# ESTABLISHING PSYCHOMETRIC PROPERTIES OF EDUCATIONAL SHORT FILMEVALUATION SCALE

**Dr. M. Balamurugan**[1]
[1]*Assistant Professor in Education*
*Government College of Education, Pudukkottai, Tamil Nadu, India*
**DOI**: https://doi.org/10.34293/eduspectra.v5i1.02

## Abstract

*This paper attempted to establish the psychometric properties of the Educational Short Film Evaluation Scale (ESFE Scale) through six methods like Content validity, Cronbach's alpha reliability, Intrinsic Validity, item-total correlations, Factor Analysis, and Structural Equation Modelling (SEM). The data were collected from 100 (male 34 and female 66) First Year B.Ed. teacher trainees in the Government College of Education, Pudukkottai, Tamil Nadu, India were asked to submit short education-based short film assignments. This scale was constructed by the investigator and contains four major aspects i.e. content aspects, design aspects, technical aspects, and audience aspects as their dimensions. The investigator discussed six methods of validation intending to standardize the educational short film evaluation scale. In the present educational survey research scenario, the most common six methods of validation were illustrated. Understanding such validation methods will be surely useful for the research scholars who were in the fundamentals of tool construction for their research.*

***Keywords:*** *Confirmatory Factor Analysis, Construct Validity, Convergent Validity, Discriminant Validity, Exploratory Factor Analysis, Item Analysis, KMO, SEM*

## Introduction

A short film is any motion picture, which is short enough in running time not to be considered a feature film. The academy of motion picture arts and sciences defines a short film as "an original motion picture that has a running time of 40 minutes or less, including all credits". All films at the beginning of the cinema were very short, sometimes running only a minute or less. Filminute, the international one-minute film festival, has presented and promoted a collection of one-minute films across multiple media since September 2006 (Wikipedia, 2020). This was the inspiration of the investigator to assign a one-minute video containing any concept you like, which may have a strong moral message, and the same was instructed to the B.Ed. trainees.

For this assignment students' concerns were seriously taken into consideration and the barriers were discussed and suggested ways to overcome them. The investigator strongly believes that this assignment will allow his students one step further in the moral pathway. The investigator evaluated these short films based on educational context with the help of the evaluation scale discussed below. Moreover, this is not about giving other ratings or certifications like parental guidance, adult, horror, violence, etc. This scale was applied to all the 100 participants who submitted their short films through google forms specially made to collect such assignments.

**Objective**

This paper attempted to establish the psychometric properties of the Educational Short Film Evaluation Scale (ESFE Scale) through six different methods namely, content validity, Cronbach's alpha reliability, Intrinsic Validity, item-total correlations, Factor Analysis, and Structural Equation Modelling (SEM).

**Educational Short Film Evaluation Scale (ESFE Scale)**

Before jumping into the psychometric properties of the tool, let's understand its constructs and items. An evaluation scale was specially designed for assessing short film assignments submitted by first-year B.Ed. trainees. This scale was constructed by the investigator and contains four major aspects i.e. content aspects, design aspects, technical aspects, and audience aspects as their dimensions or theoretical constructs. The dimensions and the items were self-explanatory. However, the investigator felt the explanation of a few terms like film structure, storyboarding, BGM, Scene transitions effects, animations, avoiding copyright issues, and Hang-over. Film structure is the planning of the small chunks, whether concept or scenes. Storyboarding is the film script work i.e. clarity of the film sequence. BGM is the background music if any. Scene transition effects, like in MS PowerPoint there are some effects in between two slides, in the same way in effects in between two scenes or images or videos, etc. Animation is an advanced level of scene transition effects. Avoiding copyright issues means, not including the video or images having copyright issues or can be taken either by the candidate or from the creative commons, it must be free from plagiarism. Hang-over is the impact of the short film on our mind like it cannot be forgetful and last long in our mind for a few days at least. The detailed list of items was shown in Table 1. The duration range of the short film was calculated as 1 min and 26 sec, with a minimum of 0.42 sec to a maximum of 2 min 08 sec in length.

