

Validation of the Eyelash Satisfaction Follow-up Questionnaire (ESFQ) for Follow-up Self-Assessment of Eyelash Satisfaction

Somali M. Burgess, PhD¹; John G. Walt, MBA¹; Min Yang, MD, PhD²; Jan E. Hansen, PhD¹; Frederick C. Beddingfield III, MD, PhD^{1,3}; Geoffrey C. Hammond², PhD; Jason C. Cole, PhD⁴

¹Allergan, Inc., Irvine, CA; ²QualityMetric, Lincoln, RI; ³David Geffen School of Medicine at UCLA, Los Angeles, CA; ⁴Independent Consultant, Torrance, CA

Abstract

Purpose: The Eyelash Satisfaction Questionnaire (ESQ), a static measure, has been previously validated for the assessment of eyelash-specific patient-reported outcomes (PROs) in subjects receiving a therapeutic treatment for hypotrichosis of eyelashes. The current study aimed to validate a similar questionnaire, the Eyelash Satisfaction Follow-up Questionnaire (ESFQ), a 39-item dynamic measure designed specifically to self-assess eyelash-specific PROs.

Methods: The ESFQ was initially examined in a 909-person validation sample consisting of online respondents. Confirmatory factor analysis (CFA) was initially used to test the measurement structure of the questionnaire. Item- and scale-level psychometric properties such as item-total correlations, internal consistency, and convergent and discriminant validities were reviewed. The ESFQ was then used in a clinical population receiving bimatoprost ophthalmic solution 0.03%, a product that improves eyelash prominence, during a 16-week randomized, controlled, masked, clinical trial.

Results: Initial CFA results revealed a “good fit” based on hypothesized similarity with the factor structure identified in the ESQ. A model using 9 indicators (3 per factor), similar to the structure of the ESQ, was chosen as the optimal method. The final model showed good internal consistency along with convergent and discriminant validities. However, the alpha-removed statistic for the length, fullness, and overall satisfaction (LFOS) construct showed a significant increase in scale consistency with the removal of item 8 (Compared to your first visit, overall, how satisfied are you with your eyelashes now?). Despite this, the model structure was retained to better facilitate comparisons between the ESQ and ESFQ and to provide sufficient measures to properly evaluate the construct.

Conclusions: The 9-item ESFQ appears to be a valid measure to assess changes in eyelash-specific PROs. The factor structure showed near equivalence with the ESQ. Future research will involve further validation of the ESFQ in response to variations in hypotrichosis etiologies and clinical treatment paradigms.

Introduction

- Eyelash prominence has been described as having a positive psychological effect^{1,2}
 - Prominent eyelashes are widely considered to be a desirable physical attribute
 - Eyelash prominence is defined by length, thickness or fullness, and darkness based on a validated measure³
- LATISSE® (bimatoprost ophthalmic solution) 0.03%; Allergan, Inc. Irvine, CA) is the first treatment approved in the United States for hypotrichosis of the eyelashes by increasing their growth, including length, thickness, and darkness⁴
- The Eyelash Satisfaction Questionnaire (ESQ) was used in the bimatoprost ophthalmic solution 0.03% phase 3 pivotal trial
 - It is a 23-item valid and reliable patient-reported outcome (PRO) measure, which determines a patient's perception of changes in physical and subjective attributes of eyelashes and eyelash care at the time of their eyelash assessment (ie, required no recall)
 - It demonstrated that patients^{5,6}
 - Reported significant improvement in eyelash length, fullness, and overall satisfaction (LFOS)
 - Felt significantly more confident, attractive, and professional (CAP) as a result of their eyelashes
 - Had an improved daily routine (DR) in making their eyelashes look presentable
 - The results from the PRO measure correlated with a bimatoprost ophthalmic solution 0.03%-dependent increase in eyelash length, thickness, and darkness
- However, to date, no follow-up PRO measure has been validated to assess eyelash prominence, satisfaction with eyelashes, and the impact of treatment on the patient's perception of self-esteem, including feeling good about general appearance, and daily eyelash routine

