Comparison of Outcomes with Non-extraction and Extraction in Borderline Class I and II Malocclusions

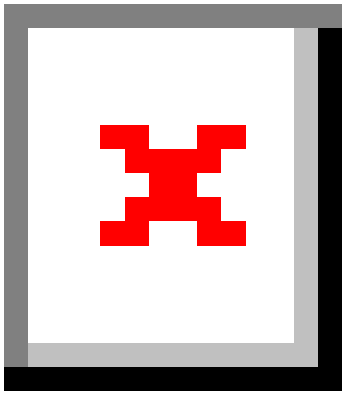
2023 Research Aid Awards (RAA)

Dr Wendy Vu

wendy.vu@mail.utoronto.ca
O: 403-400-4822

FollowUp Form

Award Information



In an attempt to make things a little easier for the reviewer who will read this report, please consider these two questions before this is sent for review:

- *Is this an example of your very best work, in that it provides sufficient explanation and justification, and is something otherwise worthy of publication? (We do publish the Final Report on our website, so this does need to be complete and polished.)*
- *Does this Final Report provide the level of detail, etc. that you would expect, if you were the reviewer?*

Title of Project:*

Comparison of Outcomes with Non-extraction and Extraction in Borderline Class I and II Malocclusions

Award Type

Research Aid Award (RAA)

Period of AAOF Support
July 1, 2023 through June 30, 2024

Institution
University of Toronto

Names of principal advisor(s) / mentor(s), co-investigator(s) and consultant(s)
John Daskalogiannakis, James Posluns

Amount of Funding
$4,579.60

Abstract
(add specific directions for each type here)

See attached.

Respond to the following questions:

Detailed results and inferences:*
If the work has been published, please attach a pdf of manuscript below by clicking "Upload a file".
OR
Use the text box below to describe in detail the results of your study. The intent is to share the knowledge you have generated with the AAOF and orthodontic community specifically and other who may benefit from your study. Table, Figures, Statistical Analysis, and interpretation of results should also be attached by clicking "Upload a file".

Figures, Tables, Statistics.pdf
1) Using Discriminant Analysis, a borderline sample of 30 non-extraction and 30 extraction Class I and mild Class II cases were isolated (Figure 1). Comparison of cephalometric and model values revealed the two groups were similar from the outset (Table 1). The different ethnicities of the samples were noted, and a Fisher's Exact Test showed there was no significant difference in the distribution of ethnicities in the extraction and non-extraction groups (p = .285). Certain cephalometric values were statistically different from the outset in the non-extraction and extraction borderline sample. The facial angle was more acute in the extraction cases by 2.26º, which may be accounted for by a measurement error. The A-point in our extraction cases was more retrusive when measured linearly in relation to the HP line through nasion by 2.14mm. This is reflective in the slightly lower values of the SNA angle in the extraction cases, although this was not statistically significant. The extraction cases also began with more vertical growth, as the Y-axis was increased by 3.15º. The FMA was also increased in the extraction cases, although not statistically significant. The maxillary incisors were slightly more protrusive in the extraction cases by 1.42mm, in relation to the NA

line. However, other angular measurements of maxillary incisors were not statistically significant between the extraction and non-extraction cases.

2) The treatment duration for the non-extraction group was 24.63 months ± 6.16 months, 4.9 months faster than the extraction group, which completed treatment in 29.53 months ± 9.94 months (Table 1).

3) Discriminant Analysis also identified maxillary crowding, overjet, lower incisor position, and facial angle as key factors influencing the decision to extract in Class I and mild Class II cases (Table 2). When Class I cases were considered in isolation, mandibular crowding replaced facial angle (Table 3).

4) Regarding occlusal outcome, the mean post-treatment overbite for the non-extraction cases was 1.55 ± 0.57mm, and 1.98 ± 0.66mm for extraction cases (Table 4). When compared to the ideal value of 2mm defined by Profitt et al., extraction cases statistically resulted in more ideal overbite; however, possibly clinically negligible, as this difference was less than 1.0mm. Examination of the histogram of distributions of post-treatment overbite values for non-extraction and extraction cases revealed extraction cases displayed a greater skew towards increased overbite compared to non-extraction (Figure 2). There were no differences between the treatment modalities in terms of overjet and molar occlusal relationship. Non-extraction resulted in more cases finishing with ideal premolar relationship, with the definition of ideal defined as values within 1mm of the Class I relationship according to the ABO-OGS (Table 5).

5) The pre-treatment and post-treatment soft tissue profile outlines of non-extraction and extraction cases were generated from cephalometric radiographs, and an online survey was conducted using a 100mm Visual Analog Scale (VAS) to obtain subjective preferences from orthodontists, general dentists, and laypeople (Figure 3). All raters preferred the post-treatment soft tissue profiles, with laypeople and general dentists favouring the non-extraction profiles by 1.90mm (95%CI [-0.81, 4.60]) and 3.68mm (95%CI [0.43, 6.92]) of a 100mm VAS, respectively, and orthodontists favouring the extraction profiles by 5.52mm (95%CI [1.64, 9.40]) of a 100mm VAS. When analyzed using a repeated measures regression model, this difference in preference for the orthodontists and general dentists were statistically significant (Table 6, Figure 4). The preferences of general dentists and laypeople were more similar to each other than to the orthodontists' preferences (Table 7, Figure 5).

6) Extraction treatment resulted in more retraction of lips compared to non-extraction treatment, with the extraction group resulting in retraction of the upper lips by 1.58mm; 95% CI [0.79, 2.37] and lower lips by 3.41mm; 95% CI [2.61, 4.21] in relation to the E-Line more than non-extraction group (Table 8). Extraction cases exhibited a statistically significant increase in the nasolabial angle by 3.55° ± 5.99°, while the non-extraction cases saw a statistically significant decrease of 2.67° ± 5.74° (Table 8). Linear regression of VAS and initial lip protrusion and nasolabial angle showed there is a significant correlation between pre-treatment upper and lower lip to E-Line and nasolabial angle with the VAS scores of the different types of respondents (Table 9). The correlation coefficient (slope) and intercept for the line of best fit were plotted separately for non-extraction and extraction cases and separated by type of respondents (Figure 6-8). When analyzing these plots, as the initial lip protrusion increases, there is a bend point (intersection) where extraction treatment resulted in higher VAS scores than non-extraction. For orthodontists, this bend-point was less (-5.31mm for the lower lip and -2.36mm for the upper lip) than for laypeople (-1.10mm for the lower lip and 1.42mm for the upper lip), and larger for general dentists (4.76mm for the upper lip and 0.96mm for the lower lip) indicating a much higher preference for more lip protrusion (Figure 6, Figure 7). Similarly, when analyzing the nasolabial angle, the bend-point for general dentists was 84.49°, for laypeople was 104.62°, and for orthodontists was 130.33°, indicating a higher tolerance for acute nasolabial angles for general dentists than for laypeople and orthodontists (Figure 8).

7) Assessment of rater preference for incisor inclination was obtained by altering a random borderline case using Invisalign's ClinCheck Web 1.4 software to simulate varying anterior incisor inclinations (Figure 9). The occlusal plane was used as a reference plane in these simulations due to the isolated view of the incisors. The range for U1-OP° and for L1-OP° was determined from the cephalometric analysis of the borderline sample and varied in 10° increments (Table 10). All rater groups preferred the appearance of more upright incisors and ranked the most proclined incisors as least esthetic, when assessed in isolation from the soft tissue and

face. The maxillary and mandibular incisor inclinations of 65º and 68º (to the occlusal plane) were ranked as most esthetic, by 81 respondents (out of 170 respondents) and second most esthetic by 73 respondents. The maxillary and mandibular inclinations of 75º and 78º were ranked second most esthetic by 72 respondents and most esthetic by 68 respondents. Most of the respondents (157 out of the 170) had ranked maxillary and mandibular incisor inclinations of 45º and 48º as least esthetic. The Kruskall-Wallis test was conducted to examine the differences between the rankings of the incisor inclination groups, which showed significance (Table 10, Figure 11). When subgrouping the type of raters, orthodontists, general dentists, and laypeople show similar trends.

