

Factor Structures of the Family Environment Scale (Amharic Version)

Mitiku Hambisa¹

Abstract

Based on the fact that the family is a primary social institution which impacts the development of various outcomes in its members and that the starting point for improving family environment (FE) is availing a psychometrically sound tool that assesses this environment, the present study sought to examine the 90 items Form R FES in the Ethiopian context. At the outset, FES items were translated to the working language of Ethiopia (i.e., Amharic) by a team of experts. Then, the content and context relevance of the items were evaluated by 8 experts. The items were further examined in a pilot study by gathering data from a sample of 159 (77 male and 82 female) adolescents attending one government and one private secondary school in Addis Ababa City. Lawshe's procedures of examining content validity and reliability analysis using Cronbach were used to analyze the data. Depending on the results, 19 of the FES items were dropped, the other items were improved, and eventually, 71 better quality items were made ready for the main study. The main study was conducted on 477 adolescents (214 males and 255 females, 8 missing cases) attending two government and one private secondary schools in Addis Ababa City. Descriptive statistics, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to analyze the data. Although further efforts are needed to improve the psychometric qualities of the FES items in the Ethiopian context, 64 of the 90 items were found to be promising to assess the 10 subscales of the FES. Implications of the findings are discussed.

Keywords: *family; family environment; family environment scale; adolescent; factor analysis*

¹ Assistant Professor of Social Psychology, School of Psychology, CEBS, AAU. Email lovemitiku@yahoo.com

Introduction

The environment to which people are exposed throughout their lives influences people's emotions, thoughts and behaviors. Family environment (FE) is one of those environments that may exert the most significant influence on family members as its effects last throughout one's lifetime. FE is the family's overall pattern of interpersonal relationships, organization, and emphasis given to the personal growth of its members. Not surprisingly, while a negative family environment stifles a person's growth, a positive family environment facilitates personal development. Yet, compared to the structure of attributes of persons, the structure of attributes of environments has not been given much empirical attention (Kidist & Sandhu, 2020; Kurock, Gruchel, Bonanati & Buhl, 2023; Saucier, Wilson & Warka, 2007).

Researchers have developed different FE assessment tools, including the Family Functioning Index, Family Assessment Measures and Family Climate Inventory. Of the various FE assessment tools, Family Environment Scale (FES) is the most widely used because it is a well-established tool with more dimensions than the other instruments. This makes the FES a tool of choice for tapping FE in a more comprehensive manner (Charalampous, Kokkinos & Panayiotou, 2013; Kurock, Gruchel, Bonanati & Buhl, 2023). The three dimensions of the FES (i.e., relationship, *R*; personal growth, *PG*; system maintenance, *SM*) are composed of a total of 10 sub-dimensions (Moos & Moos, 2009). Each of the subscales are described in Table 1.

The development of valid and reliable measurement instruments is essential to the advancement of science. Based on the Social Climate Model, the initial English version of the FES was developed using the empirical approach. Then, the scale was translated into more than 22 languages, including French, German, Arabic, Chinese,

Korean, Japanese and Hebrew (Rivers & Sanford, 2019). The translation of the FES into different languages indicates the practical and research applications of the scale in different cultures. For instance, using the FES, family environment was shown to predict adolescents' emotional intelligence in India (Irfan & Kausar, 2020), academic achievement and delinquent behaviors in China (Zhao & Zhao, 2022; Huang & McKeown, 2022) and mediate the relations between poverty-reduction intervention and mental health in Uganda (Karimli, Sewamala & Neilands, 2023). Despite its continued use in different cultures, the psychometric characteristics of the FES have remained to be an issue (Montalescot, Speyer, Legrand, Ayav, Combe, Stengel & Untas, 2022).

Notwithstanding its conceptual and empirical appeal, researchers have raised doubts about the FES's factorial structure in general and the reliability and validity of its subscales in particular. For instance, in spite of the conceptually derived three dimensions, other studies found only two dimensions of FES. Using a varimax-rotated maximum likelihood factor analysis, Fowler (1981) found cohesion versus conflict and, organization-control dimensions for Form R of FES. Other empirical studies have also confirmed these two major factor solutions (Boake & Salmon, 1983; Chipuer & Villegas, 2001). Likewise, Oliver, Handal, Enos and May (1988) found only eight subscales of the FES failing to replicate the 10 factors initially conceptualized. In addition to construct validity problems, estimates of internal consistency (using Cronbach's Alpha) for five subscales of the FES were found to be lower than those reported in the scale's manual and below what is considered to be an acceptable level (i.e., 0.70) for research purposes (Roosa & Beals, 1990a).

In his response to Roosa and Beals's (1990a) critique of FES, Moos (1990) attributed the reliability and validity problems of the scale to variability in the samples used in different contexts. According to him, the sample utilized for the analysis greatly influences how a tool's factor structure is determined. He claimed that this is so because samples with greater levels of heterogeneity are more likely to produce more factors. In their Greek translation of the FES, Charalampous, Kokkinos and Panayiotou (2013) empirically demonstrated that Moos's contention that heterogeneous samples produce more factors was false. They concluded that even though the first-order factor structure of FES might not be the same across different cultural contexts, it is likely that the FE constructs conceptualized as three dimensions (R, PG and SM) hold true regardless of language or context. However, some studies (e.g., Zulaifah, 2012) indicate that cross-cultural validity of the items of the FES is questionable. In her dissertation project, Zulaifah attempted to adapt the FES to the Indonesian cultural context and found inadequate reliability coefficients for most of the 10 subscales and impossible (negative) coefficients for Expressiveness, Achievement and Moral-Religious subscales. Concluding that the Indonesian version of the FES does not exhibit sound internal consistency, Zulaifah was compelled to reconstruct the items to better fit the cultural context. More recently, Montalescot et.al (2022) found low Cronbach's alpha (.40) for Expressiveness and inadequate Average Variance Extracted (AVE) values for Cohesion, Expressiveness and Conflict subscales of the FES. Overall, these findings indicate that FES has a sound conceptual base whose cross-

Table 1

Descriptions of the Ten Subscales of the Family Environment Scale

Dimension	Subscale	Description
Interpersonal Relationships	Cohesion	The degree of perceived commitment, support and help family members provide for each other.
	Expressiveness	The degree to which family members are encouraged to express feelings and problems.
	Conflict	Amount of openly expressed anger, aggression and conflict among family members.
Personal Growth	Independence	The extent to which family members are assertive, make own decisions, and are self-sufficient.
	Achievement Orientation	The extent to which school and work activities are cast as indices of achievement or areas of competition.
	Intellectual-Cultural Orientation	The extent to which family members show interest in political, social, intellectual, and cultural activities.
	Active-Recreational Orientation	The extent to which family members emphasize participation in social and recreational activities.
	Moral-Religious Emphasis	The extent to which family members emphasize ethical and religious issues and values.
System Maintenance	Organization	The extent to which the family endorses clear organization and structure in planning family activities and responsibilities.
	Control	The extent to which rules and procedures are followed and enforced by family members.

