

Development of the Arabic family relation test among school students

Marwa Nasser Alrajhi^{1,2}, Kamisah Osman², Hussain Ali Alkharusi³

¹Student Counseling Center, Sultan Qaboos University, Muscat, The Sultanate of Oman

²Faculty of Education, Universiti Kebangsaan Malaysia (The National University of Malaysia), Bangi, Malaysia

³College of Education, Sultan Qaboos University, Muscat, The Sultanate of Oman

Article Info

Article history:

Received Oct 31, 2024

Revised Nov 18, 2025

Accepted Jan 1, 2026

Keywords:

Arabic
Computerized
Family-relation
School-students
Test

ABSTRACT

Family relationships are highly valued; therefore, their assessment is essential. Documented literature has emphasized the significance of family relationships in students' different outcomes. However, there is a lack of psychometrically robust measures of family relationships in Arabic-speaking environments. This study aimed to develop the Arabic family relation test (AFRT) and examine its psychometric properties. The test explores six main dimensions of parent-child relationships: restrictiveness, affection, vulnerability, justice, acknowledgment, and trust. The cross-sectional quantitative design was utilized. The study sample included 601 school students (47.6% males) from grades 5, 7, 9, and 11, selected through a cluster-random sampling approach. The findings revealed a valid 60-item AFRT with a two-second-order-factor, six-first-order-factor factorial structure. Measurement invariance of the AFRT was achieved in the father-child model across groups of age and gender. However, measurement invariance was only supported across age groups in the mother-child model. The AFRT demonstrated good reliability, as indicated by Cronbach's alpha, test-retest, split-half, and composite reliability. The researchers concluded that the AFRT is an appropriate tool for family assessment; however, further psychometric investigation is recommended. In addition, recommendations for implications and suggestions for future research were provided.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Marwa Nasser Alrajhi
Student Counseling Center, Sultan Qaboos University
Muscat, A'seeb, 123, The Sultanate of Oman
Email: malrajhi@squ.edu.om

1. INTRODUCTION

Since the creation of mankind, societies have sought different techniques to raise and socially rear their children. Raising children in a way that helps them realize life changes and behavioral values is crucial. Even though children are born with existing biological and inherited behavioral patterns, they still need someone to guide them in adapting to their surrounding environment. Family is an essential element of society. Each family member has a significant role in the group's growth [1]. Rather than digging into individual intrapsychic problems, psychologists now direct their attention to the family context where the interactions between family members occur. As they attempt to understand problems, psychologists focus on how a family member interacts with others and is influenced by other family members [2].

A core element in family interventions is accurately assessing family relationships. Research has emphasized the importance of collecting data on the quality of family relationships, using children's perspectives, and employing reliable measures. The measures should be appropriate for the child's current

developmental stage, use suitable and understandable language, and have an attractive design that can grab the child's attention [3]. The current research aims to develop the Arabic family relation test (AFRT) and examine its psychometric properties in Omani school children aged 10 to 17.

2. LITERATURE REVIEW

2.1. Theoretical and conceptual framework of the Arabic family relation test

The theoretical framework of the current study is supported by the research of Oud and Welzen [4] in which the original Nijmegen Family Relation Test (NFRT) (i.e., Nijmeegse Gezinsrelatie Test or NGT) was developed. The NFRT was designed in the Dutch language and used by other researchers in different cultures [3], [5], [6]. The test contains 67 items distributed into six subscales: restrictiveness, affection, vulnerability, justice, acknowledgment, and trust. Some studies found these subscales to form higher second-order factors: control and support [3], [5].

NFRT is an objective test designed to examine 9-12-year-old children's perspectives on the relationships between all family members. This test assesses relationships regarding mother-child, father-child, husband-wife, and child-sibling dyads. However, the NFRT has been repeatedly used in previous studies to measure the mother-child and father-child dyads [3], [5], [6]. Moreover, the NFRT has been utilized for different purposes, such as detecting psychological problems in family relations. Other interventions, such as family therapy for children, can be developed based on this test.

Two main theories were manipulated to develop the NFRT: Stierlin's binding theory [7] and Boszormenyi-Nagy and Spark loyalty theory [8]. The two factors, restrictiveness and affection, were derived from Stierlin's psychoanalytic binding theory, which focuses on different transactions between parents and children. Stierlin elaborated that a family is where affection and love should be present, emotional bonds are formed, and children's needs are met. Hence, binding is a transaction in which parents and children participate [9].

The remaining factors—vulnerability, justice, acknowledgment, and trust—are derived from the loyalty theory [8]. In this theory, family members' perceptions of loyalty are assumed to be a product of trust and justice. Children feel that their family members show justice when treated fairly within the family rules and are trustworthy when treated with love and appreciation.

The present study used the structural equation modeling (SEM) to test the proposed conceptual framework. In the development of the NFRT, Oud and Welzen [4] concluded that the test consisted of six first-order factors, each including about 11 to 12 items. In contrast, Skoczeń *et al.* [3] and Fitriana [5] postulated that the two versions, the Bandung Family Relation Test (BFRT) [5] and the Computerized Family Relations Test (CFRT) [3], supported the higher second-order factor model. This model consisted of two higher-order factors: control and support. The control factor contains restrictiveness and justice. The support factor consists of vulnerability, trust, acknowledgment, and affection.

Although the first-order model includes the six factors and their respective items as observed variables [4], Skoczeń *et al.* [3] and Fitriana [5] used parcels as observed variables rather than individual items. Parceling is a technique researchers use in factor analysis when the test consists of a large number of items. It comprises averaging or adding subsets of items for a latent component and using the resulting mean scores (or sum scores)—“parcel scores”—as indicators for its latent construct. Using parcel-level factor analysis ensures having a suitable sample-size-to-model-size ratio. Parceling strategies vary between random, empirical, and content-driven methods [10]. Figures 1 and 2 demonstrate the proposed first-order and second-order models of the developed AFRT.

Two types of models were examined for the AFRT in the current study using confirmatory factor analyses (CFA): the first- and second-order models. Three parcels of items were included under each first-order factor, into which the items were randomly distributed based on the content-driven method. Content-driven allocation is when items are randomly distributed between parcels depending on the content the items share within each sub-scale. This means that each item should have an equal chance to be selected in one of the parcels under that domain [10]. It should be noted that each model was examined twice: the mother-child dyad and the father-child dyad. Thus, four models were examined in total.

In addition to assessing the factor structure of the AFRT, the measurement invariance technique using CFA was conducted as well to examine the invariant structure of the AFRT across groups of children's gender (males and females) and ages (grades: 5, 7, 9, and 11). This technique primarily aims to assess the invariance or stability of measurement across informants or over time to avoid bias against particular groups [11], [12]. Three levels of measurement invariance were tested in this study at the level of first-order factor model: Configural invariance (i.e., assumes that the same structure is measured across groups), metric invariance (i.e., factor loadings are constrained to be equal across models), and scalar invariance (i.e., indicators intercepts, or latent variable's means, are set to be equal across models) [13]. Moreover,

structural weights invariance was tested at the second-order factor level (i.e., assuming equal factor loadings in the second-order structure) [3]. A pivotal point is that the measurement invariance of first-order level models is a prerequisite for any measurement invariance at higher-order level models [13]. Then, loadings at second-order models can be meaningfully compared across groups, which allows the researchers to test second-order factor loadings across groups to determine whether their structure is similar [3], [13].

A commonly used measure of group difference is the Chi-square difference ($\Delta \chi^2$) test; however, this test often rejects models, especially when the sample size is large, even when violations are minor [13], [14]. Hence, Chen [14] suggested using differences in goodness-of-fit indices (Δ GFIs) between the compared models as an alternative criterion. Precise criteria of Δ GFIs are provided by Chen [14] for small sample size ($N \leq 300$) and large sample size ($N > 300$). The rule of thumb associated with ($N > 300$) is of interest for the current study.

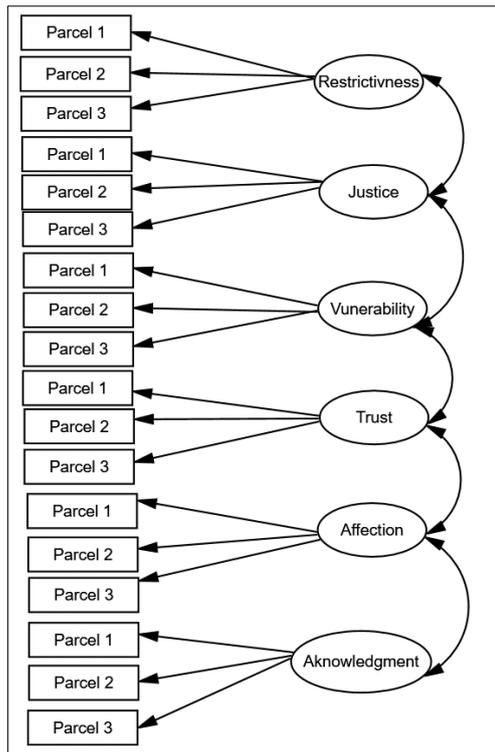


Figure 1. The proposed first-order model of the AFRT

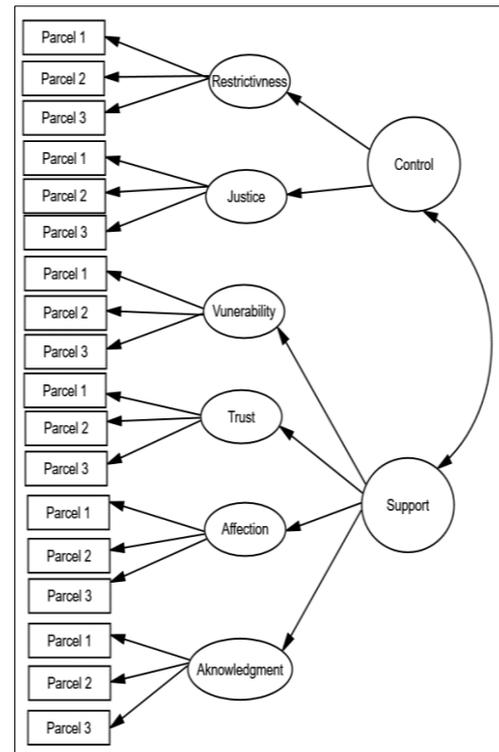


Figure 2. The proposed second-order model of the AFRT

2.2. Previous studies on the family relation test

Various studies have been documented in the literature regarding the different versions of the NFRT. In the original NFRT, Oud and Welzen [4] utilized both exploratory factor analysis (EFA) and CFA to examine the factor structure, which confirmed that the six-factor solution provided the best model fitting for the mother and father scores. Kuder Richardson (K-R 20) and test-retest were used to assess reliability. The findings showed that the K-R values ranged between 0.67 (vulnerability) and 0.76 (trust) for mothers and between 0.63 (vulnerability) and 0.76 (acknowledgment) for fathers. Regarding the test-retest, reliability values ranged between 0.56 (trust) and 0.70 (acknowledgment) for mothers and between 0.64 (vulnerability) and 0.74 (trust) for fathers.

