

Bias in superimposition of lateral cephalograms in anterior openbite patients

Sarah R. Ovard

A thesis

**submitted in partial fulfillment of the
requirements for the degree of**

Master of Science in Dentistry

University of Washington

2021

Committee:

Greg Huang

Geoff Greenlee

Brian Leroux

Program Authorized to Offer Degree:

Orthodontics

©Copyright 2021
Sarah R. Ovard

University of Washington

Abstract

Bias in superimposition of lateral cephalograms in anterior openbite patients

Sarah R. Ovard

Chair of the Supervisory Committee:

Greg Huang

Department of Orthodontics

Introduction

Anterior open bite malocclusion can be treated non-surgically using fixed appliances, clear aligners, or TADs. Proponents of clear aligners and TADs often attribute bite closure to molar intrusion and counterclockwise rotation of the mandibular plane. These changes may be supported by superimpositions. However, the process of creating a superimposition is subjective, and therefore may be influenced by practitioner bias.

Methods

The initial and final lateral cephalograms from 30 adult anterior openbite patients were used in this study. Ten of these patients were treated with fixed appliances, 10 with clear aligners, and 10 with temporary anchorage devices (TADs). We asked six orthodontic graduate students and six orthodontic practitioners to complete superimpositions using these radiographs in three separate sessions. In the first session, the raters were told that all patients received treatment with fixed appliances only. In the second session, the raters were told that all patients were treated with clear aligners only, and at the third session, they were told all patients were treated with fixed appliances and TADs. Superimpositions were performed using Dolphin software, and each superimposition was saved as a PDF image. Change in the mandibular plane was the primary outcome and was assessed categorically (closed, no change, opened). Cephalometric values were measured and used to investigate the dental and skeletal changes associated with treatment.

Results

Although the raters demonstrated a slight tendency toward the mandibular plane closing or staying the same when told the treatment was clear aligners or TADs, these differences were not statistically significant. A high degree of intra- and inter- rater variability in the mandibular plane change was present from all 3 superimposition sessions. The measurements from the lateral cephs showed significant changes for overbite and for incisor vertical and angular movements. Almost no change was observed in anterior facial height, mandibular plane angle, or vertical movement of the 1st molars.

Conclusions

A significant amount of superimposition bias was not observed in this study. However, there was considerable intra- and inter-rater reliability. This suggests that the same initial and final cephs may be interpreted very differently, based on the subjective superimposition of the raters.

Introduction and literature review

Anterior open bite (AOB) is a condition where there is no contact between the upper and lower incisors when the posterior teeth are in occlusion. This vertical discrepancy can be related to, or caused by: growth, habits (i.e. digit sucking or tongue posture), respiratory issues, neuromuscular issues, condylar changes, or any combination of these (Cangialosi, 1984; Ng et al., 2008). Patients with an AOB primarily of dentoalveolar nature exhibit relatively normal facial proportions but intruded incisors. Patients with an AOB of skeletal etiology may present with an increased mandibular plane angle, gonial angle and anterior facial height, decreased posterior facial height, and divergent cephalometric planes. Facially, these factors often lead to a dolichofacial appearance (Cangialosi, 1984; Subtelny and Sakuda, 1964; Lopez-Gavito, et al., 1985; Nahoum, et al., 1972). Due to the multifactorial etiology and high relapse potential, AOB treatment continues to be one of the most challenging malocclusions for successful orthodontic treatment (Cangialosi, 1984).

Several non-surgical methods have been used for the treatment of AOB. Standard fixed appliances and elastics can be used to extrude incisors or be paired with premolar or molar extractions to allow greater retraction and lingual tipping of the incisors through the “drawbridge effect” (Aras, 2002; Freitas, et al., 2018; Janson, et al., 2006). The Multiloop Edgewise Archwire (MEAW) technique, developed by Kim in 1987, has also shown successful treatment of AOB (Kim, 1987; Kim, et al., 2000; Chang and Moon, 1999).

In some cases, temporary anchorage devices, such as mini-screws and mini-plates, may be used with the goal to intrude the posterior teeth (Nanda, 2005; Kuroda et al., 2007). Several studies have utilized bone-borne anchorage through mini-screws or mini-plates to achieve posterior intrusion and decreased facial height due to increased rotation of the mandible (Freitas,

et al., 2018; Sherwood, et al., 2002; Scheffler, et al., 2014; Deguchi, et al., 2011). However, according to a recent systematic review, there is currently only weak evidence suggesting that molar intrusion with TADs may cause mandibular counterclockwise autorotation (Alsafadi, et al., 2016).

As clear aligner therapy continues to rise in popularity, several case series have reported that open bite closure and treatment effects are similar to that achieved by fixed appliances (Guarneri et al., 2013; Schupp, et al., 2010; Garnett et al., 2019). It is thought the double thickness of the aligner plastic on the occlusal surfaces, in combination with the force of mastication, are able to produce an intrusive force on the posterior dentition (Guarneri et al., 2013; Boyd, 2007). Another advantage is the ability to virtually plan treatment and produce isolated and selective tooth movement, reducing the incidence of unwanted incisor extrusion (Boyd, 2008). However, recent studies have shown AOB correction with clear aligners is primarily achieved through incisor extrusion (Khosravi, et al., 2017), with no significant differences in the magnitude of overbite correction or changes in mandibular plane angle when compared to fixed appliances (Garnett, et al., 2019).

Typically, dental and skeletal treatment changes are assessed by the superimposition of pre- and post-treatment lateral cephalometric radiographs. Identifiable points are marked, and landmarks are traced (Bjork and Skieller, 1983; Bjork and Skeiller, 1977; Melsen 1974; Enlow, 1964; Buchang, 1986; Doppel, 1994; Nielsen, 1989). While many of these landmarks are stable, there are still errors due to projection and identificaiton (Baumrind and Frantz, 1971).

Practitioner bias may also be introduced when completing manual superimpositions due to the knowledge of the treatment provided and the anticipated dental and skeletal changes as a

manifestation of the Pygmalian, or Rosenthal, effect which refers to a person's expectations influencing performance (Rosenthal 1994).

While case reports demonstrate mandibular plane closure due to molar intrusion, a systematic review reports less convincing findings, suggesting that practitioner bias may be impacting the interpretation of results (Alsafadi, et al., 2016). The purpose of this study is to evaluate bias in manual cephalometric superimposition when assessors are aware of the type of treatment provided. This was accomplished by having assessors perform repeat superimpositions on the same set of 30 patients, but deceptively informing them that the patients were treated with different techniques at each of three superimposition sessions. Cephalometric values were also used to investigate the dental and skeletal changes associated with treatment.

Material and Methods

This study was approved by the University of Washington Human Subjects Division IRB and informed consent was obtained from all participants prior to gathering data. Twelve participants were recruited from June 2020 to July 2020 at the University of Washington to include 6 practitioners and 6 graduate orthodontics students. A power calculation was used to detect differences in the proportion of superimpositions indicating mandibular plane closure when participants are informed that treatment involved clear aligners or TADs, compared to conventional fixed appliances. With a sample of 30 radiographs and 12 examiners, the power was at least 80% to detect a 25% difference in cases with closure of the mandibular plane angle. Increasing the number of participants further would only have a small effect on the power.

Inclusion criteria for the participants included: practitioners must have been graduates of an accredited orthodontic program and have been in practice for >5 years and orthodontic

graduate students must have completed the cephalometrics course and be in their second or third year of residency. Anyone who was aware of the goals of this study was excluded from participation.

