Journal of Professional & Applied Psychology *Original Article*

Urdu Translation & Validation of Executive Skills Questionnaire-Revised: A Self-Report Tool of Executive Functioning

Umme Farwah[1*], Attiya Inam[2], Muhammad Abiodullah[3]

Abstract

Executive functions (EF) include skills such as working memory, inhibition and cognitive flexibility that are required for efficient functioning in all disciplines of life, be it personal, educational or professional. The assessment of EF skills using task-based instruments or self-report forms is a debatable topic in research arena. The current study focused on translation and psychometric analysis of a comprehensive yet short and freely available tool of EF skills; Executive Skills Questionnaire-Revised (Strait et al., 2019) in Urdu language. The study followed a rigorous tool translation procedure including forward and back translation. Data was collected from 294 undergraduate students; both boys (*n*=142) and girls (*n*=152) aged 18-21 years (*M*= 19.2, *SD*=0.98) enrolled in various public sector higher educational institutes of Lahore, Pakistan. Data collection was followed by analysis which included use of descriptive statistics to analyze demographic characteristics and calculation of Cronbach's alpha of sub-scales to determine internal consistency. Item total correlation was then calculated. Analysis further included Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Results revealed the internal consistency of all sub-scales in acceptable range having Cronbach alpha value ranging from 0.72-0.80. CFA confirmed the 5-factor structure of the translated tool similar to the original tool. All fit indices were found to be excellent. Results revealed the ESQ-R Urdu as a valid and reliable tool to assess EF skills in student population.

Keywords: Confirmatory Factor Analysis, Executive Functions, Executive Skills Questionnaire

Received: 01 June 2024; Revised Received: 24 June 2024; Accepted: 25 June 2024

[1*]Lecturer, Department of Home Economics, Government Graduate College for Women, Gulberg, Lahore, Pakistan.
[2]Associate Professor, Department of Home Economics, Government Graduate College for Women, Gulberg, Lahore, Pakistan.
[3]Associate Professor, Department of Educational Research & Evaluation, University of the Punjab, Lahore, Pakistan.

***Corresponding Author Email:**
ummefarwah@live.com.pk

Introduction

Executive functions (EF) include a set of cognitive processes and abilities enabling individuals to act in thoughtful, planned ways in order to organize, prioritize, and perform tasks effectively. These functions are critical for goal-directed behavior, problem solving, decision-making, and self-regulation in a variety of settings, including academics, work, and personal lives (Benson et al., 2013; Dohle et al., 2018; Rabinovici et al., 2015). Blair (2016) described EF as an umbrella term used for various cognitive abilities allowing people to concentrate on a particular task, redirect attention and focus on multiple tasks, deter irrelevant thoughts causing distractions, as well as plan and direct behavior towards the achievement of a goal. Multiple definitions of EF have been quoted in the literature and despite the variations, all agree on EF being a multi-faceted higher order construct in which the integrity of efficient EF is founded on intact foundation cognitive skills (Burlak, 2019).

Having their neural basis within the prefrontal cortical region of the brain (Johann & Karbach, 2022), these skills include three distinct, yet somewhat correlated skills including working memory (WM), inhibition and cognitive flexibility

(Friedman & Miyake, 2017). Diamond (2013) suggested that these core EF skills provide foundation for advanced higher-order EF including planning, reasoning and problem-solving.

Deficits in EF can significantly impact an individual's capacity to function independently and adapt to different situations as this may cause significant reduction in one's ability to successfully perform essential daily life activities such as academic tasks, social activities, and self-care (Serrani, 2013; Slick et al., 2006). Assessing EF is essential for understanding an individual's cognitive strengths and weaknesses, diagnosing neurodevelopmental and neuropsychiatric disorders, and designing effective intervention strategies (Suchy et al., 2017; Souissi et al., 2022). Over the years, researchers from various domains of psychology, education and human development have developed a profound interest, in exploring and developing scales and measures to evaluate EF across different age groups and populations (Berg et al., 2020; McCoy, 2019). The current practice is to employ two different approaches for assessment of EF skills including "performance-based tests" and "rating scales". Performance-based tests also called psychometric or neuropsychological tests are designed for assessment of specific EF skills in strict laboratory conditions in the context of clinical or research settings (Berardi et al., 2021; Miranda et al., 2015; Souissi et al., 2022). Some popular and widely used tests include Digit Span Test to measure WM (Wechsler, 2014); Five Digit Test (Lang, 2002) or the Stroop Test (Stroop, 1992) to measure inhibition; Children's Color Trails Test (Llorente, 2003) as well as the Wisconsin Card Sorting Test (Grant & Berg, 1948) to assess cognitive flexibility and Tower of Hanoi (Goel & Grafman, 1995) to measure planning etc. Along with these individual tests, there are certain neuropsychological batteries consisting of a range of tests assessing various components of EF. These include the Cambridge Neuropsychological Automated battery (CANTAB: Luciana, 2003), Behavioral Assessment of the Dysexecutive Syndrome for Children battery (BADS-C: Emslie et al., 2003) and Delis-Kaplan Executive Function System (D- KEFS: Delis et al., 2004), to name a few. The condition of conducting these assessments in strict laboratory settings puts a question mark on the ecological validity and generalizability of performance-based tests (Soto et al., 2020). Further these tests require a lot of financial and time investment, making them a difficult approach to be used with larger samples (Toplak et al., 2013).