A team of researchers at Padjadjaran University (Indonesia) has developed the Indonesian version of the NFRT, the BFRT [15]. The sample consisted of 529 students aged 9-12. The findings showed good Cronbach's alpha for restrictiveness, affection, and acknowledgment, with overall coefficients of at least 0.80. On the other hand, vulnerability, justice, and trust showed alpha coefficients lower than 0.80. In addition, CFA findings showed that four subtests had a good fit in the fathers' ratings: vulnerability, justice, acknowledgment, and trust. However, for mothers' rating, five subtests showed a good fit: restrictiveness, vulnerability, justice, acknowledgment, and trust. Two subtests did not fulfill the criteria regarding the father's model, restrictiveness and affection, and one subtest, affection, did not fulfill the criteria concerning the mother's model.

Other researchers have utilized the NFRT or its different versions for psychometric or descriptive studies [3], [5], [6]. For example, Fitriana [5] explored some of the psychometric properties of the BFRT among 529 (9-12 years old) students. The study confirmed a second-order model with two higher-order factors, support and control. Restrictiveness and justice were loaded with opposite signs to the second-order factor, control, and the other four factors were loaded with equal signs to the second-order factor, support. However, the researcher highlighted that vulnerability and justice were bifactorial. Vulnerability has “small validity” for control, and justice shared “moderate validity” for support. In addition, alpha coefficients ranged between 0.58 (justice) and 0.82 (acknowledgment). Split-half reliability showed that acknowledgment has the highest reliability coefficient (0.82) while justice has the lowest (0.58).

Similarly, the initial analysis of the Polish version, the CFRT, was demonstrated by Skoczeń and Ciecuch [9]. A sample of 96 children (aged 9-13) participated in this initial analysis. A short version of the CFRT, including 24 items with four items per scale, was created. The findings of this study showed good reliability coefficients, measured by Cronbach’s alpha. For the mother scores, reliability ranged from 0.57 (injustice) to 0.76 (affection). For the father scores, reliability ranged from 0.68 (vulnerability) to 0.90 (trust). Regarding CFA, the mother-child model demonstrated acceptable fit indices, yet the researchers had to merge affection and trust factors into one factor for the father-child model. Also, two correlations between errors (i.e., uniqueness) were allowed. After these modifications, the model achieved good fit indices.

Psychometric properties of the CFRT long version were published by Skoczeń *et al.* [3]. The sample in this study included 404 primary school Polish children (aged 8 to 13 years old). The internal consistency of CFRT was assessed using McDonald’s omega and test-retest coefficients. Omega coefficients were highest in acknowledgment and trust for mother’s $\omega=0.84$ and father’s models $\omega=0.86$, respectively. Mother ratings showed lower test-retest stability, $r=0.76$, than father ratings, $r=0.71$. Furthermore, Skoczeń *et al.* [3] revealed that the fit indices of the first-order model were acceptable, but there were high intercorrelations between some factors. Therefore, the second-order model was tested with two main factors: control and support. The second-order model demonstrated better-fit indices and intercorrelations.

In addition, the researchers located only two studies that utilized analysis of measurement invariance to examine the construct validity of the different versions of the NFRT. A case in point is Skoczeń *et al.* [3], who examined the measurement invariance of the CFRT across children’s ratings of the mothers and the fathers. The results indicated that the measurement’s invariance across mothers and fathers was supported as the changes in the comparative fit index (CFI) were less than 0.01, the changes in root mean square error of approximation (RMSEA) were less than 0.015, and the changes in standardized root mean square residual (SRMR) were less than 0.03. Moreover, Toharudin *et al.* [6] tested the measurement invariance of the BFRT across three social economic strata (SES) groups (low, medium, and high) in a sample of 394 students from Indonesia. The findings showed that the BFRT’s measurement invariance was not supported for the different SES groups.

2.3. Context of the study

The current study is conducted in an Arabic context. The Sultanate of Oman is an Arabic-speaking country in Western Asia and one of the Gulf states. For a long time, Omani families have been known for strong relationships among their members, strengthened by Islamic and cultural principles. However, with the development that accompanied the Omani Renaissance in 1970, many social, political, and economic changes occurred that affected Omani families, particularly the roles of husbands and wives [16].

Omani researchers conducted several studies to explore the common parenting styles in Omani families based on Baumrind’s framework [17] (i.e., authoritarian, authoritative, and permissive parenting styles) [18]. Most of these studies have agreed that the most prevalent parenting styles in Omani families are the authoritative and the authoritarian, while the permissive style is relatively less common. However, no studies were found in the literature on other aspects of Omani family relationships. Aldhafri [19] has argued that there are insufficient empirical findings regarding parenting practices in Omani families. Therefore, the current study will attempt to develop AFRT to provide a new validated tool for examining various aspects of family relationships.

2.4. The research problem

The quality of family relationships has a significant role in children’s outcomes, such as children’s competence, adaptation, academic performance [20], depression [21], bullying [22], and mental health problems [23]. Hence, assessing family relationships is crucial. The current research problem is mainly presented by the lack of tests that measure family relationships in Arabic contexts. Searching through the literature, one well-known test adapted to the Arabic environment and used in many studies is the parenting authority questionnaire (PAQ) [24]. The PAQ was developed to measure only three types of parenting styles: authoritative, authoritarian, and permissive styles.

Besides parenting styles, few studies attempted to capture other types of family relationships in the Arab context [25]. However, based on the current researcher's best knowledge, no frequent or common uses of these tests have been utilized by other researchers. They were self-developed/adapted questionnaires explicitly used by those researchers only. Additionally, the Bene-Anthony Test is one of the family relations tests adapted to the Arabic culture [26]. Nevertheless, it was criticized for its lack of psychometric appropriateness, difficulty using the test, interpreting the results, and its specific use in clinical contexts [26].

Therefore, the NFRT was found by the researcher to be an appropriate measure to test family relationships for its objective format, suitability to a normal sample (i.e., not necessarily clinical), easy interpretation of the results, replicated utilization of the tests across different cultures, and inclusion of family relations' dimensions that were not assessed by Arab researchers previously. Furthermore, the NFRT is not designed to measure the family as one unit. It measures the relationships between different individual members of the family. Finally, the test was developed to capture different dyads inside the families, such as parents-children, husband-wife, and child-siblings, which allows for wide utilization of the test in family assessment [4].

The present study aims to develop the AFRT and examine its psychometric properties. It mainly attempts to answer the following research questions:

- i) What are the construct validity indicators of AFRT among Omani school students (grades 5, 7, 9, and 11) using CFA?
- ii) What are the construct validity indicators of AFRT among Omani school students (grades 5, 7, 9, and 11) using the measurement invariance method across gender and age?
- iii) What are the reliability indicators of AFRT among Omani school students (grades 5, 7, 9, and 11) using Cronbach's alpha, test-retest, split-half, and composite reliability (CR)?

3. METHOD

3.1. Study design

This study employs a quantitative approach with a descriptive cross-sectional design. Descriptive studies help explore a phenomenon, comprehend relationships between variables, and examine the characteristics of those variables among participants [27]. Hence, this design helps to achieve the objectives of the current study.

3.2. Population and sample

The population of this study comprises all public schools, which consist of grades 5, 7, 9, and 11 in A'seeb district, the Sultanate of Oman. A'seeb is known for its diversity among Omani families, who usually come from different areas inside and outside Muscat (i.e., the capital city), which ensures population diversity that contributes to the variation in students' characteristics and family environments. Given the large number of students and the inaccessibility of the list of all students' names, the cluster random sampling method was employed, using schools as clusters. SPSS software was utilized for a random cluster sampling technique to select the schools' sample. The total sample was 601 students selected from 10 schools. They were enrolled in grades 5, 7, 9, and 11, with ages ranging between 10 and 17 years old ($M=13.11$) (i.e., 47.6% males, 22.62% grade 5, 24.45% grade 7, 27.62% grade 9, and 25.29% grade 11).

At the time of data collection, all participants were Omanis and lived with both parents. Although the AFRT offered the option of "mothers only" or "fathers only" for the students who lived with one parent, the researchers had to remove these cases. Those were a few cases—fewer than 100—which could not be compared to the other two-parent cases. Furthermore, extreme values were found in some of the individual-parent cases. All students in each class were invited to a separate learning resources room for data collection process. The MoE and the headmaster's consent were obtained. MoE found that the study poses no risks to the participants involved and therefore complies with Oman's ethical research regulations. The students were asked to answer the study's questions willingly. Furthermore, it was ensured that there were no orphans among the participating students and that all of them were living with both parents or at least one parent. One of the researchers and a research assistant supervised the data collection process.