The lateral cephalometric radiographs were selected from records housed at the University of Washington Department of Orthodontics. Thirty pairs of de-identified radiographs were selected based on the following inclusion criteria: (1) all patients 18 years or older (2) good quality pre and post treatment radiographs; (3) millimetric ruler available on both x-rays to calibrate magnification; (4) 10 patients treated with clear aligners; (5) 10 patients treated with fixed appliances only; (6) 10 patients treated with TADs. Radiographs were excluded if the patient underwent orthognathic surgery or were of poor quality.

Pre- and post-treatment radiographs were imported into Dolphin Imaging software (version 11.0; Dolphin Imaging and Management Solutions, Chatsworth, California). The following landmarks were identified by two calibrated examiners: A point, ANS, anatomical gnathion, articulare, B point, basion, bridge of the nose, condyilion, constructed gonion, gonion, menton, mid ramus, nasion, orbitale, PNS, PT point, porion, ramus point, ruler point 1, ruler point 2, soft tissue glabella, soft tissue A point, soft tissue B point, soft tissue gnathion, soft tissue nasion, soft tissue pogonion, sella, sigmoid notch, soft tissue menton, stomion inferius, stomion superious, upper and lower lip, subnasale, throat point, tip of nose, L6 mesial,

Cephalometric measurements
SNA (°)
SNB (°)
MP - SN (°)
Anterior Face Height (NaMe) (mm)
Posterior Facial Height (mm)
U1 - NA (°)
U1 - NA (mm)
L1 - NB (°)
L1 - NB (mm)
IMPA (L1-MP) (°)
Overbite (mm)
Overjet (mm)
U6 - PP (UPDH) (mm)
L6 - MP (LPDH) (mm)
U1 - PP (UADH) (mm)
L1 - MP (LADH) (mm)

Figure 1. Cephalometric measurements collected for initial and final tracings from 30 pairs of radiographs

distal and occlusal; U6 mesial, distal, and occlusal; L1 root, tip, labial and lingual gingival borders; U1 root, tip, labial and lingual gingival borders. The cephalometric measurements listed in **Figure 1** were automatically calculated for both time periods with the Dolphin Imaging software based on a custom analysis designed for this study.

One investigator made master line tracings using existing landmarks of all 30 selected pairs of pre- and post-treatment radiographs. The 30 tracings were duplicated into three sets, each of which included 5 duplicate tracings to assess intra-rater reliability, for a total of 35 tracings per set. The sets were labeled A, B, and C, and the order of the 35 tracings was randomized within each set. Each participant was randomly assigned a set order of either ABC, BAC, ACB, CAB, BCA, CBA. The participants had three separate tracing sessions, all about one month apart. At each tracing session the participant was asked to complete one set (A, B, or C) of all 35 manual superimpositions using the Tracing Superimposition feature on the Dolphin Imaging software. The investigator pulled up each pair of tracings and asked the participant to complete the superimposition. Then, the investigator would save a PDF image of the completed superimposition.

For the first tracing session, consent was obtained from the participants and they were deceptively informed about the purpose of the study. Specifically, they were told, “This study is being conducted to evaluate the reliability of manual superimposition methods.” They also were told that all patients were adults and received treatment with fixed appliances only. At this session, the participants also filled out a short survey to collect demographic information as well as information on their preferred reference structures for completing superimpositions. For the second tracing session, the participants were told all adult patients received treatment with clear aligners only. For the third and final session the participants were told all adult patients received

treatment that included fixed appliances and TADs. The participants were informed some duplicated cephs were included in each session to assess intra-rater reliability.

The change in mandibular plane angle was considered to be the primary outcome measurement. Comparison was made between the participants' three tracing sessions. The change in mandibular plane angle was evaluated categorically by assessing if the mandibular plane angle opened, closed, or stayed the same when looking at the completed superimposition of the initial and final tracing (**Figure 2**).

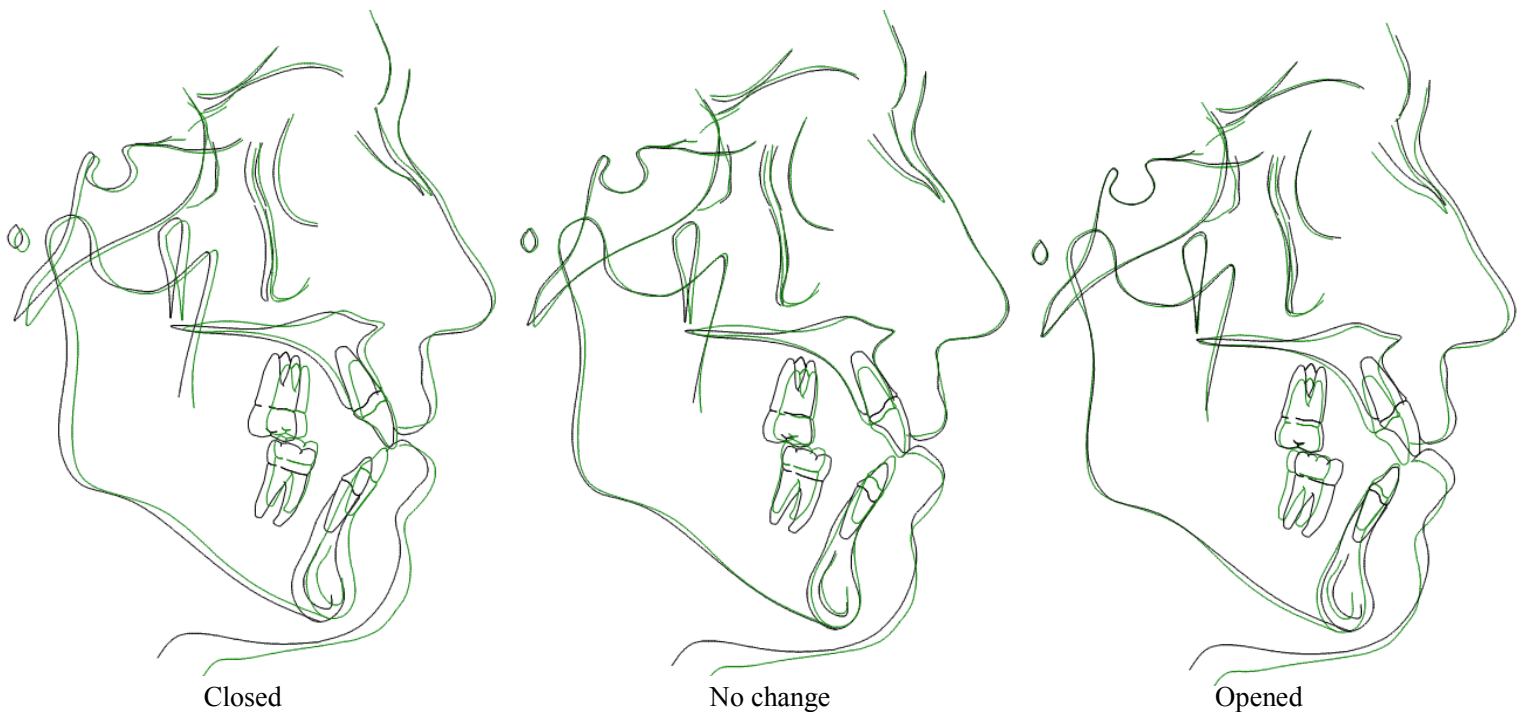


Figure 2. Example of closed, opened, and no change in mandibular plane angle on the same patient from one rater's three separate tracing sessions. Initial tracing is in black, final tracing is in green.