Source: Moos and Moos (2009)

Cultural reliability and validity need to be investigated. Accordingly, the major purpose of the present study was to examine the psychometric properties of the FES's full version (90 items, Form R) in the Ethiopian context. Specifically, the study was intended to: (1) explore the factor structures of the FES, (2) examine the reliability of the subscales of the FES, (3) identify better quality items that measure dimensions of the FE in the Ethiopian context.

Methods

Study Approach and Site

This study employed a quantitative approach to research. Particularly, correlational research design was used. The selection of this approach and the design was guided by the nature of the objectives of the study. The present study sought to examine reliability and factor structures of the FES which are commonly addressed by the correlational research design in studies that adapt the FES into new cultural contexts (Montalescot et al., 2022; Zulaifah, 2012).

Addis Ababa, where the present study was conducted, was founded in 1886 (Bahiru, 2002). This city has been serving as the capital city and cultural hub of Ethiopia for about 130 years. It is, therefore, one of the oldest cities in Africa. Because it is the capital city of a multicultural society, Addis Ababa is also the home of people of diverse ethnicity, languages and religions.

Sample and Sampling Techniques

Data for this study was collected from two government and one private secondary schools (Grades 9-12) that were purposively selected. Owing to COVID-19 restrictions and makeup classes that resulted from interruptions of academic calendar caused by this pandemic during the time of data collection, it was difficult to get an adequate number of willing schools as per the principles of the proportionate random sampling technique. Thus, it was decided to collect data from two government and one private secondary schools (Grades 9-12) which permitted the collection of the data. The government secondary schools were Addis Ketema (in Addis Ketema Subcity) and Kechene Debere Selam (in Gullele Subcity), while the private school was Bistrate Gebriel (in Nifas Silk Lafto Subcity). Data obtained from the record offices of the

respective schools showed that, at the time of data collection, there were 2905 (Male =1391; Female =1514) regular students in Addis Ketema, 2388 (Male = 1023; Female =1365) regular students in Kechene Debere Selam, and 455 (Male =213; Female =242) regular students in Bisrate Gebriel secondary schools. Thus, depending on these data, 5748 (i.e., 2905 +2388 +455 =5748) was considered as the accessible population of the study. Whereas target population is the population to which research results are generalized ideally, generalization to accessible population is more realistic (Gay, Mills & Airasian, 2012).

The minimum sample size commonly used in simple SEM studies is 200 (Blunch, 2013; Kline, 2016). On the other hand, Jackson's (2003) rule of sample size in terms of the proportion of the number of participants (N) to the number of parameters to be estimated (q) (N: q; 20:1) is also suggested in the Structural Equation Modeling (SEM) literature. From the models examined in the present study, the Confirmatory Factor Analysis (CFA) Model of the Second-Order Factors (see Figure 1) had the maximum number of parameters to be estimated:21 (i.e., 10 error variances,10 loadings and one covariance). Using Jackson's rule for this model, we obtained $(N: q; 20:1) = 20 \times 21 = 420$ participants. Anticipating the number of students who might be unwilling or unable to participate in the study, and responses that might be incomplete or inappropriate and, therefore, would be discarded, 80 more participants were added to the 420. This increased the sample size to 500. Because the number of students in the government schools (5293) was by far higher than the number of students in the private school (455). Therefore, in order not to extremely underrepresent the private school student population, based on number of students in government and private schools (Addis Ababa Education Bureau, 2018), it was decided to sample 74% of the 500

students from government schools and 26% of the 500 students from the private school (see Table 2). Such a disproportionate stratified sampling technique is used when the strata being investigated are vastly different in their sizes (Daniel, 2012). For a preliminary analysis of the FES items, pilot study was conducted on a sample of 159 (77 male and 82 female) adolescents attending one government (i.e., Dejach Balcha Abanefso) and one private (i.e., Abunegorgorios) secondary school in Addis Ababa City.

To select participants in accordance with this sample size, a proportionate stratified random sampling method (by considering sex and grade level as stratum) was used. To decide the number of participants that was to be selected from each stratum, the following formula for proportional stratified sampling was used (Brown, 2007).)

$$K_s = n \left(\frac{N_s}{N} \right)$$

Where K_s = Number of sample participants selected from stratum S ; n = sample size for stratum S ; N_s = Number of participants in stratum S ; N = Overall population size. This formula allocates sample sizes according to the number of participants in a stratum.

Table 2 shows the demographic characteristics of the participants of the main study. Although all of the participants (500) were present, able and willing to participate in the study and therefore filled in and returned the questionnaire (response rate = 100%), data screening indicated that 23 (14 males, 9 females; 17 from government schools, 6 from the private school) did not fill in the questionnaire appropriately (e.g., some acquiescence response sets and zigzag response patterns were observed). Response sheets of these 23 participants were dropped. As a result, a total of 477 participants remained in the analysis (see Table 2). While the educational level of

the fathers or male guardians of the respondents varied from *No Education* (n=13; 2.70%) through *Grades 11-12* (n=102; 21.40%) to *PhD and above* (n=9; 1.90%), that of the mothers or female guardians ranged from *No Education* (n =24; 5.00%), through *Grades 11-12* (n=99; 20.80%) to *PhD and above* (n =3; .60%). In a similar manner, the family's reported overall monthly income varied from *Less than 500 Birr* (n =17; 3.60%) through 10,001-15,000 Birr (n=69;14.50%) to Greater than 15,000 Birr (n =150;31.40%) (Birr is Ethiopian currency). The participants reported living in family sizes ranging from two to 16 members (Mean = 5.88; SD = 2.18). Likewise, the age of the participants ranged from 13 to 22 (Mean =16.86; SD=1.41). These characteristics indicate that the sample is eligible for providing valid and reliable data which facilitate generalizations of the findings to the accessible population of the study.

