior analysis) to 4.4 mm (needle-pointed dividers, overall analysis.

Table 10 gives the results of the t-test for related samples. Fourteen of the 16 analyses demonstrated no statistically significant differences (at the $P<0.05$ confidence level) between the two measurements made by the individual investigators.

The Pearson correlation coefficient (Table 11) demonstrates statistically significant correlations (at the $P<0.05$ confidence level) between the two measures made by the same investigators for four of the eight analyses performed with the needle-pointed dividers and for six of the eight analyses performed with the Boley gauge.

Table 12 gives the results of the interclass correlation coefficient (for measures made by different investigators). Every analysis made with both instruments demonstrated a statistically significant correlation between the four investigators.

Discussion

Analyzing the proportionality of the maxillary teeth to the mandibular teeth before treatment begins could allow clinicians to plan treatment so that unforeseen yet significant problems in achieving an ideal occlusion are avoided. The analysis advocated by Bolton, if it were reliable

Table 11
Pearson's correlation coefficients and probability values of two analyses made by the same investigator at least two weeks apart (performed on data given in Tables 1-4).

Investigator	1	2	3	4
Boley gauge				
Overall analysis	0.71835 P=0.0026	-0.15099 P=0.5912	0.83919 P=0.0001	0.45937 P=0.0850
Anterior analysis	0.79527 P=0.0004	0.60533 P=0.0168	0.75379 P=0.0012	0.59544 P=0.0192
Needle-pointed dividers				
Overall analysis	0.69471 P=0.0040	0.25333 P=0.3623	0.44155 P=0.0994	0.34313 P=0.2105
Anterior analysis	0.54536 P=0.0355	0.49711 P=0.0549	0.59393 P=0.0204	0.77240 P=0.0007

for every occlusion, would be an ideal method to use.

Tables 1-4 list the results of two Bolton analyses performed on the same casts by four different investigators. A careful evaluation of these tables reveals that many of the analyses yielded different results on the same casts. Tables 5-8 were calculated to demonstrate the size of the errors in measurement when the same casts were measured by the same investigator with the measurements separated by a 2-week time interval. Perhaps the most informative way to evaluate these errors is to compare them with a clinical standard for significance of a Bolton analysis. A popular orthodontic textbook states, "A tooth size discrepancy of less than 1.5 mm is rarely significant, but larger discrepancies create treatment problems and must be included in the orthodontic problem list."[5] When this standard is used, it can be seen that every investigator made at least one error in measurement that was greater than a clinically significant value for tooth-size excess. This means that even if the patient's maxillary and mandibular teeth were perfectly matched, it would be possible to make an irreversible error in treatment based on the measurement error alone. There was a considerable range of individual variation in the frequency of measurement errors that were large enough to be clinically significant, ranging from 7% to 53% of the repeated measures performed. Possibly the most useful data, in this situation where the clinician would like to reduce irreversible errors, is to describe the largest errors made. These are listed in Table 9. To take an example for discussion, the average of the largest errors when the anterior analysis was performed with the needle-pointed dividers was 2.2 mm. Therefore, if a clinician's repeatability of the Bolton analysis is average, calculations of tooth-size discrepancy should be viewed as ±2.2 mm. With this information in mind, it may be best to view the Bolton analysis as a screening tool. Because of its ease and rapidity of use relative to performing a diagnostic wax-up, it could be used first to determine the possible range of discrepancy. If the entire range would be treated in the same manner (by reducing the width of the teeth), treatment could proceed and the degree of tooth reduction could be tailored to the occlusion as treatment is nearing completion. If, on the other hand, the range of discrepancy indicates two treatment alternatives (reducing the width of the teeth verses extracting a tooth), it would be wise to pursue the more time-consuming diagnostic wax-up. Because the variation between investigators was large for both the number of clinically significant measurement errors and for the size of the measurement errors, and because investigators differed on the instrument which produced fewer errors, it may be wise to evaluate the repeatability of each individual who will be performing the tooth-size analysis.