The PDF images of each completed superimposition were evaluated and scored based on the change in the mandibular plane angle. A score of 1 was given if the mandibular plane opened, -1 if the mandibular plane closed, and 0 if the mandibular plane angle stayed the same. The criteria that was used to assess the change in mandibular plane was the anterior portion of

the lower mandibular border. If that area of the superimposition had clear separation of the lines, the mandibular plane was determined to be opened or closed. If the entire mandibular lower borders were parallel, the mandibular plane angle was assessed to be the same. All 1,260 PDF images were ranked in this manner, which included the duplicate patients that were used to assess intra-rater reliability. However, only the participants' first encounter with a duplicate was used for the analysis of the overall results. The second encounter with the duplicate was only used for assessment of intra-rater reliability.

A counts table was created from the rankings to evaluate the overall results. The 30 patients in each session were divided by their actual treatment method, then the number of open/closed/no change were added. Totals for each of these categories were calculated per session, per rater, and per actual treatment type (**Appendix 2**).

Statistical Analysis

Changes from initial to final cephalometric measurements were assessed using an effect size calculation: $|(Mean_{T2} - Mean_{T1})|/SD_{T1}$. Effect size was calculated for each measurement in Figure 1 for each treatment group and the overall patient group. Resulting effect sizes more than 0.5 were considered moderate, and effect sizes greater than or equal to 1 were considered large.

Each rater's duplicate superimpositions were used to assess rater reliability. Cohen's kappa coefficient (κ) and intra-class correlation coefficient (ICC) was used to measure the intra-rater reliability. Inter-rater variability within each session was also assessed using the ICC.

Results

Six female orthodontic graduate students with a mean age of 29 and range of 27-32 were tested. Four students were in their third year and two were in their second year (These two

classes had all female compositions). The primary structures used for reference when completing their superimpositions were the cranial base at the anterior Sella. Only one graduate student said they referenced the cranial base using the cranial fossa structures. The mean number of days between their tracing sessions was 38.4 days (**Table 1**).

Graduate students' demographics	
Age	29 (27-32)
Gender	
Male	0
Female	6
Year in Residency	
3	4
2	2
Orthodontic Program	
Univ. of Washington	6
Method for Superimposition	
Cranial Base - Anterior sella	5
Cranial Base - Cranial fossa structures	1
Days between T1-T2	39.7 (34-48)
Days between T2-T3	37.2 (32-43)
Mean days between sessions	38.4 (33-44)

Table 1. demographics of orthodontic graduate students that participated.

Six orthodontic practitioners with a mean age of 47 and range of 36-58 were tested. Three of the practitioners were male and three were female. The average number of years in practice was 16.5 with a range of 6-30 years. The primary structures used for reference when completing their superimpositions were the cranial base at the anterior sella. One practitioner said they referenced the cranial base using the cranial fossa structures. One practitioner said they referenced the Sella-Nasion line. And one practitioner said they aimed for “best fit”. The mean number of days between their tracing sessions was 40.3 days (**Table 2**).

Practitioners' demographics	
Age	47 (36-58)
Gender	
Male	3
Female	3
Years in practice	16.5 (6-30)
Orthodontic Program	
Univ. of Washington	2
New York University	1
Univ. of Southern California	1
Univ. of Tennessee	1
Univ. of Texas School of Dentistry	1
Method for Superimposition	
Cranial Base - Anterior sella	3
Cranial Base - Cranial fossa structures	1
Sella-Nasion	1
Best Fit	1
Days between T1-T2	42 (28-63)
Days between T2-T3	38.7 (28-49)
Mean days between sessions	40.3 (28-49)

Table 2. demographics of orthodontic practitioners that participated.

When looking at the graduate student's results for Session 1, when the raters were told the patients were treated with fixed appliances, 82 superimpositions were evaluated to show no change in the mandibular plane angle. Forty-two were shown to have opening of the mandibular plane angle and 56 were shown to have closure of the mandibular plane angle. In Session 2, when told the treatment was done with clear aligners, 80 were evaluated to show no change, 59 closed, and 41 opened. Finally, in Session 3, when told the treatment involved TADs, 84 superimpositions were rated to have no change, 57 closed, and 39 opened in the mandibular plane angle (**Table 3**).

		Session1 - Fixed	Session2 - Aligners	Session3- TADs
Graduate Students Total per session	Open	42	41	39
	No Change	82	80	84
	Closed	56	59	57

Table 3. Graduate student results from three tracing sessions and evaluation of if superimposition showed mandibular plane angle opening, closure, or no change.

When looking at the orthodontic practitioner’s results for Session 1, 62 superimpositions were evaluated to show no change, 70 showed closure, and 48 showed opening of the mandibular plane angle. In Session 2, 67 were evaluated to show no change, 65 closed, and 48 opened. In Session 3, 67 superimpositions were rated to have no change, 70 closed, and 43 opened in the mandibular plane angle (**Table 4**).

		Session1 - Fixed	Session2 - Aligners	Session3- TADs
Practitioners Total per session	Open	48	48	43
	No Change	62	67	67
	Closed	70	65	70

Table 4. Orthodontic practitioner results from three tracing sessions and evaluation of if superimposition showed mandibular plane angle opening, closure, or no change.

When comparing graduate students to practitioners, the graduate students had 9.3% more instances where they showed no change in the mandibular plan in the superimpositions than the practitioners did. Instead, the practitioners tended to show more superimpositions with mandibular plane angle closure, 6.1% on average.

In Session 2 (aligners), the graduate students showed slightly more closure of the mandibular plane than the other sessions. The orthodontic practitioners, though, had five less superimpositions in Session 2 showing mandibular plane closure than sessions 1 and 3. However, there were no statistically significant difference between sessions for either residents

($p=0.73$ for comparison of Session 2 versus Session 1, $p=0.78$ for Session 3 versus Session 1) or for faculty ($p=0.40$ for Session 2 versus Session 1, $p=0.49$ for Session 3 versus Session 1), suggesting any changes were due to rater variability rather than true influence of rater bias.

Even when the two groups' results are combined, the overall results do not show any statistical significance between any of the sessions, ($p=0.94$ for Session 2 versus Session 1, $p=0.54$ for Session 3 versus Session 1). Most raters' superimpositions showed no change in the mandibular plane angle, followed by closure of the mandibular plane in each session. Opening of the mandibular plane angle was demonstrated the least in each session and though slight, the trend of mandibular plane angle opening decreased over each subsequent session. Like the individual groups, the combined results suggest any variability was due to rater variability instead of rater bias (**Table 5**). When looking at the individual results, most raters showed a small trend in the expected direction of our hypothesis. One rater showed statistically significant differences in the opposite direction ($p < 0.01$). The trends were not statistically significant for all other raters ($p > 0.05$).

		Session1 - Fixed	Session2 - Aligners	Session3- TADs
Overall Total per session	Open	90	89	82
	No Change	144	147	151
	Closed	126	124	127

Table 5. Combined graduate student and practitioner results from three tracing sessions and evaluation of if superimposition showed mandibular plane angle opening, closure, or no change.