Non-extraction treatment resulted in proclination and protrusion of the maxillary and mandibular incisors, while extraction treatment resulted in retroclination and retraction of the incisors. In the non-extraction group, the maxillary incisors exhibited a statistically significant increase in inclination of 2.38° ± 6.01° when measured from the palatal plane. In contrast, the extraction cases experienced a retroclination of the maxillary incisors by 7.69° ± 8.08° when measured from the palatal plane. In the non-extraction group, there was a significant increase in lower incisor inclination to mandibular plane of 4.67° ± 7.47°. Conversely, in the extraction group, the lower incisor inclination to mandibular plane angle was significantly reduced by 6.41° ± 8.33° (Table 12).

8) Conflicting preferences emerged among general dentists regarding the general desirability for more protrusive lips and acute nasolabial angles, typically achieved through non-extraction, while at the same time, preferring more upright incisors, typically achieved through extractions. The laypeople's preference towards non-extraction profiles were not statistically significant; however, their preferences were similar to that of general dentists.

Cohen's d effect values, which is a standardized mean difference, were calculated for the orthodontists' and general dentists' preferences for the extraction and non-extraction profiles, respectively, and their preferences for more upright incisors. The standardized mean difference was not calculated for laypeople, as their preference for non-extraction profiles was not statistically significant. The preference of the general dentists for the non-extraction profiles over the extraction profiles yielded a small effect (dcohen=0.35), and of the orthodontists for the extraction profiles over the non-extraction profiles yielded an intermediate effect (dcohen=0.44). The incisor rankings of the general dentists, orthodontists, and laypeople yielded a much larger effect, dcohen=2.70, dcohen=2.25, dcohen=3.24, respectively. These results indicate that general dentists and orthodontists prefer the more upright incisor inclinations more than their preference for the profiles.

The desire for more upright incisors was more pronounced than for a fuller profile, and these discrepancies in preferences highlight the importance of effective communication of treatment objectives and expectations between orthodontists, patients, and general dentists.

Were the original, specific aims of the proposal realized?*

Yes, the original specific aims of the proposal were realized.

Were the results published?*

No

Have the results of this proposal been presented?*
Yes

To what extent have you used, or how do you intend to use, AAOF funding to further your career?*
I have used the AAOF funding for this project, to ensure that the result and conclusions are statistically robust.
Hopefully, further residents can use this borderline sample for further research in this topic.

Accounting: Were there any leftover funds?
$2,800.00

Not Published

Are there plans to publish? If not, why not?*
Yes, there are plans to publish this work.

Comment: *The AAOF commends your execution and completion of this project and encourages you to pursue publication in the future. We look forward to your continued contributions and hope that you may seek support from the AAOF in the future.*

Presented

Please list titles, author or co-authors of these presentation/s, year and locations:*
Title: Evaluation of Profile and Incisor Inclination Preferences in Orthodontically Treated Borderline Non-extraction and Extraction Cases
Authors: Wendy Vu1, Dimitrios Konstantonis2, Natoosha Nargaski1, James Posluns1, John Daskalogiannakis1
1Orthodontics and Dentofacial Orthopedics, Faculty of Dentistry, University of Toronto, Canada
2Orthodontics, Faculty of Dentistry, National and Kapodistrian University of Athens, Greece
Year: 2024
Locations: EOS 2024

Was AAOF support acknowledged?
If so, please describe:

Yes, AAOF was acknowledged in the acknowledgement section of the poster.

Internal Review

Reviewer comments

Reviewer Status*

Approved

File Attachment Summary

Applicant File Uploads

- Figures, Tables, Statistics.pdf

Figures and Tables

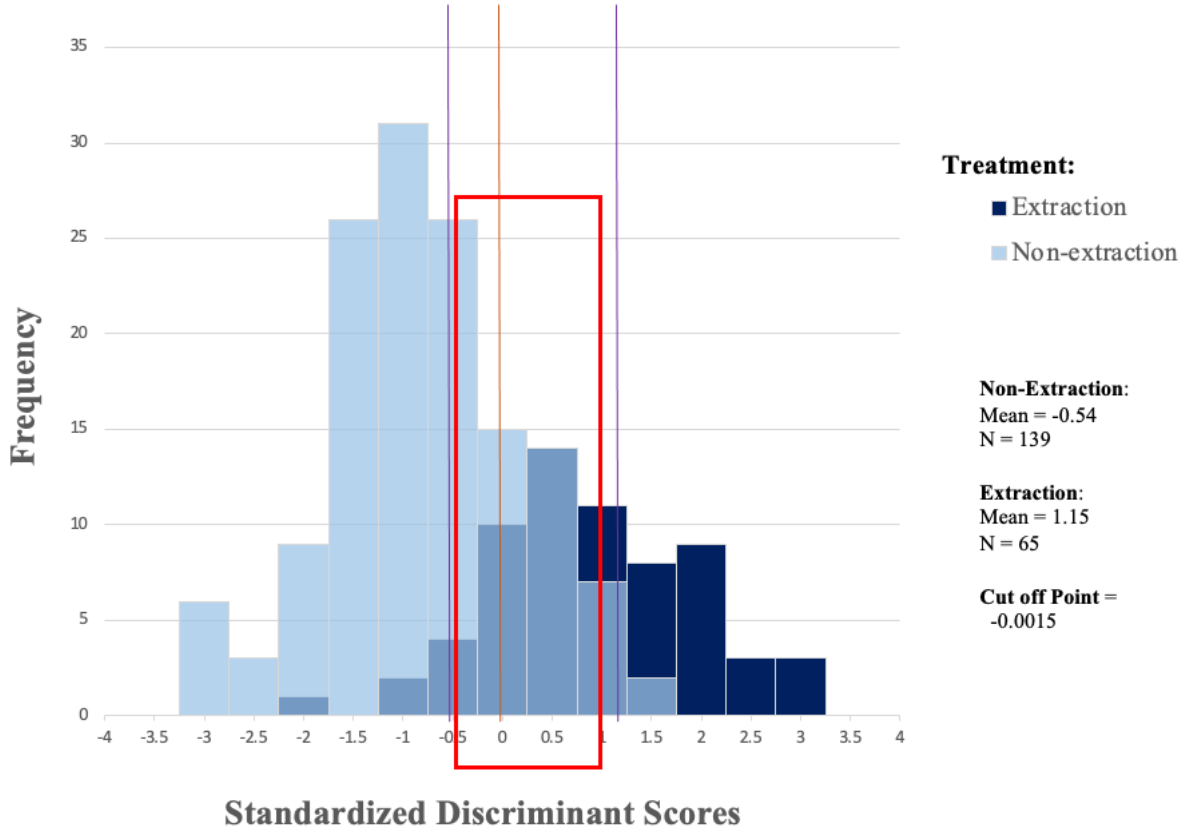


Figure 1. Histogram of standardized discriminant scores (SDS) of the 204 Class I and mild Class II cases treated at the University of Toronto. Purple lines denote the mean SDS of the non-extraction (light blue) and extraction (dark blue). Orange line denotes the weighted mean of the two centroids (or cut off point). Red box denotes the overlapping borderline extraction and non-extraction cases used.