**Table 1 Educational Short Film Evaluation Scale– Items**

| Q.N | Items/ Questions | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Content Aspects** | | | | | | | | | | | |
| Q1 | Title/ Concept/ Message aptness | | | | | | | | | | | |
| Q2 | Introduction/ Beginning | | | | | | | | | | | |
| Q3 | Pronunciation/ Text / Word | | | | | | | | | | | |
| Q4 | Concept accuracy/ Film structure | | | | | | | | | | | |
| Q5 | Closure/ Ending/ Climax | | | | | | | | | | | |
| | **Design Aspects** | | | | | | | | | | | |
| Q6 | Scripting/ Storyboarding | | | | | | | | | | | |
| Q7 | Content flow/ Link | | | | | | | | | | | |
| Q8 | Video Pacing/ Timing | | | | | | | | | | | |
| Q9 | Scenes/ Images/ Dress code | | | | | | | | | | | |
| Q10 | Language/ BGM aptness | | | | | | | | | | | |
| | **Technical Aspects** | | | | | | | | | | | |
| Q11 | Sound clarity (Voice/ BGM)/ HD quality | | | | | | | | | | | |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Q12 | Visual stability/ Video clarity | | | | | | | | | | | |
| Q13 | Scene transitions effects/ Animations | | | | | | | | | | | |
| Q14 | Audio – Video sink | | | | | | | | | | | |
| Q15 | Originality/ Avoiding copyright issue | | | | | | | | | | | |
| | **Audience Aspects** | | | | | | | | | | | |
| Q16 | Entertainment value/ Interesting | | | | | | | | | | | |
| Q17 | Motivating/ Inspiring | | | | | | | | | | | |
| Q18 | Impactful/ Hang-over | | | | | | | | | | | |
| Q19 | Creativity/ Innovative | | | | | | | | | | | |
| Q20 | Realistic/ Connectedness | | | | | | | | | | | |

**Sample and Sampling of the Study**

In this study, purposive sampling was used. The data comprises 100 (male 34 and female 66) First Year B.Ed. teacher trainees in the Government College of Education, Pudukkottai, Tamil Nadu, India were asked to submit short education-based short film assignments through a google form.

**Scoring of the Scale**

After obtaining the short film assignments, they were scored on 20 items on 11 point scale, from 0 to 10 based on that these short films were assessed by the investigator and one of his colleagues. The maximum score possible was 200 and the minimum score was 0.

**Methods of Reporting Reliability and Validity of a Scale**

The investigator decided to discuss the commonly used methods applied in reporting the reliability and validity of the tool in the research with the help of this data.

**First Method of Validation – Adopting Validation from the Reported Form**

If the scale is already reported and standardized, the basic method is, you can refer to the same reliability and validity in your research paper, dissertation, or thesis. This was considered to be the easiest way of reporting reliability and validity. The ESFE scale was not adopted as such or reported anywhere before. Thus this validation method cannot be applied to this research study.

**Second Method of Validation – Content and Construct Validation**

The investigator prepared this scale. In that case, one can report its face (first impression or item layout) and content (item indicators) validation with the help of experts in the field of education and media, which can be extended to construct validity in case of a tool has dimensions in it. The researcher can do in-depth content and construct analysis, with the suggestions and modifications written in front of each item concerning the construct if any Along with that, any reliability measure can be determined. The investigator decided to apply Cronbach's Alpha Coefficient using SPSS (ver. 26) which was found to be 0.942for ESFE scale, as shown in Table 2.This method of reliability and validity is very prominent in

qualitative education research. Let's discuss the other quantitative methods of reliability and validity.

**Third Method of Validation – Intrinsic Validity Coefficient**

As discussed earlier, Cronbach's Alpha Coefficient was calculated to establish the reliability of any tool. We can report statistical intrinsic validity by taking the square root of Cronbach's Alpha Coefficient. The intrinsic validity coefficient was established by taking the square root of the reliability coefficient0.942, which is equal to 0.970 as shown in Table 2. Thus from these two coefficients, it may be inferred that this ESFE Scale is highly reliable and valid (Balamurugan, 2005).