The overall percentage of intra-examiner agreement was 68.3%. The unweighted kappa, based on comparing percent agreement with the expected percent agreement, had a result of 0.50. However, this result treats all disagreements the same even though some are worse than others (e.g. duplicate one rated as open and duplicate two rated as closed is worse than duplicate one rated as open and duplicate two rated as no change), which makes reliability look worse than

it is. The weighted kappa statistic, which accounts for the magnitude of the disagreements, was 0.60 which is considered moderately good (0.8 or higher is considered excellent). Inter-rater variability using an intra-class correlation coefficient (ICC) was 0.70, 0.62, and 0.69 for Sessions 1, 2, and 3, respectively, indicating moderate agreement.

The initial mean measurements were compared for the 30 patients' lateral cephs and used to evaluate how similar the patients were to each other (**Table 6**). The ten patients actually treated with TADs had the least severe open bite of the thirty patients, at only -1.6 mm vs -2.4 and -2.3 mm. The patients treated with clear aligners had the steepest mandibular plane angle and patients treated with TADs had the lowest mandibular plane angle to start with.

	Mean T1			
	<u>Fixed</u>	<u>Aligners</u>	<u>TADs</u>	<u>Overall</u>
<i>SNA (°)</i>	80.1	78.8	81.5	80.1
<i>SNB (°)</i>	76.4	74.9	79.6	77.0
<i>MP - SN (°)</i>	39.8	41.0	33.2	38.0
<i>Anterior Face Height (NaMe)</i>	121.5	124.5	121.7	122.6
<i>Posterior Facial Height (mm)</i>	45.7	45.1	50.0	47.0
<i>U1 - NA (°)</i>	21.9	23.8	25.9	23.9
<i>U1 - NA (mm)</i>	5.3	4.9	5.7	5.3
<i>L1 - NB (°)</i>	32.5	30.2	27.6	30.1
<i>L1 - NB (mm)</i>	7.6	7.4	6.2	7.1
<i>IMPA (L1-MP) (°)</i>	96.4	94.3	94.8	95.2
<i>Overbite (mm)</i>	-2.4	-2.3	-1.6	-2.1
<i>Overjet (mm)</i>	2.7	3.1	1.9	2.5
<i>U6 - PP (UPDH) (mm)</i>	24.2	25.2	25.5	25.0
<i>L6 - MP (LPDH) (mm)</i>	33.9	34.7	33.7	34.1
<i>U1 - PP (UADH) (mm)</i>	29.2	30.8	28.7	29.6
<i>L1 - MP (LADH) (mm)</i>	41.4	42.5	41.6	41.9

Table 6. Summary of mean T1 measurements for 30 patients.

Large effect sizes were seen across all the groups for change in overbite. U1-NA and L1-NB had large effect sizes for the clear aligner group. L1-NB and IMPA had large effect sizes for the TADs treatment group. Several dental categories showed moderate changes as well, most notably, vertical change measurements of U1 and L1 had near moderate change in all categories except L1-MP in the aligners group. No skeletal categories or vertical molar categories showed moderate or large effect size outcomes (**Table 7**).

	Fixed	Aligners	TADs	Overall
<i>SNA (°)</i>	0.2	0.3	0.2	0.2
<i>SNB (°)</i>	0.1	0.2	0.1	0.0
<i>MP - SN (°)</i>	0.0	0.0	0.1	0.1
<i>Anterior Face Height (NaMe)</i>	0.0	0.0	0.0	0.0
<i>Posterior Facial Height (mm)</i>	0.0	0.1	0.0	0.0
<i>U1 - NA (°)</i>	0.1	1.0	0.4	0.4
<i>U1 - NA (mm)</i>	0.1	0.6	0.1	0.2
<i>L1 - NB (°)</i>	0.5	0.9	1.1	0.8
<i>L1 - NB (mm)</i>	0.1	0.5	0.5	0.4
<i>IMPA (L1-MP) (°)</i>	0.5	0.6	0.9	0.7
<i>Overbite (mm)</i>	3.4	1.8	2.2	2.2
<i>Overjet (mm)</i>	0.2	0.5	0.3	0.1
<i>U6 - PP (UPDH) (mm)</i>	0.1	0.1	0.0	0.0
<i>L6 - MP (LPDH) (mm)</i>	0.1	0.1	0.1	0.0
<i>U1 - PP (UADH) (mm)</i>	0.5	0.5	0.3	0.4
<i>L1 - MP (LADH) (mm)</i>	0.3	0.1	0.4	0.3

Table 7. Effect size calculation for cephalometric measurements.

Discussion

In this study, graduate students and orthodontic practitioners were asked to complete superimpositions on the same group of 30 patients at three time points. Deception was employed to blind them from the actual goals of the study, and we expected that the superimpositions they produced would be biased by the treatment information provided to them. Surprisingly, we saw

very little evidence of bias, which may have been explained by some of the study parameters. For instance, the raters were not asked to form any conclusions about the findings or validity of their completed superimpositions. The raters also were not involved with the treatment of the patients. Therefore, it is possible that our raters expressed less bias than might have been expected if the superimpositions were performed by raters who had strong interests in demonstrating specific dental and skeletal changes. Instead, our raters were told that the study was investigating superimposition reliability, and therefore may have been trying to complete the superimpositions as “correctly” as possible, based on their education and experience. In retrospect, implementing an aspect of self-evaluation of the raters’ superimpositions and expected outcomes for each treatment method, or conducting a study testing proponents/non-proponents of clear aligners and TADs for molar intrusion and mandibular plane closure might allow us to better evaluate this potential bias.

Though the results of this study did not show any statistically significant bias, some interesting anecdotal observations were made during the tracing sessions. One rater stated, “This superimposition could be done differently depending on what you think the treatment results were.” This statement illustrates how superimpositions do not have a single answer but are a judgment call that can be manipulated within a range of what may be considered acceptable. This statement also suggests a subjective bias may in fact exist but was not demonstrated in our results due to not having the raters evaluate or interpret their own superimpositions.

Another observation was made during the third tracing session, when the raters were told the patients were all treated with TADs. The majority of the raters expressed, “This group is a lot harder to superimpose!” The raters interpreted the task to be more challenging when they thought the patients were treated with TADs, and their superimpositions may not have demonstrated the

effects they were expecting. Again, however, since the raters were not asked to evaluate their superimpositions, the true reason for the raters' perceived difficulty was not fully explored. Also, since the raters were under the impression they were being asked to focus on cephalometric reliability rather than treatment outcomes, their superimposition bias may have been underestimated.

What we did find in this study was significant variability, both within and between raters. While the summary data in Table 5 show only a handful of changes between the sessions, many more changes actually occurred, but cancelled each other out when they were totaled.

To put the variability in perspective, some examples were compiled. The first one assesses intra-rater variability between sessions. Using Session 1 (fixed appliances) as the baseline, we tallied the number of times each raters' superimpositions changed for each patient in Session 2 (**Table 8**). Forty-two (29%) of the 144 superimpositions judged to have no change in mandibular plane angle in Session 1 were judged differently in Session 2. The 42 patients that changed were divided nearly in half to show open or closed mandibular plane angles in Session 2. For patients' superimpositions that showed opening or closing in Session 1, 23% changed in the second session. Most commonly, these superimpositions changed to show no change in mandibular plane angle rather than the extreme opposite. However, in 6 patients, the superimposition changed from one extreme to the other between the two sessions.

	Session 1	Session 2				
	Session 1	# Changed	% Changed	MP Same	MP Opened	MP Closed
Same	144	42	29%	102	19	23
Opened	90	21	23%	17	69	4
Closed	126	29	23%	27	2	97

Table 8. Intra-rater variability of mandibular plane response between Session 1 and Session 2.

