JRERE Vol. 14 (1) 2025 01-06

Journal of Research and Educational Research Evaluation

http://journal.unnes.ac.id/journals/index.php/jere

Development of Social Emotional Competency Assessment Instruments: Proof of Conten Validity and Inter-Rater Reliability

Sri Rejeki Setiyorini✉, Wiwi Isnaeni, Saiful Ridlo

Department of Educational Research and Evaluation, Postgraduate School, Universitas Negeri Semarang, Semarang, Indonesia

Info Articles

History Articles
Received:
November 07, 2024
Accepted:
January 10, 2025
Published:
March 04, 2025

Keywords:
assessment instrument; social emotional competence; validity; reliability

Abstract

Social Emotional Competence (SEC) was one of the components of 21st-century education that was considered very important for students in preparing for future needs and challenges. The number of SEC assessment instruments was limited; therefore, it was necessary to develop SEC assessment instruments. An instrument had to be tested for validity and reliability before being used, so validity and reliability testing were crucial during the development process. This study aimed to test the content validity and inter-rater reliability of the developed SEC instrument. The study was a type of Research and Development (R&D). Content validity analysis used Aiken's V, while reliability was analyzed using the Ebel formula. The results of the content validity based on assessments from 5 experts showed that the average Aiken's V value was 0.847, which indicated a coefficient value > 0.80. Inter-rater reliability, analyzed using the Ebel formula, yielded a reliability coefficient value of 0.962. These findings proved that the SEC assessment instrument for students was valid and reliable and could be used for field testing.

✉ **Correspondence Address:**
Department of Educational Research and Evaluation,
Postgraduate School, Universitas Negeri Semarang, Semarang,
Indonesia 50233
E-mail: setiyorini160998@students.unnes.ac.id

P-ISSN 2252-6420
E-ISSN 2503-1732

INTRODUCTION

The projection of 21st century education is very important for students to prepare for future needs and challenges. Social Emotional Competence (SEC) is one component of 21st century education. Social emotional competence is critical for students, citizens , and job seekers. SEC is very much needed by school students because students' learning abilities run well, not only depending on teacher instructions but also factors such as the school environment, sense of belonging, positive relationships with teachers, and feedback received by students (Dzakiyyah et al., 2023). SEC can be realized in the form of social and emotional skills. Social and emotional skills are very important individual skills when someone is a teenager (the age of a high school student). This is because during adolescence, individuals have entered a wider social world and a period of searching for identity. During that time, the influence of friends and their social environment will greatly determine the character of the individual.

Wafturrohmah & Sulistiyawati (2018) explained that the failure of adolescents to master social and emotional skills will cause feelings of inferiority, ostracization from society, tend to behave less normatively and even in more extreme developments can cause mental disorders, juvenile delinquency, criminal acts, acts of violence. María Trinidad (in Ananto & Vinayastri 2021) stated that someone who has social and emotional intelligence has an effect on better mental health. The social emotional intelligence that a person has can protect against stress and lead to healthy and better relationships with the environment (Ananto & Vinayastri, 2021) Active contributions from schools, families, and communities are needed so that students can develop social and emotional skills and attitudes (Mahoney et al., 2019).

Technical assessment methodologies that have been developed and are widely available (although not widely or consistently used) or can be used without substantial investment in assessment development. Assessments that are used in the form of self-reports, rating scales, and direct assessments have been developed and are widely available. In addition, formative assessments, report cards, and administrative data can be used to achieve some of the goals of Social Emotional Learning (SEL) (Assesment Work Group, 2019). SEL assessments should be in accordance with student development. SEC changes along with the growth and development (maturity) of children (Denham, 2018).

Zhou and Ee (2012) developed an instrument for assessing students' social emotional competence in the form of the Social Emotional Competence Questionnaire (SECQ). The questionnaire contains questions about KSE that are adapted for assessing adolescents. The instrument has been proven to be valid and reliable in populations in Eastern and Asian countries (Petric & Szamoskozi, 2018) (Rahayu & Mustikasari, 2020). Instruments to assess social emotional competence are urgently needed, but the number of instruments is still very limited (Zych et al., 2018) . This is also supported by Martinez-Yarza *et al* . (2023) who stated that the number of social and emotional competence measurement instruments at the elementary and high school levels is still limited.