To address these issues, various questionnaires and rating scales have been developed; postulating the idea that the items contained in these scales tend to measure behaviors associated with processes gauged by EF psychometric tests (Souissi et al., 2022). These are further divided into a Parent Rating Scale (PRS), Teacher Rating Scale (TRS), or a Self-Report Form (SRF) based on the characteristics of the respondents whose EF skills have to be measured. Among the popular rating scales available to assess EF include Behavior Rating Inventory of Executive Function (BRIEF: Gioia et al., 2000); Behavioral Assessment System for Children, Third Edition (BASC3: Reynolds & Kamphaus, 2015), Childhood Executive Functioning Inventory (CHEXI: Thorell & Nyberg, 2008) and the Delis Rating of Executive Functions (D-REF: Delis, 2012) etc. Rating scales tend to be less time-consuming and are shown to have convergent validity. They are believed to measure EF skills employed by individuals in daily life situations and have concurrent and predictive validity as well (Gross et al., 2015; Muris et al., 2008). However; according to some researchers, these scales have questionable content and construct validity. They claim that rating scales tend to measure externalizing behaviors rather than focusing on cognitive aspect of EF skills (Souissi et al., 2022; Toplak et al., 2013). Also, the correlation of EF SRFs with task-based instruments gauging

similar EF skills is either weak or inconstant (Gross et al., 2015; Soto et al., 2020).

With the ever-increasing importance of EF skills for success in personal and professional lives, researchers are now focused upon finding ways to improve executive functioning skills of general population, especially students. The first step towards this is exploring the current levels of EF and finding deficits in EF. Despite all the criticism, SRFs are becoming a popular way to measure EF skills of various population segments. However, most of these scales have been developed in high income countries for an English-speaking population.

Pakistan is a Lower Middle-Income country (LMIC) where several cultures exist and many regional languages are spoken. Urdu; however, is the national language of Pakistan, widely spoken and understood throughout the country. It is also a widely spoken language in South East Asia including many states of India. To our knowledge, there is no free tool available in Urdu language to measure EF Skills in various population segments. Though previously a widely used tool of EF i.e., BRIEF (Gioia et al., 2015) was translated and used in Urdu language; however, the high cost of scale, need of a trained person to administer test and lack of research funding limits the applicability of the scale in LMICs like Pakistan. It is imperative therefore to translate and validate a freely available EF scale in Urdu language that can be used with the young adults. The following research aimed at translating and validating a freely available tool to gauge EF skills of young adults in Urdu language.

Executive Skills Questionnaire-Revised

To fill the gap of availability of a comprehensive yet affordable SRF of EF; Dawson and Guare (2010, 2012, 2018) developed multiple versions of the Executive Skills Questionnaire (ESQ) including a scale for adults (36 items), teenagers (33 items), students (33 items) as well as children (33 items). All of these covered assessment of the 11 skill areas of EF including WM, response inhibition, goal-directed behaviors, and continued attention.

The original version developed by Dawson and Guare was later revised by Strait et al. (2019) who designed 25 Items Executive Skills Questionnaire-Revised (ESQ-R) in an attempt to provide a valid, reliable as well as cost and time efficient measure for the young adult population, particularly students as most of the items concern academic success. EF was operationalized as an integration of core EF skills given by Miyake et al. (2000) and higher order EF skills theorized by Diamond (2013) that were particularly academic focused. Number of items in the ESQ-R was reduced to 25 items centering on five factors namely plan management (PM), emotional regulation (ER), behavioral regulation (BR), time management (TM) and material organization (Org). PM (11 items) involves determining a course of action to attain certain goals. TM (4 items) is an EF skill that requires conscious effort on part of an individual to plan and schedule his/her activities so as to achieve maximum output in a short span and to meet the deadlines. ORG (3 items) refers to an individual's capacity to keep track of his/her resources/materials and organize them efficiently to achieve his/her goals. ER (3 items) involves effectively dealing with one's emotions; not letting them hinder efficiency. BR (4 items) on the other hand implicates displaying self-control and considering consequences before any actions. The multidimensional ESQ-R is found to comprehensively measure EF skills of participants despite the fact that it puts less cognitive and physical burden on the subject due to simple interface and less number of items.

The current study had the following objectives.

1. To translate and validate the ESQ-R for young adults enrolled in higher educational institutes (HEIs).
2. To determine the psychometric properties of the Urdu version of ESQ-R by using EFA and CFA.

Though the tool is freely available to use for academic purposes, researcher sought permission from the author of the tool before starting with the translation process.