In the second example, the number of times a rater agreed with themselves over all three sessions for a particular patient was evaluated (**Table 9**). The highest percentage of agreement was 83% percent and the lowest was 43%, by each a graduate student and practitioner. The mean percent agreement was 63%, which is similar to the overall intra-rater agreement as well. For one rater, of the ten patients actually treated with clear aligners, they were in complete agreement with themselves for all three sessions.

	Graduate Students						Practitioners					
Fixed (n=10)	7	4	5	8	6	4	5	8	6	7	4	5
Aligner (n=10)	9	4	8	5	7	6	4	7	6	10	6	7
TADs (n=10)	9	5	6	7	7	6	4	7	7	8	7	7
Total (n=30)	25	13	19	20	20	16	13	22	19	25	17	19
% Agree	83%	43%	63%	67%	67%	53%	43%	73%	63%	83%	57%	63%

Table 9. Intra-rater agreement over all 3 sessions.

Two additional assessments of inter-rater variability were also performed. First, we assessed the number of patients who had at least one rater producing a superimposition showing mandibular plane angle opening, closing, or no change within the same session (**Table 10**). This occurred in each of the three session about 30%. In addition, we tallied the number of patients whose mandibular planes were superimposed to demonstrate the same treatment effect by all 12 raters within a session (**Table 10**). These were assessed in each session and as a total overall. Only 20% of the patients across all three sessions had superimpositions agreed upon by all 12 raters. The least agreement was in Session 2 (aligners) with 13% and the most agreement was in Session 1 (fixed) with 30%.

	Session 1	Session 2	Session 3	Total
Patient rated open/closed/no change by at least 1 rater (n, %)	8 (27%)	9 (30%)	8 (27%)	25 (28%)
All raters agree on MP response (n,%)	9 (30%)	4 (13%)	5 (17%)	18 (20%)

Table 10. Inter-rater variability and agreement

In general, it appeared that if patients had clear mandibular plane opening or closing, based on the consensus of three of the study investigators, participants also were most likely to produce superimpositions consistent with these changes. On the other hand, in patients that had minimal changes of the mandibular plane angle, again, based on the consensus of the study three investigators, variability could be in either the closing or opening direction amongst the participants. **Figures 3, 4, and 5** illustrate the variability that can be seen in resulting superimpositions for one patient from three different raters during a single session.

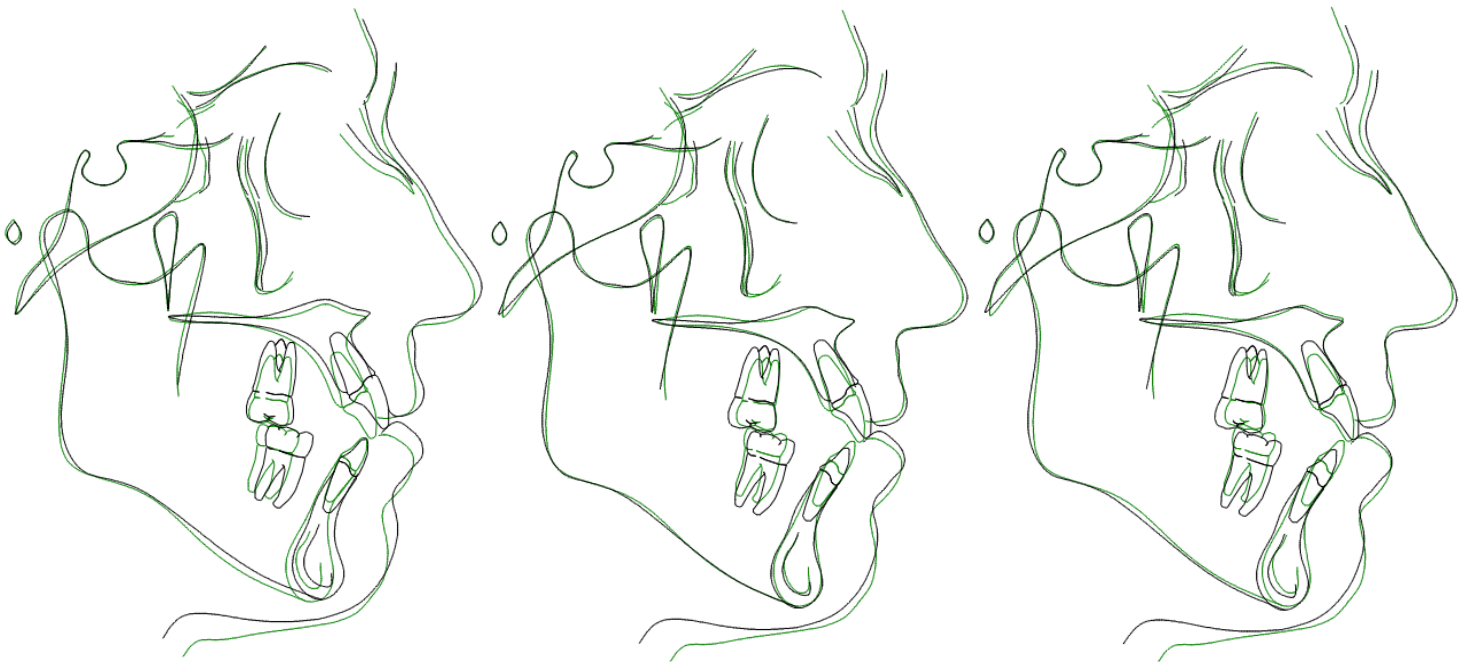


Figure 3. Session 1 superimpositions from three different raters showing mandibular plane angle opening, no change, and closing for the same patient.



Figure 5. Session 2 superimpositions from three different raters showing mandibular plane angle opening, no change, and closing for the same patient.