Table 1. Descriptive Statistics and Pre-treatment Comparisons of the Extraction and Non-Extraction Borderline Sample. P-value based on independent sample t-test (2-sided p value).

	Non-Extraction		Extraction		p-value
Patient Demographics					
Age (Mean)	17.65 yrs (SD: 8.28 yrs) Min: 12.00 yrs; Max: 52.42 yrs		16.99 yrs (SD: 5.97 ys) Min: 11.25 yrs; Max: 42.94 yrs		.723
Male	14		14		
Female	16		16		
Malocclusion	Class I	Class II	Class I	Class II	
	19	11	18	12	
Ethnicity	Caucasian: n = 10 (33.3%) Black: n = 6 (20.0%) Asian: n = 2 (6.7%) Middle Eastern/ Other: n = 12 (40.0%)		Caucasian: n = 4 (13.3%) Black: n = 7 (23.3%) Asian: n = 5 (16.7%) Middle Eastern/ Other: n = 14 (46.7%)		
Extraction Pattern			U4/L4	U4/L5	
			22	8	
Treatment Duration (months)	Mean = 24.63	SD = 6.16	Mean = 29.53	SD = 9.94	**.025***
Model Values					
Maxillary Crowding	Mean	-3.91	-4.37		.524
	SD	2.83	2.72		
Mandibular Crowding	Mean	-5.85	-6.21		.568
	SD	2.29	2.60		
Overbite (%)	Mean	41.23	40.15		.823
	SD	19.48	17.97		
Overbite (mm)	Mean	3.05	3.01		.927
	SD	1.48	1.32		
Overjet (mm)	Mean	3.96	4.18		.616
	SD	1.55	1.77		
Right Molar (mm)	Mean	0.77	0.87		.642
	SD	0.78	0.87		
Right Premolar (mm)	Mean	1.70	2.20		.270
	SD	1.47	2.01		
Left Molar (mm)	Mean	0.82	0.75		.771
	SD	0.85	0.92		
Left Premolar (mm)	Mean	1.97	2.43		.145
	SD	1.19	1.26		
Cephalometric Values					
SNA(°)	Mean	82.47	80.87		.123
	SD	3.73	4.21		
SNB(°)	Mean	77.77	76.70		.284
	SD	3.43	4.15		
ANB(°)	Mean	4.70	4.17		.389
	SD	2.49	2.20		

Facial Angle (FH-NPg°)	Mean	87.59	85.33	**.009***
	SD	3.11	3.40	
N-A (HP) (mm)	Mean	-0.58	-2.64	**.029***
	SD	3.47	3.66	
N-B (HP) (mm)	Mean	-8.42	-10.35	.236
	SD	5.66	6.76	
N-Pg (HP) (mm)	Mean	-9.65	-11.57	.323
	SD	6.42	8.33	
Convexity Angle (N-A-Pg°)	Mean	9.35	8.24	.471
	SD	6.31	5.60	
Wits Appraisal (mm)	Mean	1.75	1.60	.832
	SD	2.52	2.81	
Mx Length (Co-A) (mm)	Mean	80.56	78.35	.096
	SD	5.74	4.24	
Md Length (Co-Gn) (mm)	Mean	104.83	102.92	.261
	SD	6.99	6.04	
Unit Length Diff (mm)	Mean	24.27	24.57	.781
	SD	4.40	4.01	
Upper Face Ht (N-ANS) (mm)	Mean	48.22	47.99	.805
	SD	3.40	3.79	
Lower Face Ht (ANS-Me) (mm)	Mean	62.45	62.6	.915
	SD	5.96	4.87	
Facial Height Ratio (N-ANS / ANS-Me)	Mean	77.55	76.87	.658
	SD	5.74	6.10	
SN-OP(°)	Mean	16.23	16.67	.707
	SD	4.26	4.70	
SN-GoGn(°)	Mean	34.21	35.82	.299
	SD	6.23	5.68	
Y-Axis (SGn-FH°)	Mean	60.51	63.66	**.004***
	SD	4.49	3.46	
FMA(°)	Mean	26.62	29.36	.068
	SD	6.64	4.59	
U1-PP(°)	Mean	112.48	114.69	.262
	SD	6.45	8.57	
U1-SN(°)	Mean	104.60	106.07	.478
	SD	5.72	9.74	
U1-NA(°)	Mean	22.12	25.21	.139
	SD	6.46	9.24	
U1-NA (mm)	Mean	4.66	6.08	**.039***
	SD	2.42	2.77	
U1-PP (mm)	Mean	27.17	27.47	.687
	SD	2.65	2.96	
U6-PP (mm)	Mean	21.22	21.88	.280
	SD	2.30	2.39	
Interincisal (U1-L1°)	Mean	122.67	119.54	.301
	SD	11.26	11.92	
IMPA(°)	Mean	96.36	96.45	.965

	SD	9.02	7.96	
L1-NB(°)	Mean	30.52	31.06	.795
	SD	8.60	7.56	
L1-Apg(°)	Mean	25.87	27.00	.494
	SD	6.26	6.48	
L1-NB (mm)	Mean	7.36	7.84	.560
	SD	3.30	3.03	
L1-Apg (mm)	Mean	4.54	5.33	.291
	SD	3.08	2.64	
L1-MP (mm)	Mean	39.42	39.28	.887
	SD	4.10	3.29	
L6-MP (mm)	Mean	29.00	28.73	.772
	SD	3.86	3.20	
Chin prominence (Pg-NB) (mm)	Mean	0.09	0.15	.908
	SD	2.17	1.56	
Holdaway Ratio (%)	Mean	0.38	0.08	.348
	SD	1.59	0.30	
Upper Lip to E-Line (mm)	Mean	-0.82	-0.46	.668
	SD	3.13	3.21	
Upper Lip to B-Line (mm)	Mean	5.25	5.75	.463
	SD	2.65	2.66	
Basic Upper Lip Thickness (mm)	Mean	12.81	12.88	.871
	SD	1.93	1.70	
Upper Lip Thickness (Vermillion) (mm)	Mean	11.35	11.28	.898
	SD	2.26	2.17	
Upper Lip Strain (mm)	Mean	1.46	1.61	.774
	SD	2.09	1.93	
Lower Lip to E-Line (mm)	Mean	1.41	2.10	.458
	SD	3.61	3.59	
Lower Lip to B-Line (mm)	Mean	5.07	5.88	.335
	SD	3.30	3.09	
Basic L-Lip Thickness (B-Point) (mm)	Mean	10.59	12.06	**.003***
	SD	1.58	2.02	
Lower Lip Thickness (Vermillion) (mm)	Mean	13.42	14.06	.298
	SD	1.97	2.71	
Lower Lip Strain	Mean	-2.83	-2.00	.178
	SD	2.24	2.47	
Nasolabial Angle (°)	Mean	108.86	105.85	.339
	SD	13.81	10.05	
Soft Tissue Profile (G'-Sn-Pg'°)	Mean	159.13	159.95	.617
	SD	6.94	5.68	
Soft Tissue Face Height	Mean	1.03	1.06	.433
	SD	0.10	0.13	

Table 2. Stepwise Discriminant Analysis of 204 Class I and II cases

Step	Variable	Wilks lambda	Standardized Canonical Coefficient	Significance
1	Maxillary Crowding (mm)	0.830	-1.098	**<.001***
2	Overjet (mm)	0.649	0.685	**<.001***
3	L1-Apg (mm)	0.628	0.664	**<.001***
4	Facial Angle (FH-NPg°)	0.614	0.263	**<.001***

Table 3. Stepwise Discriminant Analysis of 109 Class I cases alone

Step	Variable	Wilks lambda	Standardized Canonical Coefficient	Significance
1	Mandibular Crowding (mm)	0.854	-0.415	**<.001***
2	L1-Apg (mm)	0.706	0.826	**<.001***
3	Maxillary Crowding (mm)	0.679	-0.801	**<.001***
4	Overjet (mm)	0.613	0.598	**<.001***

Table 4. Comparison of the post-treatment values and the change in overbite and overjet values between non-extraction and extraction groups from model analysis. *P-value from one-sample t-test to compare post-treatment overjet and overbite to the 2mm ideal and paired sample t-test to compare pre-treatment and post-treatment. ** P-value from independent sample t-test to compare non-extraction with extraction.

