Research Article https://doi.org/10.12973/eu-jer.11.1.445

European Journal of Educational Research
Volume 11, Issue 1, 445 - 456.
ISSN: 2165-8714
https://www.eu-jer.com/

Development of Affective Self-Assessment Instrument of Chemistry for High School Student as the Daily Assessment Guideline

Farintis Jihadul Alivi
Universitas Negeri Yogyakarta, INDONESIA

Widihastuti*
Universitas Negeri Yogyakarta, INDONESIA

Received: August 5, 2021 ▪ Revised: November 3, 2021 ▪ Accepted: December 10, 2021

Abstract: The study objectives were (1) developing a valid and reliable Affective Self-assessment Instrument of Chemistry for High School Student and (2) discovering the chemistry affective domain ability trend of high school students based on gender. The current development study utilized 10 non-test instrument development procedures from Mardapi. The study population was all high school students in Yogyakarta Special Region. The sample size was 405 students categorized into two stages and sampling techniques, i.e., the trial stage using cluster random sampling and the measurement stage using simple random sampling. The data analysis techniques were validity test using the Aiken index and construct validity and reliability using the second-order Confirmatory Factor Analysis model. The study findings were (1) the Affective Self-assessment Instrument of Chemistry for High School Student had 15 valid and reliable items and 15 available items to be utilized by teachers to measure students' affective in the learning process and (2) the chemistry affective domain ability trend of male high school students was dominated by the "good" category and "very good" category for female students.

Keywords: *Affective domain, reliability, validity.*

To cite this article: Alivi, F. J., & Widihastuti. & (2022). Development of affective self-assessment instrument of chemistry for high school student as the daily assessment guideline. *European Journal of Educational Research*, *11*(1), 445-456. https://doi.org/10.12973/eu-jer.11.1.445

Introduction

Chemistry learning requires evaluation to discover the learning indicator achievement. The learning indicator achievement is observed from the learning results in the cognitive, affective, and psychomotor aspects (Bloom et al., 1956). The observation result performed on several high schools in Yogyakarta Special Region demonstrated that the validity and reliability of the applied affective assessment instrument were unknown. The instrument utilized by teachers was insufficient to conduct an assessment (Syamsudin et al., 2016). A study illustrated that the affective assessment by teachers had not used the assessment regulation following the determined curriculum (Muslich, 2014). Another fact depicts teachers merely assessing based on interactions between students and teachers since they assess students as individuals (Kusumaningtyas et al., 2018). Also, another problem was discovered where teachers rarely measure the affective domain and utilize it to advise students as the learning process improvement (Saxon et al., 2008). School facts show that teachers understand the importance of affective assessment; however, they do not have an appropriate instrument for affective measurement (Riscaputantri & Wening, 2018).

The validity and reliability of the affective domain measurement instrument should be known to facilitate its use (Mardapi, 2018). Affective assessment instrument development is crucial as the assessment guideline for teachers in the learning process to accountably measure students' affective ability. Conversely, without a clear assessment guideline, the assessment process is uncontrollable, and each student's score equality is doubted (Setiadi, 2016). Besides, the affective domain is the supporting factor for learning success in the cognitive domain (Setiawan et al., 2013). A study revealed an insignificant difference in the cognitive domain between male and female students (Lilleholt, 2019). Another study showed that male and female students had different problem-solving approaches. Males tend to use daring strategies, while females tend to use procedural and familiar strategies (Cimpian et al., 2016). Gender difference affects student motivation during learning (Marshman et al., 2018). This condition indicates that several factors influence the affective measurement in the learning process. Contrasting the cognitive domain that is measurable and concrete, the affective domain is challenging to be defined and assessed due to its abstract nature

* **Corresponding author:**
Widihastuti, Universitas Negeri Yogyakarta, Department of Fashion Technology Education, Faculty of Engineering, Indonesia.
✉ widihastuti@uny.ac.id

(Setiawan & Mardapi, 2019). There are five crucial affective types, i.e., attitude, self-esteem, value, moral, and interest (Mardapi, 2017).

One of the instruments used to measure affective ability is a questionnaire in case studies, where students are asked to respond (Mardapi, 2012). The instrument encouraging students to answer following the condition is using self-assessment, where students are presented with questions and answer options with possible events. Complex students' behaviors and limited affective assessment instrument availability, particularly in chemistry, are reasons for this affective instrument development. Therefore, the developed affective self-assessment instrument can measure students' affective in the chemistry learning process.

Literature Review

Affective Domain

Learning result assessment based on the Regulation of the Minister of Education and Culture of the Republic of Indonesia No. 23 of 2016 article 3 mentions that educational assessment includes attitude, behavior, and skill aspects. The attitude assessment in question is teachers' activities in acquiring descriptive information concerning students' behaviors. The assessments may take the form of observation or other types. The affective domain is defined as interests, attitudes, and values (Bloom et al., 1956). Affective is also defined as emotional attitudes or reactions against an object, behavior on objects, and faith in objects (Hart, 1989). Moreover, the affective domain components comprise emotions, attitudes, and behaviors (Beltrán-Pellicer & Godino, 2020). Meanwhile, Binet, Simon and Tyler (Anderson & Bourke, 2013) argued that the affective character variations are attitudes, interests, values, and self-esteem. These characteristics reflect students' attention and persistence in class. It is reinforced by Mardapi (2017) that there are five crucial affective types, i.e., attitude, self-esteem, interest, value, and moral.

Attitude: The attitude term has several meanings related to the topic. Attitude can be conceptualized as the tendency to respond positively or negatively to particular objects, situations, concepts, or people (Aiken, 1980). Students' attitudes on chemistry can be measured via scientific attitudes. Scientific attitudes are portrayed when students work on an experiment completely and scientifically (Wiwin & Kustijono, 2018). Scientific attitudes include a) curiosity, b) respect to data and facts, c) diligence, d) creativity and innovation, e) cooperation, f) open-minded, g) respect people's opinions, and h) sensitive to the environment (Harlen, 1985). Scientific attitudes are vital to be applied in the chemistry learning process.