Objectives

- To develop and validate the ESFQ, a 39-item, **dynamic** (requires recall) PRO instrument used to identify and measure a patient's perception of changes in physical and subjective attributes of eyelashes and eyelash care compared to their baseline visit

Methods

Description of Samples

Validation Sample

- The same Internet sample that was used for the online psychometric investigation⁸ of the ESQ (N=928) was used in this study
 - Male respondents were excluded from analyses due to insufficient sample size
 - The final sample contained 909 subjects

Treatment Sample

- The treatment sample data were collected from a multicenter, double-masked, randomized, parallel study that assessed the safety and efficacy of a once-daily application of bimatoprost ophthalmic solution 0.03% compared with vehicle in increasing overall eyelash prominence
- Data was collected at day 1 (baseline), weeks 1, 4, 8, 16, and a posttreatment assessment at week 20

Data Analysis

- The latent structure of the questionnaire was assessed using confirmatory factor analysis (CFA) based on the hypothesized questionnaire structure that was consistent with the original conceptual model of the ESQ⁸
 - Three first-order factors were thought to impact the way each patient scored on the respective scale items. This model was modified based on fit statistics and correspondence with the hypothesized latent structure of the measure
- Subsequent analyses reviewed the questionnaire's item and scale measurement properties
 - Item-level psychometric evaluation.** The following analyses were conducted:
 - Item-total correlations for each scale
 - Variances for each scale
 - Alpha-removed statistics
 - Scale-level psychometric evaluation.** Psychometric properties of the scales was examined for internal consistency reliability using 3 indices:
 - Cronbach's alpha
 - Spearman-Brown adjusted alpha to a 10-item scale
 - Average inter-item correlation

Results

ESFQ

- The initial ESFQ consisted of 39 candidate items and was designed to measure **the change from baseline (recall required)** of self-perceived eyelash quality and satisfaction on 3 distinct constructs: LFOS, CAP, and DR. Responses are rated on 5-point Likert-type scales

Sample composition

- The validation and clinical samples consisted of 909 and 278 subjects, respectively, with a mean age (SD) of 39.1 (15.6) and 49.8 (11.4) years, respectively (**Table 1**)

Table 1. Demographic Characteristics of the Validation and Treatment Samples

	Validation Sample (N=909)	Clinical Sample (Baseline) (N=278)
Age, y		
Mean	39.1	49.8
Standard deviation	15.6	11.4
Range	17-86	22-78
Marital Status, n (%)*		
Single	264 (29.0)	
Married	494 (54.3)	
Otherwise	151 (16.7)	
Race, n (%)*		
African American	25 (2.8)	1 (0.4)
Asian American	22 (2.4)	34 (12.2)
Caucasian	833 (91.8)	225 (80.9)
Native American, other, or mixed	27 (3.0)	18 (6.5)
Education, n (%)*		
High school or less	233 (25.7)	
Some college	331 (36.4)	
Bachelors or Associates degree	288 (31.7)	
Higher education	56 (6.1)	
Declined to answer	1 (0.1)	
Missing	2 (0.2)	
Annual Income, n (%)*		
<\$35,000	321 (35.3)	
>\$35,000 but <\$65,000	328 (36.2)	
>\$65,000 but <\$100,000	130 (14.3)	
>\$100,000	92 (10.1)	
Declined to answer	38 (4.1)	

*Marital status, education, and income information was collected for the Internet sample. Socioeconomic data were not collected for the clinical sample.