**Table 2 Cronbach's Alpha Reliability and Intrinsic Validity**

| Scale/ Research Instrument | Cronbach's Alpha Reliability Coefficient | Intrinsic Validity |
|---|---|---|
| Educational Short Film Evaluation Scale 20 items | 0.942 | 0.970 |

**Fourth Method of Validation – Item Analysis**

Now let's discuss the fourth method, where Item Total Correlation (ITC) was considered for item validity. In ESFE Scale there are 20 items, the researcher calculated the total of these 20 items in a separate column, and determine the value of 'r' is the Person's Product Moment Correlation (PPMC) for each item using SPSS (ver. 26). Generally less than 0.3 was not advisable to select as a rule of thumb or the investigator can decide the selection criterion with the rationale of the study. In ESFE Scale 19 items had significant r values that were above 0.5 as shown in Table 3. Whereas in one item Q15 had nearly zero correlation, and that too was insignificant at 0.05 level of significance. The investigator decided to leave this particular item. Thus here we have ESFE Scale having 19 items. After validation, the reliability coefficient was determined to be 0.949 by calculating Cronbach's Alpha Reliability Coefficient (Balamurugan & Kumaran, 2008).

**Table 3 Item Total Correlation**

| Items | r value | Sig. (2-tailed) |
|---|---|---|
| Q1 | .797** | 0.001 |
| Q2 | .756** | 0.001 |
| Q3 | .699** | 0.001 |
| Q4 | .824** | 0.001 |
| Q5 | .774** | 0.001 |
| Q6 | .852** | 0.001 |
| Q7 | .878** | 0.001 |
| Q8 | .749** | 0.001 |
| Q9 | .713** | 0.001 |
| Q10 | .718** | 0.001 |

| Items | r value | Sig. (2-tailed) |
|---|---|---|
| Q11 | .612** | 0.001 |
| Q12 | .700** | 0.001 |
| Q13 | .513** | 0.001 |
| Q14 | .557** | 0.001 |
| Q15 | -0.002 | 0.981 |
| Q16 | .762** | 0.001 |
| Q17 | .708** | 0.001 |
| Q18 | .720** | 0.001 |
| Q19 | .761** | 0.001 |
| Q20 | .634** | 0.001 |

**Fifth Method of Validation - Exploratory Factor Analysis (EFA)**

The fifth method of validation can be divided into two namely, basic and modified exploratory factor analysis. The basic EFA where the principal component analysis (PCA) was run using SPSS, as it is, with Eigenvalue more than one, however, modified EFA can be applied with a fixed number of factors. Both were discussed in the following sections.

**Basic EFA Method of Validation**

Before hopping into the next method of validation the investigator had to test the sampling adequacy for factor analysis, this can be done using the Kaiser-Meyer-Olkin (KMO)measure of sampling adequacy and Bartlett's test of Sphericity. The KMO measure of sampling adequacy provides an index (between 0 and 1) of the proportion of variance among the variables that might be common variance (i.e., that might be indicative of underlying or latent common factors). For the present study, the KMO measure of sampling adequacy was employed using SPSS (ver. 26), which was found to be 0.914. For the *df* of 171,approx. chi-square value for Bartlett's test of Sphericity was identified as 1381.955, which was found to be significant at 0.001 level. These estimations proved to be appropriate for the factor analysis (Balamurugan, 2013).

In this stage as the investigator expected, there were four factors obtained in exploratory factor analysis (EFA), but the factors were not related to the conceptual construct or the dimensions allotted for the items. To identify the factors of the ESFE Scale the Principal Components Analysis (PCA) with Varimax Rotation was executed by taking 19 items (one item Q15 was left out using item analysis), which accounted for 70.87% of the total scale variance.