3.3. Instruments

3.3.1. Theoretical basis of the AFRT items

The AFRT was developed in the current study based on the work of Oud and Welzen [4]. Personal communication with Efi Fitriana (one of the authors of the BFRT) and Kai Welzen, the original authors of NFRT, revealed that some modifications were made to the items of the Indonesian version (BFRT) to enhance validity. Based on their recommendations, the researcher decided to adapt the English version of the BFRT to develop the Arabic version. However, in the Arabic version, the researcher has carefully focused on the content, language, and cultural aspects. The BFRT includes 67 items on a 5-point Likert scale.

3.3.2. The new AFRT

The International Test Commission (ITC) procedures [28] were first followed to generate the item pool. The researchers translated BFRT items from English to Arabic (i.e., forward translation). Then, revision was done by translators and language professionals in English and Arabic. After that, items were translated from Arabic to English and sent to the original authors of the previous versions to check the backward translation, namely Welzen and Fitriana. In addition, based on the researchers' review of previous literature on family relationships, especially in the Arabic context, and deep understanding and analysis of the translated items, four items were initially deleted from the 67-item version, resulting in an initial version of 63 items. The deletion of these items was mainly due to their unsuitability for Omani culture or redundancy with other items.

After that, those items were sent for face validity testing by eight experts in psychology, education, evaluation, and measurement. The content validity ratio (CVR) proposed by Lawshe in Romero-Jeldres *et al.* [29] was used as a major criterion for retaining or deleting an item in the face validity process. CVR refers to the number of panelists who stated that an item is essential (i.e., out of three options: essential, useful but not essential, and not necessary), divided by the total number of panelists.

As Lawshe in Romero-Jeldres *et al.* [29] recommended, eight experts require a minimum value of $CVR=0.75$ to retain an item. Moreover, experts' feedback and suggestions on all items were acknowledged. Based on this step, three items were deleted because they had a CVR score below 0.75. Other content modifications were made in some items based on experts' feedback, and some items were moved to a different dimension than their original ones. This step yielded a version of 60 items of the new AFRT. In the current study, the test was administered in a computerized format. The only difference between the computerized and paper versions is the "digitalization" of displaying the items and the response scale. The computerized format was used to integrate several features that cannot be included in a paper-based format, such as a continuous data scale, a recorded voice reading the items to the children, and male and female silhouettes representing the father and the mother. Also, the authors believed that the computerized format is more attractive and engaging to the student. This format is similar to the CFRT version of the test that was developed by Skoczeń and Ciecuch [9]. However, the current study did not intend to compare the paper-based and computerized formats of the test in terms of psychometric properties.

The test starts with a front page that explains the test's aim and the data's confidentiality. Then, it included four main parts: i) the participants' demographic information; ii) the AFRT; iii) the anxiety scale (TAS); and iv) the behavioral problems scale (BPS). The last two scales were used to examine criterion-related validity; their results are outside the current study's scope and have been reported in another study by the researchers [30].

The second section of the test is the main section, which includes the AFRT items. The section began with an animated guide that provided an example item and explained the instructions to participants. One item was displayed on the screen each time. A picture of a male silhouette representing the father on the left and a female silhouette representing the mother on the right appeared with each item. A recorded voice was included on the screen as an option for the students to listen to the items. The respondent has to answer each question once for the mother and once for the father. A continuous response scale (i.e., in a thermometer style)- ranging from 1 (totally disagree) to 2 (totally agree)- was used to answer the questions. The continuous scale is believed to give more sensitive responses and does not restrict the respondents to discrete values like the regular Likert scale [3]. Furthermore, the assumptions of many statistical techniques, including CFA and the full information maximum likelihood, are met by data collected using this approach. This type of response scale solves the problem of information loss and applies sophisticated and robust statistical analysis [31]. Figure 3 presents an item example of the AFRT; the test was administered in Arabic.

A pilot study was conducted on the computerized 60-item version of the AFRT. A total of 106 school students (50.9% males) were included in the pilot study. The students were enrolled in grades 5, 7, 9, and 11. For the test and retest data collection, 78 students (43.6% female) were selected from the same pilot sample. The pilot study findings showed acceptable preliminary psychometric properties. For all dimensions, Cronbach's alpha values (α) were within the accepted range of 0.60–0.70 [32]. Moreover, the test-retest reliability of the AFRT was examined within a two-week interval. All test-retest values showed excellent stability of the participants' answers, as all (r) values were >0.50 . Corrected item-total correlations showed that most of the values fall within the accepted range ($r \geq 0.30$) [33] except for four items relating to the justice dimension.

Further inspection of "Cronbach's alpha if item deleted" suggested that deleting these items would not enhance the internal consistency of the Justice dimension or the corrected item-total correlations. Justice was the only dimension that included reverse-coded items. The researchers noticed that the students were uncertain about those items and faced difficulty understanding how to answer them. Therefore, these items were rephrased to be on one side (i.e., all expressing justice only; the injustice items were rephrased) with no

reverse-coded items. The pilot study procedures resulted in some modifications in the language and content of the items; however, no items were deleted. That is, the 60-item version was utilized for CFA.

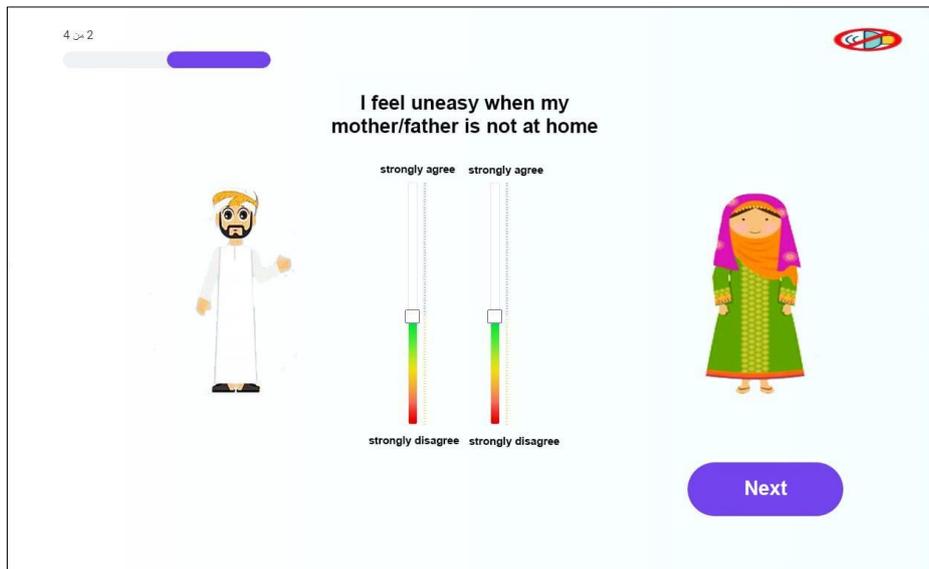


Figure 3. An item example of AFRT

4. RESULTS

4.1. Preliminary analysis

First, all data were screened for errors in data entries and distribution using frequencies and percentages. Frequencies of responses on all items ranged between 1 to 10. Of the total sample (n=690), 622 students indicated that they live with both parents, 57 with their mothers only, and 11 with their fathers only. These cases were deleted because they constitute a small number and cannot be compared with both parents' cases. In addition, a complete data set was necessary at the CFA analysis level to examine modification indices, which AMOS cannot offers in the presence of missing data. In addition, data were screened for outliers and extreme cases; as a result, 21 more cases were deleted. The final sample size consists of 601 respondents.

Second, the normality of data distribution was examined at the item level utilizing skewness and kurtosis values. All 60 items from both parents' ratings met the accepted levels of skewness (≤ 2) and kurtosis (≤ 7), indicating normality [34]. Third, the descriptive statistics of the six main factors (i.e., means and standard deviation) and the bivariate correlations using Pearson correlation coefficients were explored. Table 1 shows the six factors' descriptive statistics and correlation values.

Table 1. Descriptive statistics and Pearson correlation coefficients between the six factors of the AFRT

Factors	M	SD	r				
			Restrictiveness	Affection	Vulnerability	Justice	Acknowledgment
Fathers							
Restrictiveness	4.34	1.92					
Affection	7.90	1.68	-0.31				
Vulnerability	7.58	1.51	-0.23	0.71			
Justice	7.92	1.72	-0.39	0.78	0.74		
Acknowledgment	8.09	1.59	-0.36	0.79	0.72	0.88	
Trust	7.24	1.81	-0.27	0.70	0.66	0.78	0.76
Mothers							
Restrictiveness	4.30	1.90					
Affection	8.31	1.45	-0.26				
Vulnerability	7.91	1.40	-0.17	0.70			
Justice	8.17	1.57	-0.34	0.75	0.69		
Acknowledgment	8.34	1.45	-0.31	0.78	0.69	0.83	
Trust	8.00	1.59	-0.26	0.66	0.66	0.74	0.75

M=Means; SD=standard deviation; r=Correlation coefficient. Note: all correlations are significant at the $\alpha \leq 0.01$ level

It is apparent from Table 1 that the restrictiveness factor of the AFRT has received the lowest means for mothers ($M=4.30$) and fathers ($M=4.34$). However, higher means of affection, vulnerability, justice, acknowledgment, and trust were reported. Furthermore, the Pearson correlation coefficients ranged between -0.23 (between restrictiveness and trust) and 0.88 (between justice and acknowledgment) for fathers. For mothers, it ranged between -0.26 (between restrictiveness and affection and between restrictiveness and trust) and 0.83 (between justice and acknowledgment). Pallant [33] and Tabachnick and Fidell [35] indicated that correlations between 0.80 and 0.90 suggest multicollinearity. Table 1 clearly shows that all correlations are under 0.80 except for those between justice and acknowledgment in the mothers' and the fathers' ratings. Different correlation coefficients are obtained using different methods [36]. Hence, the researchers decided to proceed to the CFA before making any decision. Also, using the same data, justice and acknowledgment were found to correlate distinctly with children's externalizing and internalizing behavioral problems. In addition, significant gender and age differences were reported in these two factors [30], indicating possibly two distinct factors.