- The descriptive statistics for the questionnaire items found to be most strongly associated with the hypothesized latent constructs (9 of the 39 items) in both the validation and the treatment sample on LFOS-follow-up (F), CAP-F, and DR-F domains are provided in **Table 2**
- In the clinical trial treatment sample, a reduction in score for all 3 domains was observed from week 1 to week 16, which indicates improvements in all domains

Table 2. Item-Level Statistics for the Final 9-items of the ESFQ for the Validation and Treatment Samples (at Baseline and Endpoint)

ESFQ	Validation Sample N=909			Clinical Sample Week 1 N=261		Clinical Sample Endpoint, Week 16 N=258	
	Range	Mean (SD)		Range	Mean (SD)	Range	Mean (SD)
LFOS-F							
Change in eyelash length (Q2)	1-5	2.39 (0.65)		2-3	2.94 (0.23)	1-3	2.06 (0.81)
Change in eyelash fullness/thickness (Q3)	1-5	2.32 (0.70)		2-3	2.95 (0.23)	1-4	2.29 (0.78)
Satisfaction with eyelashes, overall (Q8)	1-5	2.29 (0.82)		2-5	3.52 (0.83)	1-5	2.48 (1.10)
CAP-F							
Feel confident in looks* (Q20)	1-5	2.65 (0.78)		2-5	3.82 (0.77)	1-5	3.00 (1.22)
Feel confident about professional appearance* (Q22)	1-5	2.70 (0.79)		2-5	3.90 (0.75)	1-5	3.20 (1.11)
Feel attractive* (Q23)	1-5	2.62 (0.82)		2-5	3.95 (0.75)	1-5	3.05 (1.15)
DR-F							
Less time spent applying mascara since treatment (Q33)	1-5	2.61 (0.83)		1-5	3.84 (0.81)	1-5	3.09 (1.28)
Less time spent removing mascara at night (Q34)	1-5	2.61 (0.85)		1-5	3.81 (0.81)	1-5	3.21 (1.13)
Convenience since treatment (Q36)	1-5	2.80 (0.76)		2-5	3.86 (0.71)	1-5	3.21 (1.13)
Summary							
Overall LFOS-F score [†]	3-15	7.01 (1.89)		6-11	9.42 (1.00)	3-12	6.84 (2.48)
Overall CAP-F score [†]	3-15	7.97 (2.23)		6-15	11.66 (2.15)	3-15	9.25 (3.36)
Overall DR-F score [†]	3-15	8.01 (2.41)		4-15	11.51 (2.07)	3-15	9.51 (3.34)

[†]Question was asked in the context of not wearing mascara.
^{††}For all 3 domains, higher scores indicated increased impairment or increased burden on the daily routine.
 ESFQ=Eyelash Satisfaction Follow-up Questionnaire; SD=standard deviation; LFOS-F=length, fullness, and overall satisfaction-follow-up; CAP-F=confidence, attractiveness, and professionalism-follow-up; DR-F=daily routine-follow-up.

Confirmatory Factor Analysis – Validation Sample

- Four different models were assessed. Model fit statistic summaries for the final model are provided in **Table 3** for the validation and treatment samples
 - The final model of the data provided a good fit (goodness of fit index [GFI] = 0.98, comparative fit index [CFI] = 0.96, nonnormed fit index [NNFI] = 0.96, and root mean square error of approximation [RMSEA] = 0.08 [90% CI = 0.07, 0.09])
 - Similar results were reported for the treatment sample at weeks 1 and 16, validating the questionnaire structure with the treatment population (baseline: GFI = 0.97, CFI = 0.95, NNFI = 0.95, and RMSEA = 0.09 [90% CI = 0.07-0.12]; week 16: GFI = 0.98, CFI = 0.97, NNFI = 0.97, and RMSEA = 0.10 [90% CI = 0.07-0.12])

Table 3. Fit Statistics for the Final Latent Model

Model	χ^2	df	GFI	CFI	NNFI	RMSEA	RMSEA 90% CI
Validation Sample	169.5	24	0.98	0.96	0.96	0.08	0.07-0.09
Treatment Sample-Week 1	77.66	24	0.97	0.95	0.95	0.09	0.07-0.12
Treatment Sample-Week 16	80.02	24	0.98	0.97	0.97	0.10	0.07-0.12