Self-esteem: Self-esteem aims to discover and utilize students' abilities and drawbacks. It strengthens Mardapi's (2017) statement that self-esteem is necessary for future career paths where students can understand their strengths and weaknesses to select an appropriate career. Student strengths can be measured using a self-esteem scale comprising physique, moral, family, personality, social, and academic or occupation (Fitts & Roid, 1964). Self-esteem can improve students' academic achievements (Hosseini et al., 2016). Therefore, information regarding self-esteem becomes teachers' reference in improving the learning process.

Interest: Interest is defined as attention to specific objects or topics and tends to linger for a long period (Harackiewicz et al., 2016). Interest is emphasized as a liking for a thing or an activity without commands. It can be measured through interest, attention in learning, learning motivation, and knowledge (Slameto, 2010). Interest is considered a strong motivator for teenagers; thus, teachers utilize this to improve student engagement in the learning process (Chen & Wang, 2017). Therefore, interest is considered essential in the learning process success.

Value: The role of teachers is required to reinforce individual values to impact positively. A value is defined as a social or religious rule or cultural etiquette. A value is also defined as students' awareness of good behaviors (Bhardwaj, 2016). The affective domain in measuring values includes honesty, integrity, fairness, freedom, and commitment (Mardapi, 2017).

Moral: Moral is defined as the idea regarding human behavior (good or bad) based on specific situations (Abadi, 2016). Furthermore, moral is also defined as ideas commonly accepted concerning good, appropriate, and reasonable actions (Hudi, 2017). Moral indicators are keeping promises, helping others, respecting others, being honest, and behaving honestly (Mardapi, 2017).

These five affective types are then utilized as aspects in the affective assessment instrument. The affective assessment in scientific learning processes is often neglected where cognitive development is prioritized (Ramma et al., 2018). It is due to two reasons related to unsystematic affective assessment, i.e., many educators perceive the affective domain as "feelings" and not becoming an educational business where the affective domain is the natural (final) result of the cognitive domain not requiring solution during the learning process (Tyler, 1973). Affective domain had effect on the teaching and learning processes, caused different models of its components and conditioning factor (Beltrán-Pellicer & Godino, 2020). Affective assessment is applicable through observation and self-report methods (Andersen, 1980).

Therefore, affective assessment also used self-assessment method (Setiawan & Mardapi, 2019). Observation can be performed on behaviors involving one's feelings and emotions. These behaviors should follow one's behavioral thought type (Andersen, 1984).

Methodology

Research Design

This study aimed to (1) develop a valid and reliable Affective Self-assessment Instrument of Chemistry for High School Student and (2) discover the chemistry affective domain ability trend of high school students based on gender. The study was a development study with a descriptive quantitative approach. The development procedure employed 10 non-test instrument development procedures from Mardapi (2018), i.e., (a) determining the instrument specification, (b) writing the instrument, (c) determining the instrument scale, (d) determining the scoring system, (e) studying the instrument, (f) conducting trials, (g) analyzing the instrument, (h) assembling the instrument, (i) measuring, and (j) interpreting the measurement outcome.

Sample and Data Collection

The study population was all XI Science class of high school students in Yogyakarta Special Region. The study samples were divided into two, i.e., trial and measurement samples. The trial sample determination using cluster random sampling is presented in Table 1. The measurement sample determination using simple random sampling is presented in Table 2. The sample size in trial and measurement was adjusted to a small sample size comprising 200 students (Li, 2016).

Table 1. Trial Sample

Location	School's Name	Number of Students
Bantul	MAN III Bantul	59
Sleman	SMA Angkasa	55
Kota Yogyakarta	MAN II Yogyakarta	86
Total		200

Table 2. Measurement Sample

School's Name	Number of Students
SMAN 7 Yogyakarta	75
SMAN 11 Yogyakarta	60
SMAN 6 Yogyakarta	50
Total	205

The affective self-assessment instrument of chemistry for high school student was developed using five aspects with 15 indicators. Each aspect is represented by three indicators. These indicators were developed following the number of indicators, i.e., 15 question items using the Likert scale with three solution alternatives (A, B, and C) (Kokom & Edwina, 2020) . The available solution alternatives are plausible events in a student's life. The multiple-choice assessment scale was transformed into an interval scale based on the gradation of assessments 1-3. The affective assessment concept note in this study is depicted in Figure 1.

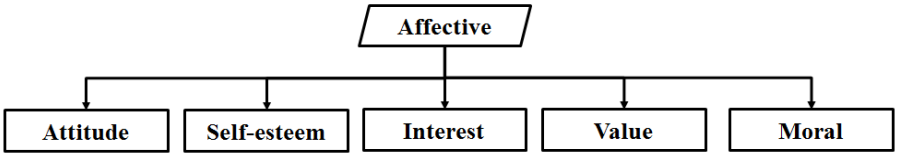

Figure 1. The Affective Assessment Concepts

The affective self-assessment instrument of chemistry for high school student have five aspects. First, attitudes comprising curiosity, respect for data and facts, and creativity and innovation. Second, self-esteems comprising physique, morals, and family. Third, interests comprising learning interest, learning attention, and learning motivation. Fourth, values comprising honesty, integrity, and freedom. Fifth, moral comprising keeping promises, respecting others, and being honest.