**Table 4 Principal Components Analysis (PCA) with Varimax Rotation**

| Items | 1 (Design) | 2 (Content) | 3 (Language and Sound) | 4 (Technical) |
|---|---|---|---|---|
| Q6 | 0.773 | | | |
| Q17 | 0.744 | | | |
| Q8 | 0.722 | | | |
| Q19 | 0.709 | | | |

| Items | 1 (Design) | 2 (Content) | 3 (Language and Sound) | 4 (Technical) |
|---|---|---|---|---|
| Q7 | 0.695 | | | |
| Q5 | 0.678 | | | |
| Q9 | 0.667 | | | |
| Q3 | | 0.804 | | |
| Q2 | | 0.711 | | |
| Q4 | | 0.708 | | |
| Q1 | | 0.683 | | |
| Q18 | | 0.621 | | |
| Q16 | | 0.580 | | |
| Q20 | | 0.540 | | |
| Q11 | | | 0.779 | |
| Q10 | | | 0.697 | |
| Q13 | | | | 0.815 |
| Q12 | | | | 0.677 |
| Q14 | | | | 0.630 |

This resulted in four factors being given names with slight modifications as design for factor 1, content for factor 2, Language and Sound for factor 3, and technical for factor 4 concerning the nature and tone of the items as shown in Table 4.

**Table 5 Cronbach's Alpha Reliability after EFA**

| Scale/ Sub Scale | Cronbach's Alpha Reliability Coefficient |
|---|---|
| Design Aspects7 items | 0.921 |
| Content Aspects7 items | 0.905 |
| Language and Sound Aspects2 items | 0.742 |
| Technical Aspects3 items | 0.765 |
| Educational Short Film Evaluation Scale 19 items | 0.949 |

After this stage, the investigator has to establish reliability for the scale as well as for the sub-scales(Balamurugan, 2013) as shown in Table 5. The investigator decided to go for fixed number factor analysis, by taking into account that only two items may not be sufficient for the determining third-factor 'language and sound'.

**Modified EFA Method of Validation with Forced Choice Extraction**

The KMO measure of sampling adequacy and Bartlett's Test of Sphericity were found to be the same as above and significant at 0.001 thus proved to be appropriate for the factor analysis as already discussed in the basic EFA method of validation section, the same was applicable here also.

**Table 6 PCA with Varimax Rotation with Forced Choice Extraction**

| | 1 (Content) | 2 (Design) | 3 (Technical) |
|---|---|---|---|
| Q3 | 0.830 | | |
| Q2 | 0.769 | | |
| Q4 | 0.762 | | |
| Q1 | 0.750 | | |
| Q16 | 0.663 | | |
| Q18 | 0.652 | | |
| Q20 | 0.580 | | |
| Q10 | 0.482 | | |
| Q6 | | 0.697 | |
| Q9 | | 0.661 | |
| Q5 | | 0.595 | |
| Q7 | | 0.574 | |
| Q8 | | 0.560 | |
| Q17 | | 0.558 | |
| Q19 | | 0.482 | |
| Q14 | | | 0.761 |
| Q12 | | | 0.750 |
| Q13 | | | 0.738 |
| Q11 | | | 0.470 |

To identify the factors of the ESFE Scale fixed number factor analysis was done. The results of factor analysis which is the Principal Components Analysis (PCA) with Varimax Rotation and a fixed number of three factors were executed by taking 19 items which accounted for 67.04% of the total scale variance. The factors were given the names, content aspects for factor 1, design aspects for factor 2, and technical aspects for factor 3 as shown in Table 6 concerning the theoretical dimensions discussed earlier in Table 1. After this stage, the investigator has to establish reliability for the scale as well as for the sub-scales (Balamurugan, 2013) as shown in Table 7. Most of the validation studies end here, but the investigator was desperate in applying SEM in ESFE Scale.

**Table 7 Cronbach's Alpha Reliability after EFA with Forced Choice**

| **Scale/ Sub Scale** | **Cronbach's Alpha Reliability Coefficient** |
|---|---|
| Content Aspects8 items | 0.908 |
| Design Aspects7 items | 0.921 |
| Technical Aspects4 items | 0.759 |
| Educational Short film Evaluation Scale 19 items | 0.949 |

**Sixth Method of Validation – Structural Equation Modelling (SEM)**

After the KMO measure and factor analysis, the CFA using AMOS (Analysis of a MOment Structures) software (ver. 23) can be done for confirmation of factors obtained. For that factor loading smaller than 0.3 were generally not included in the Structural Equation Modelling (SEM) to obtain a better model fit.