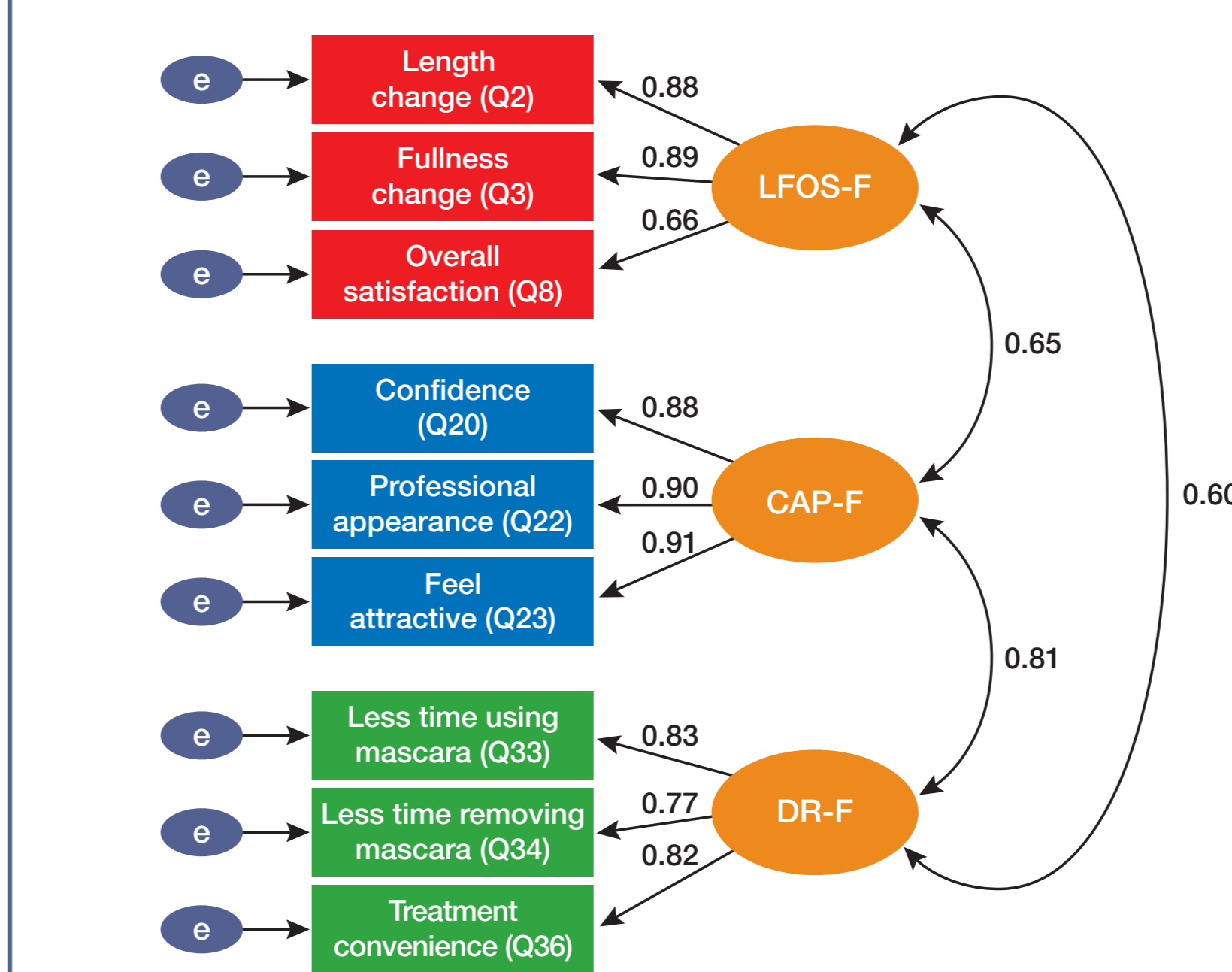
df=degree of freedom; GFI=goodness of fit; CFI=comparative fit index; NNFI=nonnormed fit index; RMSEA=root mean square error of approximation; CI=confidence interval.