Analyzing of Data

Content validity shows that elements in a measurement instrument are relevant and representative of the measurement objective (Haynes et al., 1995). The items developed in this affective self-assessment instrument were validated by five expert judges. The experts were selected based on their fields, i.e., chemistry and educational research and evaluation. The experts assessed the instrument, scored the items, and delivered critics and suggestions to the items via the available comment column. The score from validators was then analyzed using the Aiken formula in equation 1 (Aiken, 1980). The valid criteria based on the Aiken table and suggestions from validators were then utilized to improve the affective self-assessment instrument. Content validity analysis aims to improve the instrument. Subsequently, it was subjected to a trial for construct validity.

$$V = \frac{\Sigma n_i}{n(c-1)} \text{ .... [1]}$$

Construct validity explains the tested items' consistency against the theories. It is mentioned by Azwar (2018) that construct validity proves the high correlation between measurement results from the items with the theoretic construct underlying the instrument arrangement. Construct validity emphasizes a logical analysis using statistical analysis results. Good validity for constructs or latent variables is achieved when (a) loading factor is ≥ 0.30 (Sarstedt & Mooi, 2014); (b) t-value from the loading factor is over the critical value (t-value ≥ 1.96) (Doll et al., 1994).

Reliability was calculated using composite reliability (CR) in equation 2 (Ghadi et al., 2012). Good construct reliability is achieved when it is ≥ 0.7 (Lance et al., 2006).

$$\rho_C = \frac{(\Sigma \lambda)^2}{(\Sigma \lambda)^2 + \Sigma var(\varepsilon)} \text{ .... [2]}$$

The confirmatory factor analysis result acquired the fit criteria, referring to the fulfillment of two from three fit criteria models, i.e., Root Mean Square Error of Approximation $(\text{RMSEA} \leq 0.8)$, $\text{p} - \text{value} > 0.05$, and $\text{Goodness of Fit Index} \geq 0.90$ (Suranto et al., 2014). Two from these three criteria show that the developed instrument model is fit to the data.

Categories of Affective Domain Ability

The affective domain ability was categorized using the ideal average and ideal standard deviation. The categories are presented in Table 3 (Mardapi, 2017).

Table 3. The Categories Affective Domain

Score	Category
$X \geq \bar{X} + 1.5SB_x$	Very Good
$\bar{X} + 1.5SB_x > X \geq \bar{X}$	Good
$\bar{X} > X \geq \bar{X} - 1.5SB_x$	Poor
$X < \bar{X} - 1.5SB_x$	Very Poor

Statistical Test

The statistical test employed to discover the difference of chemistry affective domain ability of high school students was one-way Anova used SPSS. F-test one-way Anova rely to the fulfillment by three criteria, i.e., samples were independent, identically distributes residuals are normal, and had equal variances between groups or homogeneous (Delacre et al., 2019). The significance level was 0.05. H_0 is accepted if $\text{Sig} > 0.05$ (Dewi et al., 2017) and $F_{calculation} > F_{table}$ (Mulyono et al., 2018).

Findings

Affective Self-Assessment Instrument

The developed affective self-assessment instrument has five aspects, i.e., attitude, self-esteem, interest, value, and moral. Each aspect has three indiactors and three question items. The instrument was presented using the Likert scale with three solution alternatives. The instrument quality was determined based on content validity, construct validity, and reliability.

Content Validity

The content validity verification used the Aiken method. The assessment was performed by two chemistry experts and five educational research and evaluation experts. The content validity analysis result shows that 15 items analyzed using the Aiken method obtained a validity range of $0.76 - 0.95$. This score was compared to Right-Tail Probabilities (ρ) for Selected Values of the Validity Coefficient (V). It was discovered that the rater number was seven and the highest score was three; therefore, the validity acceptance score was ≥ 0.76 (Aiken, 1985). Based on this result, all items developed in the study instrument were valid.

Construct Validity

The construct validity verification using the second-order CFA method is presented in Figure 2. Before discovering the construct validity result, a statistical analysis model verification was performed. Based on Figure 2, (a) $\text{RMSEA} = 0.068 \leq 0.8$, fulfilling the fit criteria, (b) $\text{p} - \text{value} = 0.00000 < 0.05$; not fulfilling the criteria, and (c) $\text{Goodness of Fit Index (GFI)} = 0.90 \geq 0.90$; fulfilling the fit criteria. From the three criteria, the model fit obtained two fulfilling criteria. Therefore, it can be concluded that the second-order CFA analysis model fulfilled the goodness of fit statistical criteria, allowing the analysis to continue.

Construct validity verification is portrayed in Figure 2 and Table 4. Figure 2 obtains the first proof where five aspects and 15 items had a $loading\ factor \geq 0.30$. Table 4 delivers the second proof that 15 items acquired a $t - value \geq 1.96$. Hence, it was discovered that 15 valid items were present to measure the chemistry affective ability of students.