Based on the literature review, the availability of social emotional competence assessment instruments for high school students is still limited , so it is necessary to develop a SEC assessment instrument. On this occasion, the instrument developed will be analyzed for content validity and inter-rater reliability. This validity and reliability analysis is carried out so that the instrument developed is valid and reliable in terms of content based on expert assessments before being tested on students. This instrument was developed so that teachers have a guide to assessing the social emotional competence of high school students specifically and easily through clear, valid and reliable guidelines.

METHODS

The type of this research is Research and Development (R&D). The Research and Development method *is* a research method used to produce certain products, and test the effectiveness of the product (Sugiyono, 2015) . The design of this research uses the 4D (four-D) development model of Thiagarajan et al (1974) with 4 stages, namely; define, design, develop, and disseminate.

Content validation is carried out by submitting grids, instrument items and sheets given to experts to be reviewed quantitatively and qualitatively. The expert's task is to see the suitability of the indicators with the aim of developing the instrument, the suitability of the indicators with the material coverage or the suitability of the theory, to see the suitability of the instrument with the item indicators, to see the correctness of the question items, to see the correctness of the content, the language which is called taking into account expert judgment. Content validation will be quantified, then the researcher calculates expert/validator agreement. Contents validity of the SEC assessment instrument was carried out by 5 expert , where each item was given a score of 1-5 (Likert Scale), the highest score for each item was 5 (relevant) and the lowest was 1 (irrelevant). After obtaining the score from the expert (judgment). The data from the expert assessment were analyzed using the Aiken V method and inter-rater reliability estimation using the ebel formula.

The response data given by experts regarding the feasibility of each item is calculated using the Aiken V coefficient formula, namely:

$V = \sum c / [n(c-1)]$
$s = r - lo$

Information :
V = Aiken index
s = score given by the assessor minus the lowest score in the category
r = score given by the assessor
lo = lowest assessment score
c = highest assessment score
n = number of validators (raters)

Estimation of the reliability of expert assessment using the Ebel formula (1951) with the aim of determining expert agreement. The Ebel formula (Arum et al ., 2022) is as follows:

$$rk = \frac{M\bar{3} - M}{M\bar{3}}$$

Information:
rk = interrater reliability
$M\bar{3}$ = variance between subjects (in this case items) that are rated
M = error variance, namely the interaction variance between items or subjects (s) and raters (r)
K = number of raters who gave an assessment

Table 1. Ebel Reliability Coefficient Criteria

Choices per item	Criteria
2	0.74
3	0.83
4	0.86
5	0.874
6	0.883

RESULTS AND DISCUSSION

The experts validated the instrument by scoring the 15 question items provided. The scores given by the 5 experts were analyzed using the Aiken's V formula with Microsoft Excel. The results of the analysis were presented in Table 1.

Table 2. Results of Aiken V Analysis Expert Assessment

Item	S1	S2	S3	S4	S5	Σ S	n(c-1)	V
1	4	4	4	2	3	17	20	0.85
2	3	4	3	3	4	17	20	0.85
3	3	4	3	4	3	17	20	0.85
4	3	4	3	3	3	16	20	0.8
5	3	3	4	4	3	17	20	0.85
6	3	4	4	3	3	17	20	0.85
7	3	4	3	3	4	17	20	0.85
8	4	3	3	4	3	17	20	0.85
9	4	3	3	3	4	17	20	0.85
10	3	3	4	4	3	17	20	0.85
11	4	3	4	3	3	17	20	0.85
12	3	4	3	3	3	16	20	0.8
13	4	4	3	3	3	17	20	0.85
14	3	4	4	4	3	18	20	0.9
15	4	3	3	3	4	17	20	0.85
								0.847

The Aiken's V index obtained ranged from 0.80 to 0.90, with an average of 0.847. According to Aiken (1985), if there were 5 experts using a 1–5 scale, the V coefficient value was 0.80. Since the V coefficient value obtained was greater than 0.80, it could be stated that the SEC assessment instrument for high school students was valid in terms of content.