4.2. Confirmatory factor analyses results of AFRT

To answer the first research question, CFA was used to examine the factor structure of the AFRT. An essential assumption of CFA is that the sample size should be adequate. An appropriate and representative sample size was selected based on statistical tables [37]. Moreover, it is proposed that 10 respondents in each CFA model are required for each parameter [32]. Like Skoczeń *et al.* [3] and Fitriana [5], item parceling was used to achieve an acceptable ratio of parameter estimates to sample size. Also, item parceling is preferred when each scale (i.e., factor) includes more than five items since it simplifies the model [3]. Three parcels were formed under each factor using the content-driven method [10]. To ensure that each subscale was unidimensional before grouping items into parcels, EFA was performed for each subscale once for the father-child model and once for the mother-child model. The findings confirmed the unidimensionality of each subscale in both models. Eigenvalue, total variance explained, and range of loadings were respectively as: restrictiveness-father (3.69, 36.93, 0.32-0.70); restrictiveness-mother (3.69, 36.90, 0.48-0.73); affection-father (4.85, 44.12, 0.40-0.76); affection-mother (4.33, 39.42, 0.34-0.74); vulnerability-father (3.46, 34.59, 0.32-0.70); vulnerability-mother (3.27, 32.76; 0.33-0.72); justice-father (4.38, 48.77, 0.60-0.77); justice-mother (4.12, 45.77, 0.59-0.72); acknowledgment-father (4.92, 44.73, 0.56-0.76); acknowledgment-mother (4.73, 43.00, 0.50-0.75); trust-father (4.00, 44.49, 0.42-0.76); trust-mother (3.67, 40.80, 0.43-0.70).

Considering the size of the item loading, Zyl and Klooster [38] noted that an item loading ≥ 0.35 is acceptable. The first-order model consisting of the six factors: restrictiveness, affection, vulnerability, justice, acknowledgment, and trust, and three parcels under each factor was explored. Then, the second-order factor model was examined, involving two second-order factors: control, including restrictiveness and justice, and support, including the other four factors, as demonstrated in the conceptual models (i.e., Figures 1 and 2); AMOS software version 23 was utilized.

4.2.1. First-order confirmatory factor analyses model

The first-order factor model for the father-child dyad and the mother-child dyad demonstrated good fit indices that fall within the acceptable range (i.e., $\chi^2/df < 5$; normal fit index (NFI), incremental fit index (IFI), Tucker-Lewis index (TLI), CFI, GFI > 0.90 , RMSEA, and SRMR < 0.08) [38], [39]. In addition, all factor loadings were high, ranging from 0.65 to 0.89 for the father-child model and from 0.60 to 0.88 for the mother-child model. However, one remarkable problem was uncovered in the first-order CFA models of fathers and mothers. Correlations between some factors were very high, which was unacceptable and suggested multicollinearity. Those high correlations between some factors suggest that the students may have experienced difficulty differentiating between these constructs. Similar high correlations were found in previous NFRT studies [3], [5]. Therefore, the researchers decided to proceed to a second-order factor model with Support and Control as higher-order factors.

4.2.2. Second-order confirmatory factor analyses model

The second-order factor model was explored for the father-child and the mother-child dyads. Control included restrictiveness and justice, while support comprised affection, vulnerability, acknowledgment, and trust. The models showed acceptable model fit indices but some unacceptable factor loadings. Also, the correlations between the two second-order factors, support and control, were higher than 0.80 . Consequently, the researchers followed two steps: first, using modification indices suggested by AMOS, and second, considering the bi-factorial structure of the test's factors.

It was found that only justice and acknowledgment showed a bi-factorial tendency of factor loadings. Thus, the researchers concluded that AFRT has a second higher-order factor model with some

correlations allowed between errors. Moreover, acknowledgment and justice showed a bifactorial structure. Figures 4 and 5 displays the final higher second-order factor models of the father-child and the mother-child dyads.

It is evident from the father-child model in Figure 4 and the mother-child model in Figure 5 that the justice and acknowledgment factors loaded on both control and support factors. That is, although justice was theoretically supposed to load on the control factor, in the current results, it showed higher validity (i.e., higher loading) into the support factor (loading=0.60 in the father-child model, and 0.66 in the mother-child model) and lower validity (i.e., lower loading) into the control factor (loading=-0.48 in the father-child model, and 0.41 in the mother-child model).

Regarding acknowledgment, it showed higher loading (i.e., higher validity) on its higher-order factor, support (loading=0.73 in the father-child model and 0.80 in the mother-child model), which aligns with the hypothesized conceptual model. However, it shared small validity with the control factor (loading=0.32 in the father-child model and 0.24 in the mother-child model). In addition, in both models, only a few correlations between errors were allowed to improve model fit, with three in the father-child model and two in the mother-child model. All correlations were within the second-order factor support, which is logical. Moreover, the correlations between the two higher-order factors, support and control, were -0.77 in the father-child model and -0.71 in the mother-child model, suggesting the absence of multicollinearity.

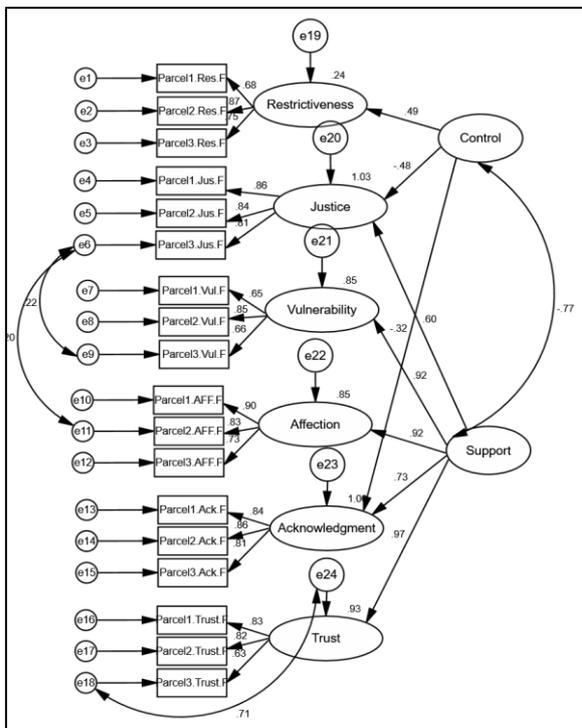


Figure 4. Final second-order factor model: father-child dyad

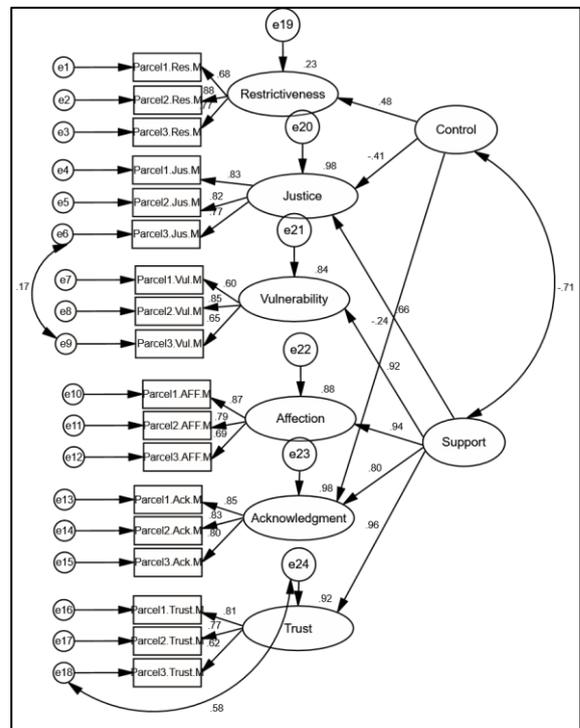


Figure 5. Final second-order factor model: mother-child dyad

Table 2 presents the fit indices for the final second-order factor models of the father-child and mother-child dyads, all within the accepted range [38], [39]. In addition, Table 3 presents the first-order factor loadings on their respective second-order factors, along with the average variance extracted (AVE) for both first- and second-order factors. All factor loadings were higher than 0.35, suggesting acceptable loadings [39]. The only low factor loadings were the acknowledgment loading on the control factor for the father-child and mother-child models. Nevertheless, this loading cannot be ignored, as it constitutes a significant part of the model structure and is the best model the researchers can obtain using the current data. Moreover, AVE values for the first- and second-order factors were above 0.50, except for control in the mother-child model, which was exactly 0.50 and still acceptable [32]. This provides evidence for the AFRT's convergent validity.

In addition, discriminant validity was examined through Fornell and Larcker's criteria [40]. In the father-child model, the correlation between support and control was -0.77, yielding a square correlation of 0.59. As can be seen in Table 3, AVE for support (0.70) and control (0.60) were higher than the squared correlation, which suggests good discriminant validity. Regarding the mother-child model, the correlation between support and control was -0.71, yielding a square correlation of 0.50. It is clear from Table 3 that the AVE for support (0.74) exceeds the square correlation, whereas the AVE of Control is the same, 0.50; yet it can still be acceptable. These results indicate good discriminant validity of the AFRT.

Overall, the AFRT demonstrated a good factorial structure, similar to that found in previous test versions. Using CFA, the AFRT yielded a valid 60-item scale with a two-second-order-factor, six-first-order-factor factorial structure. The first second-order factor was control, which consisted of restrictiveness, moderate variance of justice, and small variance of acknowledgment. The other second-order factor was support, which included affection, vulnerability, trust, high variance of justice, and high variance of acknowledgment.