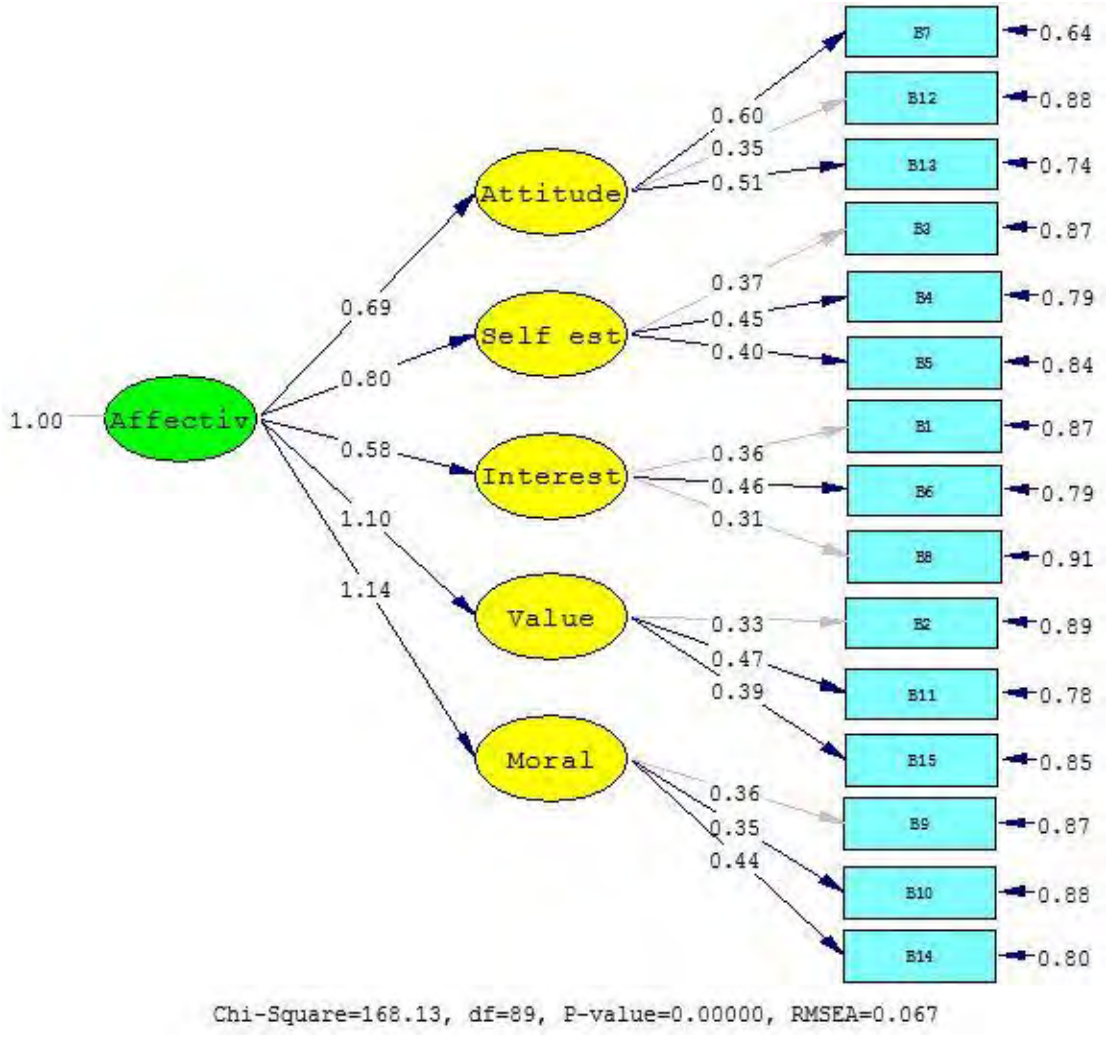

Figure 2. The Result of the Second Order-CFA Analysis of Affective Self-assessment Instrument

Note: B1 : Learning interest; B2 : Integrity; B3 : Family; B4 : Ethics-moral; B5 : Physique; B6 : Learning attention; B7 : Curiosity; B8 : Learning motivation; B9 : Being honest; B10 : Respecting others; B11 : Honesty; B12 : Respect for data and facts; B13 : Creativity and innovation; B14 : Keeping promises; B15 : Freedom

Table 4. The Result of the Second Order-CFA Analysis of Affective Self-assessment Instrument

Indicator	Loading Factor	t value	R^2	Notes
B1	0.36	-	0.13	*Reference var*
B2	0.33	-	0.11	*Reference var*
B3	0.37	-	0.13	*Reference var*
B4	0.45	3.42	0.21	*Indicator fit*
B5	0.40	3.26	0.16	*Indicator fit*
B6	0.46	3.19	0.21	*Indicator fit*
B7	0.60	4.66	0.36	*Reference var*
B8	0.31	-	0.095	*Reference var*
B9	0.36	-	0.13	*Reference var*
B10	0.35	3.41	0.12	*Indicator fit*
B11	0.47	3.63	0.22	*Indicator fit*
B12	0.35	-	0.12	*Reference var*
B13	0.51	4.55	0.26	*Indicator fit*
B14	0.44	3.85	0.20	*Indicator fit*
B15	0.39	3.35	0.15	*Indicator fit*

Note: *Reference var*: reference variable, t-value is not estimated, t-value target ≥ 1.96.

Reliability

Reliability verification using composite reliability (ρ_c) obtained $\rho_c = 0.75$. This score shows that the effective self-assessment instrument was reliable and could properly measure in different conditions. The reliability score of > 0.7 concludes that the measurement reliability is good (Lance et al., 2006).

Chemistry Affective Ability Trend of Students

The measurement process aimed to discover the affective domain ability trend of high school students based on gender. The chemistry affective domain measurement result of high school students in Yogyakarta Special Region is presented in Figure 3. Based on Figure 3, it was discovered that the chemistry affective domain of students was dominated by good and very good categories with percentages of 56% and 42%, respectively.

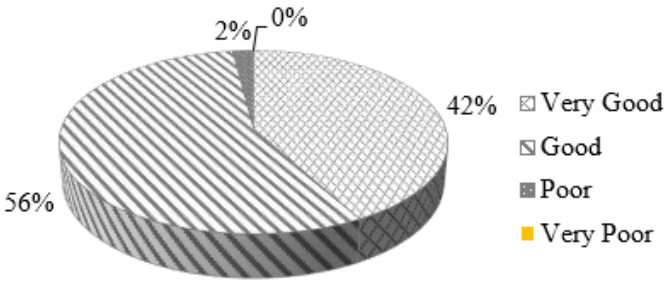

Figure 3. The Chemistry Affective Domain Measurement Result

Chemistry affective abilities of students in the attitude aspect are presented in Table 5. Table 5 shows that students had good attitudes in the chemistry learning process, marked by the very good category frequency being the most dominant. Therefore, it is concluded that the affective ability of students in the attitude aspect was very good.