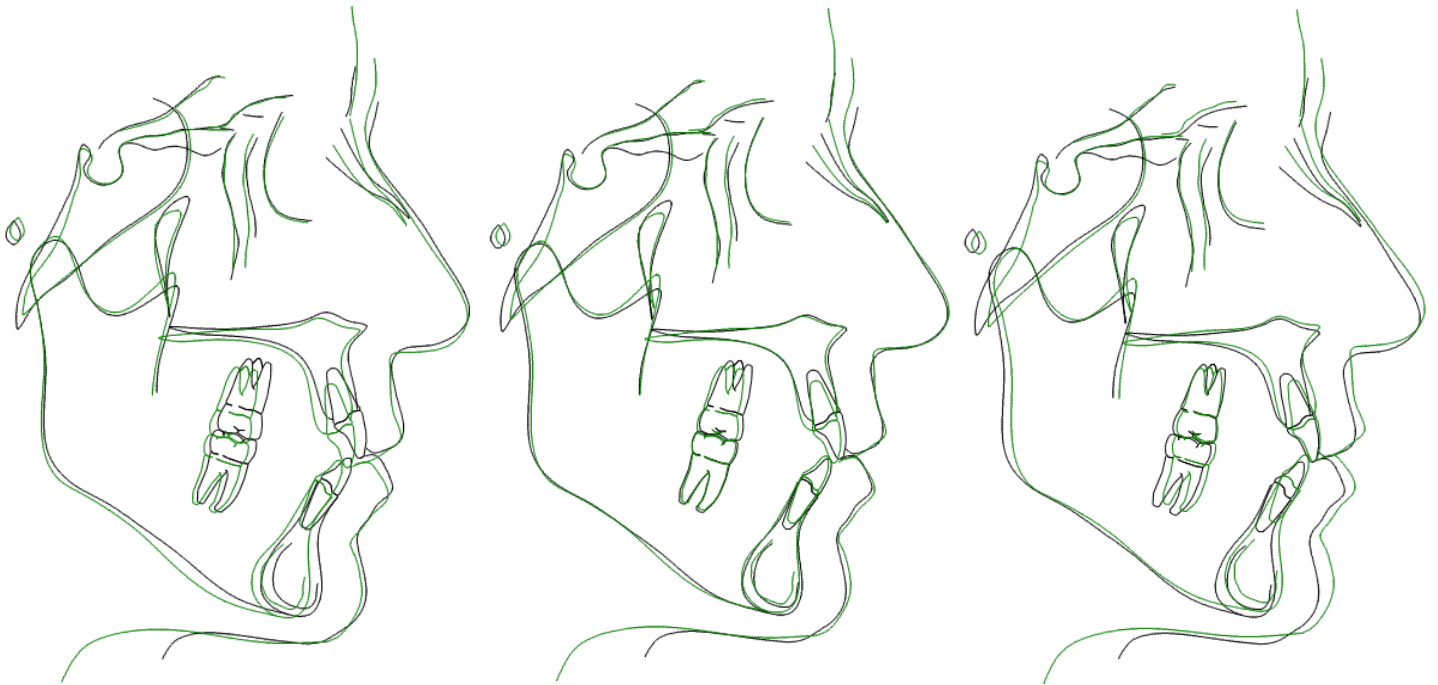


Figure 4. Session 3 superimpositions from three different raters showing mandibular plane angle opening, no change, and closing for the same patient.

Superimpositions allow orthodontists to identify areas in which change has occurred. Depending on the method used for the superimposition, these perceived changes may reflect differing dental or skeletal changes. Baumrind, et al. 1976 stated that visual comparisons of superimpositions are more sensitive than direct comparison of angular measurements due to the risk of measurements error with angular measurements. They are also an easy visual tool to show treatment outcomes.

The concept of the lateral cephalometric superimposition implies that a single, correct relationship exists. However, in reality, this “true” relationship may be impossible to know. The act of superimposing is an act of judgement, and even under the best conditions will always exhibit some degree of inaccuracy, and perhaps some degree of bias. Prior to the superimposition, random error can arise from head positioning during image acquisition, landmark identification, and tracing of the cephalometric image (Baumrind, et al. 1976).

The main errors seen in superimpositions are either rotational or translational. Rotational errors produce a larger portion of total error than translational errors. Baumrind, et al. 1976 and Gliddon, et al. 2006 both concluded that error will be increased for landmarks farther from the center of rotation. For example, a landmark 100 mm from the center of rotation with 1 degree of error will show 1.74 mm of displacement. The variability seen in the results of this study appear to demonstrate this trend as well. If sella is treated as the center of rotation, then menton, gonion, and pogonion are all the farthest landmarks and will therefore show the greatest variability in positional change due to rotational error (Baumrind, et al. 1976; Gliddon, et al. 2006). In turn, this rotational error can result in an apparent visible change in the mandibular plane angle, which may be used to justify that a particular technique is effective in producing its desired effect.

The effect size calculation for the lateral cephalometric measurements showed the greatest effect was on dental measurements, the largest being change in overbite across all categories. There were no moderate or large effect size outcomes for any skeletal measurement, including change in mandibular plane angle. These results support the idea that these appliances/treatment methods have little effect on the mandibular plane or skeletal components, but rather a greater effect on the dentition. In other words, non-surgical anterior open bite correction is largely achieved with incisor extrusion. This result has also been demonstrated in previous studies (Khosravi et al., 2017; Garnett et al., 2019; Alsafadi et al., 2016).

Our study did have some limitations. First, the groups did have some initial differences, with the TAD group having lower MP angles. Second, the raters were not asked to evaluate their superimpositions, and therefore, their subjective bias was not assessed. We employed deception to distract the raters from our true aims, but this may have also led to the participants focusing on trying to complete the task “correctly”. If the participants were asked to assess their superimpositions, this may have led to different results. Additionally, our sample only consisted of non-growing patients, so these results cannot be applied to growing patients. Finally, though our sample size had adequate power for our study, the individual treatment groups were too small to make conclusions that could be applied to all patients treated with fixed appliances, clear aligners, or TADs.

Though our study had some limitations, it also had some strengths. Only adult patients were used for the superimpositions, which removed the potential error and bias that may have been introduced with growth between the initial and end-of-treatment time points. Treatment effect size is greater in adolescent patients and will likely have skeletal change in addition to dental changes. Another advantage to this study is that the raters were blinded to the actual

treatment rendered and to the patients themselves. None of the raters treated these patients and were only exposed to the cephalometric tracings. Finally, this is one of the first studies to investigate operator bias in completing manual superimpositions.

In adults, if many patients truly have minimal changes in their mandibular plane angle during treatment, then our findings would indicate that manual superimpositions of many patients could easily be constructed to illustrate opening or closing of the mandibular plane. Studies with large samples, in which landmarks are identified and then measurements are automatically performed using cephalometric software, should theoretically result in less biased findings than manual superimpositions.

Conclusions

In conclusion, we did not observe a significant amount of superimposition bias in this study. However, several factors may have resulted in the bias of our raters not being fully expressed. We did observe considerable intra- and inter-rater variability in the superimpositions, indicating that orthodontists might use the same tracings and come to very different conclusions about the effects of treatment. The correction of overbite in this relatively small sample of adult openbite patients seemed to have been accomplished primarily by dental movements – eg, extrusion and retroclination of incisors.

Acknowledgements

This project was supported financially by the University of Washington Alumni Association. I would like to thank Greg Huang, Geoff Greenlee, and Brian Leroux for their collaboration as well as the graduate orthodontic students and practitioners that participated in this project.

References

- Alsafadi AS, Alabdullah MM, Saltaji H, Abdo A, Youssef M. Effect of molar intrusion with temporary anchorage devices in patients with anterior open bite: a systematic review. *Prog Orthod*. 2016 Mar 23;17:9.
- Aras A. Vertical changes following orthodontic extraction treatment in skeletal open bite subjects. *Eur J Orthod*. 2002 Aug;24(4):407–16.
- Baumrind S, Frantz RC. The reliability of head film measurements. *Am J Orthod*. 1971 Aug;60(2):111–27.
- Baumrind S, Miller D, Molthen R. The reliability of head film measurements. 3. Tracing superimposition. *Am J Orthod*. 1976 Dec;70(6):617–44.
- Bjork A, Bjork A, Skieller V. Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years [Internet]. Vol. 5, *The European Journal of Orthodontics*. 1983. p. 1–46. Available from: <http://dx.doi.org/10.1093/ejo/5.1.1>
- Björk A, Skieller V. Growth of the Maxilla in Three Dimensions as Revealed Radiographically by the Implant Method [Internet]. Vol. 4, *British Journal of Orthodontics*. 1977. p. 53–64. Available from: <http://dx.doi.org/10.1179/bjo.4.2.53>
- Boyd RL. Esthetic Orthodontic Treatment Using the Invisalign Appliance for Moderate to Complex Malocclusions. *Journal of Dental Education*. 2008 Aug;72(8):948–67.
- Boyd RL. Complex orthodontic treatment using a new protocol for the Invisalign appliance. *J Clin Orthod*. 2007 Sep;41(9):525–47; quiz 523.
- Buschang PH, LaPalme L, Tanguay R, Demirjian A. The technical reliability of superimposition on cranial base and mandibular structures. *Eur J Orthod*. 1986 Aug;8(3):152–6.
- Cangialosi TJ. Skeletal morphologic features of anterior open bite. *Am J Orthod*. 1984 Jan;85(1):28–36.
- Chang YI, Moon SC. Cephalometric evaluation of the anterior open bite treatment. *Am J Orthod Dentofacial Orthop*. 1999 Jan;115(1):29–38.
- Deguchi T, Kurosaka H, Oikawa H, Kuroda S, Takahashi I, Yamashiro T, et al. Comparison of orthodontic treatment outcomes in adults with skeletal open bite between conventional edgewise treatment and implant-anchored orthodontics. *Am J Orthod Dentofacial Orthop*. 2011 Apr;139(4 Suppl):S60–8.
- Doppel DM, Damon WM, Joondeph DR, Little RM. An investigation of maxillary superimposition techniques using metallic implants. *Am J Orthod Dentofacial Orthop*. 1994