Table 2. Fit indices of the final second-order factor models

No.	Fit indices	Father-child	Mother-child
1	χ^2 (<i>p</i> value)	348.14 (0.000)	415.96 (0.000)
2	χ^2/df	2.83	3.32
3	NFI	0.96	0.94
4	IFI	0.97	0.96
5	TLI	0.96	0.95
6	CFI	0.97	0.96
7	RMSEA	0.05	0.06
8	SRMR	0.03	0.04

Table 3. Factor loadings of the final second-order factor models and AVE

Models	Factor loadings	AVE	Factor loading	AVE
Father-child model	Control	0.60	Support	0.70
Restrictiveness	0.49	0.60		
Justice	-0.48	0.70	0.60	
Vulnerability		0.53	0.92	
Affection		0.68	0.92	
Acknowledgment	-0.32	0.70	0.73	
Trust			0.97	
Mother-child model	Control	0.50	Support	0.74
Restrictiveness	0.48	0.60		
Justice	-0.41	0.81	0.66	
Vulnerability		0.70	0.92	
Affection		0.62	0.94	
Acknowledgment	-0.24	0.68	0.80	
Trust		0.73	0.96	

4.3. Measurement invariance results of AFRT

To answer the second research question, AMOS software was also used to test the measurement invariance structure of the AFRT for both models, the father-child and the mother-child, as part of the construct validity. Measurement invariance was used to examine whether the test's factorial structure is the same across different groups by gender and age (i.e., by grade: 5, 7, 9, and 11). First, configural, metric, and scalar invariance were examined in the first-order models of the father-child and mother-child dyads. Then, structural weights (i.e., factor loadings) in the second-order factor model were fixed to test whether the second-order factor model has the same meaning across gender and age groups.

Moreover, the differences in goodness of fit indices (Δ GFIs) between the compared models were utilized as criteria to test whether the models were invariant. Particularly, the cut of criteria followed were: a change of ≥ 0.010 in CFI, combined with a change of ≥ 0.015 in RMSEA or a change of ≥ 0.030 in SRMR, indicated non-invariance [15]. More stress was put on the CFI because the model complexity and sample size influence RMSEA and SRMR [14]. Table 4 shows the results of the measurement invariance tests for the father-child model, while Table 5 presents the measurement invariance results of the mother-child model. Although χ^2 value is not used as a criterion, it is presented here to specify the models' information.

For the father-child model, configural invariance was supported. The results showed that all parcels loaded on their factors with high factor loadings, and all models manifested good fit indices. Moreover, as shown in Table 4, metric invariance at the first-order model was achieved across gender and age groups, as changes in CFI (0.001), RMSEA (0.001), and SRMR (0.003) between the metric and configural levels

were all below the cut-off criteria. It is also clear that scalar invariance was achieved, as changes in CFI (0.009), RMSEA (0.004), and SRMR (0.012) between the scalar and metric levels were all below the cut-off criteria. This indicates that the first-order model with the six AFRT factors is invariant across gender and age groups in the father-child model. Table 4 highlights that measurement invariance in the second-order models was not achieved across genders or ages. Changes between the structural weight level and the scalar level in CFI (0.016) (across genders); and changes in CFI (0.013) and SRMR (0.048) (across ages) were higher than the cut-off criteria.

Similarly, in the mother-child model, all models exhibited the same configural (i.e., factorial) structure, with acceptable factor loadings and fit indices, thereby achieving configural invariance. However, it is clear from Table 5 that measurement invariance at both the metric level and the scalar level was not met across models of gender since differences in CFI (0.016), RMSEA (0.042), and SRMR (0.085) were all higher than the cut-off criteria between the metric level and the configural level. Also, the change in CFI (0.019) between the scalar and metric levels was greater than 0.010, indicating that measurement invariance was not met at the scalar level across gender groups. Regarding second-order models across gender, measurement invariance at the structural weight level was not met, as changes in CFI (0.037) and SRMR (0.084) exceeded the cut-off criteria.

Regarding models across age groups, Table 5 shows that configural, metric, and scalar invariance was achieved in the first-order models, as changes in CFI, RMSEA, and SRMR were all below the cut-off criteria. In addition, measurement invariance at the structural weight level was also met in the second-order model, as the changes in CFI (0.006) and RMSEA (0.000) were within accepted criteria. Although the difference in SRMR was higher than 0.030, the accepted difference in CFI, combined with the difference in RMSEA, are sufficient to achieve measurement invariance across models.

Table 4. Measurement invariance results for the father-child model

Model	χ^2 (df)	CFI	CFI difference	RMSEA	RMSEA difference	SRMR-S	SRMR difference
Invariance across gender							
Configural invariance	558.61 (240)	0.959		0.047		0.043	
Metric invariance at the first order level	590.47 (258)	0.958	0.001	0.046	0.001	0.046	0.003
Scalar invariance at the first order level	674.96 (273)	0.949	0.009	0.050	0.004	0.058	0.012
Structural weights (equality of loadings at the second order part)	530.76 (254)	0.965	0.016*	0.042	0.008	0.050	0.008
Invariance across ages							
Configural invariance	971.18 (480)	0.939		0.041		0.067	
Metric invariance at the first order level	1047.48 (534)	0.936	0.003	0.040	0.001	0.092	0.025
Scalar invariance at the first order level	1145.87 (579)	0.930	0.006	0.040	0.000	0.099	0.007
Structural weights (equality of loadings at the second order part)	896.47 (460)	0.943	0.013*	0.040	0.000	0.051	0.048*

*the values did not meet the criteria of measurement in variance

Table 5. Measurement invariance results of the mother-child model

Model	χ^2 (df)	CFI	CFI difference	RMSEA	RMSEA difference	SRMR-S	SRMR difference
Invariance across gender							
Configural invariance	619.38 (240)	0.945		0.051		0.042	
Metric invariance at the first-order level	753.27 (258)	0.929	0.016*	0.057	0.042*	0.127	0.085*
Scalar invariance at the first-order level	904.44 (279)	0.910	0.019*	0.061	0.004	0.129	0.002
Structural weights (equality of loadings at the second-order part)	619.87 (254)	0.947	0.037*	0.050	0.011	0.045	0.084*
Invariance across ages							
Configural invariance	493.33 (480)	0.937		0.040		0.064	
Metric invariance at the first-order level	1026.99 (534)	0.933	0.004	0.039	0.001	0.090	0.026
Scalar invariance at the first-order level	1124.38 (579)	0.925	0.008	0.040	0.001	0.106	0.016
Structural weights (equality of loadings at the second-order part)	1024.34 (518)	0.931	0.006	0.040	0.000	0.072	0.034*

*the values did not meet the criteria of measurement in variance

4.4. Reliability indicators of AFRT

To answer the third research question, reliability indicators were investigated for the six first-order factors. Three methods were used to check the reliability of AFRT factors: Cronbach's alpha (α), test-retest, and split-half. A test-retest was done for 78 (43.6% females, 56.4% males) students from the pilot study sample, as described in the instrument section. Table 6 demonstrates the reliability indicators of the six factors in the AFRT for both the father-child and mother-child models. Cronbach alpha coefficients were all >0.70 , ranging from 0.87 (vulnerability) to 0.87 (justice and acknowledgment) in the father-child model and between 0.75 (vulnerability) and 0.87 (acknowledgment) in the mother-child model. Not very different: all six factors showed robust test-retest correlations ranging from 0.63 (trust) to 0.79 (affection) in the father-child model and from 0.69 (trust) to 0.84 (affection) in the mother-child model. Regarding split-half reliability, all Guttman coefficients were excellent. The coefficients ranged between 0.70 (restrictiveness) and 0.83 (affection and acknowledgment) in the father-child model and between 0.69 (restrictiveness and vulnerability) and 0.81 (acknowledgment) in the mother-child model. Considering CR values, all were ≥ 0.70 , indicating good CR.

Table 6. Reliability indicators of the final version of AFRT

Factors	Father-child model				Mother-child model			
	Cronbach alpha	Test-retest	Guttman split-half coefficient	CR	Cronbach alpha	Test-retest	Guttman split-half coefficient	CR
Restrictiveness	0.80	0.76	0.70	0.81	0.80	0.77	0.69	0.82
Affection	0.86	0.79	0.83	0.86	0.82	0.84	0.78	0.82
Vulnerability	0.77	0.65	0.73	0.76	0.75	0.74	0.69	0.74
Justice	0.87	0.70	0.81	0.87	0.85	0.81	0.79	0.85
Acknowledgment	0.87	0.76	0.83	0.87	0.86	0.78	0.81	0.86
Trust	0.84	0.63	0.74	0.80	0.82	0.69	0.72	0.77

5. DISCUSSION

This research focused on developing and validating the AFRT. The test, designed in computerized form, involves six main factors reflecting the interactions between the parents and their children: restrictiveness, affection, vulnerability, acknowledgment, justice, and trust. The researchers focused primarily on the test's psychometric properties and applicability in the Arab context, specifically Oman.

5.1. Factor structure of the AFRT

CFA was used to check the factorial structure of the AFRT. The results indicated that the final models for the father-child and the mother-child dyads were very similar. The final factor structure of AFRT was revealed as a second higher-order model with control including restrictiveness, medium variance of justice, and small variance of acknowledgment, and support consisting of affection, vulnerability, trust, high variance of justice, and high variance of acknowledgment. Thus, justice and acknowledgment showed a bifactorial structure. Furthermore, all first-order factors loaded significantly with high factor loadings into their respective second-order factors. Very good fit indices were reported for both the father-child and the mother-child models. In addition, correlations between control and support were acceptable and did not indicate multicollinearity. Convergent validity was also supported by AVE values, and discriminant validity was supported through Fornell and Larcker's criteria [40].

It was clear from the CFA analysis that restrictiveness was a distinctive factor compared to the rest of the five factors since it showed low and negative correlations with them in both the father-child and the mother-child models. This type of correlation is typical, as restrictiveness is the only factor that reflects negative practices in parent-child relationships, as it conveys parents' dominant rules and firm control over their children [15]. On the other hand, all other factors —affection, acknowledgment, vulnerability, justice, trust —reflect positive practices in family relationships, with some controlling practices conveyed by justice and acknowledgment, as suggested by the test's factor structure. Hence, restrictiveness positively loaded onto the second-order factor control, while justice and acknowledgment negatively loaded onto control.