Table 5. Chemistry affective of the attitude aspect

Score	Category	Frequency	Percentage (%)
$X \geq 7.5$	Very good	105	51
$7.5 > X \geq 6$	Good	83	41
$6 > X \geq 4.5$	Poor	10	5
$X < 4.5$	Very poor	7	3
Total		205	100

Chemistry affective abilities of students in the self-esteem aspect are presented in Table 6. Table 6 shows that students had good self-esteem in the chemistry learning process, marked by the good category frequency being the most dominant. Therefore, it is concluded that the affective ability of students in the self-esteem aspect was good.

Table 6. Chemistry Affective of the Self-esteem Aspect

Score	Category	Frequency	Percentage (%)
$X \geq 7.5$	Very good	76	37
$7.5 > X \geq 6$	Good	107	52
$6 > X \geq 4.5$	Poor	17	8
$X < 4.5$	Very poor	5	2
Total		205	100

Chemistry affective abilities of students in the interest aspect are presented in Table 7. Table 7 shows that students had good interests in the chemistry learning process, marked by the good category frequency being the most dominant. Therefore, it is concluded that the affective ability of students in the interest aspect was good.

Table 7. Chemistry Affective of the Interest Aspect

Score	Category	Frequency	Percentage (%)
$X \geq 7.5$	Very good	45	22
$7.5 > X \geq 6$	Good	123	60
$6 > X \geq 4.5$	Poor	24	12
$X < 4.5$	Very poor	13	6
Total		205	100

Chemistry affective abilities of students in the value aspect are presented in Table 8. Table 8 shows that students had good values in the chemistry learning process, marked by the very good category frequency being the most dominant. Therefore, it is concluded that the affective ability of students in the value aspect was very good.

Table 8. Chemistry Affective of the Value Aspect

Score	Category	Frequency	Percentage (%)
$X \geq 7.5$	Very good	170	83
$7.5 > X \geq 6$	Good	28	14
$6 > X \geq 4.5$	Poor	5	2
$X < 4.5$	Very poor	2	1
Total		205	100

Chemistry affective abilities of students in the moral aspect are presented in Table 9. Table 9 shows that students had good morals in the chemistry learning process, marked by the very good category frequency being the most dominant. Therefore, it is concluded that the affective ability of students in the moral aspect was very good.

Table 9. Chemistry Affective of the Moral Aspect

Score	Category	Frequency	Percentage (%)
$X \geq 7.5$	Very good	143	70
$7.5 > X \geq 6$	Good	57	28
$6 > X \geq 4.5$	Poor	5	2
$X < 4.5$	Very Poor	0	0
Total		205	100

Chemistry affective domain abilities of high school students based on gender are illustrated in Figure 4. Figure 4 demonstrates that Chemistry affective abilities of high school students based on gender had different abilities. It is demonstrated by the "good" affective ability category domination in male students and the "very good" category domination in female students. This difference was obtained from one-way Anova. The data must fulfillment three assumpsion, i.e., 1) the data collection technique used simple random sampling showed that the data were independent (Wong, 2015), 2) normality indicates that the two data are normally distributed, indicated by the $\text{Sig. Kolmogorov} - \text{Smirnov}\ (boy = 0.716\ \&\ girl = 0.060) > 0.05$ (Sari & Hidayat, 2019), and 3) the chemistry affective domain ability data had $Sig.\ Lavene\ Statistic = 0.230 > 0.05$, which can be concluded to have homogeneous variance (Akbarini et al. al., 2018). The result of an $F(12.218) > 3.89$ (Mulyono et al., 2018) and $\text{Sig.}\ (0.001) < 0.05$ (Dewi et al., 2017), concluding a statistically significant difference of Chemistry affective domain abilities in male and female students.

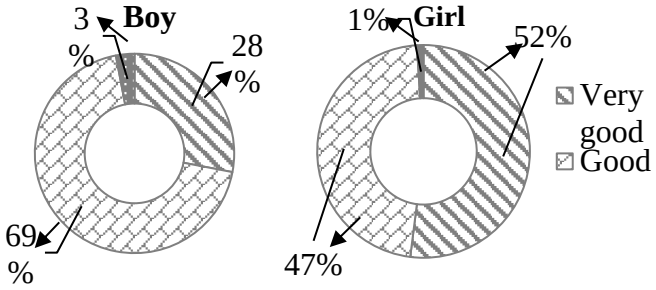

Figure 4. Chemistry Affective Domain Abilities of High School Students Based on Gender

Discussion

This study aimed to develop a affective self-assessment instrument of chemistry for high school student with known quality and the chemistry affective domain ability trend of high school students. The chemistry affective was measured from attitude, self-esteem, interest, value, and moral. The chemistry affective instrument quality was observed from content validity, construct validity, and reliability. Content validity using the Aiken index obtain a range of $0.76 - 0.95$. It shows that all items in the Affective self-assessment instrument of chemistry were valid with a content validity value of ≥ 0.76 (Aiken, 1985). Before discovering the construct validity result, a statistical analysis model verification was performed using confirmatory factor analysis. The fit test obtained $(\text{RMSEA} = 0.067) < 0.8$, $(\text{p} - \text{value} = 0.000) < 0.0.5$, and $(\text{Goodness of fit ststistic} = 0.90) > 0.90$. The fit test revealed that one requirement, i.e., $\text{p} - \text{value}$, did not fulfill the fit criteria. Hence, it fulfill two from three criteria (Suranto et al., 2014). Construct validity shows that 15 items fulfilled the valid criteria with a $\text{loading factor} \geq 0.30$ (Sarstedt & Mooi, 2014) and $\text{t} - \text{value} \geq 1.96$ (Doll et al., 1994). Reliability $(\rho_c = 0.75) > 0.7$ shows good reliability (Lance et al., 2006). Based on the analysis result of the instrument quality, 15 items could be used to measure chemistry affective ability. The analysis process follows Hidayati et al. (2021), in which standard instruments are prepared through a validation process by expert-judgment to be refined, followed by an empirical validation process to determine the validity and reliability of the developed instrument. The existence of an assessment process will improve students' attitudes towards chemistry learning (A'izzah et al., 2017). In addition, another study revealed a relationship between affective and cognitive, where students with good affective abilities tend to have good cognitive abilities (Rosa, 2015). It shows that affective assessment with a good quality instrument can provide benefits for evaluating the cognitive domain.