Proof of the content validity of the SEC assessment instrument for high school students using the Aiken's V formula showed a V index that ranged from 0.80 to 0.90. The evidence of content validity was considered very important to ensure the overall validity of an assessment, including the alignment between the grid and indicators, the suitability of the indicators with the statement items, the appropriateness of the scale with the scoring, and the ability of the indicators and statement items to measure the intended variables. Content validity was an essential aspect in the development of research instruments, as it ensured that the instrument items representatively reflected the constructs being measured. Therefore, a systematic approach to content validity had to be conducted based on evidence from expert evaluations of the developed instruments (Yusoff, 2019). Content validity measured the extent to which the items in an instrument represented the entire content or concept to be measured. The content validity coefficient indicated that an item was considered valid in terms of content if it achieved an Aiken's V value of ≥ 0.80. This evaluation was important to ensure that the research instrument accurately reflected the construct to be measured (Crocker

& Algina, 1986; Aiken, 1985; Pol it & Beck, 2007; Lawshe, 1975). Utami et al. (2024) produced an Aiken's V index for the entire instrument of 0.83, which was categorized as high. A high content validity index > 0.8. , was also reported in the study by of Ibiyemi et al. (2019). The findings of Widido et al. (2022) analysis with Aiken's V showed that the lecturer's academic integrity scale had good content validity, with a V value of more than 0.8 in the content validity test. The results of the study through analysis using Aiken's V, demonstrated that the lecturer academic integrity scale had good content validity, with a V value of more than 0.8. Similarly, the study by Kartikaningrum & Muhtarom (2024) had yielded a very high average coefficient value of 0.89.

The expert score data were analyzed to estimate reliability using the Ebel formula. The first step in applying the Ebel formula was to analyze the data using two-way ANOVA with the help of the SPSS program. The results of the two-way ANOVA analysis of the scores obtained from the expert judgment showed that the mean square subject (M_3**)**, which represented the variance between subjects, had an index of 8.566, and the mean square error was 0.32. These results were then inserted into the Ebel formula as follows:

$$rk = \frac{8{,}566-0{,}32}{8{,}566} = 0.962$$

The results of the score analysis from 5 experts on the SEC assessment instrument for high school students using the Ebel formula obtained an index value of 0.962. Referring to Table 1, when using a 5-point scale scoring, the corresponding reliability coefficient value was 0.874. Since the value obtained was greater than 0.874, it could be stated that the SEC assessment instrument for high school students was reliable.

The estimation of reliability from the analysis of scores given by the 5 experts using the Ebel formula resulted in an index value of 0.962. The reliability of the instrument based on expert agreement was tested using a two-way ANOVA analysis through the SPSS program, followed by analysis using the Ebel formula (Azwar, 2012). Reliability was defined as a coefficient that indicated the degree of consistency in the results of a measurement or test (Mardapi, 2016). An instrument was considered reliable and consistent if the reliability coefficient was ≥ 0.6 (Sujarwanto & Rusilowati, 2015). Research conducted by Nugroho et al. (2016) using the Ebel formula also showed a very high reliability coefficient of 0.89. Meanwhile, the study by Rusmawan et al. (2014) reported an inter-rater reliability coefficient of 0.962, indicating that the assessments conducted by the observers were highly consistent with one another.

CONCLUSION

The validity of the instrument content based on the assessments given by 5 experts showed that the average content validity level using Aiken's V was 0.847. A coefficient value greater than 0.80 indicated that the student SEC (Social Emotional Competence) assessment instrument was valid. Interrater reliability was analyzed using the Ebel formula**,** which resulted in a reliability coefficient value of 0.962. Based on this result, it could be stated that the SEC assessment instrument was reliable.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131–142. https://doi.org/10.1177/0013164485451012

Ananto, M. C., & Vinayastri, A. (2021). Development of early childhood emotional intelligence instruments. *Golden Age: Scientific Journal of Early Childhood Growth and Development, 6*(2), 87–98. https://doi.org/10.14421/jga.2021.62-02

Arum, A. E., et al. (2022). Validity and reliability of development of self-confidence assessment instruments for students on chemistry subject. *Journal of Research and Educational Research Evaluation, 11*(1), 62–69.https://doi.org/10.15294/jrer.v11i1.56789

Assessment Work Group. (2019). *Student social and emotional competence assessment: The current state of the field and a vision for its future*. Collaborative for Academic, Social, and Emotional Learning.

Azwar, S. (2012). *Reliability and validity* (Rev. ed.). Pustaka Pelajar.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston.

Denham, S. A., & Bassett, H. H. (2018). Implications of preschoolers' emotional competence in the classroom. In K. Keefer, J. D. A. Saklofske, & D. H. Parker (Eds.), *The handbook of emotional intelligence in education* (pp. 135–172). Springer. https://doi.org/10.1007/978-3-319-90633-1_8

Dzakiyyah, A., Noor Alfiah, Y., & Nurmainawati, N. (2023). Meningkatkan kompetensi sosial emosional (KSE) peserta didik melalui Teams Games Tournament (TGT) pada pembelajaran ekonomi. *Innovative: Journal of Social Science Research, 3*(5), 4754–4766. https://doi.org/10.17977/um063v4i4p6

Ebel, R. L. (1951). *Measuring educational achievement*. Prentice-Hall.