Rationale of the Study

The contribution of EF skills to holistic development in personal, academic and professional lives (Benson et al., 2013; Dohle et al., 2018; Rabinovici et al., 2015) has made it imperative to provide teachers and researchers with a comprehensive yet cost and time effective tool for measurement of the construct. The current study filled that gap by attempting to translate and validate a freely available, concise tool of EF.

The current study was significant as the translation of tool into Urdu language improved the accessibility of the scale to a broader audience, making it relevant to varied populations and ensuring inclusivity. This is of particular relevance for educators, particularly in public sector institutes who can then plan interventions for executive skills deficits (Nasir et al., 2021) thus supporting students' academic outcomes.

Method

The translation and validation of any tool involves adaptation of an existing scale from one language or cultural context to another and verification of its reliability and validity for the target population (Serrani, 2013). It is imperative to conduct a thorough review of existing literature and available scales in both the source and target languages in order to identify the most suitable scale for translation based on factors such as reliability, validity, comprehensiveness, and relevance to the target population (Cesari et al., 2024). Once a suitable scale is selected, the translation process begins, involving multiple steps to ensure linguistic and conceptual equivalence between the original and translated versions (Kleist et al., 2021). This includes forward translation by bilingual experts, reconciliation of discrepancies, back translation by independent translators, and pilot testing with the target population to assess comprehension and cultural appropriateness. After the translation phase, the validation process focuses on establishing the translated scale's validity and dependability among the intended audience. While validity evaluates if the instrument is measuring what it intends to measure and whether it effectively represents the construct, reliability means the consistency and dependability of the scale's measurements over time and between raters (Garcia-Barrera et al., 2015).

Instrument

The scale chosen for current study, as mentioned above, is ESQ-R revised by Strait et al., (2019). It consisted of 25 items divided into 05 sub-scales. Scale is scored on 4-point likert scale. Options were based on frequency ranging from never or rarely (0) to very often (3).

Sample

Sample included students enrolled in undergraduate programs in public sector colleges, purposively selected from 4 administrative towns of Lahore, namely Gulberg Town, Samanabad Town, Iqbal town and Lahore Cantonment. Gender, age, type of family (nuclear, joint or single-parent), parental age, education and occupation, monthly income etc. were some of the demographic characteristics asked from the respondents.

Procedure

The study was quantitative in nature having a cross-sectional research design. It was phased out in two stages. Stage I was translation and pilot-testing of the selected tool while stage II involved establishing psychometric properties using advanced statistical procedures.

Stage I

Stage I comprised of three steps described below.

Forward Translation

As per the recommended process (Tsang et al., 2017) the initial translation from the base language (English) to the required language (Urdu) was done by two independent bilingual translators. One of the translators had a background in Psychology and was aware of the concepts;

the questionnaire intended to measure. The other one, however was naïve of the objective of the questionnaire and more of a language expert.

Backward Translation

After forward translation was completed, both Urdu versions were independently back-translated (translated back into original language, English from target language, Urdu) by two bilingual translators who had command on both languages. Each item of these back-translated versions was compared to the original version of the tool and the statements with identical meaning were shortlisted for final translated version.

Pilot-Testing

Before using the tool with intended population, the pre-final version was pilot-tested on a small sample of 30 students as per recommendation of Perneger et al. (2015). The respondents were then verbally asked about the items they felt any difficulty in. Majority were unable to understand the translation of Item no. 24 which stated "I miss the big picture" and was translated as "میں وسیع النظر نہیں ہوں-"

Few had to ask the meaning of چھٹی حس in Item no. 19 that was the literal translation of phrase "go with my gut". The item stated "I go with my gut when making decisions" and was translated as " ہوں

کوئی بھی فیصلہ کرتے وقت میں اپنی چھٹی حس پر انحصار کرتا

After mutual discussion with the experts, it was decided that there was no better translation of the phrase "go with my gut" than چھٹی حس. So, it was left unchanged in the final version. However, the translation of Item no. 24 was changed in the final version to a simpler statement. It was rephrased as

" میں کسی بھی صورتحال میں اصل مقصد کی بجائےغیر ضروری تفصیلات پر غور کرنے لگتا ہوں-"

Before moving to stage II, the final translated tool was given for review to another expert with a PhD degree in Psychology who assessed and compared it with the English version, judging it for face validity i.e., to what extent a test seems like measuring the construct it claims to measure (Johnson, 2021). Despite being considered a weak form of validity due to involvement of subjectivity; face validity still gives a starting point in development of a tool or measure.

Stage II

The translated tool was thus administered to the target population i.e., students aged 18-21 years enrolled in higher educational institutes. Data collection also included demographic information including gender, type of family (nuclear, joint or single parent), household income. The sample consisted of students (n=294) age range between 18-21 years (M= 19.2, SD=0.98). Sample was representative of both males (n=142, 48.3%) and females (n=152, 51.7%). The other socio-demographic characteristics of participants are listed in Table 1.