The negative correlation between parents' restrictiveness and other parenting styles coincides with previous studies' results [3]–[5], [9]. In a similar vein, authoritarian parenting in Arabic studies negatively correlated with other styles, such as authoritative and permissive [18]. These results show that restrictive parents apply firmer practices with their children than other parents. Nevertheless, the other five factors (affection, acknowledgment, vulnerability, justice, and trust) were highly correlated, suggesting they were not very distinguishable by the children as they shared very similar characteristics. Previous researchers also identified the high correlations between these factors, so they had to either merge some factors or look for the possibility of a higher-order factor model [3], [5], [10].

The Loyalty theory [8] underlines how these aspects of family relationships are interconnected. According to this theory, children feel their parents are just with them when they are treated fairly within the family's rules. In addition, when parents provide acknowledgment, love, and support to their children, children can feel trustworthy and vulnerable toward their parents. On the other hand, by closely attending to other family patterns, such as uncertainty and injustice, children in this family atmosphere feel lost. They are treated good or bad (i.e., rewarded or punished) for the same behavior, by uncertain parents. Consequently, affection, vulnerability, justice, and trust are also connected but negatively. Children do not find a clear pattern of family interaction within these families. They are often confused. They may be acknowledged once and not at other times; they might trust their parents at some events but not at others, and so on.

The high correlations between these factors were not very different from those reported in the original NFRT developed by Oud and Welzen [4], who reported some correlations exceeding 0.80. Therefore, this study adopted the second-order higher factor structure to obtain a more precise, unique-factor structure for the AFRT. This implies that the five factors- affection, acknowledgment, vulnerability, justice, and trust- measure one higher-order factor called support. All these factors express support, love, and affection as practiced by parents toward their children. However, restrictiveness reflects control practices by parents, along with some medium validity of Justice and small validity of acknowledgment.

The current findings align with Fitriana [5], who proposed that vulnerability and justice shared some validity with both control and support. However, the current study context, characteristics of the sample, and culture have yielded different bi-factorial structures in which justice and acknowledgment loaded onto control with low to medium validity. On the other hand, they loaded with high validity on the support factor.

A justification for the bi-factorial tendency of the justice factor can be offered. In applying the AFRT, all Justice items were reworded to measure only positive practices; that is, the negatively phrased items referring to injustice in the original pool were rephrased to represent justice in an affirmative form. During the pilot testing, children demonstrated noticeable difficulty comprehending and responding accurately to the reverse-coded items. Such items increased cognitive load and led to misunderstandings, particularly among younger respondents with developing reading and reasoning abilities. After rephrasing these items into straightforward, positively worded statements, both item comprehension and response consistency improved. Psychometrically, this adjustment resulted in higher item-total correlations and clearer factor loadings, indicating improved construct coherence. Consequently, the justice subscale reflected a more homogeneous and interpretable dimension, which also accounted for its stronger association with the support factor than with control.

The acknowledgment factor was highly correlated with the support factor, as theoretically expected, and showed a tiny correlation with control. The bi-factorial tendency of this factor requires further investigation. That is, the validity shared by acknowledgment to control was minimal. Theoretically, the low extreme of acknowledgment items reflects non-acknowledging rather than controlling practices. However, the high correlation among the five first-order factors initially identified could explain why the acknowledgment items were pooled with the justice items into the control factor. In addition, the statistical validity of some loadings may not always reflect practical and theoretical validity, especially when validity is low. Hence, future research must check the bi-factorial tendency of both justice and acknowledgment.

However, one unnoticeable justification for this result in Arabic cultures is that some controlling practices may be positive and lead to positive outcomes for children. Some Omani studies [18] found that the most prevalent parenting styles among Omani families are authoritarian and authoritative. Where authoritative parents are high in demandingness and responsiveness, authoritarian parents are high in demandingness but low in responsiveness. This suggests that most Omani parents tend to use demanding practices with their children, yet the demandingness may be expressed positively or negatively. Authoritative demandingness is associated with responsiveness; parents care for their children's feelings, allow autonomy and decision-making, and give space for discussion of the role. Conversely, authoritarian demandingness is more controlling, does not allow any discussion of the rules, and expects obedience only. Furthermore, various Eastern studies [18], [41], [42] found positive effects of authoritarian parenting styles on children's outcomes such as academic achievement and moral intelligence. As highlighted by Ghandour *et al.* [43], unlike European and American studies, which reported negative impacts of authoritarian parenting, Arab and Asian studies found inconsistent results. In some studies, authoritarian parents are not always perceived negatively and are not connected with adverse psychological outcomes. Briefly, the limited validity of acknowledgment on control cannot be ignored, as it can also theoretically suggest positive parental control practices towards their children.

5.2. Measurement invariance of the AFRT across groups of gender and age

As part of examining the construct validity of the AFRT, measurement invariance was assessed across gender and age groups (i.e., grades 5, 7, 9, and 11). Measurement invariance was examined for both

the father-child and the mother-child models. At the first-order factor model, configural, metric, and scalar invariance were tested. For the second-order model, structural weight invariance was tested.

The results showed that, in the father-child model, measurement invariance was achieved at the configural, metric, and scalar levels of the first-order model across age and gender groups. However, the structural weight level of the second-order model was found to be non-invariant across gender and age groups. On the other hand, in the mother-child model, measurement invariance was not supported at the metric and scalar levels across gender groups; however, it was supported across age groups. In addition, the second-order model's structural weight level was invariant across age groups.

Few studies have explored measurement invariance in previous versions of the family relation tests [3], [6]. There is a scarcity of research investigating the measurement invariance of family relationship tests across groups of children of different ages and genders. Consequently, it remains unclear whether the findings of the present study are consistent with those of previous versions of the family relationship test regarding measurement invariance. As Brown [11] explained, measurement invariance means measurement stability across different informants or periods. Hence, before making further comparisons between groups in any traits, skills, or anything, it is crucial to ensure that all group members perceive or understand the test similarly.

The current results showed that students' understanding of the test items regarding their relations with their fathers was equal. To illustrate, the first six-order factors measured by the AFRT were perceived similarly by all students regardless of age or gender. This is an advantage of the test, allowing further comparisons between these groups to explore how they differ in their perceptions of their relationships with fathers, without worrying about differences in the construct itself.

In contrast, measurement invariance was not supported in mother-child relationships across gender groups, but it was endorsed across age groups. Hence, it can be said that male and female children did not comprehend the items equally in their relationships with their mothers; thus, any differences in this regard cannot be attributed to their different personalities or natural characteristics, but only to their perceptions of the test itself. One plausible rationale for the non-invariance in gender groups in mother-child relationships might be differences in how males and females perceive their relationships with their mothers. Previous studies showed evidence of differences between male and female children in their relationships with their mothers [43], [44]. However, it is essential to highlight that the differences here are related to males' and females' understandings of the items themselves. To elaborate, the item "My father/mother is happy when I embrace him/her" could be interpreted differently by male and female children. For instance, unlike female children, male children may not care much about embracing their mother or feel shy about it; they would not pay much attention to this item or consider it irrelevant to their relationships with their mother. However, when it comes to female children, the case is different. Female children appreciate the value of embracing and emphasize its importance in their relationships with their mothers.

On the other hand, the narrow age gaps between the age groups included in this study (i.e., skipping only one school grade) have contributed to the achievement of measurement invariance in both the father-child and the mother-child models. Florean *et al.* [45] emphasized the role of the social, physical, and cognitive characteristics of children included in the comparison groups. As long as these differences are minimal, the measurement invariance is more likely to be achieved. In their study, Florean *et al.* [45] argued that the developmental characteristics of the participants' preadolescent and adolescent groups may have contributed to the inequality in intercepts across some items. Confounding variables other than age alone, such as developmental or cognitive characteristics, might lead to a variant scale structure. Hence, upcoming studies must attempt to examine the measurement invariance of the AFRT in the father-child and the mother-child models across other groups of age, who are assigned larger age gaps.

In addition, the researchers recommend that future studies investigate the measurement invariance of the AFRT, especially in the mother-child relationships, at all possibilities. Differences such as individual items, covariances, or errors must be discovered at parameter levels. Hence, researchers can reach an overall final factorial model equal to all underlying groups. For instance, some items, covariances, or errors might be deleted or made free by allowing some correlations. Some statistical studies utilize different techniques to improve the factorial structure and make it valid across groups of comparisons. An example of these techniques is the IRT-based linking used by some researchers [46]. However, the purpose of this study was only to reach a final structure that is almost identical or very similar across the father-child and mother-child models. Changing the structure of each model separately requires different techniques outside the current study's scope. In addition, achieving measurement invariance at all levels in the father-child model across age and gender groups, and invariant structure across age groups in the mother-child model, is considered a fair step that does not overlook the fact that AFRT is appropriate for most of the examined groups.

Concerning the second-order factor model with control and support factors, the findings revealed that this structure was variant across age and gender groups in the father-child model; yet, it was invariant

across age groups only, not gender groups, in the mother-child model. Hence, the researchers strongly recommend further examination of measurement invariance at the higher-order level of the test.

5.3. Reliability indicators of the AFRT

Reliability indicators of the AFRT were investigated using Cronbach's alpha, test-retest, split-half, and CR. The final reliability indicators, as reported by these four methods, were very good, indicating the stability of the measurement. The strong reliability indicators provided in this study align with earlier studies examining other versions of the family relation tests, such as NFRT, BFRT, and CFRT [3]–[6], [9]. The positive reliability indicators shown by the AFRT imply its applicability to other sample groups in the Arab context.

However, careful attention should be given to the test's sub-scales related to higher-order factors: support and control. In the present study, the bi-factorial structure of justice and acknowledgment hindered the assessment of each higher-order factor's reliability. Thus, some extra examination should be done with another set of data to check the replicable tendency of justice and acknowledgment to be bifactorial again. Given that the test's reliability is sample-dependent, as it is a property of test scores rather than the test itself, different groups of participants may yield different reliability results [47].