Chemistry affective abilities were dominated in the good and very good categories, while the chemistry affective domain ability trend between male and female students was statistically and significantly different. It is shown by the analysis result using one-way Anova with $\text{F}(12.218) > 3.89$ and $\text{Sig.}\,(0.001 < 0.05)$ (Mulyono et al., 2018). It is discovered that the affective domain abilities of male students were dominated by the "good" category while female students were dominated by the "very good" category. High ability is an essential variable in learning since it determines the extent of success in learning (Istiyono et al., 2020). It illustrates that the affective domain learning indicators are successfully demonstrated by the dominance of the "good" and "very good" categories. The results of this study are supported by a study showing that male and female students have different beliefs in the learning process (Rojo Robas et al., 2018). Another study found evidence that the affective abilities of male and female students were different (Adenzato et al., 2017). The results of other studies showed gender differences in students' attitudes towards science (Wan & Lee, 2017).

In studies that focus or include the affective domain, some researchers included gender as a variable of interest. Kavousi et al., (2017) showed that the differences in affective domain between male and female students, preferences of senior high school students indicated that there were differences in the level of education and ability. Meanwhile, male students have more physical activity than female students. Ayub et al., (2017) showed that female students better than boys in all domains (affective, cognitive, psychomotor). Female students have a higher average mathematical engagement than male students. High school students are teenagers, Herman-Giddens (2006) mentions that the onset of puberty is associated with changes in affective domains such as increased emotional intensity, sensation seeking, risk taking, sleep/arousal regulation, and affective disorders. Based on this description, it can be seen that there are several possibilities that affect the difference in affective domains between male and female students.

Conclusion

This study instrument was developed to measure the chemistry affective domain of high school students including attitudes, self-esteem, interests, values, and morals. The results showed 15 valid and reliable items to be used to measure the chemistry affective. Furthermore, it is expected that this affective self-assessment instrument of chemistry can be used by teachers in assessing the learning process. The affective domain ability trend of students was statistically and significantly different between male and female students. The chemistry affective domain ability of male students was dominated by the "good" category while the "very good" category dominated in female students.

Recommendations

Referring to the study results that the affective domain assessment process is critical to be carried out using instruments with known quality, this affective domain self-assessment instrument can be used to measure the affective domain in the chemistry learning process in schools. Affective aspects that are measured are attitudes, self-esteem, interests, values, and morals. The affective domain trend shows differences between male and female students; hence, is possible to conduct an affective domain study with other affecting factors based on current technological developments.

Limitations

The study had two major limitations. First, the research was conducted online so that only some students filled out this instrument at one school. Therefore, several schools are needed to assess chemistry affective abilities of students. Second, the ability estimation is carried out only in class XI science class and does not cover all levels class.

References

Abadi, T. W. (2016). Aksiologi: Antara etika, moral, dan estetika [Axiology: Between ethics, morals, and aesthetics]. *Kanal: Jurnal Ilmu Komunikasi*, *4*(2), 187-204. https://doi.org/10.21070/kanal.v4i2.1452

Adenzato, M., Brambilla, M., Manenti, R., De Lucia, L., Trojano, L., Garofalo, S., Enrici, I., & Cotelli, M. (2017). Gender differences in cognitive Theory of Mind revealed by transcranial direct current stimulation on medial prefrontal cortex. *Scientific Reports*, *7*(1), 1-9. https://doi.org/10.1038/srep41219

Aiken, L. R. (1980). Attitude measurement and research. *New directions for testing and measurement*, *7*, 1-24.

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

A'izzah, A. A., Susialaningsih, E., & Sumarti, S. S. (2017). Pengembangan instrumen penilaian (attitude toward chemistry) dengan teknik peer dan self assessment siswa SMA N 2 Salatiga [Development of assessment instruments (attitude toward chemistry) with peer techniques and self-assessment for students of SMA N 2 Salatiga]. *Chemistry in Education*, *6*(2), 29-34. https://bit.ly/3lkg0xs

Akbarini, N. R., Murtini, W., & Rahmanto, A. N. (2018). The effect of Lectora inspire-based interactive learning media in vocational high school. *Jurnal Pendidikan Vokasi*, *8*(1), 78-87. https://doi.org/10.21831/jpv.v8i1.17970

Andersen, E. B. (1980). *Discrete statistical models with social science applications* (No. 04; HA29, A53.). North Holland Publishing Company.

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology*, *46*(2), 294. https://doi.org/10.1037/0022-3514.46.2.294

Anderson, L. W., & Bourke, S. F. (2013). *Assessing affective characteristics in the schools*. Routledge. https://doi.org/10.4324/9781410605443

Ayub, A. F. M., Yunus, A. S. M., Mahmud, R., Salim, N. R., & Sulaiman, T. (2017). Differences in students' mathematics engagement between gender and between rural and urban schools. *AIP Conference Proceedings 1795*, 1–6. https://doi.org/10.1063/1.4972169

Azwar, S. (2018). *Reliabilitas dan validitas* [Reliability and validity]. Pustaka Pelajar.