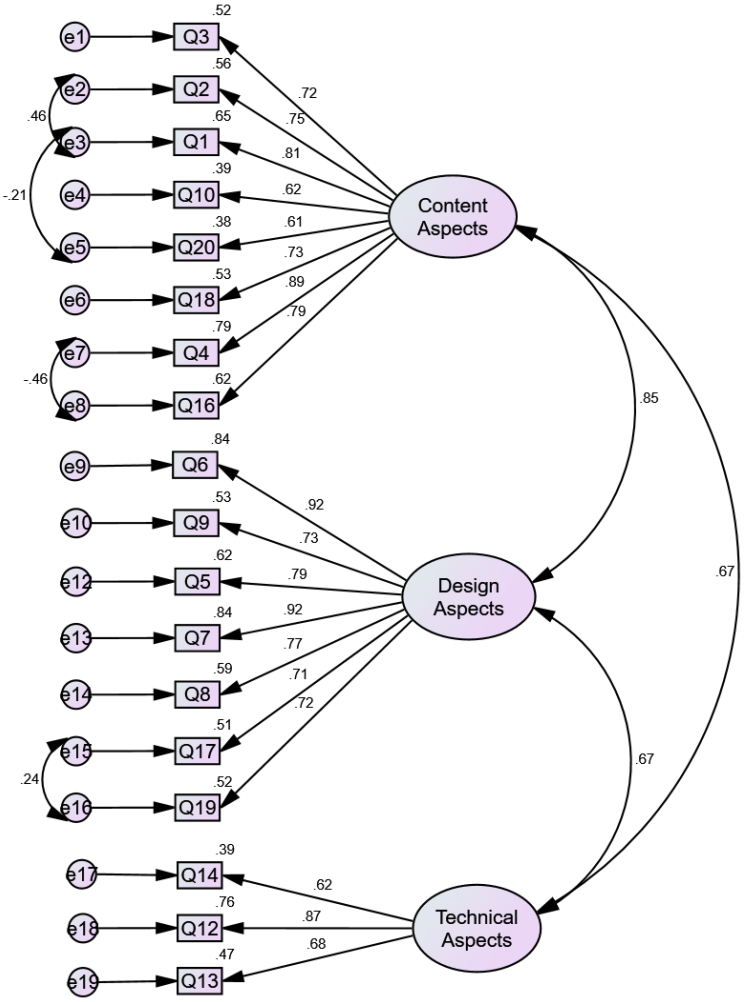


**Figure 1 Structural Equation Measurement Model**

SEM of data-driven measurement model with 3 constructs (Content, Design, and Technical Aspects) obtained from AMOS achieved absolute model fit, the model is recursive with 21 exogenous variables and 18 endogenous variables(18 items: Q15 was dropped by the application of item analysis and Q11 was dropped due to model fit suggestion in AMOS) as shown in figure 1, with CMIN = 194.723, DF = 128, CMIN/DF = 1.521, CFI = 0.944, SRMR = 0.057, RMSEA = 0.073and PC lose = 0.040(Hair et al., 2006)and was presented in Table 8.The opinion that a value of about 0.08 or less for the RMSEA would indicate a reasonable error of approximation, thus ESFE Scale is valid.(Browne & Cudeck, 1993).

**Table 8 AMOS Model Fit Estimates**

| Measure | Estimate | Threshold | Interpretation |
|---|---|---|---|
| CMIN | 194.723 | -- | -- |
| DF | 128 | -- | -- |
| CMIN/DF | 1.521 | Between 1 and 3 | Excellent |
| CFI | 0.944 | >0.95 | Acceptable |
| SRMR | 0.057 | <0.08 | Excellent |
| RMSEA | 0.073 | <0.06 | Acceptable |
| PClose | 0.04 | >0.05 | Acceptable |

**Table 9 Convergent and Discriminant Validity**

| **Factors** | **CR** | **AVE** | **Content Aspects** | **Design Aspects** | **Technical Aspects** |
|---|---|---|---|---|---|
| Content Aspects | 0.908 | 0.59 | **(0.77)** | | |
| Design Aspects | 0.924 | 0.52 | 0.62 | **(0.72)** | |
| Technical Aspects | 0.874 | 0.51 | 0.55 | 0.55 | **(0.71)** |