Translation of the FES Items

In the present study, a questionnaire composed of the translated Amharic version of Form R of the Family Environment Scale (FES) was used to assess the family's social functioning as perceived by the adolescents. The elements of a multi-step translation process framework known as Translate, Review, Adjudicate, Pretest, and Document (TRAPD) (Willis et al., 2010) were employed in translating the FES items. According to these authors, in contrast to the traditional back translation process, TRAPD is a team-based approach that is capable of addressing both linguistic and socio-cultural elements in the translation process.

Since it was assumed that multi-point response options would yield more reliable scores than the original dichotomous (i.e., True-False) response options of the FES (Greene & Plank,1994; Moos, 1990), a four-point scale that ranged from

definitely true for my family (4) to *definitely not true for my family* (1) was used. Twenty-three negatively worded items were also reverse-scored.

Data Gathering Procedures

The cooperation request letter that the School of Psychology, Addis Ababa University, provided was given to the administrators of the selected schools. After realizing that the data collection would not harm the instructional time and the students in any way, the administrators gave permission for data collection. . From each school, the administrators assigned one teacher (data collection assistant) to facilitate data collection. The assistants were oriented about the purposes of the study and the procedures of responding to the questionnaire. Then, the name list of the students, which served as a sampling frame, was obtained from the schools' record offices. In cooperation with the administrators and the assigned teachers, free periods and a separate hall

Table 2

Demographic Characteristics of the Main Study Participants (N = 477)

Variable		N	%
School Type	Government	353	74.00
	Private	124	26.00
Sex	Male	214	44.90
	Female	255	53.50
	<i>Missing Values</i>	8	1.70
Grade	9	105	22.00
	10	113	23.70
	11	110	23.10
	12	149	31.20
Duration of Living in the Family	From Birth Date	371	77.80
	From 10-15 Years	27	5.70
	From 6-9 Years	27	5.70
	From 2-5 Years	29	61.60
	From 5-12 Months	8	1.70
	Less than 5 Months	9	1.90
Family Structure: Living With:	<i>Missing Values</i>	6	1.30
	Father and Mother (Nuclear Family)	294	61.6
	Only Mother (Single Parent Family)	74	15.5
	Only Father (Single Parent Family)	15	3.10
	Step Father and Mother (Blended Family)	10	2.10
	Father and Step Mother (Blended Family)	9	1.90
	Grandparents (Grand Father, Grand Mother or both; Extended Family)	18	3.80
	Parents and Grand Parents (Extended Family)	28	5.90
	Relatives (Uncle, Aunt)	19	4.00
	Others (with brother, sister, etc)	9	1.90
<i>Missing Values</i>	1	.20	

(e.g., library hall) was arranged in advance. In the hall, , the researcher, together with the assistants, provided the students with orientation about the study and established rapport. The participants were informed that they could stop participating in the study

once they began without incurring any undesirable consequences. The participants were also well-versed about the confidentiality of their responses.

Techniques of Data Analysis

In this study, descriptive statistics (i.e., frequency, percentage, mean and standard deviation) were used to describe the data. In the pilot study, the FES items' content validity and context relevance were assessed using Lawshe's techniques. Exploratory Factor Analysis (EFA) that employed maximum likelihood factor extraction with varimax rotation and Confirmatory Factor Analysis (CFA) (Blunch, 2013; Byrne, 2010) were used to examine the internal structures of the FES items. The analyses were carried out using version 23.0 of Analysis of Moment Structures (AMOS) and Statistical Package for the Social Sciences (SPSS). Since these software packages do not have functions that produce outputs for Construct Reliability (CR), Index of Quality (IoQ) and Average Variance Extracted (AVE) (see Results Section and Table 4), Microsoft Excel was used to compute values of these statistics. For all inferential tests, a test of significance was set at the .05 level.

In the main study, of the various types of construct validity (Hair, Black, Babin & Anderson, 2010; Kline, 2016), convergent and discriminant validities were examined. Reasoning that setting higher loading sizes (e.g., .71) would be unfeasible in Ethiopian context where there were no ample previous empirical foundations regarding the FES items and their subscales, a loading value of greater than or equal to .32 was used (Tabachnick & Fidell, 2013). As a result, in EFA, at least a loading of .32 and the corresponding explained variance of $(.32)^2 = 10.24\%$ by the factor in a single item were considered as salient. Accordingly, for a latent factor in the present

study, an AVE of at least 10.24% was considered as meaningful. In addition to loading of greater than or equal to .32, criteria of at least three items per factor (Hair, Black, Babin & Anderson, 2010) was also used to consider a factor (i.e., subscale) as meaningful.

Results

First-order Factors of the FES

The first objective of the present study was exploring the factor structures of the FES. At the outset, using all of the 71 FES items that were selected during content and context relevance analyses and the pilot study and following the common practice in the FES literature that examined unidimensionality of the 10 scales of the FES separately (e.g., Charalampous, Kokkinos & Panayiotou, 2013), EFA was run for each of the 10 scales separately. Unidimensionality assessment results are presented for each of the 10 subscales of the FES in the following paragraphs. Numbers in the parenthesis corresponding to the subscales indicate the number of items that passed quality screening in the pilot study and were re-presented for quality examination in the main study.

1. Cohesion Subscale (7 items): EFA extracted only one factor for this subscale. The one-factor CFA model of the Cohesion Subscale of the FES, which was composed of the 7 items, fitted the data well (see Table 3). All of the items except Rchn18R had standardized loadings of .32 or above (the cutoff point being used in the present study) and all loadings were statistically significant ($p < .001$). Thus, the items or indicators were generally moderately and significantly related to their latent constructs. The average variance extracted (AVE) by the Cohesion Subscale was 29.28% with

Construct Reliability (CR), Index of Quality (IoQ) and Cronbach Alpha of .73. Stringent statisticians consider only an AVE of 50% and a CR of .70 or above as adequate evidence for convergent validity (Hair, Black, Babin & Anderson, 2010). Other scholars consider an overlapping variance as low as 20% as fair, leaving the selection of the level suitable for a specific study to the researcher (Tabachnick & Fidell, 2013). Given that the AVE by the Cohesion Subscale was above the level considered to be fair in some literature (i.e., 20%) and that the AVE by the factors was above the criterion set for the present study (10%), it appeared that the convergent validity of the one-factor CFA model for Cohesion Subscale would be acceptable.