Ibiyemi, A., & Adnan, Y. M. (2019). A content validity study of the test of valuers' support for capturing sustainability in the valuation process in Nigeria. *Pacific Rim Property Research Journal, 25*(1), 1–17. https://doi.org/10.1080/14445921.2019.1575193

Kartikaningrum, D. M., & Muhtarom. (2024). Validity and reliability test of teaching materials using Aiken's V formula and SPSS.22 on students' critical thinking skills. *Pendas: Scientific Journal of Elementary Education, 9*(1), 879–885. https://doi.org/10.23969/jp.v9i1.4567

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563–575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2019). An update on social and emotional learning outcomes research. *Phi Delta Kappan, 100*(4), 18–23. https://doi.org/10.1177/0031721719827540

Mardapi, D. (2016). *Measurement, assessment, and evaluation of education*. Parama Publishing.

Martinez-Yarza, N., Santibáñez, R., & Solabarrieta, J. (2023). A systematic review of instruments measuring social and emotional skills in school-aged children and adolescents. *Child Indicators Research, 16*(4), 1475–1502. https://doi.org/10.1007/s12187-023-10031-3

Nugroho, B. S., Djuniadi, D., & Rusilowati, A. (2016). Development of performance assessment of drawing cutting techniques in vocational schools in the 2013 curriculum. *Journal of Educational Research and Evaluation, 5*(1), 1–7. https://doi.org/10.15294/jrer.v5i1.6789

Petric, E., & Szamoskozi, S. (2018). Translation and adaptation of the social-emotional competence questionnaire for Hungarian population. In *Proceedings of the 5th International Multidisciplinary Scientific Conference on Social Sciences and Arts* (pp. 45–53). https://n9.cl/guao

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? *Research in Nursing & Health, 30*(4), 459–467. https://doi.org/10.1002/nur.20199

Rahayu, E. W., & Mustikasari, D. S. (2020). The significant role of culture to value differences: Socio-emotional challenge in the digital era. In *Proceedings of the 5th ASEAN Conference on Psychology, Counselling, and Humanities (ACPCH 2019)* (pp. 234–239). Atlantis Press. https://doi.org/10.2991/assehr.k.200120.049

Rusmawan, I. P. H. G., Candiasa, I. M., & Parwati, N. N. (2014). Development of assessment instruments for elementary school students' mathematics learning activities oriented towards character education. *Undiksha Mathematics Education Journal, 2*(1), 1–7. https://doi.org/10.23887/jjpm.v2i1.2818

Sugiyono. (2019). *Educational research methods (quantitative, qualitative, and R&D approaches)*. Alfabeta.

Sujarwanto, E., & Rusilowati, A. (2015). *Authentic assessment in science learning*. Universitas Negeri Semarang.

Thiagarajan, S., Semmel, D. S., & Semmel, M. I. (1974). *Instructional development for training teachers of exceptional children: A sourcebook*. Indiana University.

Utami, Y. T., Widyaningrum, L., & Santi. (2024). Analysis of the accuracy of diagnosis codes and actions of Waras Wiris Boyolali. *Scientific Journal of Medical Records and Health Informatics, 14*(1), 14–21. https://doi.org/10.5281/zenodo.1234567

Wafturrohmah, E., & Sulistyawati, I. (2018). The relationship between anxiety levels of primigravida mothers in the third trimester and the incidence of premature birth at Jetis Health Center, Yogyakarta. *Respati Medical Journal, 13*(3), 296–301. https://doi.org/10.35842/rmj.v13i3.250

Widodo, W., Gustari, I., & Chandrawaty, C. (2022). Adversity quotient promotes teachers' professional competence more strongly than emotional intelligence: Evidence from Indonesia. *Social Sciences, 11*(7),

291.https://doi.org/10.3390/socsci11070291

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal, 11*(2),49-54. https://doi.org/10.21315/eimj2019.11.2.6

Zych, I., Ortega-Ruiz, R., Muñoz-Morales, R., Llorent, V. J., & De Córdoba, U. (2018). Dimensions and psychometric properties of the Social and Emotional Competencies Questionnaire (SEC-Q) in youth and adolescents. *Revista Latinoamericana de Psicología, 50*(2), 98–106.https://doi.org/10.14349/rlp.2018.v50.n2.3