	Non-Extraction			**Extraction**			
	Mean	**SD**	**p-value***	**Mean**	**SD**	**p-value***	**p-value****
Post-Tx Overbite (mm)	1.55	0.57	**<.001***	1.98	0.66	.891	**.009***
Post-Tx Overjet (mm)	2.91	0.66	**<.001***	2.67	0.63	**<.001***	.167
Post Tx – Pre-Tx OB (mm)	-1.50	1.59	**<.001***	-1.03	1.43	**<.001***	.238
Post-Tx – Pre Tx OJ (mm)	-1.05	1.64	**<.001***	-1.50	1.89	**<.001***	.324

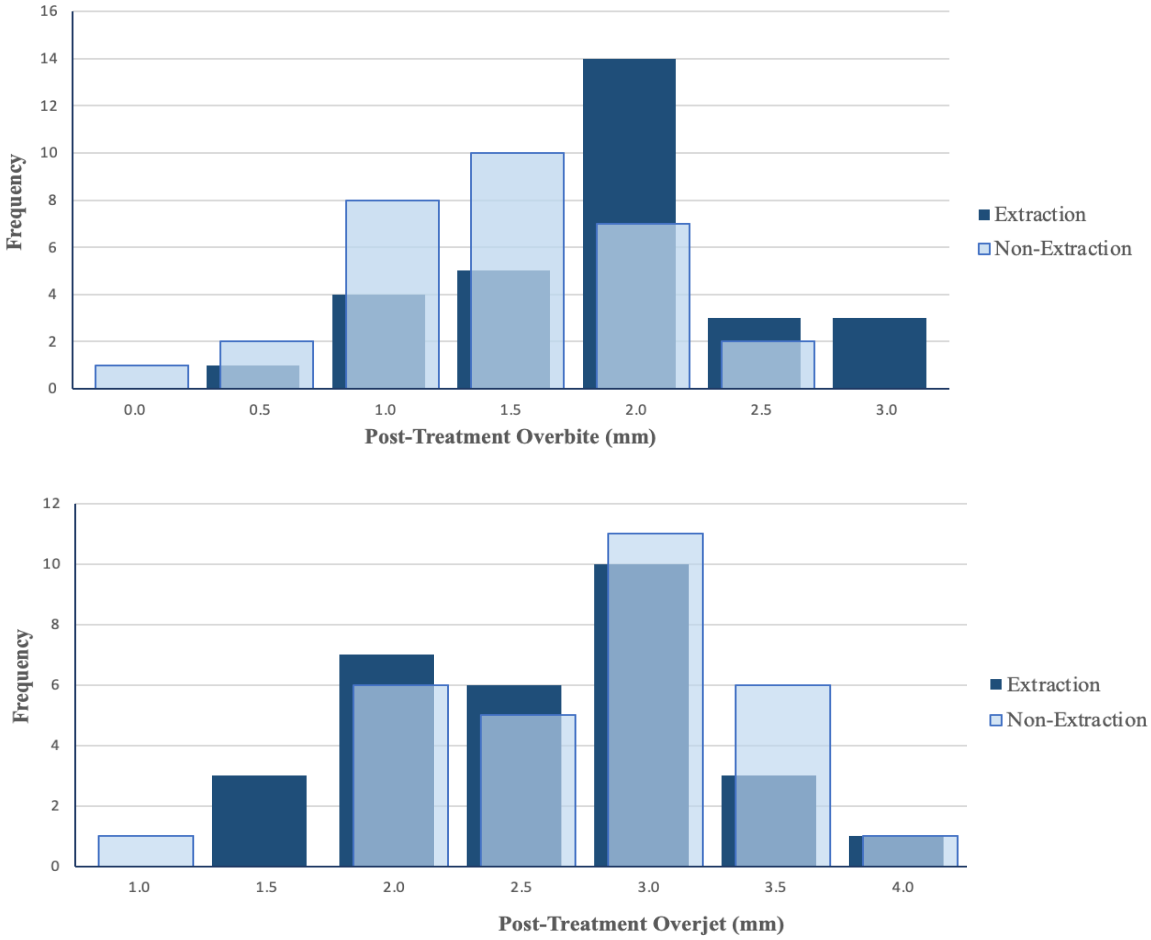


Figure 2. Histograms of the distribution of post-treatment overbite and overjet (in mm) for non-extraction and extraction cases.

Table 5. Numbers and percentages of cases where post-treatment molar and premolar relationships for left and right sides were ideal (within less than 1mm) or non-ideal (>1mm). * Chi-square test used when no cells have expected count of less than 5. If <5, Fisher's Exact Test used. P-value based on 2-sided significance.

		Non-Extraction		Extraction			
		n	**%**	**n**	**%**	X^2*	**Sig.**
Right Molar	**Ideal (within <1mm)**	**21**	**70.0%**	**23**	**76.7%**	.341	.771
	Non-Ideal (>1mm)	9	30.0%	7	23.3%		
Right Premolar	**Ideal (within <1mm)**	**13**	**56.7%**	**2**	**6.7%**	**10.756**	**.002***
	Non-Ideal (>1mm)	17	43.3%	28	93.3%		
Left Molar	**Ideal (within <1mm)**	**19**	**63.3%**	**19**	**63.3%**	.000	1.000
	Non-Ideal (>1mm)	11	36.7%	11	36.7%		
Left Premolar	**Ideal (within <1mm)**	**11**	**36.7%**	**3**	**10.0%**	**5.963**	**.030***
	Non-Ideal (>1mm)	19	63.3%	27	90.0%		

Table 6. Mean visual analog scores (VAS) in mm (with standard deviation and standard error) for non-extraction and extraction profile pairings of orthodontists, general dentists, and laypeople.