The results of the t-test for related samples (Table 10) demonstrated that, with the exception of two tests out of 16, there were no time-related trends for measurement error. The two that were found to be significantly different probably indicate a lack of reliability for that investigator with the specific instrument. Thus, for the great bulk of the tests, the clinically significant errors

made in performing the Bolton analysis were not due to any consistent difference in measurement technique between test 1 and test 2.

The correlation between the two measurements made on each set of casts by the same individual (Table 11) were significant at the P<.05 level at least half of the time. We know that clinically significant errors were made by each investigator in each test with both instruments, so these correlations might be best used in comparing the two instruments. The Boley gauge had significant correlations for two of the four investigators when used for the overall analysis, while the needle-pointed dividers had a significant correlation for only one investigator in this analysis. The anterior analysis yielded a similar trend with the Boley gauge demonstrating significant correlations for all four investigators while the needle-pointed dividers had significant correlations for three of four investigators. Thus, the Boley gauge had a slightly better correlation between two measures repeated by the same clinician.

The correlation between the investigators (Table 12) was statistically significant (at the P<0.05 confidence level) for every analysis with both instruments. This was true when the first measurements were evaluated, when the second measurements were evaluated, and when all measurements were combined. Thus, it was not possible to discern a difference between the instruments with regard to the correlation of results between investigators.

Conclusions

The results of this study demonstrate that clinically significant measurement errors can occur when the Bolton tooth-size analysis is performed on casts that have at least 3 mm of crowding. The size and frequency of these errors evidenced considerable inter-individual variation.

The Boley gauge demonstrated a higher frequency of significantly correlated repeated measures and thus may be slightly more reliable for this analysis than the needle-pointed dividers.

Table 12
Intraclass correlation coefficients and probability values for all investigators in the first measurement session, the second session, and both sessions combined (performed on data given in Tables 1-4)

Measurement session	First	Second	Both
Boley gauge			
Overall analysis	0.45 P=0.0001	0.41 P=0.0002	0.61 P=0.0001
Anterior analysis	0.80 P=0.0001	0.49 P=0.0001	0.79 P=0.0001
Needle-pointed dividers			
Overall analysis	0.55 P=0.0001	0.29 P=0.0067	0.59 P=0.0001
Anterior analysis	0.33 P=0.0020	0.67 P=0.0001	0.69 P=0.0001

Author Address

W. Craig Shellhart
Department of Growth and Development
School of Dentistry, UCHSC
4200 E. Ninth Ave., Box C284
Denver, CO 80262

W.C. Shellhart is an associate professor in the Division of Orthodontics, School of Dentistry, University of Colorado. This work was completed during his tenure at the University of Kentucky.

D.W. Lange is in the private practice of orthodontics in Ohio and was formerly a dental student at the University of Kentucky.

G.T. Kluemper is an assistant professor, Section of Orthodontics, College of Dentistry, University of Kentucky.

E.P. Hicks is an associate professor and director of the graduate program, Section of Orthodontics, College of Dentistry, University of Kentucky.

A.L. Kaplan is an associate professor, College of Dentistry, University of Kentucky.

References

1. Bolton WA. Disharmony in tooth size and its relation to the analysis and treatment of malocclusion. Angle Orthod 1958;14:67-71.
2. Lundstrom A. Intermaxillary tooth width ratio and tooth alignment and occlusion. Acta Odont Scand 1954;12:265-292.
3. Neff CW. Tailored occlusion with the anterior coefficient. Am J Orthod 1949;35:309-314.
4. Steadman SR. Predetermining the overbite and overjet. Angle Orthod 1949;19:101-105.
5. Proffit WR. Contemporary orthodontics. St. Louis: Mosby, 1993:158-160.
6. Bolton WA. The clinical application of a tooth-size analysis. Am J Orthod 1962;48:504-529.