Results

Given that the ESQ-R Urdu version has been adapted and translated from English to Urdu, EFA was applied to determine the factorial structure of the Urdu version and to validate it in context of Pakistani population. Initially, an inter-item correlation matrix for all 25 items of the scale was generated. The values for all items were above 0.2 which was within the acceptable range of 0.2 to 0.4 for any multidimensional scale (Hobart & Cano, 2009), suggesting that despite being sufficiently homogenous, items share enough variance (Piedmont, 2014).

Bartlett's test of sphericity was then applied that assesses the degree of correlation among the variables.

The results were significant (χ^2 =1362.45, p = .000) which indicated that variables share a common variance and scale is suitable for factor analysis. The next step was applying Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy which is another statistic used to assess the suitability of data for factor analysis as it suggests the extent to which variables share variance, indicating the potential for common variance among them. The value

of KMO ranges from 0 to 1 where a value closer to 1 indicates high correlation among variables and aptness of data for factor analysis. In our case, the KMO was 0.83 which is sufficiently high. Finally, communalities for all the 25 items of the scale were computed and found to be above 0.3 in some case and above 0.5 in most of the cases. Communalities explain shared variance between each variable, helping us determine how much variance in each variable can be accounted for by the identified factors. These values range from 0 to 1, with higher values indicating that a larger proportion of the variable's variance is accounted for by the underlying factors (Tabachnick & Fidell, 2019). 5 Factors having eigenvalues greater than 1 were retained for factor analysis.

Table 1
Demographic Characteristics of the Participants (N=294)

Variables	Frequency	Percentage (%)
Gender		
Male	142	48.3%
Female	152	51.7%
Age		
18yrs	86	29.3%
19yrs	94	32%
20yrs	81	27.6%
21yrs	33	11.2%
Fathers' Age		
40-49yrs	120	41%
50-59yrs	144	38.4%
60plusyrs	30	10.2%
Mothers' Age		
30-39yrs	27	9.3%
40-49yrs	211	73%
50-59yrs	50	17.36%
Fathers' Profession		
Govt. Job	70	23.8%
Personal Business	117	39.8%
Private Job	95	32.3%
Mothers' Profession		
Home Maker	265	90.1%
Job Holder	20	6.8%
Businesswoman	3	1.0%
Monthly Income		
Below 50,000	172	58.5%
50,001-100,000	88	29.9%
100,001-200,000	25	8.5%
Above 200,000	8	2.7%
Type of Family		
Nuclear Family	170	57.8%
Joint Family	100	34%
Single-Parent Family	26	8.8%

Table 2

Means, Standard Deviations, Alpha Reliability Coefficients of Scale and Sub-scales (N=294)

Sub-scales	*M*	*SD*	α	No. of Items
PM	1.98	0.55	0.72	11
TM	2.04	0.63	0.73	4
Org	1.66	0.61	0.79	3
ER	2.48	0.82	0.78	3
BR	2.24	0.62	0.80	4

Note: PM: Plan Management; TM: Time Management; Org: Materials Organization; ER: Emotional Regulation; BR: Behavior Regulation, α: Cronbach Alpha

Table 2 suggests the alpha coefficients values of 25 items of ESQ-R. Among the subscales, "PM" (M = 1.98, SD = 0.55) exhibits the lowest but acceptable internal consistency reliability (α = 0.72), while "BR" (M = 2.25, SD = 0.80) demonstrates the highest internal consistency (α = 0.80). The findings suggest an internally consistent and reliable ESQ-R scale which can be dependably used to measure the underlying construct.

Table 3

Correlation Values of Individual Items (N=294)

Subscale	Item	*r*	Subscale	Item	*r*
PM	6	.56**	**TM**	15	.65**
	7	.45**		20	.64**
	12	.55**	**Org**	3	.65**
	13	.50**		8	.73**
	14	.65**		9	.65**
	16	.61**	**ER**	4	.72**
	17	.53**		5	.69**
	18	.57**		5	.69**
	22	.55**	**BR**	1	.54**
	23	.59**		2	.65**
	24	.48**		19	.61**
TM	10	.63**		25	.59**
	11	.57**			

**$p < .01$

Table 4

Correlation among the Overall Scale and Sub-scales of ESQ-R Urdu Version (N = 294)

Variables	ESQ-R	PM	TM	Org	ER	BR
ESQ-R	-	-	-	-	-	-
PM	.89**	-	-	-	-	-
TM	.77**	.62**	-	-	-	-
ORG	.53**	.39**	.37**	-	-	-
ER	.61**	.38**	.37**	.14**	-	-
BR	.48**	.22**	.21**	.12**	.29**	-

*$p<.05$, **$p<.01$

As evident in Table 4, significant positive correlations were found between ESQ-R and each of its sub-scales PM (r = .89, p < .01), TM (r = .77, p < .01), ORG (r = .53, p < .01), ER (r = .61, p < .01), and BR (r =

.48, $p < .01$), with correlations ranging from moderate to strong.