Beltrán-Pellicer, P., & Godino, J. D. (2020). An onto-semiotic approach to the analysis of the affective domain in mathematics education. *Cambridge Journal of Education*, *50*(1), 1-20. https://doi.org/10.1080/0305764X.2019.1623175

Bhardwaj, A. (2016). Importance of education in human life: A holistic approach. *International Journal of Science and Consciousness*, *2*(2), 23-28. https://bit.ly/3rstOcQ

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Handbook I: cognitive domain*. David McKay.

Chen, A., & Wang, Y. (2017). The role of interest in physical education: A review of research evidence. *Journal of Teaching in Physical Education*, *36*(3), 313-322. https://doi.org/10.1123/jtpe.2017-0033

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open*, *2*(4). https://doi.org/10.1177/2332858416673617

Delacre, M., Leys, C., Mora, Y. L., & Lakens, D. (2019). Taking parametric assumptions seriously: Arguments for the use of Welch's F-test instead of the classical F-test in one-way ANOVA. *International Review of Social Psychology*, *32*(1), 1-12. http://doi.org/10.5334/irsp.198

Dewi, I. P. M., Suryadarma, I. G. P., Wilujeng, I., & Wahyuningsih, S. (2017). The effect of science learning integrated with local potential of wood carving and pottery towards the junior high school students' critical thinking skills. *Journal of Indonesia Science Education/ Jurnal Pendidikan IPA Indonesia*, *6*(1), 103-109. https://doi.org/10.15294/jpii.v6i1.9598

Doll, W. J., Xia, W., & Torkzadeh, G. (1994). A confirmatory factor analysis of the end-user computing satisfaction instrument. *MIS Quarterly*, *18*(4), 453-461. https://doi.org/10.2307/249524

Fitts, W. H., & Roid, G. H. (1964). *Tennessee self concept scale*. Counselor Recordings and Tests.

Ghadi, I., Alwi, N. H., Bakar, K. A., & Talib, O. (2012). Construct validity examination of critical thinking dispositions for undergraduate students in University Putra Malaysia. *Higher Education Studies*, *2*(2), 138-145. https://.doi.org/10.5539/hes.v2n2p138

Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest matters: The importance of promoting interest in education. *Policy Insights from the Behavioral and Brain Sciences*, *3*(2), 220-227. https://doi.org/10.1177/2372732216655542

Harlen, W. (1985). *Teaching and learning primary science*. Teachers College Press.

Hart, L. E. (1989). Describing the affective domain: Saying what we mean. In D. B. McLeod & V. M. Adams. *Affect and Mathematical Problem Solving* (pp. 37-45). Springer. https://doi.org/10.1007/978-1-4612-3614-6_3

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238-247. https://doi.org/10.1037/1040-3590.7.3.238

Herman-Giddens, M. (2006). Recent data on pubertal milestones in United States children: the secular trend toward earlier development. *International Journal of Andrology, 29, 241–246.* https://doi.org/10.1111/j.1365-2605.2005.00575.x

Hidayati, H. N., Hayat, B., & Rahayu, W. (2021). Assessment of the validity and reliability of mental health instruments of high school student in Indonesia. *European Journal of Educational Research*, *10*(2), 729-742. https://doi.org/10.12973/eu-jer.10.2.729

Hosseini, S. N., Alavijeh, M. M., Matin, B. K., Hamzeh, B., Ashtarian, H., & Jalilian, F. (2016). Locus of control or self-esteem; Which one is the best predictor of academic achievement in Iranian college students. *Iranian Journal of Psychiatry and Behavioral Sciences*, *10*(1), 1-6. https://doi.org/10.17795/ijpbs-2602

Hudi, I. (2017). Pengaruh pengetahuan moral terhadap perilaku moral pada siswa SMP Negeri Kota Pekan Baru berdasarkan pendidikan orangtua [The influence of moral knowledge on moral behavior in public junior high school students in pekanbaru city based on parents' education]. *Jurnal Moral Kemasyarakatan*, *2*(1), 30-44. https://bit.ly/3EOmb4s

Istiyono, E., Dwandaru, W. S. B., Setiawan, R., & Megawati, I. (2020). Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use. *European Journal of Educational Research*, *9*(1), 91-101. https://doi.org/10.12973/eu-jer.9.1.91

Kavousi, E., Khazaei, Z., Amini, A., Fattahi, E., Panahi, A., Sohrabivafa, M., Beiranvand, R., & Goodarzi, E. (2017). Promoting behaviors of healthiness in two domains of physical activity and nutrition statue in high school students. *International Journal of Pediatrics*, *5*(5), 4839-4847. https://doi.org/10.22038/ijp.2017.21416.1797

Kokom, K., & Edwina, T. N. (2020). Relationship between social support families with meaningfulness of life in the elderly in rural Kebandungan. *Jurnal Psikologi Talenta*, *5*(2), 87-96. https://doi.org/10.26858/talenta.v5i2.9859

Kusumaningtyas, P., Yusvitasari, R. E., & Majid, A. (2018). Pengembangan instrumen penilaian kinerja untuk mengukur kompetensi siswa dalam kegiatan praktikum kimia di SMA/K [Development of performance assessment instruments to measure student competence in chemistry practicum activities at SMA/K]. *Jurnal Inovasi Pendidikan Kimia*, *12*(2), 2128-2136.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational research methods*, *9*(2), 202-220. https://doi.org/10.1177/1094428105284919

Li, C. H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936-949. https://doi.org/10.3758/s13428-015-0619-7

Lilleholt, L. (2019). Cognitive ability and risk aversion: A systematic review and meta analysis. *Judgment and Decision Making*, *14*(3), 234-279.

Mardapi, D. (2012). *Teknik penyusunan instrumen tes dan non tes* [Techniques for preparing test and non-test instruments]. Mitra Cendekia.

Mardapi, D. (2017). *Pengukuran, penilaian dan evaluasi pendidikan* [Measurement, assessment and evaluation of education]. Parama Publishing.