2. *Conflict Subscale (8 items)*: EFA extracted 3 factors, none of which met the criterion of at least three items per factor. The removal of item Rcnf11 whose Squared Multiple Correlation (R^2) was very low ($R^2 = .009$) resulted in the increment of the reliability of the subscale. Examination of unidimensionality was conducted using CFA. This one factor Conflict Subscale model with seven items fit the data well (see Table 3) without further modification. Four out of the seven items of this subscale had standardized loadings of .32 or above and all of the loadings were statistically significant ($p < .001$). The AVE by the Conflict Subscale was 15.69% with CR of .55, IoQ of .62 and Cronbach Alpha of .54. Thus, because its convergent validity seemed to be acceptable, the Conflict Subscale appears to be a unidimensional construct in the present study.

3. *Expressiveness Subscale (6 items)*: EFA extracted two factors. However, one of the two factors could not fulfill the criterion of at least 3 items per factor. Thus, one-factor model was tested and items to which the latent factor contributed the least (as indicated by R^2) were gradually eliminated to search for better fitting model. Accordingly, items

Rexp65 and Rexp50 were removed one after the other. By this process, even if CFI and other fit indices improved [$\chi^2(2) = 9.000, p = .011; GFI = .991; CFI = .899; RMSEA = .086$] for the CFA model of the Expressiveness Subscale composed of four items, RMSEA worsened. A closer examination of the modification indices indicated that the largest parameter change (.096) and drop of the discrepancy between this model and the data by at least 4.198 would occur if errors associated with items Rexp2R and Rexp19R were allowed to covary. Because the relations between these two items is meaningful (both are aspects of one latent construct), they were allowed to covary. As a result, all of the fit indices of this model improved, indicating that the model fitted the data well (see Table 3). Three out of the four items of this subscale had standardized loadings of .32 or above and all of the loadings were statistically significant ($p < .01$). The AVE by the Expressiveness Subscale was 16.94% with CR of .43, IoQ of .63 and Cronbach Alpha of .42. Thus, although reliability indices using Cronbach Alpha and CR were low; IoQ, size of the loadings and AVE appeared to support the convergent validity of the one factor CFA model of this subscale.

4. Achievement Orientation Subscale (8 items): EFA extracted 2 factors one of which did not meet the criterion of at least three items per factor. Thus, unidimensionality examination was conducted using CFA. Most of the fit indices except CFI were acceptable. Following suggestions from modification indices, after allowing errors associated with two items covary, CFI was improved (see Table 3). Seven out of the eight items of this subscale had standardized loadings of .32 or above and all of the loadings were statistically significant at least at $\alpha = .05$. The AVE by the Achievement Orientation Subscale was 12.66% with CR of .52, IoQ of .59 and Cronbach Alpha of

.53. Thus, it appears that CFA supported the unidimensionality of the Achievement Orientation Subscale of the FES.

5. Independence Subscale (6 items): EFA produced two interpretable factors with three items each. Although both one and two-factor CFA model achieved acceptable fitness, one factor model was able to produce the maximum possible reliability coefficients for this subscale than each of the two factors suggesting convergence of the items. Thus, a CFA model consisting of the six items was tested for the Independence Subscale of the FES. The fitness of the model to the data appeared to be acceptable (see Table 3). All of the six items had standardized loadings of .32 or above and all of the loadings were statistically significant ($p < .001$). The AVE by this subscale was 21.80% with CR of .62, IoQ of .68 and Cronbach Alpha of .61. Thus, it appears that CFA supported the unidimensionality of the Independence Subscale of the FES.

Table 3

Fit Indices for CFA Models of the 10 Subscales of the FES

Subscale	χ^2 (df)	GFI	CFI	RMSEA
Cohesion	19.453 (14)	.988	.989	.029 [(90% CI: (.000, .056), PCLOSE = .886)]
Conflict	22.226 (14)	.986	.952	.035 [(90% CI: (.000, .061), PCLOSE = .802)]
Expressiveness	1.628(1)	.998	.991	.036 [(90% CI: (.000, .134), PCLOSE = .436)]
Achievement Orientation	32.935* (19)	.983	.921	.039 [(90% CI: (.014, .061), PCLOSE = .769)]
Moral-Religious Emphasis	31.718*(200)	.983	.938	.035 [(90% CI: (.005, .057), PCLOSE = .854)]
Independence	29.89*** (9)	.978	.917	.070 [(90% CI: (.043, .098), PCLOSE = .104)]
Intellectual Cultural Orientation	3.644(4)	.997	1.000	.000 [(90% CI: (.000, .066), PCLOSE = .850)]
Active Recreational Orientation	31.609** (14)	.981	.945	.051 [(90% CI: (.027, .075), PCLOSE = .425)]
Organization	16.994 (20)	.991	1.000	.000[(90% CI: (.000, .033), PCLOSE = .998)]
Control	.857 (2)	.999	1.000	.000 [(90% CI: (.000, .071), PCLOSE = .867)]
<i>Suggested Criteria for Fit Indices****</i>	<i>Cut-off</i> Nonsignificant χ^2	> .90	> .90	< .08

Notes: * $p < .05$; ** $p < .01$; *** $p < .001$; ****Whittaker & Schumacker (2022).

6. *Active Recreational Subscale (7 items)*: EFA produced two factors one of whose factor did not meet the criterion of at least three items per factor. Thus, the items were tested for unidimensionality. This one factor CFA model fitted the data well (see Table 3). All of the seven items had standardized loadings of .32 or above and all of the loadings were statistically significant ($p < .001$). The AVE by this subscale was 21.82% with CR of .64, IoQ of .67 and Cronbach Alpha of .63. Thus, it appears that CFA supported the unidimensionality of the Active Recreational Subscale of the FES.