The purpose of applying SEM is to estimate the convergent and discriminant validity of the tool. The reliability assessment is assessed by Composite Reliability (CR) and validity is determined by Average Variance Extracted (AVE).The result of this study showed that the three constructs: content, design, and technical aspects CR coefficient of 0.908, 0.924, and 0.874 respectively as shown in Table 9. Referring to the CR coefficient threshold of 0.70 by (Gefen et al., 2000), it is concluded that every variable in this study was reliable and feasible to use. Research studies recommended an average variance extracted (AVE) as a convergent validity measure since AVE could explain the degree to which items are shared between the construct in structural equation modeling (SEM) where AVE 0.5 or more is acceptable as convergent validity.

The result showed that the AVE values for the three constructs respectively were: 0.59, 0.52, and 0.51. As all constructs exceeded the threshold AVE value of >0.50, it is concluded that they could measure the latent variables. Hence, they fulfilled the convergent validity criteria. The results presented in Table 9 inform that the three latent constructs respectively had square roots of AVE: 0.77, 0.72, and 0.71. The square roots of the AVE of the latent constructs were greater than the inter-construct correlation. Conclusively, the three latent constructs fulfilled the criteria of discriminant validity (Sujati et al., 2020). After this,to report reliability coefficients for the namesake, they were determined and given in Table 10. This ESFE scale with 18 items classified into three constructs was considered satisfactory by the investigator and used in further analysis that was behind the scope of this research paper.

**Table 10 Cronbach's Alpha Reliability after CFA**

| Scale/ Sub Scale | Cronbach's Alpha Reliability Coefficient |
|---|---|
| Content Aspects8 items | 0.908 |
| Design Aspects7 items | 0.921 |
| Technical Aspects3 items | 0.765 |
| Educational Short film Evaluation Scale 18 items | 0.948 |

## Conclusion

The investigator discussed six methods of validation intending to standardize the educational short film evaluation scale with more emphasis on factorial validity. This research paper insisted more on the quantitative methods of reliability and validity. Qualitatively, the researcher can do in-depth content and construct analysis, with the suggestions and modifications written in front of each item regarding the construct if any. In the present scale development process, the most common six methods of validation were illustrated. Understanding such validation methods will be surely useful for the research scholars and academicians, who were in the fundamentals of tool construction for their research. Aftera statistical understanding of the tool standardization research scholars can choose any method. Simply for the sake of statistical analysis and complexity, they were sequenced, except for the content and construct validity. In this research, the investigator had not compared which one is the best or more prominent, but in turn, discussed various methods to identify and report reliability and validity in the research studies.

## References

1. Balamurugan, M. (2005). A study of Students' Affiliation and Academic Achievement. [Master of Education, dissertation, University of Madras].
2. Balamurugan, M. (2013). Structure of Student Time Management Scale (STMS). *Journal on School Educational Technology*, *8*(4), 22–28.
3. Balamurugan, M., & Kumaran, D. (2008). Development and Validation of Students' Stress Rating Scale (SSRS). *Online Submission*, *7*(1), 35–42.
4. Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Long (Eds.), Testing structural equation models. *Newbury Park, CA: Sage*, 136–162. https://www.scirp.org /(S(i43dyn45teexjx455qlt3d2q))/ reference/ References Papers. aspx?ReferenceID=213477
5. Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, *4*(1), 7. https://doi.org/10.17705/1cais.00407
6. Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate Data Analysis.* Pearson Prentice Hall. https://www.scirp.org /(S(lz5mqp 453edsnp55rrgjct55))/reference/ReferencesPapers.aspx?ReferenceID=1747215

7. Sujati, H., Sajidan, Akhyar, M., & Gunarhadi. (2020). Testing the construct validity and reliability of curiosity scale using confirmatory factor analysis. *Journal of Educational and Social Research, 10*(4), 229–237. https://doi.org/10.36941/JESR-2020-0080
8. Wikipedia. (2020). *Short film - Wikipedia*. Wikipedia. https://en.wikipedia.org/wiki/Short_film