Table 5

Model Fit Indices of CFA for ESQ-R Urdu Version (N = 294)

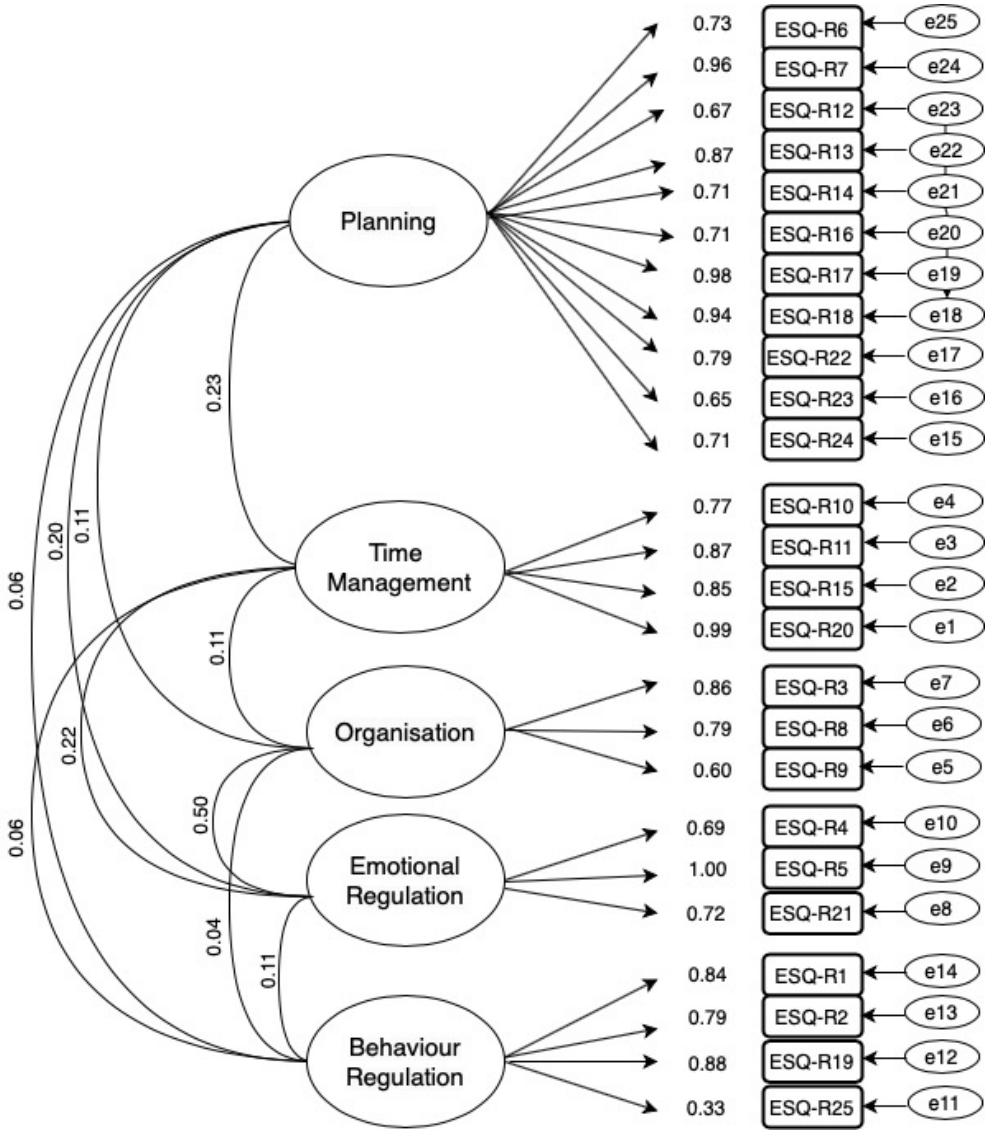
X^2	df	χ2/df	CFI	SRMR	RMSEA	PClose
440.96	270	1.633	0.945	0.063	0.046	0.766

Note. X^2 = Chi-square; df = Degree of Freedom; CFI = Comparative Fit Index; SRMR =Standardised Root Mean Square Residual; RMSEA = Root Mean Square Error of Approximation; PClose= p-value of the test of close fit

The data obtained from CFA for current research revealed significant results required for good model fit. The chi-square ratio of 1.633 and the CFI being 0.945 indicated a relatively good fit. Similarly, the SRMR and RMSEA values were 0.063 and 0.046, respectively, both suggesting reasonable fits. Moreover, the probability associated with the chi-square statistic (PClose) was 0.766, indicating that the model adequately fits the data. These findings suggest that the translated tool in Urdu exhibits a satisfactory level of fit.

Figure 1

The 5-factor Model of the Executive Skills Questionnaire-Revised Urdu Version



Discussion

Obtaining psychometrically significant results require administration of a standardized and valid research tool. However; constructing a new tool from scratch is a complex and time-consuming process (Goyal & Aleem, 2023). Thus, the study was carried out to translate and validate a freely available, suitable measure for evaluating EF skills of young adults enrolled in higher educational institutions (HEIs) in Urdu language, and to eventually establish its psychometric properties. As the previously available tools have mostly been developed for diagnostic purposes and are typically applied in clinical settings; the questionnaire used and translated in current study i.e., the ESQ-R is the perfect choice as it has been developed for typically developing population, has a self-report nature and concise items (Nasir et al., 2021).