	Non-Extraction			Extraction				
Group	**Mean**	**SD**	**SE**	**Mean**	**SD**	**SE**	**Mean Diff (95% CI)**	**p-value**
Orthodontists	3.24	40.97	1.31	8.76	42.80	1.57	5.52 (1,64,9.40)	**.005***
General Dentists	8.19	45.63	1.29	4.51	43.75	1.54	3.68 (.43,6.92)	**.027***
Laypeople	6.29	46.39	0.98	4.39	47.38	1.13	1.90 (-.81,4.60)	.169

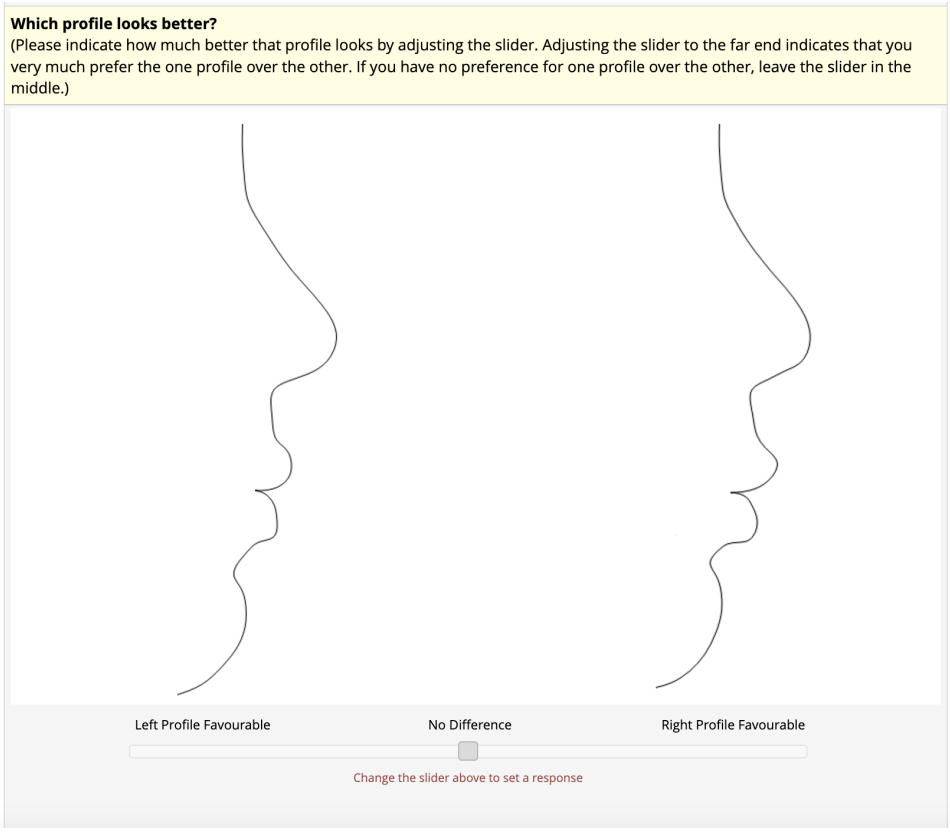


Figure 3. Sample of VAS from survey distributed to orthodontists, general dentists, and laypeople to gain their perspective on facial profile esthetics. Each set of profile consisted of a pre-treatment and post-treatment tracing (generated from cephalometric radiographs, using the horizontal plane (SN-7º) as the reference) of the same patient, randomized.

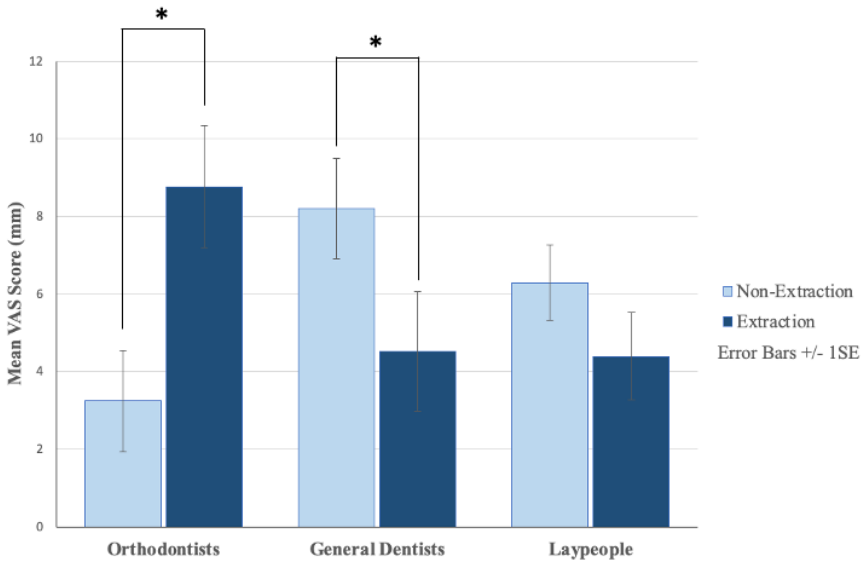


Figure 4. Graphical representation of the mean VAS scores (in mm) for extraction and non-extraction profile pairings for orthodontists, general dentists, and laypeople. *p < 0.05

Table 7. Comparison of VAS between different groups of raters for non-extraction and extraction profile pairings. P-values (2-sided) derived from a repeated measures regression model.

		Non-Extraction	**Extraction**
		p-value	**p-value**
Laypeople	**General Dentist**	.240	.949
	Orthodontist	**.061**	**.024***
General Dentist	**Laypeople**	.240	.949
	Orthodontist	**.007***	**.053**
Orthodontist	**Laypeople**	**.061**	**.024***
	General Dentist	**.007***	**.053**

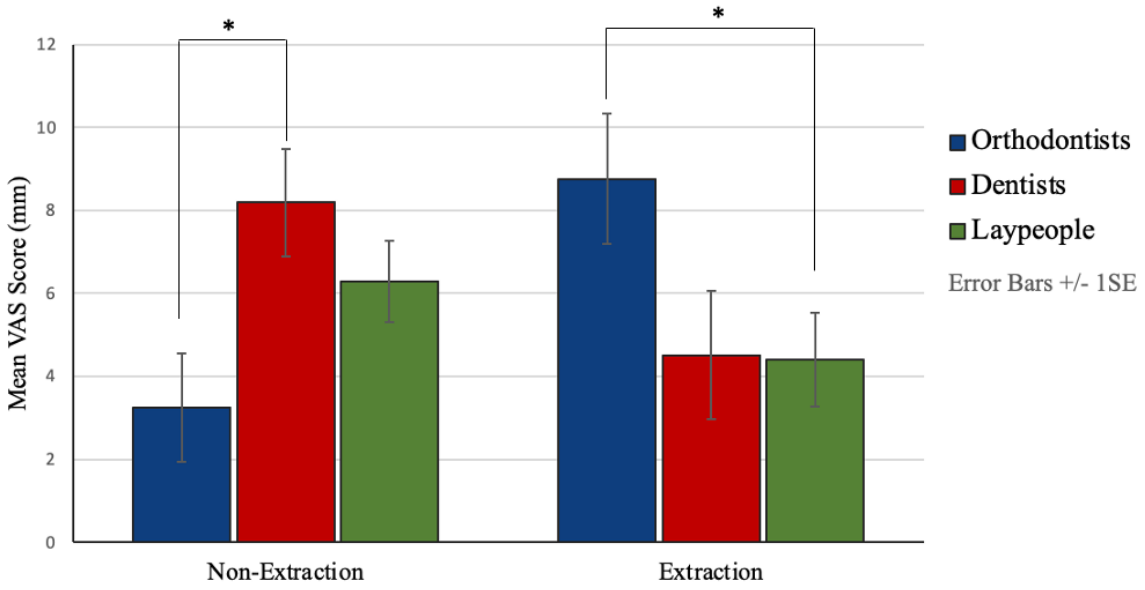


Figure 5. Graphical representation of the comparison of different rater groups preference for non-extraction and extraction profiles. *p < 0.05

Table 8. Mean difference between pre-treatment and post-treatment values in soft tissue cephalometric parameters in non-extraction and extraction cases. *P-value (2-sided) derived from paired sample T-test; **P-value (2-sided) from independent sample T-test.