Mardapi, D. (2018). *Teknik penyusunan instrumen tes dan non tes* [Techniques for preparing test and non-test instruments]. Parama Publishing.

Marshman, E., Kalender, Z. Y., Schunn, C., Nokes-Malach, T., & Singh, C. (2018). A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics*, *96*(4), 391-405. https://doi.org/10.1139/cjp-2017-0185

Mulyono, D., Asmawi, M., & Nuriah, T. (2018). The effect of reciprocal teaching, student facilitator and explaining and learning independence on mathematical learning results by controlling the initial ability of students. *International Electronic Journal of Mathematics Education*, *13*(3), 199-205. https://doi.org/10.12973/iejme/3838

Muslich, K. M. (2014). Pengembangan model assessment afektif berbasis self assessment dan peer assessment di SMA Negeri 1 [Development of the self-assessment and peer assessment based affective assessment model at SMA Negeri 1]. *Jurnal Kebijakan dan Pengembangan Pendidikan*, *2*(2), 143-148. https://bit.ly/3xH3mxu

Ramma, Y., Bholoa, A., Watts, M., & Nadal, P. S. (2018). Teaching and learning physics using technology: Making a case for the affective domain. *Education Inquiry*, *9*(2), 210-236. https://doi.org/10.1080/20004508.2017.1343606

Riscaputantri, A., & Wening, S. (2018). Pengembangan instrumen penilaian afektif siswa kelas IV sekolah dasar di Kabupaten Klaten [Development of affective assessment instruments for fourth grade elementary school students in Klaten District]. *Jurnal Penelitian dan Evaluasi Pendidikan*, *22*(2), 231-242. https://doi.org/10.21831/pep.v22i2.16885

Rojo Robas, V., Villarroel Villamor, J. D., & Madariaga Orbea, J. M. (2018). The affective domain in learning mathematics according to students' gender. *Revista Latinoamericana de Investigación en Matemática Educative*, *21*(2), 183-202. https://doi.org/10.12802/relime.18.2123

Rosa, F. O. (2015). Analisis kemampuan siswa kelas X pada ranah kognitif, afektif dan psikomotorik [Analysis of the ability of class X students in the cognitive, affective and psychomotor domains]. *Omega: Jurnal Fisika dan Pendidikan Fisika*, *1*(2), 24-28.

Sari, V. T. A., & Hidayat, W. (2019). The students' mathematical critical and creative thinking ability in double-loop problem solving learning. In *Journal of Physics: Conference Series,* 1315(1), 012024. https://doi.org/10.1088/1742-6596/1315/1/012024

Sarstedt, M., & Mooi, E. (2014). A concise guide to market research. *The Process, data, and methods using IBM SPSS statistics*, Springer. https://doi.org/10.1007/978-3-642-53965-7

Saxon, D. P., Levine-Brown, P., & Boylan, H. R. (2008). Affective assessment for developmental students, part 1. *Research in Developmental Education*, *22*(1), 1-4. https://bit.ly/3I9lkgJ

Setiadi, H. (2016). Pelaksanaan penilaian pada Kurikulum 2013 [Implementation of the assessment in the 2013 Curriculum]. *Jurnal Penelitian dan Evaluasi Pendidikan /Journal of Educational Research and Evaluation*, *20*(2), 166-178. https://dx.doi.org/10.21831/pep.v20i2.7173

Setiawan, A., & Mardapi, D. (2019). The development of instrument for assessing students' affective domain using self- and peer-assessment models. *International Journal of Instruction*, *12*(3), 425-438. https://doi.org/10.29333/iji.2019.12326a

Setiawan, D. A., VH, E. S., & Mulyani, S. (2013). Prestasi belajar dengan pembelajaran kooperatif tipe numbered heads together (NHT) lebih tinggi dari pada think-pair-share (TPS) pada materi pelajaran tata nama senyawa kimia dan persamaan reaksi kimia [Learning achievement with cooperative learning type numbered heads together (NHT) is higher than think-pair-share (TPS) in the subject matter of chemical compound nomenclature and chemical reaction equations]. *Jurnal Pendidikan Kimia,* (4), 11-18. https://bit.ly/3lheYlE

Slameto. (2010). *Belajar & Faktor-faktor yang Mempengaruhi* [Learning & Influencing Factors]. Rineka Cipta.

Suranto, S., Muhyadi, M., & Mardapi, D. (2014). Pengembangan instrumen evaluasi uji kompetensi keahlian (UKK) administrasi perkantoran di SMK [Development of an evaluation instrument for the skill competency test (UKK) for office administration in SMK]. *Jurnal Penelitian dan Evaluasi Pendidikan, 18*(1), 98-114. https://doi.org/10.21831/pep.v18i1.2127

Syamsudin, A., Budiyono, B., & Sutrisno, S. (2016). Model of affective assessment of primary school students. *REiD Research and Evaluation in Education*, *2*(1), 25-41. https://doi.org/10.21831/reid.v2i1.8307

Tyler, R. W. (1973). Assessing educational achievement in the affective domain. *NCME Measurement in Education*, *4,* 1-8. https://bit.ly/3rjFzT1

Wan, Z. H., & Lee, J. C. K. (2017). Hong Kong secondary school students' attitudes towards science: A study of structural models and gender differences. *International Journal of Science Education*, *39*(5), 507-527. https://doi.org/10.1080/09500693.2017.1292015

Wiwin, E., & Kustijono, R. (2018, March). The use of physics practicum to train science process skills and its effect on scientific attitude of vocational high school students. *Journal of Physics, 997*(1), 12040. https://doi.org/10.1088/1742-6596/997/1/012040

Wong, T. T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, *48*(9), 2839-2846. https://doi.org/10.1016/j.patcog.2015.03.009