- Figures 1 and 2** show the finalized CFA model of the questionnaire for the validation and treatment sample, including the standardized path coefficients for the latent variables on each of the items, as well as the level of correlation between factors. High standardized factor loadings were observed for most items on their respective factors

Item-Level Psychometrics

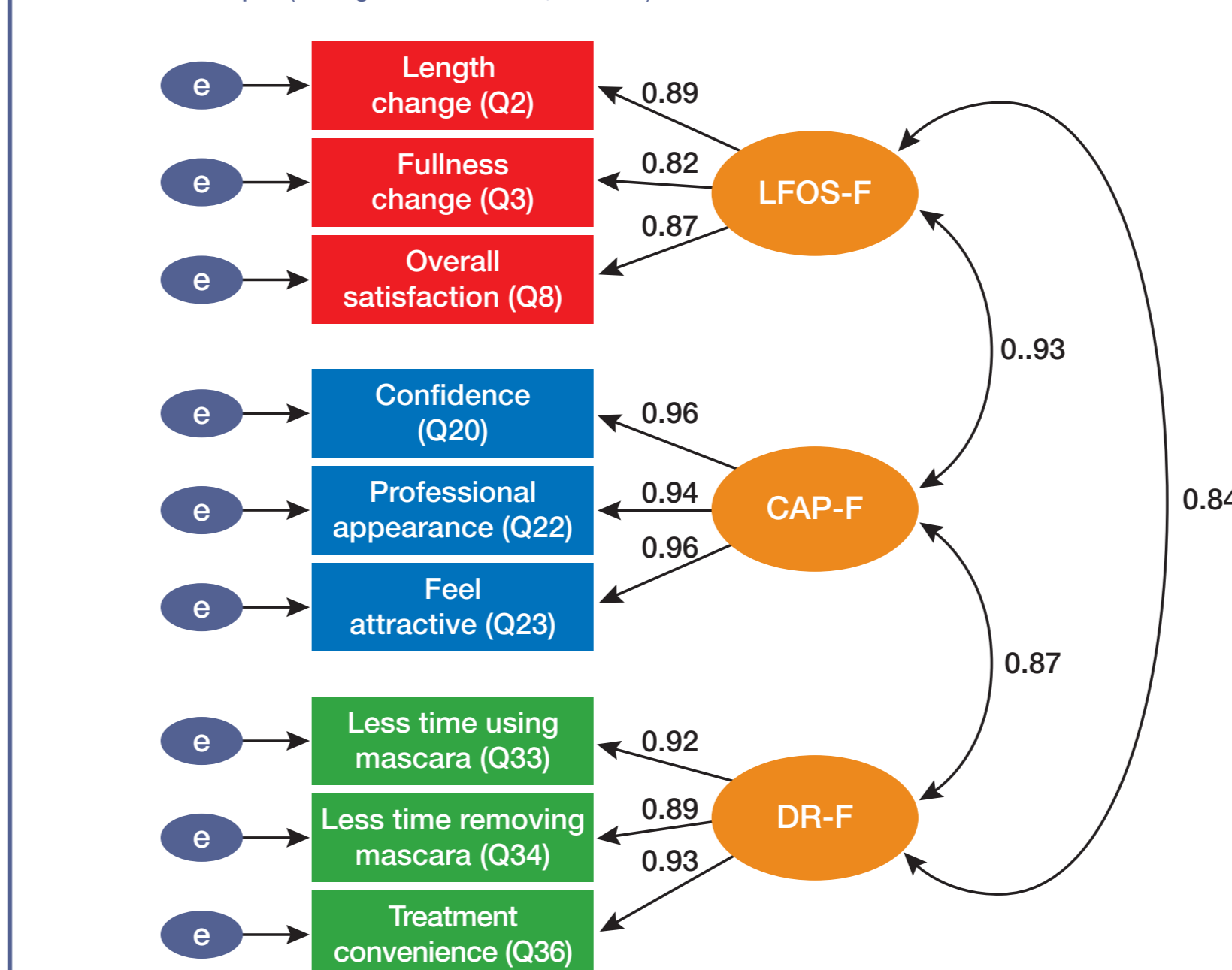
- A summary of the item-level psychometrics is presented in **Table 4**
- The average item-total correlations for each of the 3 scales were LFOS-F = 0.697, CAP-F = 0.837, DR-F = 0.717
- Domain-specific variances were compared with Hartley's F_{max} test
 - Values for LFOS-F and CAP-F vs DR-F were 1.28; for LFOS-F vs CAP-F was 1.39
 - None of the contrasts were significant, indicating sufficient variance equivalence
- Item-total correlations were compared against a criterion of 0.40
 - The smallest corrected item-total correlation for each scale was 0.60 (item 8), 0.83 (item 20), and 0.69 (item 36) for LFOS-F, CAP-F, and DR-F, respectively
- Alpha-removed statistic measures the impact for an item's removal on a scale
 - Deletion of item 8 from LFOS-F resulted in an increase of >0.02 in alpha-removed statistic; however, the structure was retained given the results on other item- and scale-level psychometrics and to retain similarity with the ESQ⁸ questionnaire