	Non-Extraction			Extraction				
	Post-Tx Mean (SD)	**Post-Tx – Pre-Tx Mean (SD)**	**p-value***	**Post-Tx Mean (SD)**	**Post-Tx – Pre-Tx Mean (SD)**	**p-value***	**Mean Difference NE-Ex (95%CI)**	**p-value****
Nasolabial Angle (°)	106.19 (12.51)	-2.67 (5.74)	**.016***	109.40 (8.60)	3.55 (5.99)	**.003***	6.22 (3.19,9.25)	**<.001***
Soft Tissue Angle (°)	159.49 (7.58)	0.36 (2.17)	.366	161.61 (5.34)	1.66 (2.18)	**<.001***	1.30 (0.18,2.42)	**.024***
Upper Lip to E-Line (mm)	-1.69 (3.25)	-0.52 (1.46)	.076	-2.36 (2.61)	-2.10 (1.44)	**<.001***	1.58 (0.79,2.37)	**<.001***
Upper Lip to B-Line (mm)	5.13 (2.56)	0.15 (1.12)	.486	4.46 (2.04)	-1.41 (1.26)	**<.001***	1.56 (0.92,2.21)	**<.001***
Lower Lip to E-Line (mm)	1.20 (3.28)	0.28 (1.37)	.295	-0.70 (2.75)	-3.13 (1.55)	**<.001***	3.41 (2.61,4.21)	**<.001***
Lower Lip to B-Line (mm)	5.33 (2.92)	0.65 (1.13)	**.006***	3.64 (2.16)	-2.49 (1.50)	**<.001***	3.13 (2.41,3.86)	**<.001***

Table 9. Correlation of VAS scores (mm) and pre-treatment upper and lower lip to E-Line. Full table of other soft tissue parameters can be found in Appendix. Pearson correlation coefficient is denoted by r. *P-value (2-sided) for Pearson's correlation; **P-value (2-sided) for repeated measured regression model.

	Ortho			GP			Laypeople		
	r	sig*	sig**	r	sig*	sig**	r	sig*	sig**
Combined									
Pre-Tx Upper Lip to E-Line (mm)	.070	**.006***	**.002***	.048	.063	.095	.091	**<.001***	**<.001***
Pre-Tx Lower Lip to E-Line (mm)	.095	**<.001***	**<.001***	.044	.084	.097	.087	**<.001***	**<.001***
Pre-Tx Nasolabial Angle (°)	-.173	**<.001**	**<.001**	-.085	**<.001**	**.009**	-.102	**<.001**	**<.001**
Non-Extractions									
Pre-Tx Upper Lip to E-Line (mm)	.045	.212	.131	.007	.850	.840	.035	.149	.098
Pre-Tx Lower Lip to E-Line (mm)	.061	.091	.069	.020	.587	.493	.013	.594	.532
Pre-Tx Nasolabial Angle (°)	-.142	**<.001**	**<.001**	-.056	.124	.099	-.037	.121	.132
Extractions									
Pre-Tx Upper Lip to E-Line (mm)	.132	**<.001***	**<.001***	.082	**.024***	**.055***	.152	**<.001***	**<.001***
Pre-Tx Lower Lip to E-Line (mm)	.158	**<.001***	**<.001***	.069	.059	.138	.179	**<.001***	**<.001***
Pre-Tx Nasolabial Angle (°)	-.199	**<.001**	**<.001**	-.118	**.001**	**.005**	-.165	**<.001**	**<.001**

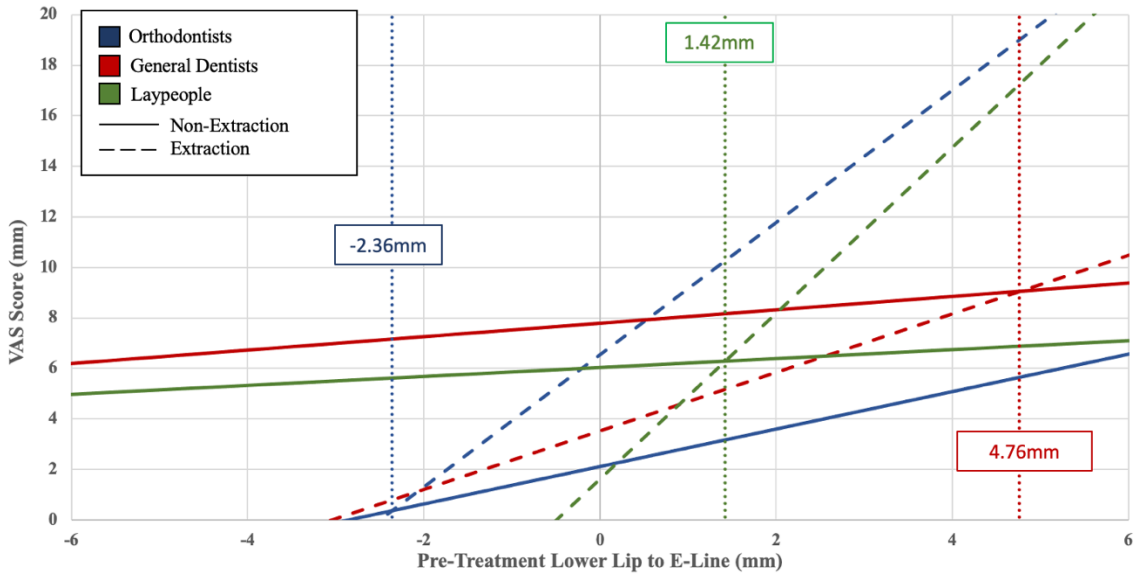


Figure 6. Plots of pre-treatment lower lip to E-Line (mm) against the VAS scores (mm) of orthodontists, general dentists, and laypeople. Bend-point where the non-extraction and extraction lines intersect are highlighted.

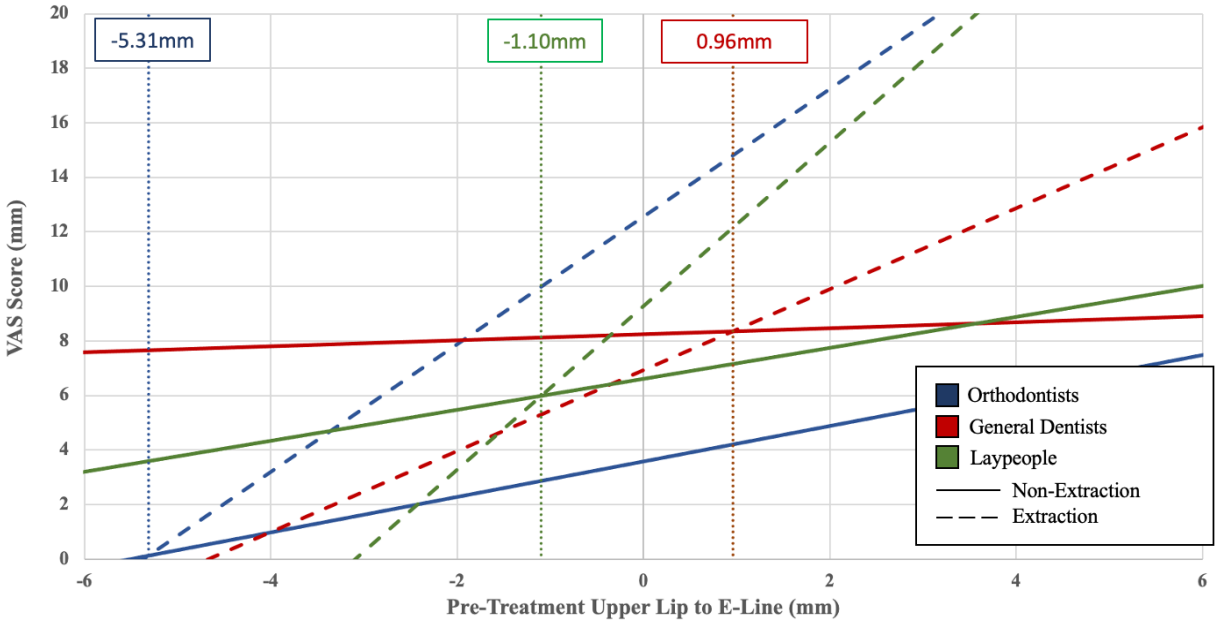


Figure 7. Plots of pre-treatment upper lip to E-Line (mm) against the VAS scores (mm) of orthodontists, general dentists, and laypeople. Bend-points where the non-extraction and extraction lines intersect are highlighted.