7. *Intellectual-cultural Orientation Subscale (6 items)*: EFA produced two factors where one of the two did not meet the criterion of at least three items per factor. Thus,

the items were tested for unidimensionality. The latent factor explained the least variance in item PGic5 ($R^2 = .014$) and its removal improved reliability. After removing this item, except for CFI, the other fit indices indicated that the CFA of the model composed of the remaining 5 items fitted the data well. Following suggestions from modification indices, after allowing errors associated with items PGic31R and PGic54 covary, CFI was improved, and the model appeared to fit the data better (see Table 3). Two of the five items had standardized loadings of above .32 and all of the loadings were statistically significant at least at .05. The AVE by this subscale was 11.48% with CR of .38, IoQ of .58 and Cronbach Alpha of .33. Thus, even though the Intellectual-Cultural Orientation Subscale had low Cronbach Alpha and CR, it appears that its CFA model supported its unidimensionality.

8. *Moral-religious Emphasis Subscale (9 items)*: EFA extracted 3 factors, two of which did not meet the criterion of at least three items per factor. Thus, unidimensional examination was conducted using CFA. At the outset, most of the fit indices except CFI were acceptable. The factor explained the least variance in item PGMre56 ($R^2 = .005$). Consequently, item PGMre56 was removed from the model and CFI improved (see Table 3). Six of the eight items in the modified model had standardized loadings of above .32 and all of the loadings were statistically significant ($p < .001$). The AVE by this subscale was 14.41% with CR of .56, IoQ of .61 and Cronbach Alpha of .54. Thus, the CFA model of the Moral-religious Emphasis Subscale appears to support its unidimensionality.

9. *Organization Subscale (8 items)*: EFA extracted only one factor, which the CFA model confirmed as fitting well to the data (see Table 3). All of the eight items in the model had standardized loadings of above .32, and all of the loadings were statistically significant ($p < .001$). The AVE by this subscale was 17.57% with CR of .61, IoQ of .64 and Cronbach Alpha of .60. Thus, it appears that CFA supported the unidimensionality of the Organization Subscale of the FES in this sample.

10. *Control Subscale (6 Items)*: EFA produced two factors, where one of the two factors did not meet the criterion of at least three items per factor. Thus, the items were tested for unidimensionality. Although the CFA composed of the six items appeared to fit the data well, reliability analysis indicated that the maximum possible reliability would be obtained if items SMctl9R and SMctl17 are eliminated. The resulting CFA-modified model composed of four items fitted the data well (see Table 3). Three of the four items in the model had standardized loadings of above .32 and all of the loadings were statistically significant ($p < .001$). The AVE by this subscale was 21.12% with CR of .50, IoQ of .67 and Cronbach Alpha of .48. Thus, it appears that CFA supported the unidimensionality of the Control Subscale of the FES in the present study

Table 4

Comparison of Reliabilities (Cronbach Alpha, α ; CR and IoQ) of the Subscales of the FES in the Present Study and Studies Conducted in Other Countries

Subscales	Ethiopia						USA		Russia		Malaysia
	Present Pilot Study		Present Main Study				Moos & Moos (2009), from Manual		Barskyet.al (2010)		Omaret.al (2010)
	K*	α	K	α	CR	IoQ	K	α	K	α	α
Cohesion	9	.69	7	.73	.73	.73	9	.78	6	.56	.70
Conflict	9	.19	7	.54	.55	.62	9	.75	8	.58	.63
Expressiveness	9	.37	4	.42	.43	.63	9	.69	6	.38	.22
Achievement Orientation	9	.38	8	.53	.52	.59	9	.64	5	.45	.24
Moral-Religious Emphasis	9	.45	8	.54	.56	.61	9	.78	7	.56	.45
Independence	9	.43	6	.61	.62	.68	9	.61	6	.36	.10
Intellectual Cultural Orientation	9	.18	5	.33	.38	.58	9	.78	-	-	.51
Active Recreational Orientation	9	.60	7	.63	.64	.67	9	.67	8	.62	.33
Organization	9	.69	8	.60	.61	.64	9	.76	8	.58	.58
Control	9	.18	4	.48	.50	.67	9	.67	5	.42	.54
Total number of Items	90		64				90		59		

Notes: *K = Number of Items; CR = Construct Reliability, IoQ = Index of Quality

Second-order Factors of the FES

The second-order factors of the FES were examined by subjecting the 10 first-order factors to Exploratory Factor Analysis (EFA). EFA produced two factors (see Table 5). The first factor was named *stability and personal-growth*. The two components that made up this factor, control and organization, came from SM and the rest came from PG. The second factor was named *Relationship* as the three original subscales (cohesion, conflict and expressiveness) loaded on this factor.

Evidence of construct validity of the two-factor model of the first order factors of the FES came from the scrutiny of their convergent and discriminant validities. The average variance extracted (AVE) by the *Relationship* dimension was 44.68% (CR = .29), while that of the *Stability and Personal Growth* was 35.43% (CR = .76). The CFA model of the two-second order

Table 5
Rotated Factor Matrix of the Subscales of the FES (N= 477)

Subscale	Factor	
	1	2
Active Recreational Orientation	.602	
Intellectual Cultural Orientation	.557	
Organization	.505	.495
Achievement Orientation	.493	.356
Moral Religious Emphasis	.441	
Control	.435	
Cohesion	.377	.814
Conflict		-.550
Independence	.477	.485
Expressiveness		.404

Notes: Bolded loadings were used to name the factors; Factor 1 = *Stability and Personal Growth (SPG)*; Factor 2 = *Relationship*.

factors (see Figure 1) fitted the data well [$\chi^2(34) = 98.884, p = .000$; GFI = .960; CFI = .952; RMSEA = .063 (.049, .078), PCLOSE = .062]. All of the subscales had standardized loadings of above .32 and all of the loadings were statistically significant ($p < .001$). Thus, it appears that CFA supported the convergence of the second-order factors of the FES.