Figure 1. Finalized CFA model structure and path coefficients for the 9-item ESFQ using the validation sample (N=909).



CFA=confirmatory factor analysis; ESFQ=Eyelash Satisfaction Follow-up Questionnaire; e=error variance; LFOS-F=length, fullness, and overall satisfaction-follow-up; CAP-F=confidence, attractiveness, and professionalism-follow-up; DR-F=daily routine-follow-up.

Figure 2. Finalized CFA model structure and path coefficients for the 9-item ESFQ using the treatment sample (change at 16 weeks, N=258).



CFA=confirmatory factor analysis; ESFQ=Eyelash Satisfaction Follow-up Questionnaire; e=error variance; LFOS-F=length, fullness, and overall satisfaction-follow-up; CAP-F=confidence, attractiveness, and professionalism-follow-up; DR-F=daily routine-follow-up.

Scale-Level Psychometrics

- A summary of scale-level psychometrics is presented in **Table 5**
- Coefficient alpha, a measure of internal consistency reliability, for the 3 scales ranged from excellent (0.924 for CAP-F) to good (0.850 and 0.830 for DR-F and LFOS-F), respectively

Table 4. Item-Level Psychometrics

Item	LFOS-F	CAP-F	DR-F	Correlations		Variance	$\alpha_{removed}$
				LFOS-F vs CAP-F Z_{diff} (P)	CAP-F vs DR-F Z_{diff} (P)		
ESFQ 2	0.74	0.53	0.47	0.18 (<.001)	0.19 (<.001)	--	0.43
ESFQ 3	0.75	0.54	0.47	0.11 (<.001)	0.26 (<.001)	--	0.49
ESFQ 8	0.60	0.53	0.47	0.46 (<.027)	0.57 (<.001)	--	0.68
ESFQ 20	0.59	0.83	0.66	0.37 (<.001)	--	0.51 (<.001)	0.61
ESFQ 22	0.55	0.85	0.66	0.45 (<.001)	--	0.64 (<.001)	0.62
ESFQ 23	0.57	0.85	0.66	0.51 (<.001)	--	0.61 (<.001)	0.68
ESFQ 33	0.49	0.64	0.74	--	0.43 (<.001)	0.22 (<.001)	0.68
ESFQ 34	0.46	0.54	0.72	--	0.65 (<.001)	0.11 (<.001)	0.73
ESFQ 36	0.48	0.70	0.69	--	0.13 (<.001)	0.34 (<.001)	0.57