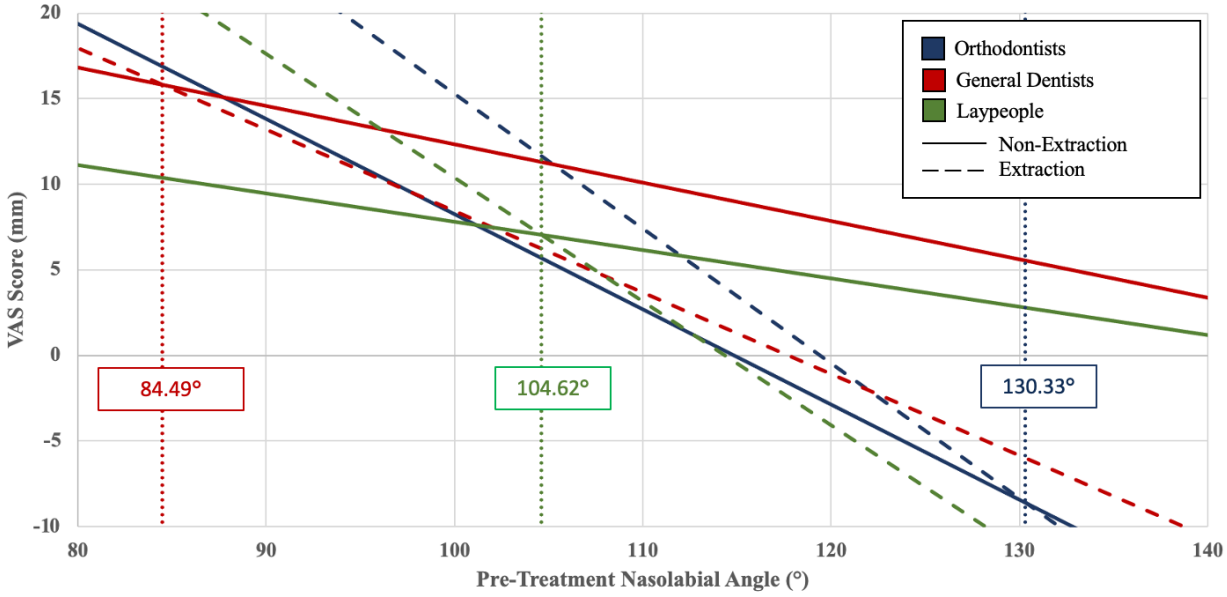


Figure 8. Plots of pre-treatment nasolabial angle (°) against the VAS scores (mm) of orthodontists, general dentists, and laypeople. Bend-points where the non-extraction and extraction lines intersect are highlighted.

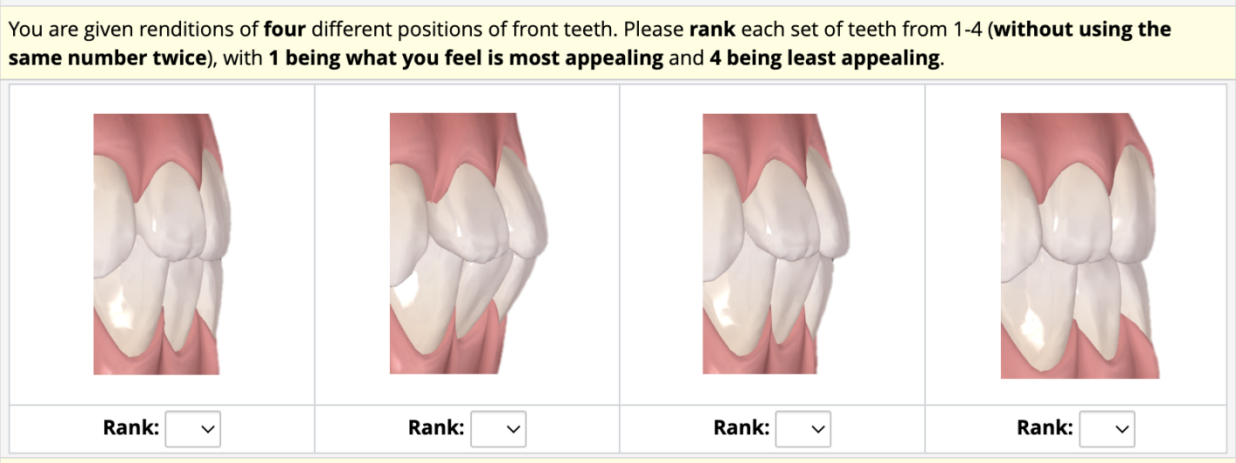


Figure 9. Final question on survey prompting the rater to indicate their preference of various incisor inclinations. From left to right (U1-OP°/L1-OP°): 65°/68°, 45°/48°, 55°/58°, 75°/78°.

Table 10. Mean, median, maximum and minimum values for U1-OP° and L1-OP° for non-extraction and extraction cases. These values were used to generate the ranges of incisor inclinations shown in the final question of the survey.

	Non-Extraction		Extraction	
	U1-OP°	**L1-OP°**	**U1-OP°**	**L1-OP°**
Mean	55.83	60.20	63.94	70.70
SD	6.39	6.43	5.98	7.11
Median	55.65	59.95	65.35	69.90
Minimum	44.40	48.60	51.60	54.0
Maximum	69.40	73.50	75.80	87.40

* Independent t-test performed comparing means of U1-OP° and L1-OP° for non-extraction and extraction cases, yielding p<.001 for both values.

Table 11. Frequency and percentage of all respondents' rankings (from 1 to 4, with 1 being most esthetic) of various incisor inclinations ranging from 45° -78°, in 10° increments.

	Rank 1		Rank 2		Rank 3		Rank 4	
U1-OP°/L1-OP°	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
45°/48°	0	0.0	3	1.8	10	5.9	157	92.4
55°/58°	23	13.5	20	11.8	121	71.2	6	3.5
65°/68°	81	47.6	73	42.9	12	7.1	4	2.4
75°/78°	68	40.0	72	42.4	27	15.9	3	1.8

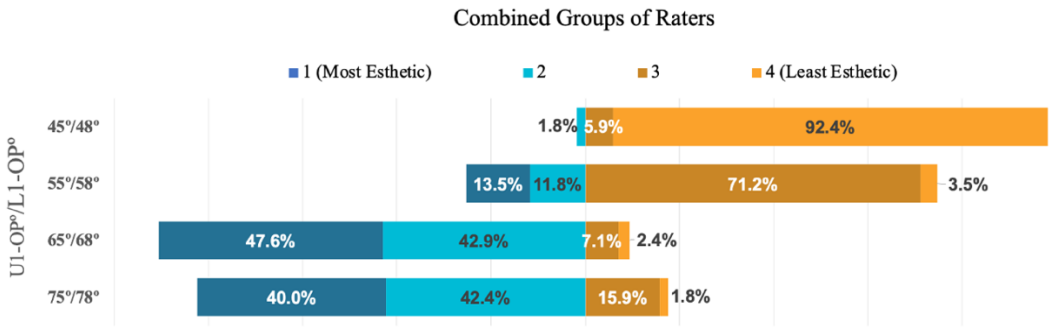


Figure 10. Graphical representation of the combined respondents' rankings (from 1-4, with 1 being the most esthetic) of various incisor inclinations, ranging from 45º -78º, in 10º increments. Kruskall-Wallis test: H(3)=437.12, P<.001.