The scale was translated by following a rigorous translation protocol which included forward translation, backward translation followed by pilot testing and final review by an expert in the field. The Urdu version was then validated through a series of statistical tests. EFA was firstly applied for this purpose which included a range of statistical tests such as inter-item correlation, Bartlett's test of sphericity,

KMO, communalities as well as eigenvalues. Results of all the tests declared the scale a suitable fit for factor analysis. Based on eigenvalues greater than 1, five factors were retained for factor analysis. Further, psychometric properties of the scale were computed. This included estimating the internal reliability of the sub-scales by computing Cronbach's co-efficient alpha. Values for internal consistency of the overall scale and subscales were adequate (Table 2) and were all above the cutoff value; ranging from 0.72 for factor 1 (PM; 11 items) to 0.80 for factor 5 (BR; 4 items). The results are slightly inconsistent from the original scale in English language where Strait et al., (2019) found the internal consistency estimates for the items in the five factors as mentioned: .89 for Factor 1 (11 items), .74 for Factor 2 (4 items), .76 for Factor 3 (3 items), .75 for Factor 4 (3 items), and .65 for Factor 5 (4 items). It was followed by computing an inter-correlation of the sub-scales. All of them were found significantly positively correlated with correlation ranging from moderate to strong.

Lastly CFA was applied to analyze if the 5-factor structure is retained in the context of current research and goodness-of fit model was drawn for the factor structure of the scale. Purpose was to confirm stability of the factor structure. The chi-square ratio was found to be within the recommended range of 1 and 3 (Hu & Bentler, 1999). Similarly, CFI value was as per benchmark recommended by Schumacker and Lomax (2010) i.e., 0.90. An SRMR of <0.08 and RMSEA of <0.06 indicate excellent model fitness (Xia & Yang, 2019). In the case of our model, they were 0.063 and 0.046 respectively showing good model fit. Results demonstrated an excellent model fit (Table 5) and supported a similar five factor structure as generated through CFA in original English version of the scale (Strait et al., 2019). Another study conducted in Malaysia to establish psychometric properties of ESQ-R by Nasir et al. (2021) suggested otherwise that a 5-factor second-order model consisting of 5 first-order factors namely PM, TM, ORG, ER and BR and a general (second-order) factor of EF is more preferred.

In summary, the current study clarifies the psychometric properties and practicality of the translated tool, establishing it as dependable instrument for assessing students' EF.

Limitations and Recommendations

The current study has its share of limitations, that can be overcome in the future researches. Limited sample size is one of them. Thus, an improved sample size and representativeness could yield slightly different results in terms of reliability estimations and factor structures. Due to time constraints, test-retest reliability couldn't be calculated as the scale was administered only once. Measurement of test-retest reliability is necessary, especially considering that the ESQ-R is particularly developed to be an intervention-focused measure. Lastly, the sample for the current study included students of HEIs only. Therefore, future studies focused on testing the psychometric qualities of the 25-item ESQ-R with a broader age group, including middle and secondary school students, should be conducted. This will make EF interventions more applicable to a wider range of student populations and greatly benefit the educators, students and parents.

Conclusion

Based on the current findings, practitioners can plan, customize, and assess the efficacy of interventions for numerous EF areas with the help of the ESQ-R, which is a promising tool. However, more research is required with an extended age and grade range to improve sample representativeness and expand psychometric evidence.

Contribution of Authors

Umme Farwah: Conceptualization, Investigation, Methodology, Data Curation, Formal Analysis, Writing – Original Draft
Attiya Inam: Methodology, Writing - Reviewing & Editing, Supervision
Muhammad Abiodullah: Methodology, Writing - Reviewing & Editing

Conflict of Interest
There is no conflict of interest declared by the authors.
Source of Funding
The authors declared no source of funding.
Data Availability Statement
The datasets of the current study are not available publicly due to ethical reasons but are available from the corresponding author [U.F.] upon the reasonable request.

References

Benson, J., Sabbagh, M., Carlson, S.M., & Zelazo, P.D. (2013). Individual differences in executive functioning predict preschoolers' improvement from Theory-of-Mind training. *Developmental Psychology, 49*(9), 1615-1627.