5.4. Limitations and suggestions for future research

The current researchers acknowledge some limitations; in turn, suggestions for future research are provided. For instance, regarding the high correlations between the first-order factors and the bifactorial tendency, other researchers can use different strategies to further investigate this result. One plausible explanation for this issue is children's inability to distinguish the meanings of some items, which may affect the scale's factor structure. The researchers suggest conducting further studies that compare the current research's results using different data collection or test administration methods. Using the computerized format in this study may influence respondents' reading quality, understanding, and perception of the items due to several factors, including screen clarity, cognitive load, and digital literacy [48]–[50]. Future research can compare the psychometric properties of AFRT in paper-based and computer-based formats. Moreover, a one-to-one method can be used, with the test administered to the children in a semi-interview format. This method was applied by previous researchers [5], [15]. Furthermore, testing the test across different age groups, such as older age groups, might yield different results. Also, the methodological and statistical techniques used in the current study can be revisited with different attempts; for example, comparing continuous response-scale results with discrete Likert-scale results, and CFA results with item parceling versus individual-item techniques [10]. Future studies are also encouraged to begin with an EFA analysis to explore possible alternative test structures. Another limitation of this study is the dependence on a self-report measure. Self-report scales are known to have some limitations, such as social desirability, recall bias, response style, and the common method bias [51], [52]. Therefore, further studies must target other informants' perspectives. For instance, other studies used reports from informants such as parents, caregivers, teachers, or even grandparents to measure family relationships [19], [53], [54]. In most of these studies, two or more perspectives are compared, and some contradictory results are found in some cases.

Another important point to consider is using a wider variety of research designs to study family relationships. This study relied on a quantitative, cross-sectional approach because its main goal was to create and validate a test suited to Arabic and Omani contexts. However, the researcher encourages future studies to adopt other approaches, such as qualitative methods (like interviews), experimental or quasi-experimental designs, and longitudinal studies that track the same participants over time. The AFRT could serve as a valuable tool in these different research settings. Using multiple research designs would not only help test and refine various psychometric properties but also allow researchers to explore family relationships from different angles and purposes. For instance, a longitudinal study could reveal how family relationships change across different ages or generations, offering deeper insights into the factors that shape them.

5.5. Implications

The results have implications for both theoretical and practical aspects. First, the AFRT is an appropriate new measurement tool for Arabic and cross-cultural family assessment. Examining this framework within the Arabic culture adds to the Arabic literature, theories, and assessment techniques. Also, it allows for comparisons between the Arabic context and other contexts. Second, the AFRT is a beneficial test for family members, counselors, or educators in various institutions in Oman, such as the Ministry of Social Development, the MoE, and counseling centers at higher education institutions or other private centers. This test can be used to assess family relationship issues and problems, and to diagnose the problems that school students face that might be attributed to the family environment. Third, the AFRT is designed to explore relationships inside families, not limited to father-child or mother-child dyads. It is also adaptable to other family relationships, such as husband-wife and sibling relationships [4].

6. CONCLUSION

The current study contributed to the family assessment literature by introducing a new measure, the AFRT, with six first-order and two second-order factors. The test can be used in future research, as well as in family relationships assessment and interventions. In addition, it can be adapted to different age groups and family relationship dyads. However, as a new test, further examination of its psychometric properties and factorial structure is essential, especially given its higher-order structure and the bifactorial nature of the two factors: justice and acknowledgment.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Marwa Nasser Alrajhi	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓
Kamisah Osman		✓				✓			✓		✓	✓		
Hussain Ali Alkharusi		✓			✓	✓				✓		✓		

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project administration

Fu : Funding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

INFORMED CONSENT

We have obtained informed consent from all individuals included in this study.

DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [MNA]. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

REFERENCES

- [1] A. Christopher, "Attributes of family relations," *Psychology and Education Journal*, vol. 58, no. 1, pp. 3030–3037, Jan. 2021, doi: 10.17762/pae.v58i1.1204.
- [2] H. Goldenberg and I. Goldenberg, "The revolution and evolution of family therapy and family psychology," in *The Wiley Blackwell Handbook of Family Psychology*, J. H. Bray and M. E. Stanton, Eds., Hoboken, NJ: John Wiley & Sons, 2009, pp. 21–36, doi: 10.1002/9781444310238.ch2.
- [3] I. Skoczeń, J. Ciecuch, J. Oud, and K. Welzen, "Development and validation of the computerized family relations test for children," *Frontiers in Psychology*, vol. 6, Nov. 2015, doi: 10.3389/fpsyg.2015.01687.
- [4] H. Oud and K. Welzen, *Nijmegen family relationship test manual*. Lisse, Netherlands: Swets & Zeitlinger, (in Dutch), 1989.
- [5] E. Fitriana, "Confirmatory factor analysis of the Bandung Family Relation Test: a simulation study comparing ML, DWLS and WLS estimation," Ph.D. dissertation, Radboud University Nijmegen, Nijmegen, Netherlands, 2011.
- [6] T. Toharudin *et al.*, "Testing measurement invariance across SES groups: application in the Bandung family relation test," *Journal of Mathematics and Computational Sciences*, vol. 11, no. 4, pp. 4314–4331, Jan. 2021, doi: 10.28919/jmcs/5587.
- [7] S. Haxhe, "Parentification and related processes: distinction and implications for clinical practice," *Journal of Family Psychotherapy*, vol. 27, no. 2, pp. 188–201, 2016, doi: 10.1080/08975353.2016.1199768.
- [8] I. Boszormenyi-Nagy and G. M. Spark, *Invisible loyalties: reciprocity in intergenerational family therapy*, 2nd ed. New York: Brunner/Mazel, 1984.
- [9] I. Skoczeń and J. Ciecuch, "Polish computerized short version of The Nijmegen Family Relation Test (NFRT)," in *The Psychology of Human Development*, D. Musial and E. Rydz, Eds., Lublin: KUL, 2012, pp. 61–78.