Next, evidence of discriminant validity of the model under discussion was examined. Although none of the AVE by the *Stability and Personal Growth* factor (.35)

and *Relationship* factor (.45) exceeded the squared correlation between the factors (.74), the correlation between the latent factors was less than .90 ($r = .859$). Moreover, the two-factor model fitted the data better [$\chi^2 (34) = 98.880$, $p = .000$; GFI = .960; CFI = .950; RMSEA = .063 (90% CI: (.049, .078), PCLOSE = .062)] than the one-factor model [$\chi^2 (35) = 127.850$, $p = .000$; GFI = .946; CFI = .931; RMSEA = .075 (90% CI: (.061, .089), PCLOSE = .060)]. From this evidence of convergent and discriminant validities, it appears that the construct validity of the two second-order factors model of the FES in the Ethiopian sample is acceptable. Moreover, the fitness of this two second-order factors model was found to be acceptable for the males' ($n = 214$) group [$\chi^2 (34) = 49.914$, $p = .038$; GFI = .956; CFI = .970; RMSEA = .047 (90% CI: (.011, .073), PCLOSE = .546)] and the females' ($n = 255$) group [$\chi^2 (34) = 84.427$, $p = .000$; GFI = .939; CFI = .936; RMSEA = .076 (90% CI: (.056, .097), PCLOSE = .018)]. These results indicate that the model is applicable in the two sub-samples.

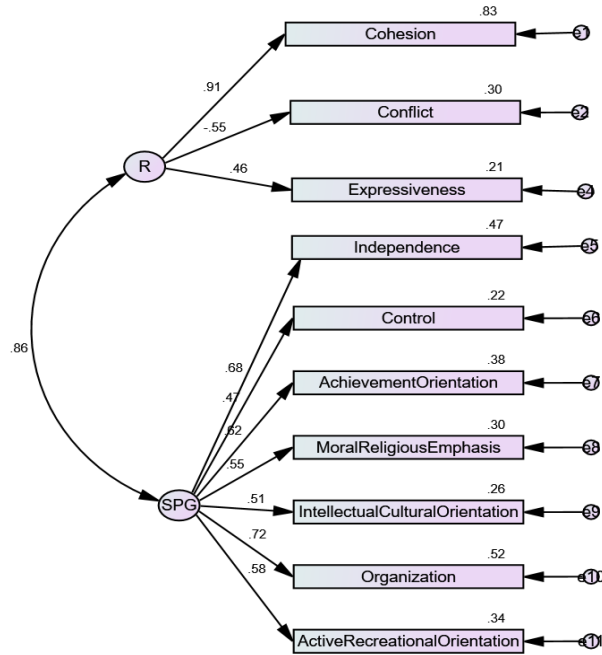


Figure 1: CFA Model of the Second-Order Factors (Relationship, R and Stability and Personal-Growth, SPG) of the FES

Discussion

The present study is the first attempt to examine the quality of the FES items in the Ethiopian context. Regarding the factor structures of the FES, it is pointed out that the empirical literature focused on second-order factors by factor-analyzing the subscales without compelling evidence that supports the existence of these 10 first-order factors (Charalampous, Kokkinos & Panayiotou, 2013). The results of the present study supported the unidimensionality (convergent validity) of the ten subscales of the FES. The number per subscale of the relatively high-quality items which were examined and gradually selected to buttress unidimensionality of the scales ranged from 4 (for Expressiveness and Control Subscales) to 8 (for Achievement Orientation, Moral Religious Emphasis, and Organization Subscales). Thus, 64 of the 90 original FES items appeared valid in the Ethiopian context. These findings indicate that FES items tap meaningful constructs corresponding to the 10 originally conceptualized first-order factors of this scale. Nevertheless, the results of the present study are only partially consistent with that of Charalampous, Kokkinos and Panayiotou because their study confirmed the existence of only seven subscales in the Greece sample.

The 26 items dropped in the present study were identified via content and context relevance analysis (i.e., Lawshe's procedures), pilot study and CFA in the main study. Indeed, the fact that some of the dropped items are culturally and contextually irrelevant for Ethiopian families is readily evident. For instance, one of the dropped items *everyone has an equal say in family decisions*, appears to be culturally not relevant as the Ethiopian culture has been hierarchical, in which case members other than the husband or the wife are rarely given opportunity in family decision making. Similarly, the dropped items *we come and go as we want to in our family, we say*

anything we want to around home and we can do whatever we want to in our family appear to be contextually not relevant as most Ethiopian families emphasize control because of which the adolescents may not have the opportunity to say or do whatever they want in their families. Other studies (e.g., Montalescot et al., 2022 in France; Zulaifah, 2012 in Indonesia) also identified problems in some of the items of the FES that can be attributed to the nature rather than the translation errors. This raises the doubt that some of the FES items might not be culturally relevant.

The existence of the second-order factors of the FES was also examined in the present study. The three conceptual dimensions of the FES (R, PG and SM) and the most recent second-order factors reported in the Greek translation of this scale (Charalampous, Kokkinos & Panayiotou, 2013) are not confirmed in the present study. Rather, a two-factor structure that resembles the two factors reported by Fowler (1981) and replicated in other studies (Boake & Salmon, 1983; Saucier, Wilson & Warka, 2007) emerged (see Table 5). The first factor was named *Relationship (R)* as the three original Relationship Dimension subscales (cohesion, conflict, and expressiveness) loaded on this factor, with magnitudes of loadings of cohesion and conflict being the highest. As in the previous studies, conflict loaded negatively on this second-order factor. The second factor was named *Stability and Personal-Growth (SPG)*. Control and organization factors of the SPG came from SM and the rest came from PG. Thus, the original distinct conceptualizations of PG and SM appear to be amalgamated in this second factor. Similar to results of the present study, in their adaptation of the FES to Bangladeshi culture, Uzzaman and Karim (2018) found two factors (Factor 1: *Achievement, Order and Culture Orientation*; Factor 2: *Emotional Atmosphere*). Factors 1 and 2 in Uzzaman and Karim's study appear to be similar to SPG and R in

the present study, respectively. These findings put into doubt Charalampous, Kokkinos and Panayiotou's assertion that, irrespective of context or language, the initial three dimensions of the FE (i.e., R, SM and PG) hold true.

Regarding the reliability of the subscales of the FES, in the present main study, seven of the 10 subscales (Cohesion, Conflict, Achievement Orientation, Moral Religious Emphasis, Independence, Active Recreational Orientation and Organization, see Table 4) had Cronbach Alphas greater than .50, a cut-off point above which some FES literature (e.g., Omar et.al., 2010) regard reliability of a scale as acceptable. Improvements in reliability indices of almost all subscales of the FES from pilot to main studies of the current study indicate that further refinements may produce subscales of more adequate reliabilities in the Ethiopian context, even with a smaller number of items.