Feb;105(2):161–8.

Enlow DH, Harris DB. A study of the postnatal growth of the human mandible. *Am J Orthod.* 1964 Jan 1;50(1):25–50.

Freitas BV, Abas Frazão MC, Dias L, Fernandes Dos Santos PC, Freitas HV, Bosio JA. Nonsurgical correction of a severe anterior open bite with mandibular molar intrusion using mini-implants and the multiloop edgewise archwire technique. *Am J Orthod Dentofacial Orthop.* 2018 Apr;153(4):577–87.

Garnett BS, Mahood K, Nguyen M, Al-Khateeb A, Liu S, Boyd R, et al. Cephalometric comparison of adult anterior open bite treatment using clear aligners and fixed appliances. *Angle Orthod.* 2019 Jan;89(1):3–9.

Gliddon MJ, Xia JJ, Gateno J, Wong HTF, Lasky RE, Teichgraeber JF, et al. The accuracy of cephalometric tracing superimposition. *J Oral Maxillofac Surg.* 2006 Feb;64(2):194–202.

Guarneri MP, Oliverio T, Silvestre I, Lombardo L, Siciliani G. Open bite treatment using clear aligners. *Angle Orthod.* 2013 Sep;83(5):913–9.

Janson G, Valarelli FP, Beltrão RTS, de Freitas MR, Henriques JFC. Stability of anterior open-bite extraction and nonextraction treatment in the permanent dentition. *Am J Orthod Dentofacial Orthop.* 2006 Jun;129(6):768–74.

Khosravi R, Cohananim B, Hujoel P, Daher S, Neal M, Liu W, et al. Management of overbite with the Invisalign appliance. *Am J Orthod Dentofacial Orthop.* 2017 Apr;151(4):691–9.e2.

Kim YH, Han UK, Lim DD, Serraon ML. Stability of anterior openbite correction with multiloop edgewise archwire therapy: A cephalometric follow-up study. *Am J Orthod Dentofacial Orthop.* 2000 Jul;118(1):43–54.

Kim YH. Anterior Openbite and its Treatment with Multiloop Edgewise Archwire. *Angle Orthod.* 1987 Oct;290–321.

Kuroda S, Sakai Y, Tamamura N, Deguchi T, Takano-Yamamoto T. Treatment of severe anterior open bite with skeletal anchorage in adults: comparison with orthognathic surgery outcomes. *Am J Orthod Dentofacial Orthop.* 2007 Nov;132(5):599–605.

Lopez-Gavito G, Wallen TR, Little RM, Joondeph DR. Anterior open-bite malocclusion: A longitudinal 10-year postretention evaluation of orthodontically treated patients. *Am J Orthod.* 1985 Mar;87(3):175–86.

Melsen B. The cranial base: The postnatal development of the cranial base studied histologically on human autopsy material: Birte Melsen *Acta Odontol. Scand.* 32:Supp. 62, 1974. *Am J Orthod.* 1974 Dec 1;66(6):689–91.

Nahoum HI, Horowitz SL, Benedicto EA. Varieties of anterior open-bite. *Am J Orthod.* 1972 May;61(5):486–92.

Nanda R UF. Management of open bite malocclusions. In: Nanda R, editor. *Biomechanical and esthetic strategies in clinical orthodontics.* Elsevier; 2005. p. 156.

Ng CST, Wong WKR, Hagg U. Orthodontic treatment of anterior open bite. *Int J Paediatr Dent.* 2008;18:78–83.

Nielsen IL. Maxillary superimposition: a comparison of three methods for cephalometric evaluation of growth and treatment change. *Am J Orthod Dentofacial Orthop.* 1989 May;95(5):422–31.

Rosenthal R. Interpersonal Expectancy Effects: A 30-Year Perspective. *Curr Dir Psychol Sci.* 1994;3(6):176–9.

Scheffler NR, Proffit WR, Phillips C. Outcomes and stability in patients with anterior open bite and long anterior face height treated with temporary anchorage devices and a maxillary intrusion splint. *Am J Orthod Dentofacial Orthop.* 2014 Nov;146(5):594–602.

Schupp W, Haubrich J, Neumann I. Treatment of anterior open bite with the Invisalign system. *J Clin Orthod.* 2010 Aug;44(8):501–7.

Sherwood KH, Burch JG, Thompson WJ. Closing anterior open bites by intruding molars with titanium miniplate anchorage. *Am J Orthod Dentofacial Orthop.* 2002 Dec;122(6):593–600.

Subtelny DJ, Sakuda M. Open-bite: Diagnosis and treatment. *Am J Orthod.* 1964 May;50(5):337–58.