- [10] S. K. Sterba and J. D. Rights, "Item parceling in SEM: a researcher degree-of-freedom ripe for opportunistic use," in *Handbook of Structural Equation Modelling*, 2nd ed. R. H. Hoyle, Ed., New York: The Guilford Press, 2023, pp. 296–315.
- [11] T. A. Brown, *Confirmatory factor analysis for applied research*, 2nd ed. New York: The Guilford Press, 2015.
- [12] J. B. Ullman and P. M. Bentler, "Structural equation modeling," in *Handbook of Psychology*, 2nd ed., J. A. Schinka, W. F. Velicer, and I. B. Weiner, Eds., Hoboken, NJ: John Wiley & Sons, Inc., 2012, pp. 661–690, doi: 10.1002/9781118133880.hop202023.
- [13] M. Rudnev, E. Lytkina, E. Davidov, P. Schmidt, and A. Zick, "Testing measurement invariance for a second-order factor: a cross-national test of the alienation scale," *Methods, Data, Analyses*, vol. 12, no. 1, pp. 47–76, 2018, doi: 10.12758/mda.2017.11.
- [14] F. F. Chen, "Sensitivity of goodness of fit indexes to lack of measurement invariance," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 14, no. 3, pp. 464–504, Jul. 2007, doi: 10.1080/10705510701301834.
- [15] M. Moeliono et al., *Quick reference guide: Bandung Family Relation Test (BFRT)*. Jakarta: PT Mediabaca Mandiri, (in Indonesian), 2018.
- [16] A. Khan, "Female socialization in the Omani Oases and the impacts of modernization after 1970," M.S. thesis, University of Central Lancashire, Preston, England, United Kingdom, 2021.
- [17] D. Baumrind, R. E. Larzelere, and E. B. Owens, "Effects of preschool parents' power assertive patterns and practices on adolescent development," *Parenting: Science and Practice*, vol. 10, no. 3, pp. 157–201, 2010, doi: 10.1080/15295190903290790.
- [18] A. Alhadabi, S. Aldhafri, H. Alkharusi, I. Al-Harthi, M. Alrajhi, and H. AlBarashdi, "Modelling parenting styles, moral intelligence, academic self-efficacy and learning motivation among adolescents in grades 7–11," *Asia Pacific Journal of Education*, vol. 39, no. 1, pp. 133–153, Jan. 2019, doi: 10.1080/02188791.2019.1575795.
- [19] S. Aldhafri, "Children's perceptions of parenting styles: ten years of research with Omani families," in *Families and Social Change in the Gulf Region*, J. E. Lansford, A. B. Brik, and A. M. Badahdah, Eds., New York: Routledge, 2020, pp. 60–79.
- [20] M. A. Shahrier, M. N. Islam, K. I. Kayesh, and M. K. Rita, "The relationship between academic performance and family relation of secondary school students," *Journal of Life Earth Science*, vol. 10, pp. 21–29, Jan. 2015.
- [21] Y. Yan et al., "The relationship between family functioning and depressive symptoms: mediating effects of psychological resilience and parent-child interactions," *Journal of Affective Disorders*, vol. 385, p. 119383, 2025, doi: 10.1016/j.jad.2025.119383.
- [22] M. Valentic, L. Simetin, D. Mayer, F. Simetin, and I. Pavic, "The impact of parental monitoring on exposure to multiple substances and bullying in Croatian students," *Children*, vol. 11, no. 11, p. 1292, Oct. 2024, doi: 10.3390/children11111292.
- [23] L. G. Speyer, H. A. Hall, Y. Hang, C. Hughes, and A. L. Murray, "Within-family relations of mental health problems across childhood and adolescence," *Journal of Child Psychology and Psychiatry*, vol. 63, no. 11, pp. 1288–1296, Jan. 2022, doi: 10.1111/jcpp.13572.
- [24] J. R. Buri, "Parental authority questionnaire," *Journal of Personality Assessment*, vol. 57, no. 1, pp. 110–119, 1991, doi: 10.1207/s15327752jpa5701_13.
- [25] M. A. Aljaffer et al., "Development and validation of an Arabic-language family stability scale for women in Saudi Arabia," *Frontiers in Public Health*, vol. 13, p. 1698456, 2025, doi: 10.3389/fpubh.2025.1698456.
- [26] A. Soliman, A.-S. G. Abdel-Salam, and M. Ahmed, "The reliability, validity and normative scores of the bene-anthony family relations test for use with Arab children," *Frontiers in Psychology*, vol. 12, p. 548493, Jan. 2021, doi: 10.3389/fpsyg.2021.548493.
- [27] S. Tripodi and K. Bender, "Descriptive studies," in *The Handbook of Social Research Methods*, 2nd ed., B. Thyer, Ed., Thousand Oaks, CA: SAGE Publications, Inc., 2010, pp. 120–130.
- [28] International Test Commission, *The ITC Guidelines for Translating and Adapting Tests (Second Edition)*. 2017. [Online]. Available: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf
- [29] M. Romero-Jeldres, E. Diaz Costa, and T. F. Nadim, "A review of Lawshe's method for calculating content validity in the social sciences," *Frontiers in Education*, vol. 8, p. 1271335, Nov. 2023, doi: 10.3389/educ.2023.1271335.
- [30] M. Alrajhi, K. Osman, and H. Alkharusi, "Criterion-related validity of the Arabic Computerized Family Relation Test," *Dirasat: Human and Social Sciences*, vol. 53, no. 6, pp. 1–20, 2026, doi: 10.35516/Hum.2026.8793.
- [31] R. B. Kline, "Assumptions in structural equation modelling," in *Handbook of Structural Equation Modeling*, R. Hoyle, Ed., New York: The Guilford Press, 2023, pp. 128–143.
- [32] J. F. Hair, B. J. Babin, W. C. Black, and R. E. Anderson, *Multivariate data analysis*, 8th ed. Andover: Cengage Learning, 2019.
- [33] J. Pallant, *SPSS survival manual: a step by step guide to data analysis using SPSS*, 4th ed. Maidenhead: McGraw-Hill, 2010.
- [34] F. J. Gravetter and L. B. Wallnau, *Essentials of statistics for the behavioral sciences*, 9th ed. Belmont, CA: Wadsworth, 2016.
- [35] B. G. Tabachnick and L. S. Fidell, *Using multivariate statistics*, 6th ed. Harlow: Pearson Education, 2014.
- [36] K. Shao, M. E. Shirvan, and A. Alamer, "How accurate is your correlation? Different methods derive different results and different interpretations," *Frontiers in Psychology*, vol. 13, p. 901412, May 2022, doi: 10.3389/fpsyg.2022.901412.
- [37] G. D. Israel, *Determining sample size*. University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS, Florida, 1992.
- [38] L. E. van Zyl and P. M. T. Klooster, "Exploratory structural equation modeling: practical guidelines and tutorial with a convenient online tool for Mplus," *Frontiers in Psychiatry*, vol. 12, p. 795672, Jan. 2022, doi: 10.3389/fpsyg.2021.795672.
- [39] S. West, W. Wu, D. McNeish, and A. Savord, "Model fit in structural equation modeling," in *Handbook of Structural Equation Modelling*, 2nd ed. R. H. Hoyle, Ed., New York: The Guilford Press, 2023, pp. 184–205.
- [40] C. M. Voorhees, M. K. Brady, G. Calantone, and E. Ramirez, "Reporting reliability, convergent and discriminant validity with structural equation modeling: An illustrative review," *Asia Pacific Journal of Marketing and Logistics*, vol. 35, no. 6, pp. 1233–1261, 2023, doi: 10.1108/APJML-03-2022-0089.
- [41] A. M. Mihret, G. S. Dilgasa, and T. H. Mamo, "Parenting style as correlates of adolescents' academic achievement motivation of Bate secondary school, Haramaya, Ethiopia," *International Journal of Education and Literacy Studies*, vol. 7, no. 2, pp. 172–176, Apr. 2019, doi: 10.7575/aiac.ijels.v.7n.2p.172.
- [42] J. Yang and X. Zhao, "Parenting styles and children's academic performance: evidence from middle schools in China," *Children and Youth Services Review*, vol. 113, p. 105017, Apr. 2020, doi: 10.1016/j.childyouth.2020.105017.
- [43] L. A. Ghandour et al., "Parenting a high and growing population of youth in the Arab Region: a scoping review for an evidence-informed research agenda," *Journal of Adolescent Health*, vol. 75, no. 1, pp. 16–25, 2024, doi: 10.1016/j.jadohealth.2024.02.033.
- [44] K. S. J. Preston, A. W. Gottfried, J. J. Park, P. D. Manapat, A. E. Gottfried, and P. H. Oliver, "Simultaneous linking of cross-informant and longitudinal data involving positive family relationships," *Educational and Psychological Measurement*, vol. 78, no. 3, pp. 409–429, Feb. 2017, doi: 10.1177/0013164417690198.
- [45] I. S. Florean, A. Dobrean, C. R. Păsăreanu, R. D. Georgescu, and I. Milea, "The efficacy of internet-based parenting programs for children and adolescents with behavior problems: a meta-analysis of randomized clinical trials," *Clinical Child and Family Psychology Review*, vol. 23, no. 4, pp. 510–528, Sep. 2020, doi: 10.1007/s10567-020-00326-0.

- [46] G. Makransky, C. W. Schnohr, T. Torsheim, and C. Currie, "Equating the HBSC family affluence scale across survey years: a method to account for item parameter drift using the Rasch model," *Quality of Life Research*, vol. 23, no. 10, pp. 2899–2907, Jun. 2014, doi: 10.1007/s11136-014-0728-2.
- [47] S. Chakrabarty, "Reliability of test battery," *Methodological Innovations*, vol. 13, no. 2, pp. 1–8, 2020, doi: 10.1177/2059799120918340.
- [48] J. Pengelley, P. R. Whipp, and A. Malpique, "A testing load: a review of cognitive load in computer and paper-based learning and assessment," *Technology, Pedagogy and Education*, vol. 34, no. 1, pp. 1–17, 2024, doi: 10.1080/1475939X.2024.2367517.
- [49] A. Ronconi *et al.*, "Effects of reading medium on the processing, comprehension, and calibration of adolescent readers," *Computers & Education*, vol. 185, p. 104520, 2022, doi: 10.1016/j.compedu.2022.104520.
- [50] W. Yu and N. Iwashita, "Comparison of test performance on paper-based testing (PBT) and computer-based testing (CBT) by English-majored undergraduate students in China," *Language Testing in Asia*, vol. 11, no. 32, pp. 2–21, 2021, doi: 10.1186/s40468-021-00147-7.
- [51] W. L. Teh *et al.*, "Measuring social desirability bias in a multi-ethnic cohort sample: its relationship with self-reported physical activity, dietary habits, and factor structure," *BMC Public Health*, vol. 23, no. 1, p. 415, Mar. 2023, doi: 10.1186/s12889-023-15309-3.
- [52] M. L. L. Zini and G. Banfi, "A narrative literature review of bias in collecting patient-reported outcomes measures (PROMs)," *International Journal of Environmental Research and Public Health*, vol. 18, no. 23, p. 12445, Nov. 2021, doi: 10.3390/ijerph182312445.
- [53] P. Braza *et al.*, "Negative maternal and paternal parenting styles as predictors of children's behavioral problems: moderating effects of the child's sex," *Journal of Child and Family Studies*, vol. 24, no. 4, pp. 847–856, 2015, doi: 10.1007/s10826-013-9893-0.
- [54] C. van Fossen, K. J. Pratt, and I. Eneli, "Family functioning assessment in a community sample of African American caregivers and children," *Contemporary Family Therapy*, vol. 44, no. 3, pp. 244–249, Sep. 2022, doi: 10.1007/s10591-021-09604-7.

BIOGRAPHIES OF AUTHORS



Marwa Nasser Alrajhi    is a research assistant in the Research and Measurement Section at the Student Counselling Centre, Sultan Qaboos University, Oman. She has a Bachelor's Degree in Education/English and a Master's in Evaluation and Measurement. She is currently a Ph.D. student in evaluation and measurement at UKM, Malaysia. Her research interests are students' personalities, beliefs, family relations, scale construction, and validation. She has authored and (co-)authored 30 referred journal articles, and she has presented in 40 international conferences. She won three research awards at national and regional levels. She has worked on four funded research projects as a principal investigator and a co-author. She can be contacted at email: malrajhi@squ.edu.om.



Kamisah Osman    is a professor at Universiti Kebangsaan Malaysia (UKM) Center of STEM Enculturation, Faculty of Education and the Director at UKM Center for Quality Assurance. The University of Manchester, United Kingdom, granted her a Master of Science in Science Education in 1996 and a Doctor of Philosophy in 2000, both from the Center of Science and Technology Education. Her area of expertise extends beyond program evaluation to STEM education, specifically assessing higher-order thinking and problem-solving abilities, as well as innovative pedagogical approaches in STEM instruction and learning. Kamisah has held the position of deputy director (audit and benchmarking) at UKM Center for Quality Assurance for the duration of her professional tenure. She is the Editor-in-Chief of the European Journal of Science and Mathematics Education and the Malaysian Journal of Education, both of which she actively supports scholarly endeavors. She can be contacted at email: kamisah@ukm.edu.my.



Hussain Ali Alkharusi    is a professor of Measurement and Evaluation in the College of Education at Sultan Qaboos University. He is currently the assistant deputy vice-chancellor for Academic Affairs and Community Service at Sultan Qaboos University. His teaching areas include educational assessment, research, and statistics. Alkharusi has so far published more than 100 refereed research papers in national, regional and international journals and conferences. He has also co-authored many books and test manuals. He is also a member of many editorial boards of research journals. His contribution to the university, community service and research has been nationally, regionally and internationally acknowledged. He can be contacted at email: hussein5@squ.edu.om.