Table 12. Mean difference between pre-treatment and post-treatment values in maxillary and mandibular incisor cephalometric parameters in non-extraction and extraction cases. *P-value (2-sided) derived from paired sample T-test; **P-value (2-sided) from independent sample T-test.

	Non-Extraction			Extraction				
	Post-Tx Mean (SD)	**Post-Tx – Pre-Tx Mean (SD)**	**p-value***	**Post-Tx Mean (SD)**	**Post-Tx – Pre-Tx Mean (SD)**	**p-value***	**Mean Difference (NE-Ex) (95% CI)**	**p-value****
U1-PP°	114.86 (8.26)	2.38 (6.01)	**.039***	107.01 (7.64)	-7.69 (8.08)	**<.001***	10.07 (6.39,13.75)	**<.001***
U1-SN°	106.55 (7.70)	1.95 (5.99)	.085	98.18 (8.68)	-7.89 (7.59)	**<.001***	9.84 (6.31,13.37)	**<.001***
U1-NA°	24.78 (8.75)	2.66 (6.11)	**.024***	17.76 (8.41)	-7.45 (8.13)	**<.001***	10.11 (6.39,13.83)	**<.001***
U1-NA (mm)	5.20 (3.18)	0.47 (2.13)	.261	1.69 (2.45)	-4.10 (2.65)	**<.001***	4.57 (3.26,5.89)	**<.001***
U1-PP (mm)	27.17 (2.64)	0.33 (1.07)	.120	27.67 (3.10)	-0.19 (1.65)	.565	0.51 (-0.24,1.27)	.180
Interincisal Angle U1-L1°	116.03 (8.93)	-6.64 (9.42)	**<.001**	134.63 (8.88)	15.09 (10.84)	**<.001***	21.73 (16.48, 26.98)	**<.001***
IMPA°	101.03 (8.36)	4.67 (7.47)	**.002***	90.05 (5.84)	-6.41 (8.33)	**<.001***	11.08 (6.99,15.17)	**<.001***
L1-NB°	34.94 (6.66)	4.43 (6.88)	**.001***	23.94 (5.99)	-7.12 (8.22)	**<.001***	11.55 (7.63,15.47)	**<.001***
L1-APg°	30.90 (5.22)	5.03 (7.57)	**.001***	21.28 (5.67)	-5.72 (8.59)	**.001***	10.76 (6.57,14.95)	**<.001***

L1-NB (mm)	8.04 (2.38)	1.03 (1.55)	**.002***	4.59 (2.40)	-3.46 (1.80)	**<.001***	4.49 (3.57,5.41)	**<.001***
L1-APg (mm)	5.52 (2.48)	1.28 (1.77)	**<.001***	2.03 (1.88)	-3.46 (2.07)	**<.001***	4.74 (3.69,5.79)	**<.001***
L1-MP (mm)	40.04 (4.46)	0.86 (1.64)	**.011***	38.26 (3.15)	-1.01 (1.88)	**.010***	1.88 (0.91,2.84)	**<.001***

Statistical Analyses:

Sample Size and Power Analysis

A power analysis and sample size calculation were conducted with the study objective of having sufficient number of raters in each group to report significant differences in preference between the non-extraction and extraction groups. Using the meta-analysis published by Konstantonis *et al.* (2008), the effect size (standardized mean difference) was reported as 0.73 for dentists and 0.29 for laypeople[9]. Based on paired sample t-test (assuming that each group is presented with equal number of non-extraction and extraction cases), the sample size of 14 dentists, 14 orthodontists, and 75 laypeople were needed with a power of 80% and level of significance of 0.05.

Comparison of Pre-treatment Parameters

To examine homogeneity between the two treatment groups, age, pre-treatment cephalometric values, and model values was analyzed with an independent t-test. The effects of sex were analyzed via cross tabulations with a chi-square test.

Comparison of Pre-Treatment and Post-Treatment Overjet, Overbite, and Occlusal Relationships in Non-Extraction and Extraction Cases

To examine the difference in pre-treatment and post-treatment values for overjet, overbite, and occlusal relationships, a paired-sample t-test was utilized. The post-treatment overjet and overbite was assessed with a one-sample t-test to determine deviation from the ideal overjet and overbite of 2mm, as defined by Proffit *et al*[1]. A histogram of overjet and overbite for non-extraction and extraction cases was used to analyze trends in distribution of values. For post-treatment molar and premolar occlusal relationships, the number of cases that lie within the ideal deviation of less than 1mm was compared with the number of cases that did not end up ideal, as defined in the ABO-OGS, was analyzed via cross tabulations in a chi-square and Fisher's Exact test.

Comparison of Pre-Treatment and Post-Treatment Soft Tissue and Incisal Measurements in Non-Extraction and Extraction Cases

Pre-treatment and post-treatment soft tissue and incisal measurements derived from cephalometric radiographs were compared using a paired-sample t-test. The comparison of mean differences of these values between non-extraction and extraction groups were completed using independent t-test. The Class I and Class II cases were also stratified into subgroups.

Comparison of Esthetic Scores Between Treatment Groups and Between Raters

General equation estimate regression models were used to account for repeated measures of esthetic scores by the raters while comparing the esthetic scores between treatment groups and groups of raters. Stratification of age, sex, years of dental experience, and history of orthodontic treatment were assessed as subgroups and as interactions in the regression model.

Correlation of Esthetic Scores to Soft Tissue Measurements

General equation estimate regression models were used to account for repeated measures of esthetic scores by the raters to correlate the esthetic scores to the soft tissue measurements. The soft tissue measurements that most correlated with the esthetic scores were assessed in the regression model. The slope and intercept from the line of best fit from the data was plotted as esthetic score against pre-treatment upper and lower lip to E-Line and nasolabial angle values for non-extraction and extraction and divided according to rater groups.

Comparison of Incisor Rankings

Kruskall Wallis test and Post-hoc analysis (with Bonferroni correction) were used to compare the rankings of the different incisor inclinations.

Comparison of Effect of Mean Difference in Profile VAS and Incisor Rankings

Cohen’s d-value (standardized mean difference) was calculated to compare the effect of the mean difference in VAS and mean difference in incisor inclination.

Inter-Rater Reliability, Intra-Rater Reliability Tests, and Error of Measurement

To assess the validity of the pre-treatment cephalometric and model analysis values given by individual residents at the time of treatment planning, three examiners (WV, MM, ME) after a calibration session, independently scored 20 cases randomly. Four weeks later, the 20 cases were re-scored by a single investigator (WV).

An Intraclass Correlation coefficient (ICC) was determined and interpreted as follows[61,62]:

0.81 – 1.0 = Excellent or Very Good Agreement

0.61-0.80 = Substantial or Good Agreement

0.41 to 0.60 = Moderate Agreement

0.21 to 0.40 = Fair Agreement

0.0 to 0.20 = Poor Agreement

The error standard deviation of the 20 cases that were re-scored by the single investigator (WV) was assessed using Dahlberg's formula: $\sqrt{(\sum d^2/2N)}$, where d is the difference between the two measurements and N is the number of cases assessed.