	FIXED					ALIGNERS					TADS					OVERALL				
	Mean T1	STDV	Mean T2	STDV	T2-T1	Mean T1	STDV	Mean T2	STDV	T2-T1	Mean T1	STDV	Mean T2	STDV	T2-T1	Mean T1	STDV	Mean T2	STDV	T2-T1
SNA (°)	80.11	3.89	79.5	3.86	-0.61	78.81	4.31	77.63	4.19	-1.18	81.49	2.94	80.96	2.61	-0.53	80.14	3.80	79.36	3.76	-0.77
SNB (°)	76.35	4.35	76.83	4.75	0.48	74.87	3.71	74.17	3.51	-0.7	79.64	3.12	79.44	3.09	-0.2	76.95	4.16	76.81	4.31	-0.14
MP - SN (°)	39.78	7.82	39.52	8.58	-0.26	40.95	7.26	40.75	7.28	-0.2	33.17	6.62	32.42	6.86	-0.75	37.97	7.82	37.56	8.24	-0.40
Anterior Face Height (Name)	121.48	8.83	121.58	7.68	0.1	124.51	5.39	124.47	5.12	-0.04	121.73	6.28	121.49	5.99	-0.24	122.57	6.89	122.51	6.29	-0.06
Posterior Facial Height (mm)	45.74	8.00	46.11	8.68	0.37	45.08	5.08	44.69	5.41	-0.39	50.04	7.08	49.85	6.61	-0.19	46.95	6.96	46.88	7.13	-0.07
U1 - NA (°)	21.91	7.00	22.61	8.22	0.7	23.77	6.86	16.7	4.83	-7.07	25.93	10.92	22.07	6.63	-3.86	23.87	8.34	20.46	7.01	-3.41
U1 - NA (mm)	5.32	2.93	5.47	3.95	0.15	4.93	2.52	3.35	2.20	-1.58	5.67	3.96	5.17	2.59	-0.5	5.31	3.10	4.66	3.06	-0.64
L1 - NB (°)	32.52	7.75	28.6	6.48	-3.92	30.16	6.77	24.04	6.80	-6.12	27.62	6.61	20.26	5.82	-7.36	30.10	7.11	24.30	7.06	-5.80
L1 - NB (mm)	7.62	2.48	7.3	2.05	-0.32	7.35	2.15	6.2	2.54	-1.15	6.24	2.55	4.87	2.23	-1.37	7.07	2.40	6.12	2.42	-0.95
IMPA (L1-MP) (°)	96.39	8.71	92.28	3.16	-4.11	94.33	9.01	89.12	4.90	-5.21	94.78	6.95	88.4	5.05	-6.38	95.17	8.03	89.93	4.63	-5.23
Overbite (mm)	-2.43	0.94	0.79	0.68	3.22	-2.31	2.13	1.46	0.67	3.77	-1.58	1.63	1.96	0.76	3.54	-2.11	1.63	1.40	0.84	3.51
Overjet (mm)	2.67	2.15	2.17	0.68	-0.5	3.05	1.33	2.39	0.59	-0.66	1.9	1.80	2.51	0.67	0.61	2.54	1.80	2.36	0.64	-0.18
U6 - PP (UPDH) (mm)	24.18	3.58	24.46	2.80	0.28	25.23	2.87	25.39	2.66	0.16	25.48	3.47	25.41	3.42	-0.07	24.96	3.25	25.09	2.91	0.12
L6 - MP (UPDH) (mm)	33.85	3.52	34.27	3.61	0.42	34.72	3.40	34.51	2.82	-0.21	33.73	2.35	33.85	1.91	0.12	34.10	3.06	34.21	2.78	0.11
U1 - PP (UADH) (mm)	29.2	3.62	30.87	3.42	1.67	30.82	3.46	32.44	3.28	1.62	28.67	4.60	29.91	4.29	1.24	29.56	3.90	31.07	3.71	1.51
L1 - MP (LADH) (mm)	41.4	4.20	42.67	3.78	1.27	42.54	3.36	42.96	3.68	0.42	41.61	2.53	42.68	2.31	1.07	41.85	3.35	42.77	3.21	0.92

Appendix 1: Summary table of cephalometric measurements for all 30 patients. Grouped by treatment type and overall for each cep measurement.

Treatment type	Resident 1			Resident 2			Resident 3			Resident 4			Resident 5			Resident 6			Total per treatment type/Residents	
	Mandibular Plane Angle	R1 Session1 - Fixed	R1 Session2 - Aligners	R1 Session3- TADs	R2 Session1 - Fixed	R2 Session2 - Aligners	R2 Session3- TADs	R3 Session1 - Fixed	R3 Session2 - Aligners	R3 Session3- TADs	R4 Session1 - Fixed	R4 Session2 - Aligners	R4 Session3- TADs	R5 Session1 - Fixed	R5 Session2 - Aligners	R5 Session3- TADs	R6 Session1 - Fixed	R6 Session2 - Aligners		R6 Session3- TADs
Open	4	4	4	4	5	1	3	6	5	5	2	2	3	5	4	4	5	3	3	2
No Change	4	3	4	5	3	8	3	2	4	7	6	6	4	4	3	5	4	6	7	7
Closed	2	3	2	1	2	1	3	2	2	1	1	1	2	1	3	1	1	1	1	1
Open	0	0	0	0	0	0	1	2	2	0	0	0	2	2	2	1	1	0	0	0
No Change	6	5	5	6	5	6	4	4	4	6	4	3	2	2	3	2	4	8	7	7
Closed	4	5	5	4	5	4	5	4	4	4	6	7	5	5	7	2	5	2	3	3
Open	3	4	3	3	2	2	3	2	2	4	2	3	4	4	4	3	2	0	2	2
No Change	4	3	3	4	4	5	5	5	5	5	5	5	3	3	3	3	5	7	6	6
Closed	3	3	4	3	3	3	4	3	3	3	3	4	4	3	4	4	3	3	2	2
Total per session	7	8	7	7	7	3	5	10	10	11	4	3	6	11	10	8	8	3	3	4
No Change	14	11	12	15	10	19	13	11	11	18	15	12	12	9	9	10	13	21	20	20
Closed	9	11	11	8	10	8	12	9	8	8	12	12	10	11	11	12	9	6	6	6
Total per rater	22	22	22	17	47	26	26	35	29	13	45	32	29	28	33	15	54	21	15	122
Open	37	37	37	26	26	29	26	35	29	13	45	32	29	28	33	15	54	21	15	246
No Change	31	31	31	26	26	29	26	35	29	13	45	32	29	28	33	15	54	21	15	172
Closed	31	31	31	26	26	29	26	35	29	13	45	32	29	28	33	15	54	21	15	172

Treatment type	Faculty 1			Faculty 2			Faculty 3			Faculty 4			Faculty 5			Faculty 6			Total per treatment type/Faculty
	Mandibular Plane Angle	F1 Session1 - Fixed	F1 Session2 - Aligners	F1 Session3- TADs	F2 Session1 - Fixed	F2 Session2 - Aligners	F2 Session3- TADs	F3 Session1 - Fixed	F3 Session2 - Aligners	F3 Session3- TADs	F4 Session1 - Fixed	F4 Session2 - Aligners	F4 Session3- TADs	F5 Session1 - Fixed	F5 Session2 - Aligners	F5 Session3- TADs	F6 Session1 - Fixed	F6 Session2 - Aligners	
Open	4	4	4	4	4	3	5	5	4	3	3	3	5	4	4	4	4	7	4
No Change	2	4	5	4	4	6	2	4	5	5	6	5	4	3	3	2	5	2	4
Closed	4	2	1	2	2	1	3	1	1	2	2	1	2	2	1	1	1	1	2
Open	0	2	0	0	0	0	3	2	2	0	0	0	0	0	0	0	1	1	0
No Change	4	2	2	4	4	4	2	2	4	5	5	5	4	4	3	3	3	4	5
Closed	6	6	8	6	6	6	5	6	4	5	5	5	6	6	7	6	6	5	5
Open	4	1	4	3	4	3	4	3	3	3	2	2	2	3	3	3	3	4	3
No Change	1	6	1	3	3	3	2	4	2	3	4	4	5	4	4	4	4	2	2
Closed	5	3	5	4	4	2	4	3	5	4	4	3	3	3	4	3	4	4	5
Total per session	8	7	8	7	7	6	12	10	9	6	5	6	7	7	7	8	12	8	12
No Change	7	12	8	11	11	15	6	10	11	13	15	14	13	11	11	8	12	8	11
Closed	15	11	14	12	12	9	12	10	10	11	10	10	10	12	15	10	10	10	12
Total per rater	23	27	27	20	37	33	31	27	32	31	17	42	31	21	32	27	31	31	27
Open	40	40	40	33	33	32	31	27	32	31	17	42	31	21	32	27	31	31	27
No Change	40	40	40	33	33	32	31	27	32	31	17	42	31	21	32	27	31	31	27
Closed	40	40	40	33	33	32	31	27	32	31	17	42	31	21	32	27	31	31	27

Appendix 2: Counts table to summarize number of mandibular planes judges to be open, closed, no change.