The internal consistency reliabilities of the subscales of the FES have become an important issue in the empirical literature. Moos and Moos (2009) argue that, in their own studies, the reliabilities of all of the subscales are in an acceptable range (.61 to .78). Nonetheless, many cross-cultural studies of the FES could not find adequate reliability coefficients for some of the subscales. For instance, Omar et al. (2010) examined the reliabilities of subscales of the FES in the Bahasa Malaysia language using adolescents aged 12-17. They obtained lower coefficients (Cronbach's alpha of between 0.10 - 0.70) than those originally reported (Cronbach's alpha of between 0.61 -0.78). Indeed, Omar et.al found Cronbach's Alphas of less than 0.5 for five subscales, which they reported as below the acceptable level for research or practical use. In the present study, Cronbach's Alpha ranged from the lowest of .33 (Intellectual Cultural Orientation Subscale) to the highest of .73 (for Cohesion Subscale). More to the point,

only three subscales (i.e., Control, Intellectual Cultural Orientation and Expressiveness) had Cronbach's Alphas of less than 0.50 in the present main study. Another cross-cultural study came from Russia which was conducted by Barsky et al. (2010) using 450 adolescents aged between 11 to 17 years. Initially, they obtained Cronbach's Alphas that ranged from the lowest of .13 (for the Achievement Orientations Subscale) to .60 (for the Active Recreational Subscale). Seven of the ten subscales of the FES had Cronbach's Alphas of below .45 in their initial analysis. After deleting items with negative item-scale correlations, they obtained Cronbach's Alphas ranging from .36 to .62. However, Cronbach's Alphas for four subscales (i.e., Expressiveness, Independence, Achievement and Control) remained below .50 while that of the Intellectual-Cultural Orientation Subscale could not be improved. Thus, it appears that even though the present main study could produce a reliability coefficient that is equal to the original FES study for one subscale (i.e., for Independence Subscale, see Table 4), internal consistency indices are generally similar to (or relatively better for some subscales, see Table 4) than that of the Barsky et al (2010) and Omar et al (2010) studies. Furthermore, the present study has examined the reliability of the subscales using factor loadings obtained from latent variable modeling (i.e., CFA). Factor loadings are the correlations of an indicator (item) with a factor. Latent variable modeling provides better estimates of reliability indices than the traditional manifest variable modeling because it accounts for measurement error in the estimation process. Cronbach's Alpha, which has been used exclusively in the FES studies, doesn't always provide the best index of reliability of a tool (Saris & Gallhofer, 2014). Indeed, despite its common use in the empirical literature in general and FES literature in particular, according to Hair, Black, Babin and Anderson (2010), Cronbach's alpha tends to

understate reliability. Besides, Cronbach's alpha does not use factor loadings (i.e., latent variable modeling) to estimate reliability. Thus, in the present study, construct reliability (CR) (Hair, Black, Babin & Anderson) and index of quality (IoQ) (Schwartz & Butenko, 2014) which employ factor loadings to estimate the reliability, were used for further examination of the reliability of the subscales. The fact that Cronbach's Alpha understates reliability is evident in the present study as CRs and IoQs of the subscales were relatively greater (see Table 4). Particularly using IoQs, all of the ten subscales had reliabilities of greater than or equal to .58. Thus, estimating reliabilities of the subscales of the FES using latent variable modeling might help to shed light on the blurred issue of psychometric property of the FES items.

Conclusions and Implications

The present study examined the factor structures of the 90 items Form R FES in the Ethiopian context. In light of the findings of the present study, the following major conclusions can be drawn.

First, while some FES items might not be relevant to the Ethiopian culture, the 10 subscales of the FES appear to exist in the Ethiopian context. Moreover, despite the initial three dimensions of the FES, the present study supports the two second-order factors that other studies found. Second, given the reliability indices found in the previous cross-cultural studies of the FES, the internal consistency reliability of the subscales of this scale that were found using latent variable modeling and empirically selected better-quality items are acceptable. Third, 64 of the 90 FES items appear to be of relatively higher quality and, therefore, promising to assess the 10 subscales of the FES in the Ethiopian context.

The implications of the findings of the present study are manifold. The present study has important theoretical implications as it supports the existence of the ten subscales and two dimensions of the FES in the Ethiopian context. This may provide fertile soil for comparative investigations of the Ethiopian FEs with FEs in other cultures (e.g., in other collectivist and individualist cultures). The present study has also important implications for future research. First, although an attempt was made to use latent variable modeling to provide additional evidence of psychometric qualities of the subscales of the FES (i.e., CR and IoQ besides the classical Cronbach's Alpha), indices of reliability for Expressiveness, Intellectual Cultural Orientation and Control Subscales using Cronbach's Alpha, and that of Expressiveness and Intellectual Cultural Orientation Subscales using CR were found to be particularly worrisome. Thus, although the IoQs of these subscales appeared to be acceptable, future researchers should pay attention to the reliabilities of these subscales in particular and to the other subscales in general. Second, although participants were selected randomly, owing to COVID-19 restrictions during data collection, the present study used purposively selected schools which may limit the generalizability of the results. Future studies are recommended to use randomly selected schools. Third, future large-scale studies should be conducted not only on adolescents in the Addis Ababa city but also on adolescents living in rural areas. Populations other than adolescents (such as husbands and wives in a family) can also be targeted for study. Fourth, the link between FEs and other variables should be examined in future studies. For instance, the relationship between adolescents' FE and self-esteem, academic achievement, well-being, emotional intelligence, social intelligence, bullying, violent behavior, altruistic behavior, nonviolent behavior, peaceful personality and other constructs can be studied.

Acknowledgment

This study emerged from an International Society for the Study of Behavioral Development (ISSBD)-sponsored Developing County Fellowship (DCF 2020-2022) research project. Thus, the writer would like to extend his heartfelt thanks to the ISSBD. He also would like to thank Professor Peter Smith (Goldsmiths, University of London, London) for his painstaking professional mentoring throughout the two years of the project.