Berardi, A., Panuccio, F., Pilli, L., Tofani, M., Valente, D., & Galeoto, G. (2021). Evaluation instruments for executive functions in children and adolescents: a systematic review. *Expert Review of Pharmacoeconomics and Outcomes Research*, *21*(5), 885–896. https://doi.org/10.1080/14737167.2021.1908889

Berg, V., Rogers, S. L., McMahon, M., Garrett, M., & Manley, D. (2020). A novel approach to measure executive functions in students: An evaluation of two child-friendly apps. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01702

Blair, C. (2016). Developmental science and executive function. *Current Directions in Psychological Science*, *25*(1), 3–7. https://doi.org/10.1177/0963721415622634

Burlak, J. (2019). A comprehensive analysis of the constituents of Executive Functioning: The utility of a paediatric model towards the clear conceptualisation of EF in healthy adults (Doctoral dissertation). Victoria University, Victoria University Research Repository. https://vuir.vu.edu.au/40047/1/

Cesari, A., Galeoto, G., Panuccio, F., Simeon, R., & Berardi, A. (2024). Evaluation instruments for executive functions in children and adolescents: An update of a systematic review. *Expert Review of Pharmacoeconomics & Outcomes Research*, *24*(4), 487–508. https://doi.org/10.1080/14737167.2024.2311872

Dawson, P., & Guare, R. (2010). *Executive skills in children and adolescents: A practical guide to assessment and intervention* (2nd ed.). Guilford Press.

Dawson, P., & Guare, R. (2012). *Coaching students with executive skills deficits.* The Guilford Press.

Dawson, P., & Guare, R. (2018). *Executive skills in children and adolescents: A practical guide to assessment and intervention* (3rd ed.). Guilford Press.

Delis, D. C. (2012). *Delis-Rating of Executive Function (D-REF)*. Bloomington, MN: Pearson.

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan executive function system: An update. *Journal of the International Neuropsychological Society, 10*(2), 301-303. doi:10.1017/S1355617704102191

Diamond A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dohle, S., Diel, K., & Hofmann, W. (2018). Executive functions and the self-regulation of eating behavior: A review. *Appetite, 124,* 4-9. https://doi.org/https://doi.org/10.1016/j.appet.2017.05.041

Emslie, H., Burden, V., Nimmo-Smith, I., Wilson, B. A., and Wilson, C. (2003). *Behavioural Assessment of*

the Dysexecutive Syndrome for Children. Bury St Edmunds: Thames Valley Test Company.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86,* 186–204. doi: 10.1016/j.cortex.2016.04.023

Garcia-Barrera, M. A., Karr, J. E., Duran, V., Direnfeld, E., & Pineda, D. A. (2015). Cross-cultural validation of a behavioral screener for executive functions: Guidelines for clinical use among Colombian children with and without ADHD. *Psychological Assessment*, *27*(4), 1349–1363. https://doi.org/10.1037/pas0000117

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behavior rating inventory of executive function professional manual*. Lutz: Psychological Assessment Resources.

Goel, V., & Grafman, J. (1995). Are the frontal lobes implicated in "planning" functions? Interpreting data from the Tower of Hanoi. *Neuropsychologia, 33*, 623–642. doi: 10.1016/0028-3932(95)90866-p

Goyal, H., & Aleem, S. (2023). Confirmatory factor analysis (CFA) and psychometric validation of healthy lifestyle and personal control questionnaire (HLPCQ) in India. *Indian Journal of Community Medicine: Official Publication of Indian Association of Preventive & Social Medicine*, *48*(3), 430–435. https://doi.org/10.4103/ijcm.ijcm_394_22

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology, 38*(4), 404–411. https://doi.org/10.1037/h0059831

Gross, A. C., Deling, L. A., Wozniak, J. R., & Boys, C. J. (2015). Objective measures of executive functioning are highly discrepant with parent-report in fetal alcohol spectrum disorders. *Child Neuropsychology. 21*, 531–538. doi: 10.1080/09297049.2014.911271

Hobart, J., & Cano, S. J. (2009). Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. *Health Technology Assessment, 13*(12), iii, ix-x, 1-177. doi: 10.3310/hta13120

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Johann, V. E., & Karbach, J. (2022). The relations between personality, components of executive functions, and intelligence in children and young adults. *Psychological Research, 86*(6), 1904–1917. https://doi.org/10.1007/s00426-021-01623-1

Johnson, E. (2021). Face validity. In F. R. Volkmar (Ed.), Encyclopedia of autism spectrum disorders (pp. 1957-1957). Springer International Publishing.

Kleist, I., Noahsen, P., Gredal, O., Riis, J., & Andersen, S. (2021). Diagnosing dementia in the Arctic: Translating tools and developing and validating an algorithm for assessment of impaired cognitive function in Greenland Inuit. *International Journal of Circumpolar Health*, *80*(1), 1948247. https://doi.org/10.1080/22423982.2021.1948247

Lang, J. A. (2002). *Validation of the Five Digit Test in a Clinical Sample: An Alternative to the Stroop Color-*

Word Task with Possible Cultural Implications. San Francisco Bay: Alliant International University.

Llorente, A. M. (2003). *Children's Color Trails Test: Professional Manual*. Lutz, FL: Psychological Assessment Resources.