The first 3 columns indicate corrected item-total correlations for the correlations in colored boxes (ie, items with their respective total scores).
 Z_{diff} =difference between Fisher's z transformation of the correlations; $\alpha_{removed}$ =alpha if item removed (compare with alpha for CAP-F = 0.924, LFOS-F = 0.830, DR-F = 0.850).
 LFOS-F=length, fullness, and overall satisfaction-follow-up; CAP-F=confidence, attractiveness, and professionalism-follow-up; DR-F=daily routine-follow-up; ESFQ=Eyelash Satisfaction Follow-up Questionnaire.

Table 5. Scale-Level Psychometrics

Scale	Reliability		
	α	α_{IB}	r_{ii}
LFOS-F	0.830	0.942	0.619
CAP-F	0.924	0.976	0.802
DR-F	0.850	0.950	0.653

α =coefficient alpha; α_{IB} =coefficient alpha with a Spearman-Brown correction to a 10-item scale; r_{ii} =average inter-item correlation; LFOS-F=length, fullness, and overall satisfaction-follow-up; CAP-F=confidence, attractiveness, and professionalism-follow-up; DR-F=daily routine-follow-up.

- All of the scales achieved a rating for their 10-item adjusted alpha: LFOS-F = 0.942, CAP-F = 0.976, and DR-F = 0.950
- Average inter-item correlation for each of the scales was appropriate
 - CAP-F showed high inter-item correlation (0.802) while LFOS-F and DR-F were lower (0.619 and 0.653, respectively)

Conclusions

- The model fit for the validation and treatment samples provided substantial evidence that the 3 domains (LFOS-F, CAP-F, and DR-F) of the ESFQ were valid
- Results from the study provide evidence that scoring the questionnaire using 3 different scales, 1 for each factor, is appropriate
- The final 9-item scale demonstrated appropriate psychometric properties, accompanied with an appropriate model fit on the CFA, a high degree of item-level reliability, and good internal consistency
- Additional psychometric analysis for future research include test-retest validity and minimally important difference
- The data demonstrate that the LFOS-F, CAP-F, and DR-F domains are valid and reliable components of the ESFQ and can be used to assess the importance of eyelash treatment when patients report health-related quality of life outcomes relating to multiple domains and complexities of their eyelash appearance

References

- Holló G. The side effects of the prostaglandin analogues. *Expert Opin Drug Saf.* 2007;6:45-52.
- Batchelor D. Hair and cancer chemotherapy: consequences and nursing care-a literature study. *Eur J Cancer Care (Engl).* 2001;10:147-163.
- Yoelin S, Wu J, Somogyi C, Beddingfield FCI. Inter-rater and intra-rater reliability of the Global Eyelash Assessment scale for assessment of overall eyelash prominence. Poster presented at the Skin Disease Education Foundation's 33rd Annual Hawaii Dermatology Seminar; February 7-13, 2009; Maui, HI.
- Latisse [package insert]. Irvine, CA: Allergan, Inc.; 2009.
- Smith S, Fagien S, Somogyi C, Whitcup SM, Beddingfield FC. Eyelash growth in subjects treated with bimatoprost ophthalmic solution 0.03%; a multicenter, randomized, double-masked, vehicle-controlled, parallel study. Poster presented at the American Academy of Dermatology's 67th Annual Meeting; March 6-9, 2009; San Francisco, CA.
- Hammond GC, Burgess SM, Cole JC, Yang M, Hansen JE, Walt JG. Development and Validation of Patient-Reported Outcomes Tool for Eyelash Characteristics. Poster presented at the International Society For Pharmacoeconomics and Outcomes Research Annual Meeting. May 16-20, 2009; Orlando, FL.