References

- Addis Ababa Education Bureau. (2018). *Education statistics annual abstract 2009 E.C. (2016/2017 G.C.)*. Addis Ababa: Author.
- Bahiru Zewde (2002). *A history of modern Ethiopia, 1855–1991* (2nd ed.). Addis Ababa: Addis Ababa University Press.
- Barsky, P. I., Gindina, E.D., Lobaskova, M.M., Malykh, S.B. (2010). Perception of Family Environment with Russian Adolescent Twins: Possible Genotype-Environment Correlation. In *Psychology in Russia: State of the Art* (pp. 412-430).
- Blunch, N.J. (2013). *Introduction to structural equation modeling using IBM SPSS statistics and AMOS*. Los Angeles: SAGE Publications Ltd.
- Boake, C., & Salmon, P.G. (1983). Demographic correlates and factor structure of the Family Environment Scale, *Journal of Clinical Psychology, 39*, 95-100.
- Brown, J. A. (2007). Stratified random sampling. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 967-971). Thousand Oaks: SAGE Publications, Inc.
- Byrne, B.M. (2010). *Structural equation modeling with AMOS* (2nd ed.). New York: Taylor and Francis Group, LLC.
- Charalampous, K., Kokkinos, C.M., & Panayiotou, G. (2013). The Family Environment Scale: Resolving psychometric problems through an examination of a Greek translation, *The International Journal of Educational and Psychological Assessment, 13* (2), 81-99.
- Chipuer, H.M., & Villegas, T. (2001). Comparing the second-order factor structure of the family environment scale across husbands' and wives' perceptions of their family environment, *Family Process, 40*, 187-198.
- Daniel, J. (2012). *Sampling essentials: Practical guidelines for making sampling choices*. New York: SAGE Publications, Inc.
- Fowler, P. C. (1981). Maximum likelihood factor structure of the Family Environment Scale. *Journal of Clinical Psychology, 37* (1), 160-164.
- Gay, L.R., Mills, G.E., & Airasian, P. (2012). *Educational research: Competencies for analysis and applications* (10th ed.). Boston: Pearson Education, Inc.
- Greene, R. C., & Plank, R.E. (1994). The short-form Family Environment Scale: Testing a different response format. *Psychological Reports, 74*, 451-464.
- Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). New Jersey: Pearson Prentice Hall, Inc.
- Huang, Y., & McKeown, S. (2022). Examining the association between family environment and adolescent delinquent behaviours in China, *Journal of Child and Family Studies, 31*, 2871–2884.
- Irfan, S., Kausar, R. (2020). Demographic characteristics and family environment as predictors of emotional intelligence in adolescents, *Journal of Behavioral Sciences, 30* (2), 61-81.

- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N: q hypothesis, *Structural Equation Modeling*, 10, 128–141.
- Karimli, L., Sewamala, F.M., & Neilands, T.B. (2023). The impact of poverty-reduction intervention on child mental health mediated by family relations: Findings from a cluster-randomized trial in Uganda, *Social Science & Medicine*, 332, 1-10.
- Kidist Tesfaye & Sandhu, T. (2020). Family environment and substance use among university students. *International Journal of Scientific and Technology Research*, 9 (2), 1530-1535.
- Kline, R. B. (2016). *Principles and practice of structural equation modelling* (4th ed.). New York: The Guilford Press.
- Kurock, R., Gruchel, N., Bonanati, S., & Buhl, M. H (2023). Family climate and social adaptation of adolescents in community samples: A systematic review, *Adolescent Research Review*, 7:551–563.
- Montalescot, L., Speyer, E., Legrand, K., Ayav, C., Combe, C., Stengel, B., & Untas, A. (2022). Reliability and validity of the French adaptation of the Family Relationship Index–short form in patients with chronic kidney disease, *Journal of Health Psychology*, 27(1), 166–175.
- Moos, R. H. (1990). Conceptual and empirical approaches to developing family-based assessment procedures: Resolving the case of the Family Environment Scale, *Fam Proc*, 29, 199-208.
- Moos, R.H., & Moos, B.S. (2009). *Family Environment Scale: Manual development, applications, research* (4th ed.). Palo Alto: Mind Garden, Inc.
- Oliver, J. M., Handal, P. J., Enos, D. M., & May, M. J. (1988). Factor structure of the Family Environment Scale: Factors based on items and subscales, *Educational and Psychological Measurement*, 48, 469-477
- Omar, K., Musa, R., Hanif, J., Muhammad, N. A., Bujang, A., & Fadhlullah, F. M. (2010). Reliability of Bahasa Malaysia version of Family Environment Scale and its measurement issues, *ASEAN Journal of Psychiatry*, 11 (1), 1-12.
- Rivers, A. S., & Sanford, K. (2019). Family Environment Scale. In J. L. Lebow, A. L. Chambers & D.C. Breunlin (Eds.), *Encyclopedia of Couple and Family Therapy* (pp. 1047-1051). Bern: Springer Nature Switzerland AG.
- Roosa, M. W., & Beals, J. (1990a). Measurement issues in family assessment: The case of the Family Environment Scale, *Family Process*, 29, 191-198.
- Saris, W. E., & Gallhofer, I. N. (2014). *Design, evaluation, and analysis of questionnaires for survey research* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.
- Saucier, G., Wilson, K.R., & Warka, J. (2007). The structure of retrospective accounts of family environments: Related to the structure of personality attributes. *Journal of Personality Assessment*, 88 (3), 295–308.
- Schwartz, S.H., & Butenko, T. (2014). Values and behavior: Validating the refined value theory in Russia, *European Journal of Social Psychology*, 44, 799-813.

- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston: Pearson Education, Inc.
- Uzzaman, M. A., & Karim, R. (2018). Assessing psychometric properties of Family Environment Scale in Bangladeshi culture, *Jagannath University Journal of Life and Earth Sciences, 4* (1): 1 – 11, 2018.
- Whittaker, T.K., & Schumacker, R.E. (2022). *A beginner's guide to structural equation modeling* (5th ed.). New York: Taylor and Francis.
- Willis et al. (2010). *Evaluation of a multi-step survey translation process*. doi: 10.1002/9780470609927.ch8.
- Zhao, L., & Zhao, W, (2022). Impacts of family environment on adolescents' academic achievement: The role of peer interaction quality and educational expectation gap, *Frontiers in Psychology, 13*, 1-17.
- Zulaifah, E. (2012). Reconstructing Family Environment Scales to fit into the Indonesian cultural context. In E. Witruk & A. Wilcke (Eds.), *Historical and Cross-Cultural Aspects of Psychology* (pp.481-496). Frankfurt am Main: Peter Lang.