Luciana, M. (2003). Practitioner review: Computerized assessment of neuropsychological function in children: Clinical and research applications of the Cambridge Neuropsychological Testing Automated Battery (CANTAB). *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 44*(5), 649–663. doi: 10.1111/1469-7610.00152

McCoy, D. C. (2019). Measuring young children's executive function and self- regulation in classrooms and other real-world settings. *Clinical Child and Family Psychology Review, 22*(1), 63–74. doi: 10.1007/s10567-019-00285-1

Miranda, A., Colomer, C., Mercader, J., Fernández, M. I., & Presentación, M. J. (2015). Performance-based tests versus behavioral ratings in the assessment of executive functioning in preschoolers: Associations with ADHD symptoms and reading achievement. *Frontiers in Psychology*, *6*(APR). https://doi.org/10.3389/fpsyg.2015.00545

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Muris, P., Van Der Pennen, E., Sigmond, R., & Mayer, B. (2008). Symptoms of anxiety, depression, and aggression in non-clinical children: Relationships with self-report and performance-based measures of attention and effortful control. *Child Psychiatry and Human Development, 39*, 455–467. doi: 10.1007/s10578-008-0101-1

Nasir, H., Tan, C. S., & Pheh, K. S. (2021). The Executive Skills Questionnaire-Revised: Adaptation and Psychometric Properties in the Working Context of Malaysia. *International Journal of Environmental Research and Public Health*, *18*(17), 8978. https://doi.org/10.3390/ijerph18178978

Perneger, T. V., Courvoisier, D. S., Hudelson, P. M., & Gayet-Ageron, A. (2015). Sample size for pre-tests of questionnaires. *Quality of Life Research*, *24*(1), 147–151. https://doi.org/10.1007/s11136-014-0752-2

Piedmont, R.L. (2014). Inter-item Correlations. In: Michalos, A.C. (eds). Encyclopedia of Quality of Life and Well-Being Research. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0753-5_1493

Rabinovici, G. D., Stephens, M. L., & Possin, K. L. (2015). Executive dysfunction. *Continuum (Minneapolis, Minn.)*, *21*(3 Behavioral Neurology and Neuropsychiatry), 646–659. https://doi.org/10.1212/01.CON.0000466658.05156.54

Reynolds, C. R., & Kamphaus, R. W. (2015). *Behavior Assessment System for Children–Third Edition, Student Observation System*. Bloomington, MN: Pearson.

Schumacker, R. E., & Lomax, R. G. (2010). *A beginner's guide to structural equation modeling* (3rd ed.). Routledge.

Serrani A. D. J. (2013). Spanish translation and validation of an Executive Battery 25 (EB25) and its shortened version (ABE12) for executive

dysfunction screening in dementia. *Neurologia (Barcelona, Spain)*, *28*(8), 457–476. https://doi.org/10.1016/j.nrl.2010.12.020

Slick, D. J., Lautzenhiser, A., Sherman, E. M. S., & Eyrl, K. (2006). Frequency of scale elevations and factor structure of the behavior rating inventory of executive function (BRIEF) in children and adolescents with intractable epilepsy. *Child Neuropsychology, 12*(3), 181-189. doi: 10.1080/09297040600611320.

Souissi, S., Chamari, K., & Bellaj, T. (2022). Assessment of executive functions in school-aged children: A narrative review. *Frontiers in Psychology, 13*(991699). https://doi.org/10.3389/fpsyg.2022.991699

Soto, E. F., Kofler, M. J., Singh, L. J., Wells, E. L., Irwin, L. N., Groves, N. B., ... & Langberg, J. M. (2020). Executive functioning rating scales: Ecologically valid or construct invalid? *Neuropsychology, 34,* 605–619. https://doi.org/10.1037/neu0000681

Strait, J. E., Dawson, P., Walther, C. A. P., Strait, G. G., Barton, A. K., & Brunson McClain, M. (2019). Refinement and Psychometric Evaluation of the Executive Skills Questionnaire-Revised. *Contemporary School Psychology, 24*, 378–388. https://doi.org/10.1007/s40688-018-00224-x

Stroop, J. R. (1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology: General, 121*(1), 15–23. https://doi.org/10.1037/0096-3445.121.1.15

Suchy, Y., Ziemnik, R. E., & Niermeyer, M. A. (2017). Assessment of Executive Functions in Clinical Settings. In E. Goldberg (Ed.), *Executive Functions in Health and Disease* (pp. 551–569). Academic Press. https://doi.org/10.1016/B978-0-12-803676-1.00022-2

Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

Thorell, L. B., & Nyberg, L. (2008). The childhood executive functioning inventory (CHEXI): A new rating instrument for parents and teachers. *Developmental Neuropsychology, 33*(4), 536–552. https://doi.org/10.1080/87565640802101516

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct*? Journal of Child Psychology and Psychiatry, 54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Tsang, S., Royse, C. F., & Terkawi, A. S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia*, *11*(Suppl 1), S80–S89. https://doi.org/10.4103/sja.SJA_203_17

Wechsler, D. (2014). *Wechsler Intelligence Scale for Children–Fifth Edition (WISC-V).* Bloomington, MN: Pearson.

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51*(2), 409-428. https://doi.org/10.3758/s13428-018-1055-2